

Mining, Visualizing, and Accelerating Science and Technology

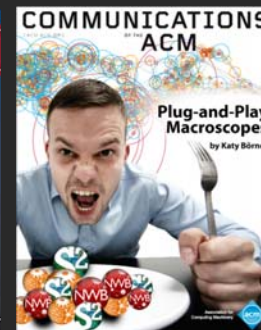
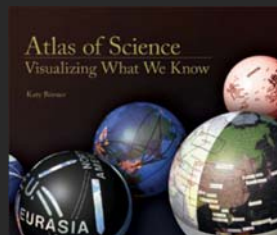
Katy Börner

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



With special thanks to the members at the Cyberinfrastructure for Network Science Center; the Sci2, NWB, and EpiC teams; and the VIVO Collaboration

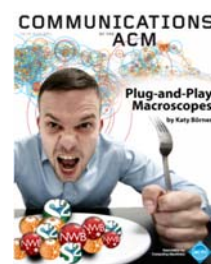
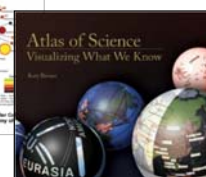
College of Information Sciences and Technology
The Pennsylvania State University
University Park, PA



April 26, 2012

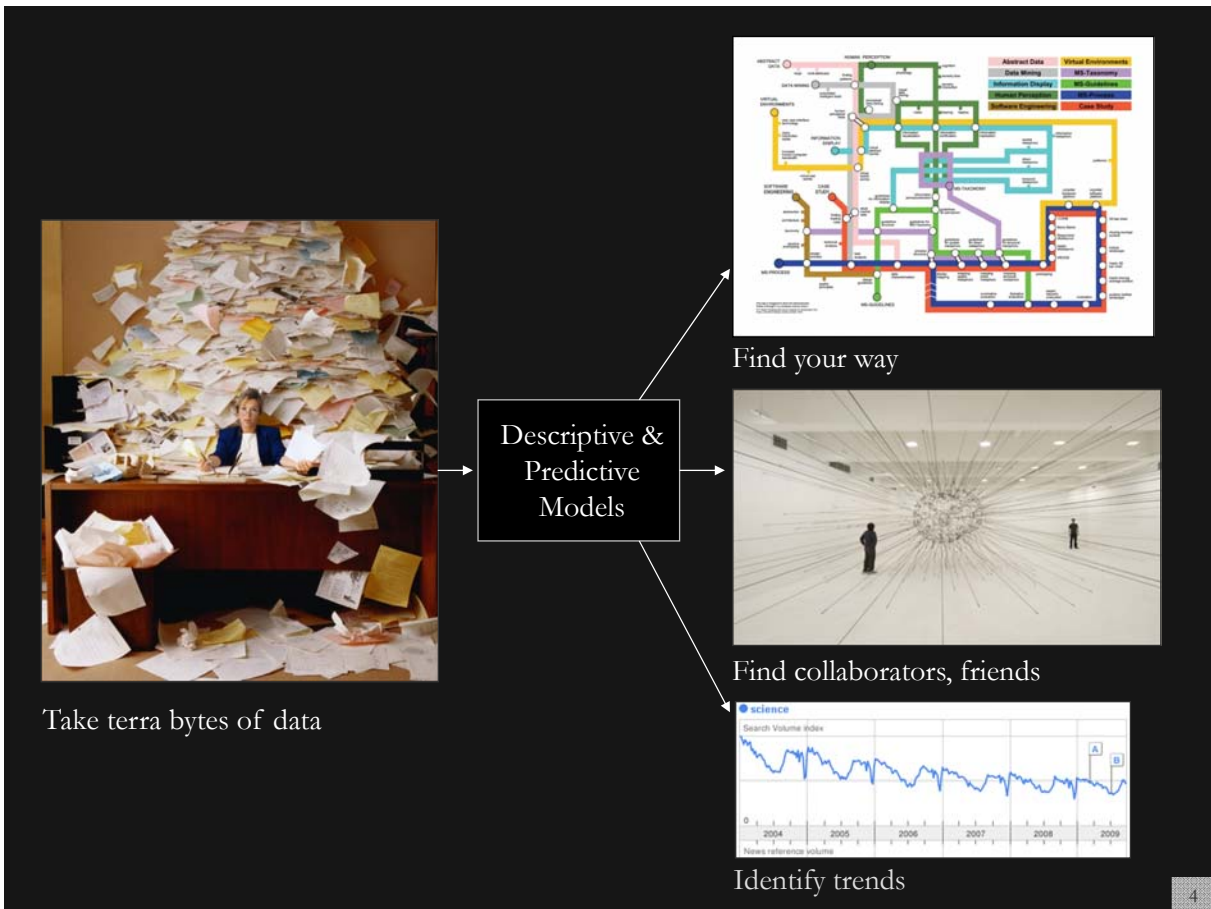
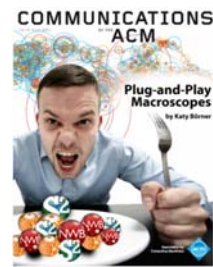
Overview

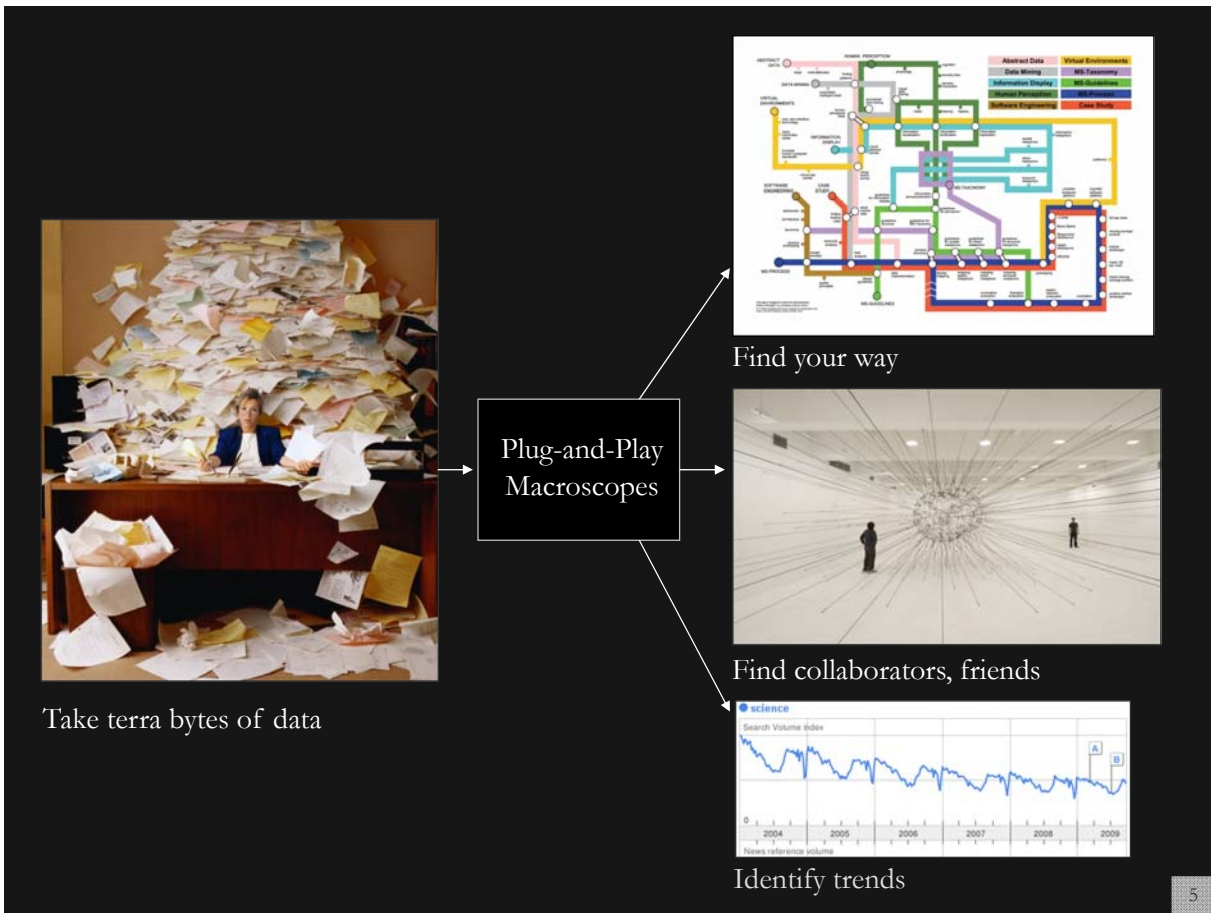
1. **Data mining and visualization research** that aims to increase our scientific understanding of the structure and dynamics of science and technology.
2. **Novel approaches and services** that improve information access, researcher networking, and research management.
3. **Data services and plug-and-play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





Type of Analysis vs. Level of Analysis

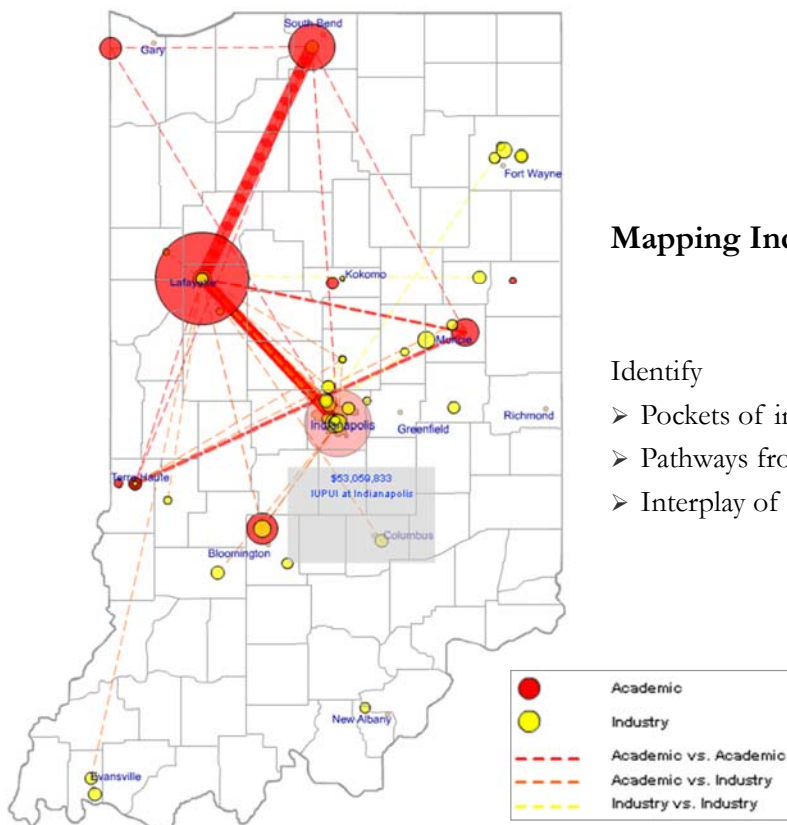
	<i>Micro/Individual (1-100 records)</i>	<i>Meso/Local (101-10,000 records)</i>	<i>Macro/Global (10,000 < records)</i>
Statistical Analysis/Profiling	Individual person and their expertise profiles	Larger labs, centers, universities, research domains, or states	All of NSF, all of USA, all of science.
Temporal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

Type of Analysis vs. Level of Analysis

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



7



Mapping Indiana's Intellectual Space

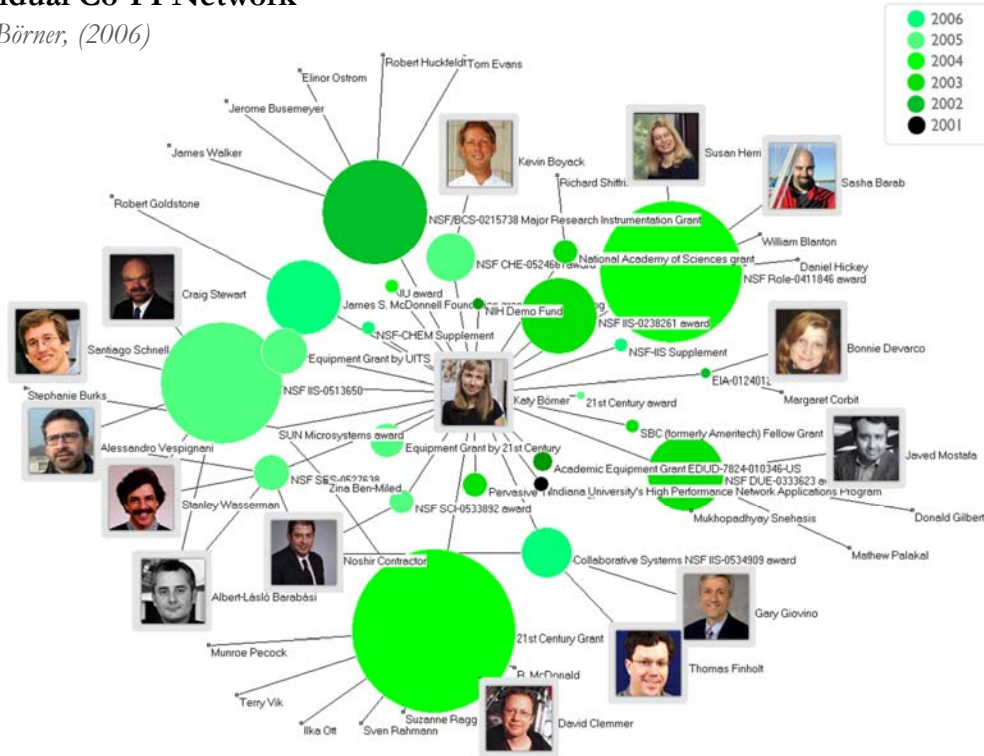
Identify

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

8

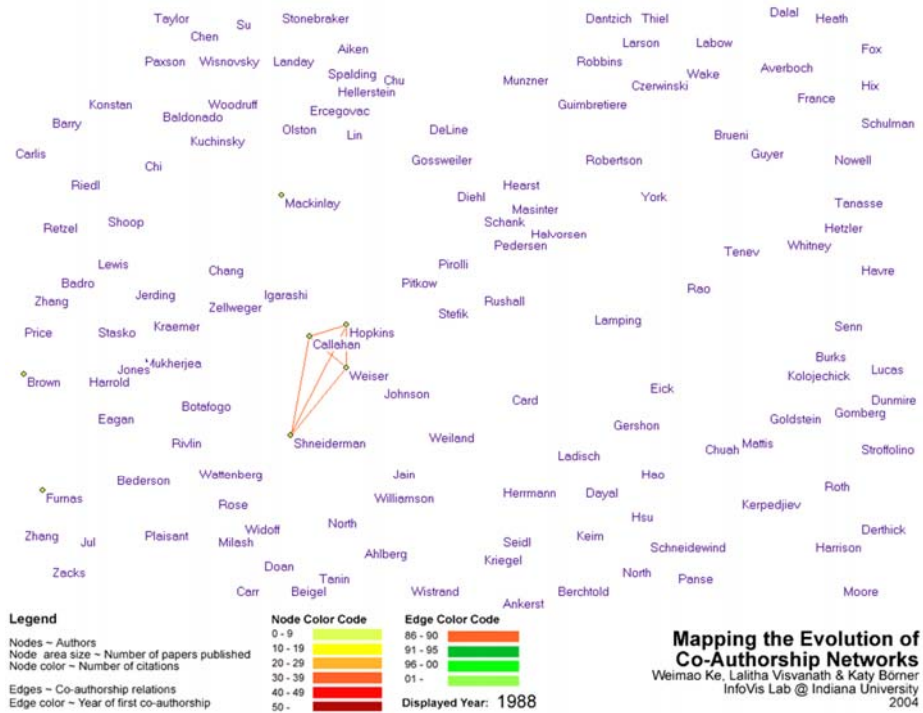
Individual Co-PI Network

Ke & Börner, (2006)



Mapping the Evolution of Co-Authorship Networks

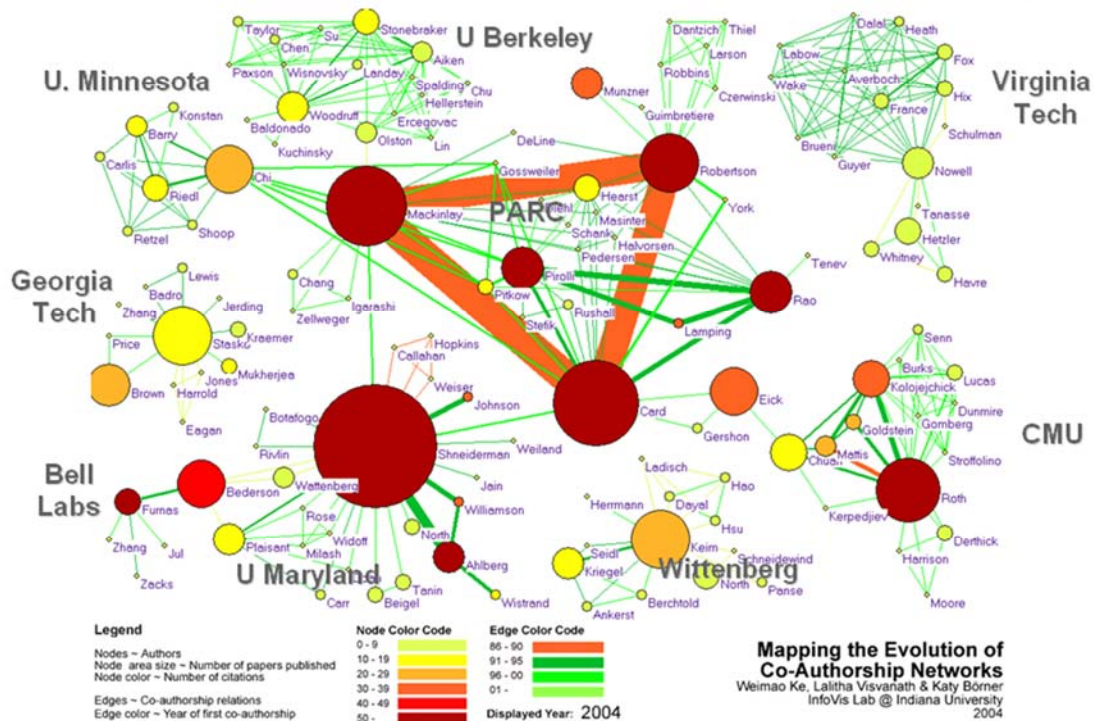
Ke, Visvanath & Börner, (2004) Won 1st price at the IEEE InfoVis Contest.



Mapping the Evolution of Co-Authorship Networks
 Weimao Ke, Lalitha Visvanath & Katy Börner
 InfoVis Lab @ Indiana University
 2004

Mapping the Evolution of Co-Authorship Networks

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



11

Studying the Emerging Global Brain: Analyzing and Visualizing the Impact of Co-Authorship Teams

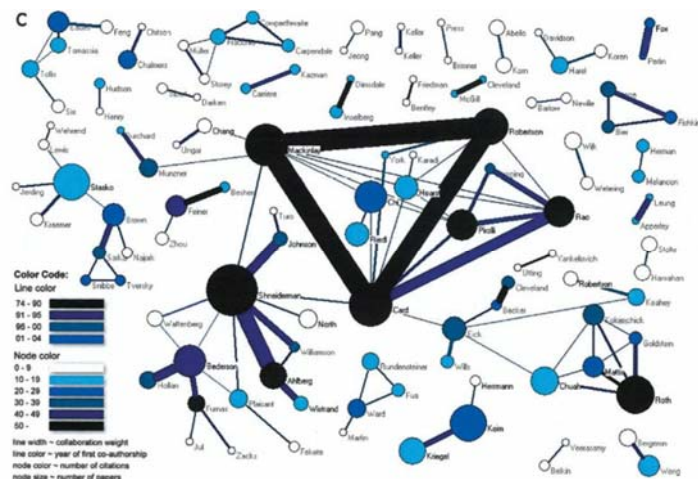
Börner, Dall'Asta, Ke & Vespignani (2005)
Complexity, 10(4):58-67.

Research question:

- Is science driven by prolific single experts or by high-impact co-authorship teams?

Contributions:

- New approach to allocate citational credit.
- Novel weighted graph representation.
- Visualization of the growth of weighted co-author network.
- Centrality measures to identify author impact.
- Global statistical analysis of paper production and citations in correlation with co-authorship team size over time.
- Local, author-centered entropy measure.



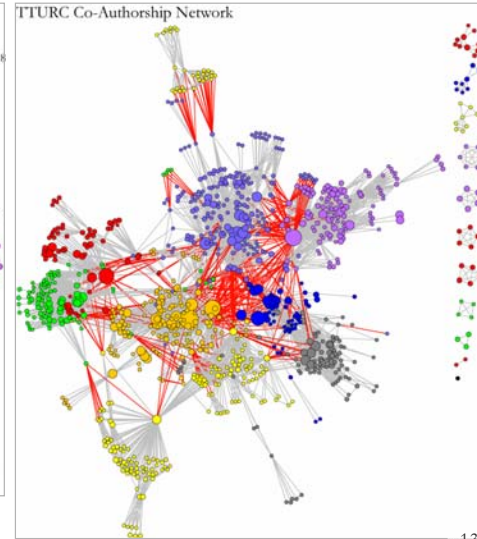
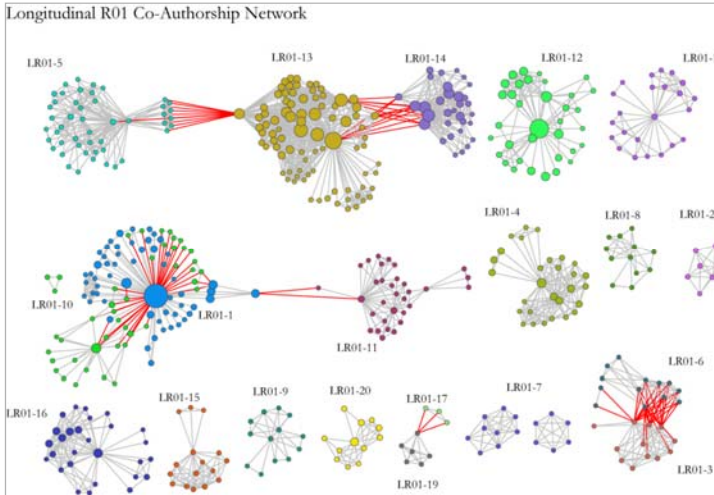
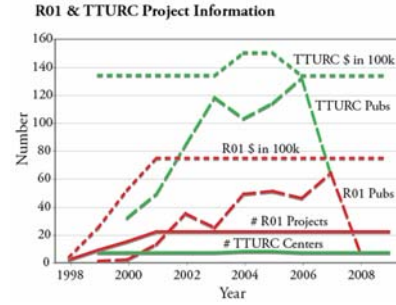
12

Mapping Transdisciplinary Tobacco Use Research Centers Publications

Compare R01 investigator based funding with TTURC Center awards in terms of number of publications and evolving co-author networks.

Zoss & Börner, *forthcoming*.

Supported by NIH/NCI Contract HHSN261200800812

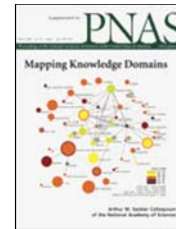
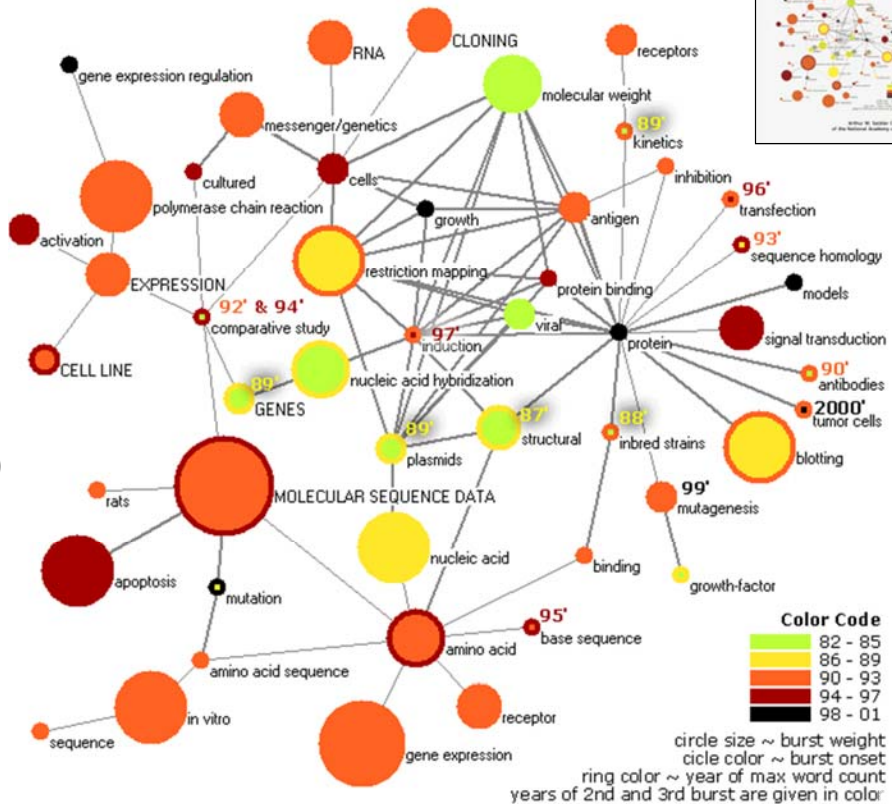


13

Mapping Topic Bursts

Co-word space of the top 50 highly frequent and bursty words used in the top 10% most highly cited PNAS publications in 1982-2001.

Mane & Börner. (2004)
PNAS, 101(Suppl. 1):
5287-5290.



14

From Spatial Proximity to Semantic Coherence: A Quantitative Approach to the Study of Group Dynamics in Collaborative Virtual Environments

Chen, Chaomei, and Börner, Katy. 2005. *PRESENCE: Teleoperators and Virtual Environments, Special Issue on Collaborative Information Visualization Environments* 14 (1): 81-103.



Questions

When, where, and why do users interact when exploring information spaces?

How does the spatial configuration of an information space correlate with spatial, semantic, and social navigation?

Major Contributions

Conceptualization and quantification of group coherence.

Novel visualizations to communicate results.

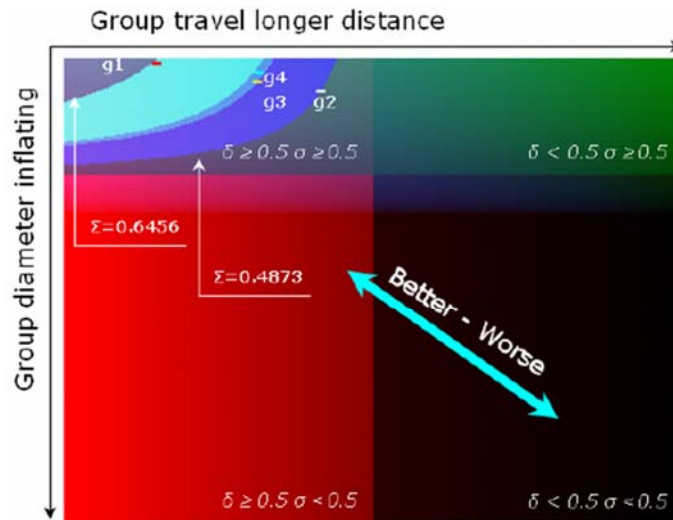


Figure 1. Group Coherence Space, colored by the different measures, showing the positions of four groups

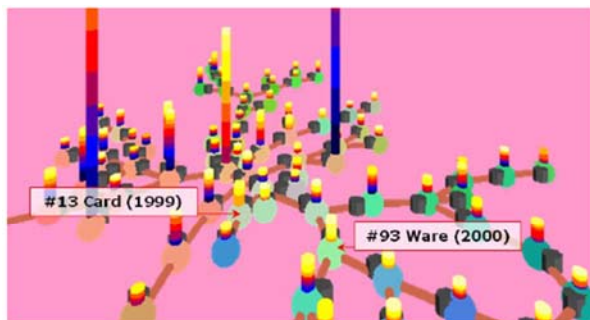


Figure 2. A visualization model of citation patterns.

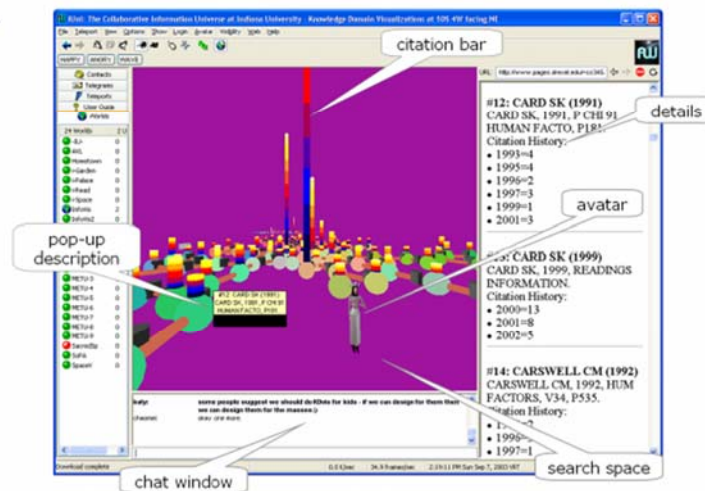


Figure 3. The user interface of the collaborative information visualization environment. 18

I think we've got the general themes question pretty close
yeah the colors seem to cover date ranges
go to dataset #2?

I think so, I wrote down part of what we discussed

ok

try and to the same so one of us may get it close
since they will be slightly different

who's this?

probably a bot for monitoring

shakes hands

or another user logged in to watch the fun

ok

more tuft

got it (I figured he'd be most cited again)

83

84

82,83,84

yep

yep

most cited seems pretty clear

looks like 82 and 83

this tallest one is tuft 82

oops, not published in 2000 though

but the publish dates are wrong

yep

nodes further from center are more recent

k

no that's not a trend

k, common themes seem to be interaction design and analysis

any idea about the most recently cited and published?

no, i've been looking for that the whole time

it's a bit difficult to grasp the patterns

seems like it shouldn't

i found a 1999

#18

good

is it near other recent publications

here



InfoVis Group 3

S5 - blue

S6 - red

Semantically indented chat sequences aim to reveal underlying connections between users' navigational movements and the contents of their discourse in a situated manner.

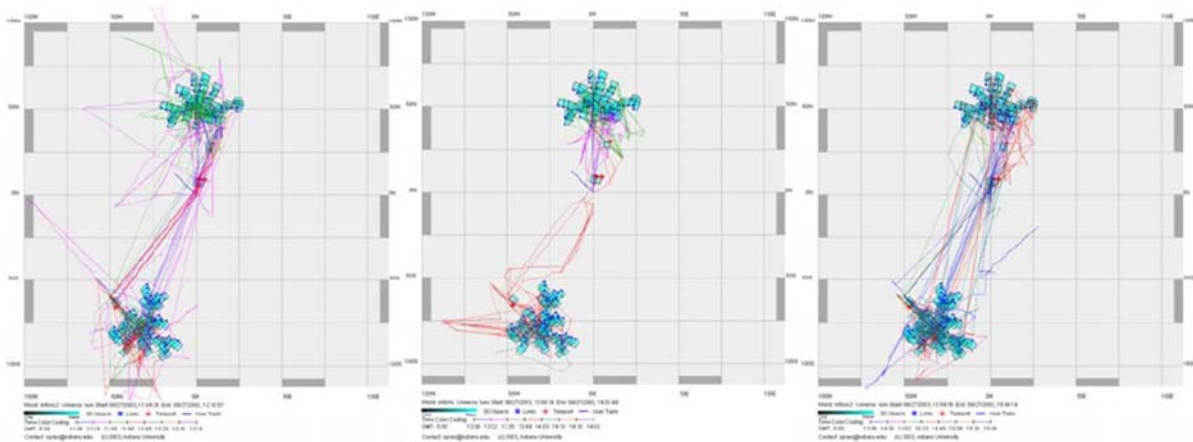


Figure 7. Group activity maps: Group 2 (left), Group 3 (middle), and Group 4 (right).

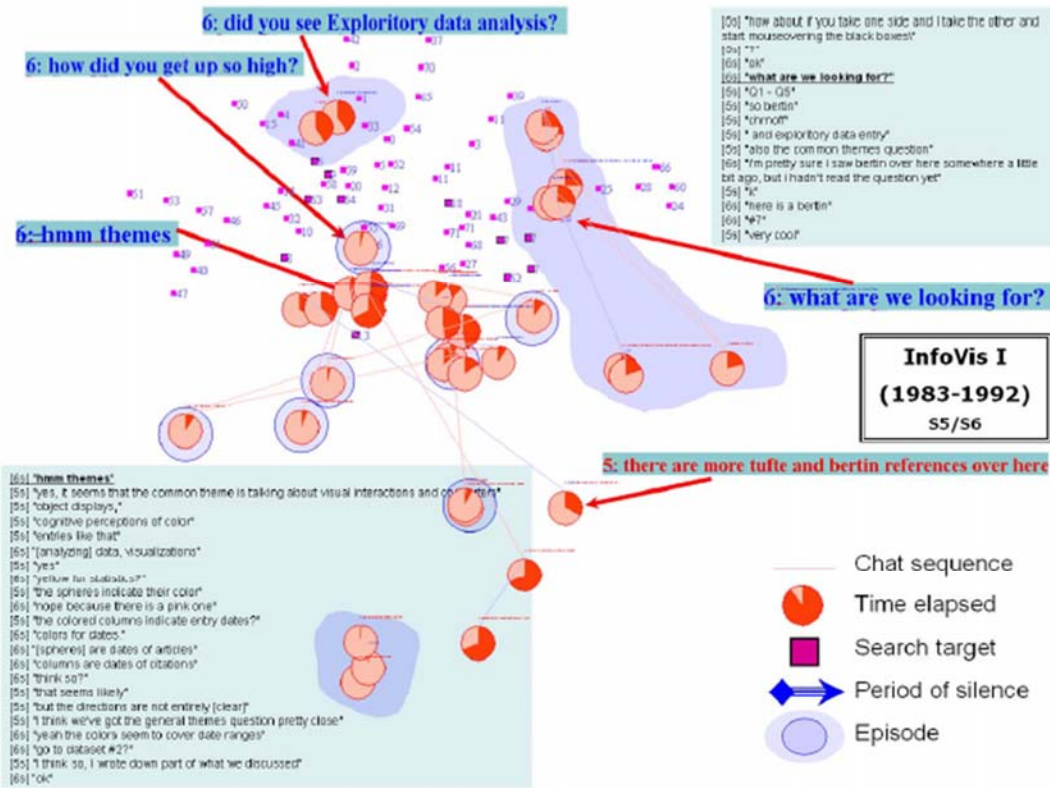


Figure 10. Episodes of collaborative search by Group 3.

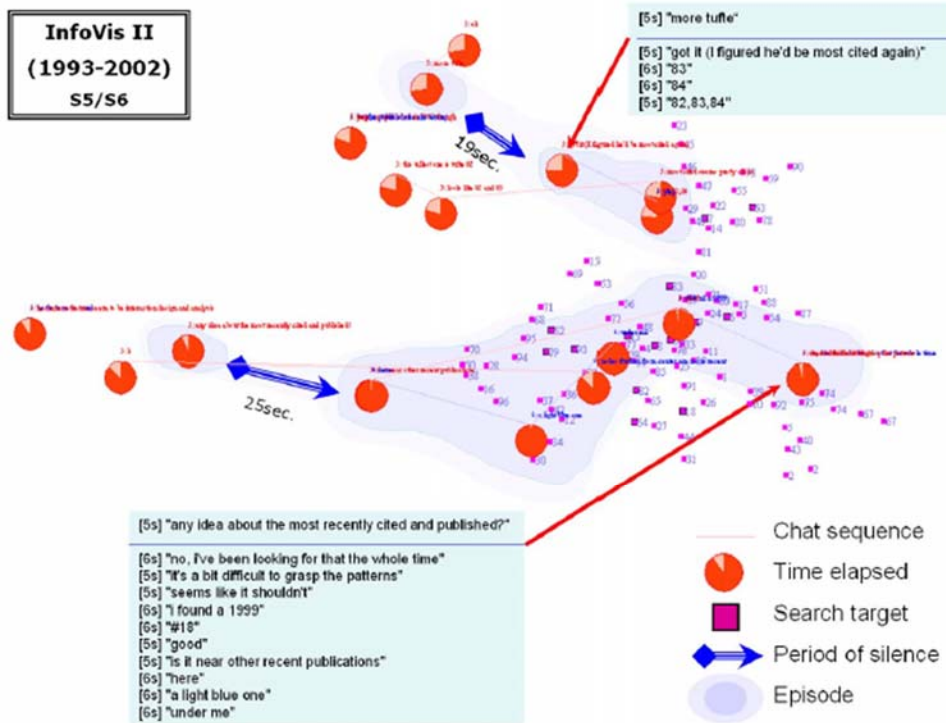


Figure 11. Episodes containing extended intervals of silence while group members being engaged in individual search.

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/

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.

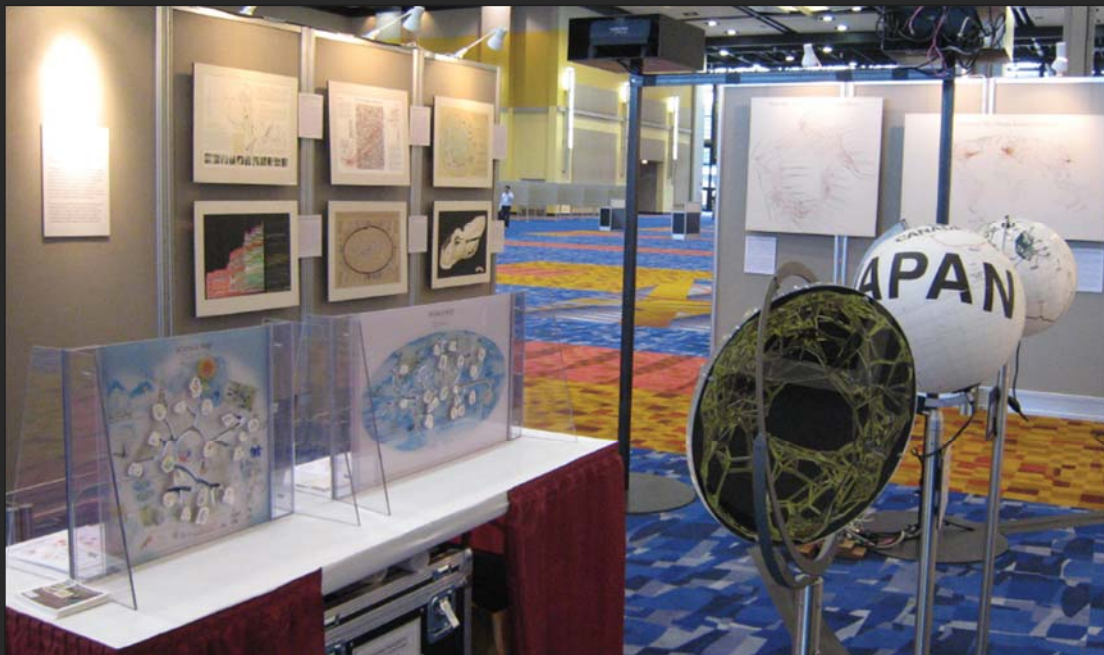


Special Issue on
“Collaborative
Information
Visualization
Environments”, MIT
Press, 14(1).



Mapping Science Exhibit – 10 Iterations in 10 years

<http://scimaps.org/>

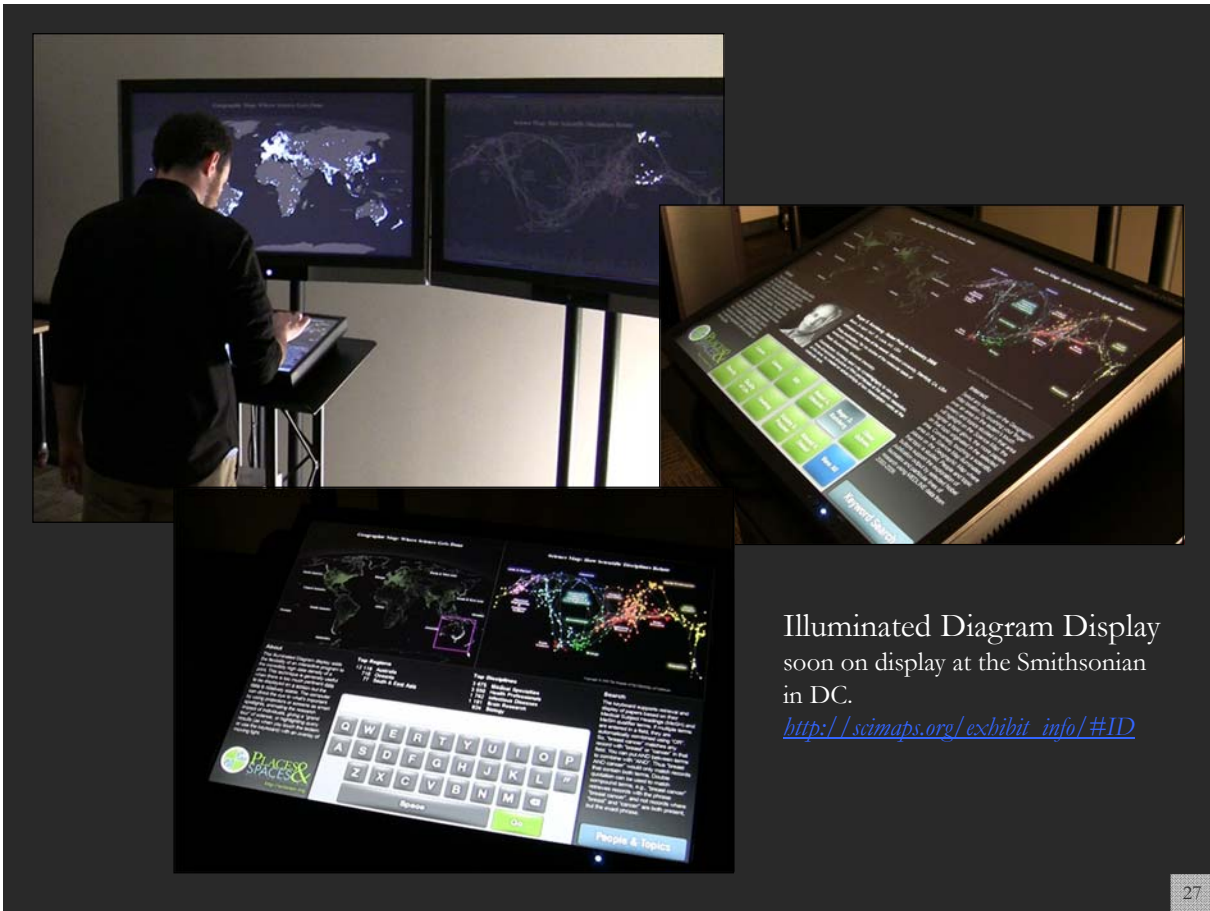




Mapping Science Exhibit at MEDIA X was on May 18, 2009 at Wallenberg Hall, Stanford University,
<http://mediax.stanford.edu>, <http://scaleindependentthought.typepad.com/photos/scimaps>



Science Maps in “Expedition Zukunft” science train visiting 62 cities in 7 months 12 coaches, 300 m long
Opening was on April 23rd, 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

Geographic Map: Where Science Gets Done

Science Map: How Scientific Disciplines Relate

Copyright © 2008 The Regents of the University of California

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, North & East Axis, 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.

<http://scimaps.org>

Cancer

Cloning

HIV

Robert G. Edwards

Roger D. Kornberg

Elinor Ostrom

Obesity

Quality of Life

Smoking

Stanley B. Prusiner

Ahmed H. Zewail

View All

Keyword Search

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 Professionals - 3,589
- Social Sciences - 2,431
- Aeronautical, Chemical, Mechanical & Civil Engineering - 2,208
- Humanities - 1,562

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.

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

People & Topics

29

12-Tokyo-Worldprocessor ⌵

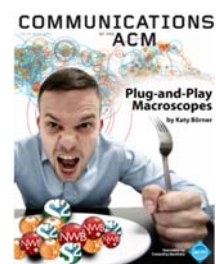
00:17
⏮
⏪
▶
⏩
⏭

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 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.



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.

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.

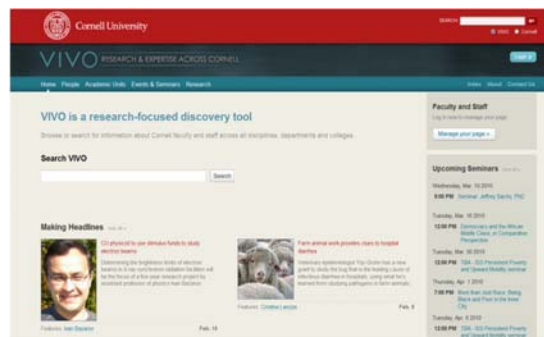
Team Leads and Science Administrators—many scholars direct multiple research projects simultaneously. Some have full-time staff, research scientists, and technicians in their laboratories and centers. Leaders need to evaluate performance and provide references for current or previous members; report 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 Raum, Brenda Stevens, Alicia Turner, Stephen Williams. **Indiana University:** Katy Borner (IU PI), William Barnett, Shanshan Chen, Ying Ding, Russell Duhon, Jon Dunn, Micah Linnemeier, Nianli Ma, Robert McDonald, Barbara Ann O'Leary, Mark Price, Yuyin Sun, Alan Walsh, Brian Wheeler, Angela Zoss. **Ponce School of Medicine:** Richard Noel (Ponce PI), Ricardo Espada, Damaris Torres. **The Scripps Research Institute:** Gerald Joyce (Scripps PI), Greg Dunlap, Catherine Dunn, Brant Kelley, Paula King, Angela Murrell, Barbara Noble, Cary Thomas, 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.

VIVO Enabling a National Network of Scientists

Home People Organizations Research Events

Davis, Vairie I | AST UNV LIBRA

Positions

- Western Science Library Outreach Librarian for Agricultural Sciences** (2002 - 2003)
- Western Science Library Staff Maintenance Supervisor** (2001 - 2002)
- AST UNV LIBRARIAN**

13 publications within the last 10 years (11 leads)

17 identifiable affiliations

Affiliation

professional title
Outreach Librarian for Agricultural Sciences

VIVO Enabling a National Network of Scientists

Home People Organizations Research Events

University of Florida

How do you want to compare?
by Publications

Who do you want to compare?
Search: [] X

Records 1 - 10 of 13

Entity Name	Publication Count	Entity Type
<input checked="" type="checkbox"/> Interdisciplinary Center for Bioremediation	18	UF Center, Agent, Center
<input checked="" type="checkbox"/> Continuing Education	24	UF Department, Agent, Non-Academic Department, Department
<input checked="" type="checkbox"/> Levin College of Law	17	Agent, UF College, College
<input checked="" type="checkbox"/> College of Agricultural and Life Sciences	14	Agent, UF College, College
<input type="checkbox"/> Hargrett College of Journalism and Communications	14	Agent, UF College, College
<input type="checkbox"/> Evelyn F. and William L. McKnight Brain Institute	8	UF Center, Agent, Center

Comparing Publications of Organizations in University of Florida

Total Number of Publications
You have selected 4 of a maximum 10 organizations to compare. **Clear**

- College of Agricultural and Life Sciences: 14
- Levin College of Law: 17
- Continuing Education: 24
- Interdisciplinary Center: 18

VIVO Enabling a National Network of Scientists

Home People Organizations Research Events

Search results for 'geriatrics'

Show only results of this type: **people activities organizations research**

AMERICAN GERIATRICS SOCIETY

- Geriatrics Education Curriculum, Residents (GEC) Program
- Evidence Based Decision Making in Geriatrics, Geriatrics, Disability

AMERICAN GERIATRICS SOCIETY

- Hartford Geriatrics Leadership Scholar
- Geriatrics and Aging Research Institute on Aging (GARI)

ASST DIR GERIATRICS ACADEMIC PROGRAMS

- US OLTH RESOURCES AND SERV ADMIN
- Suflonin Study
- 2003 Scholar, Hartford Institute of Geriatrics, Nursing Research, John A. Hartford Institute for Geriatrics, Nursing, New York University
- Genetics, Rehabilitation and Prevention of Disability
- Assistant to the Director, Illinois
- Carlini, Wilson, Boudreau, Boudreau, and MacKinnon
- AMES ACAD OF NURSING
- The Epidemiology of Stress and the Menopausal Syndrome
- Statement to a Sub-Committee

VIVO Enabling a National Network of Scientists

Home People Organizations Research Events

Welcome to VIVO

VIVO is a research-focused discovery tool that enables collaboration among scientists across all disciplines.

Browse or search information on people, departments, courses, grants, and publications.

Search VIVO

Log in

Email: []
Password: []
Remember me:

Browse by

- Grants (11,814)
- People (48,721)
- Activities (11,818)
- Courses (1116)
- Events (379)
- Organizations (20,338)
- Research (11,283)
- Locations (376)

Faculty Member (8882) []
Graduate Student (1) []
Librarian (67) []
Non-Academic (7536) []
Non-Faculty Academic (12) []
Alumn (8972) []
Professor Emeritus (802) []

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

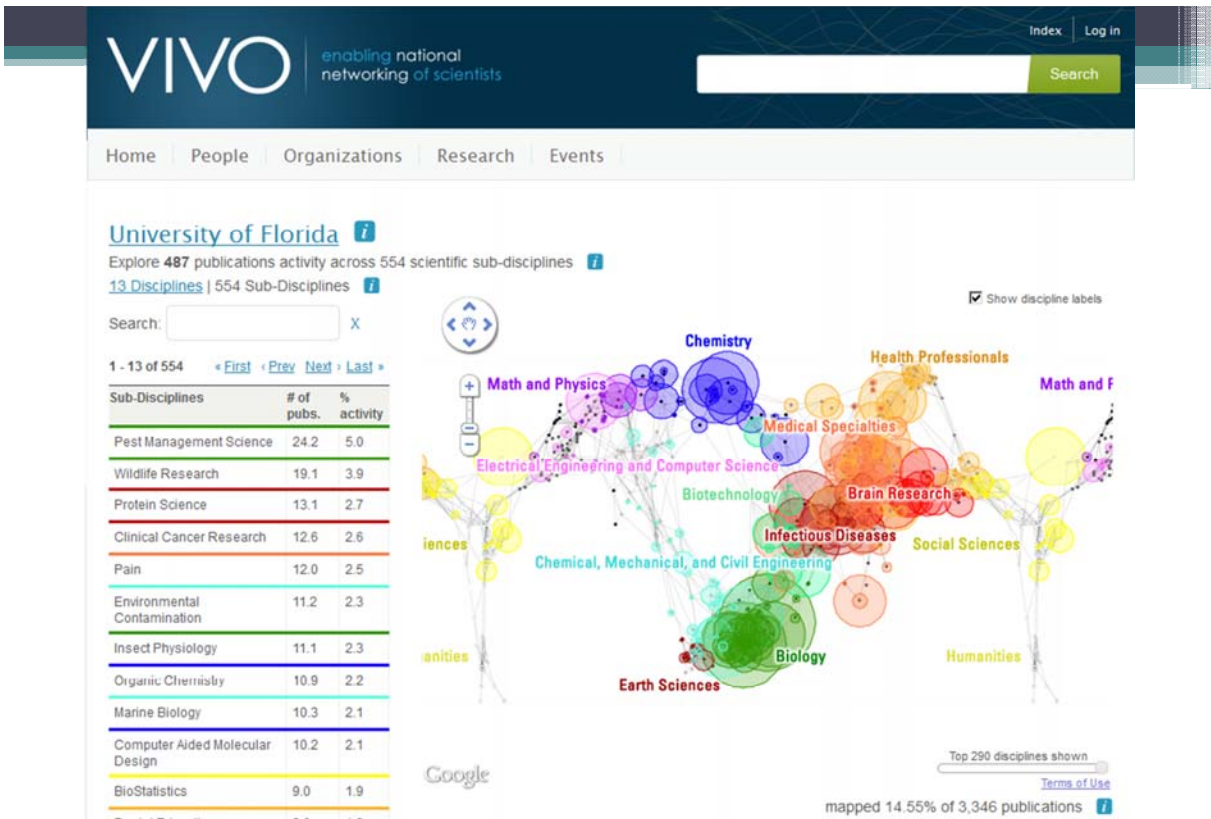
Comparing Grants of Organizations in University of Florida

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

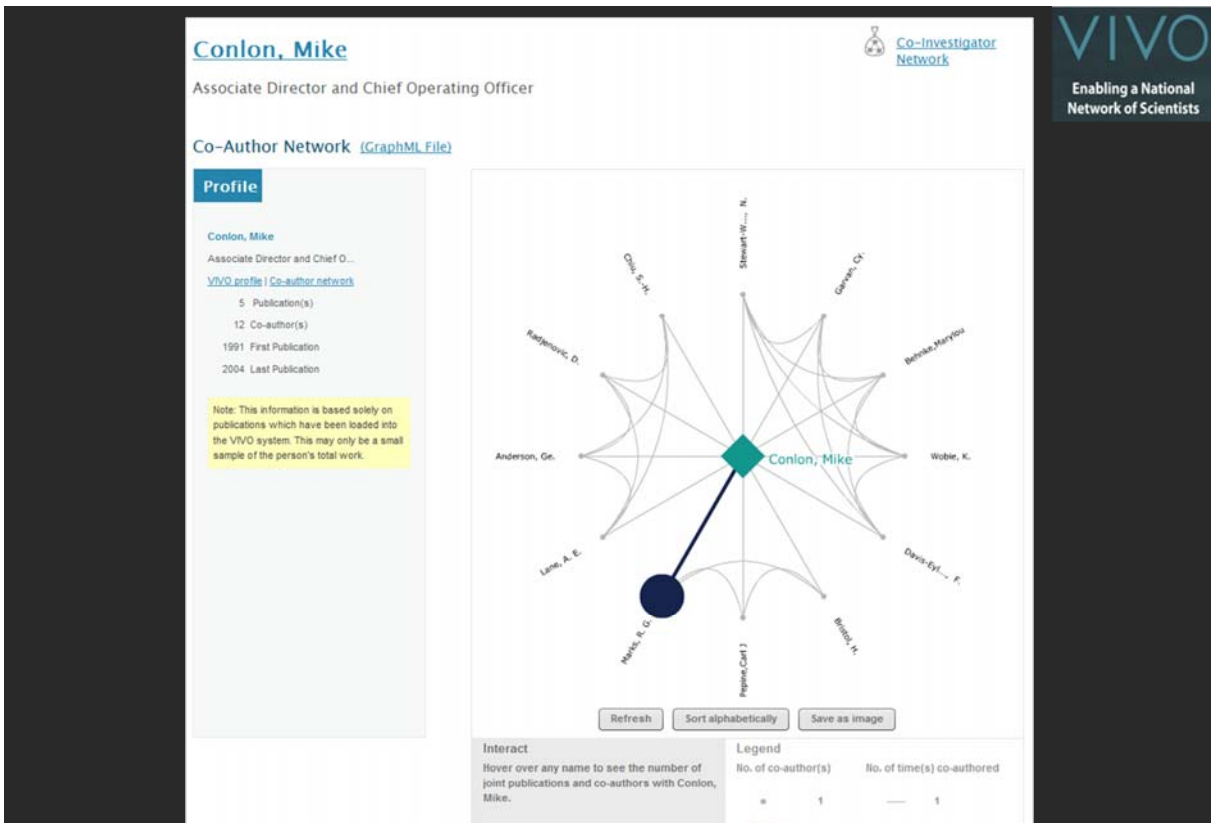
Save as CSV **Clear**

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



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

39



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

40

National Researcher Networking Visualization 1.0



<http://nrn.cns.iu.edu>

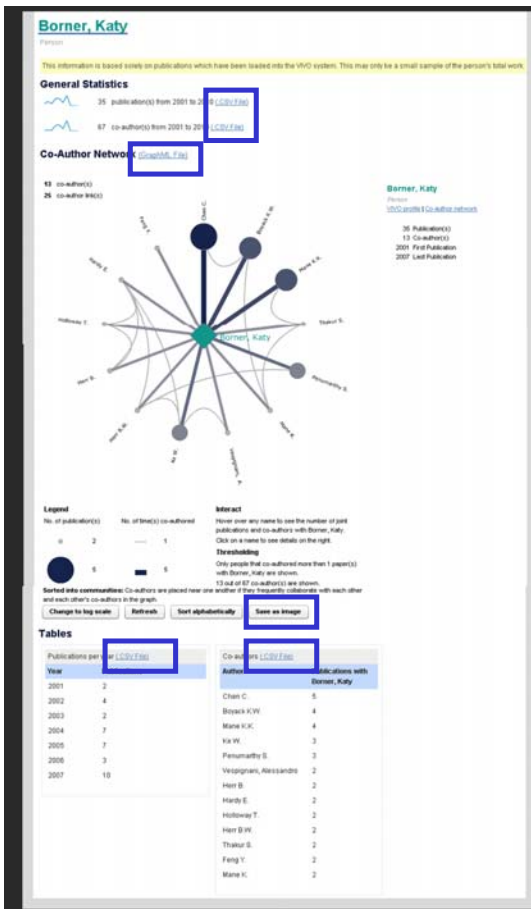
Geospatial Analysis (Where) Where is what science performed by whom? Science is global and needs to be studied globally.



VIVO On-The-Go

Overview, Interactivity,
Details on Demand
come to
commonly
used devices
and environments





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)

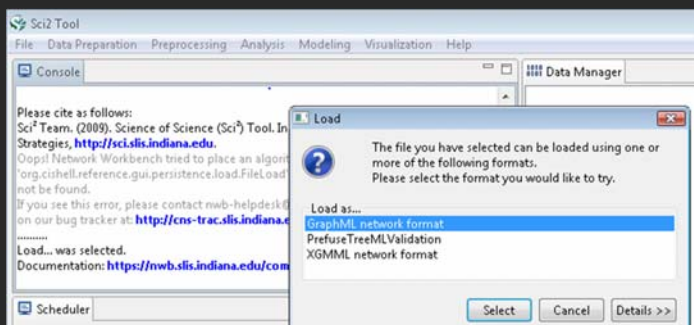
Save as Image (.PNG file)

Tables

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

<http://vivo.iu.edu/vis/author-network/person25557>

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

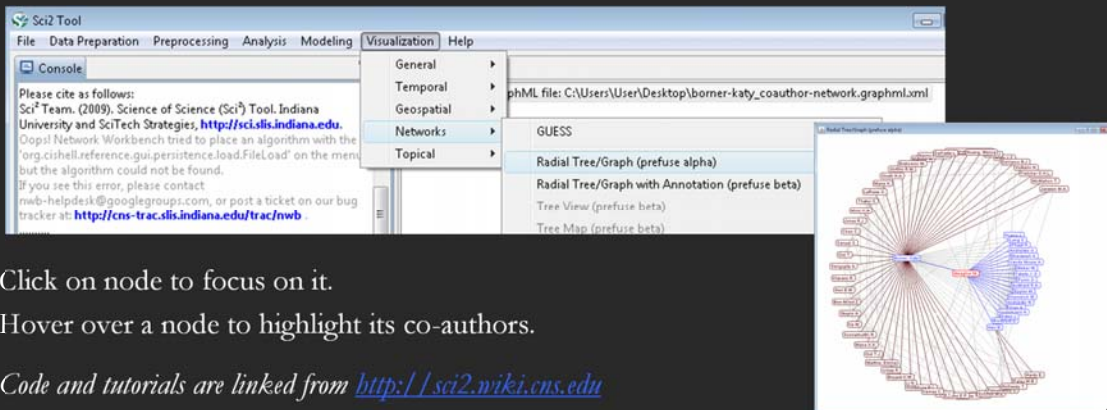


Network Analysis Toolkit

Nodes: 81

Edges: 390

Visualize the file using Radial Graph layout.



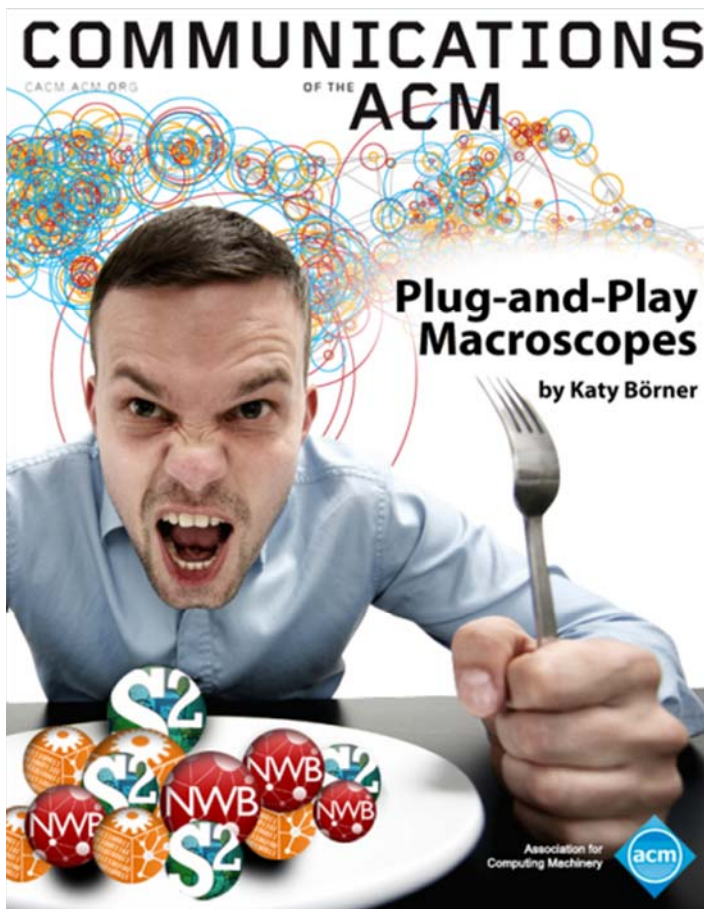
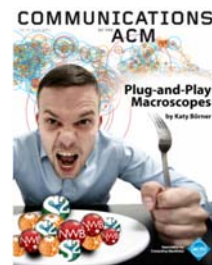
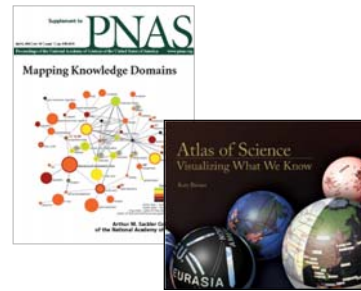
Click on node to focus on it.

Hover over a node to highlight its co-authors.

Code and tutorials are linked from <http://sci2.wiki.cns.edu>

Overview

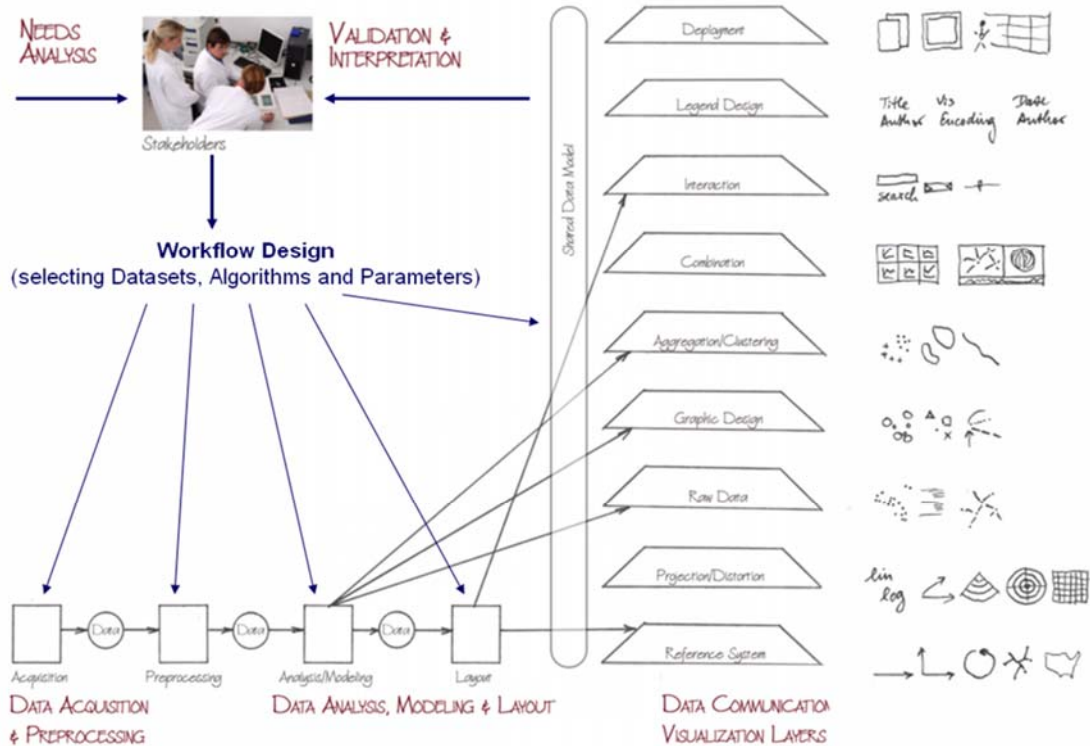
1. **Data mining and visualization research** that aims to increase our scientific understanding of the structure and dynamics of science and technology.
2. **Novel approaches and data services** that improve information access, researcher networking, and research management.
3. **Data services and plug-and-play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



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

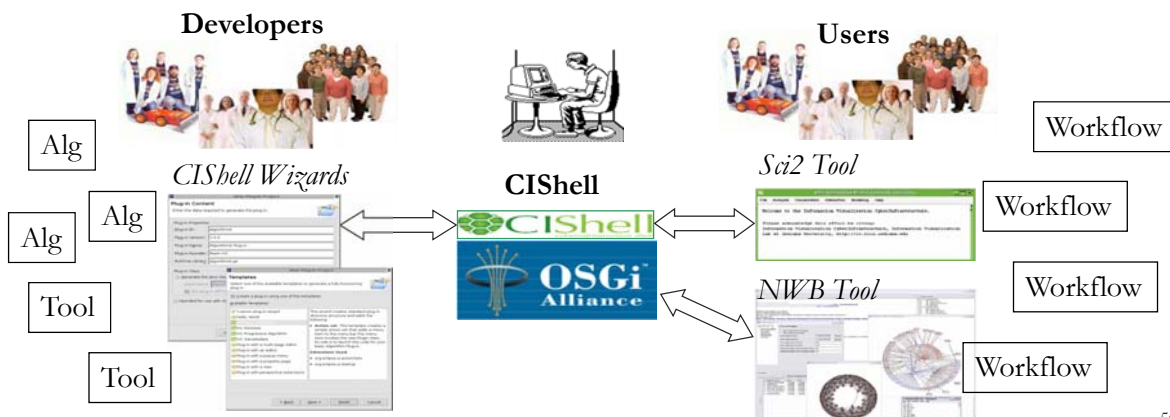
Video and paper are at <http://www.scivee.tv/node/27704>

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](#)

51



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.

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

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

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](#).

Gallery

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.

52

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

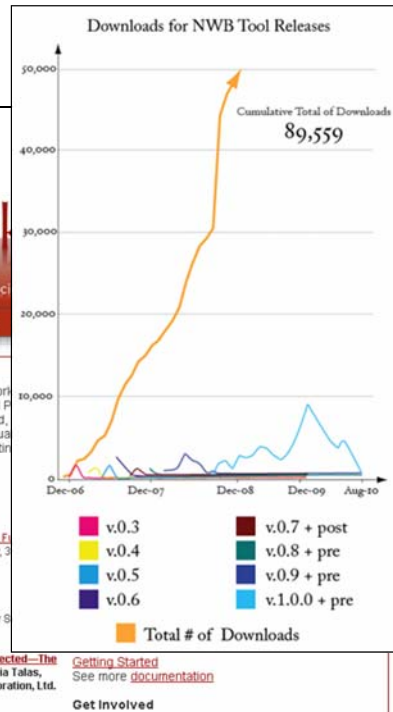
In February 2009, the tool provides more than 169 plugins that support the preprocessing, analysis, modeling, and visualization of networks.

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.



The screenshot shows the Network Workbench website with a navigation menu (Home, People, Research, Publications) and a 'Summary' section. The summary describes the tool as a large-scale network toolkit for biomedical, social science, and physics. It also includes a 'News & Updates' section with several entries from 2009, such as 'Mapping the Frontiers of Knowledge' and 'Tutorial abstract for S 2009'.



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.

53

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.

54

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).

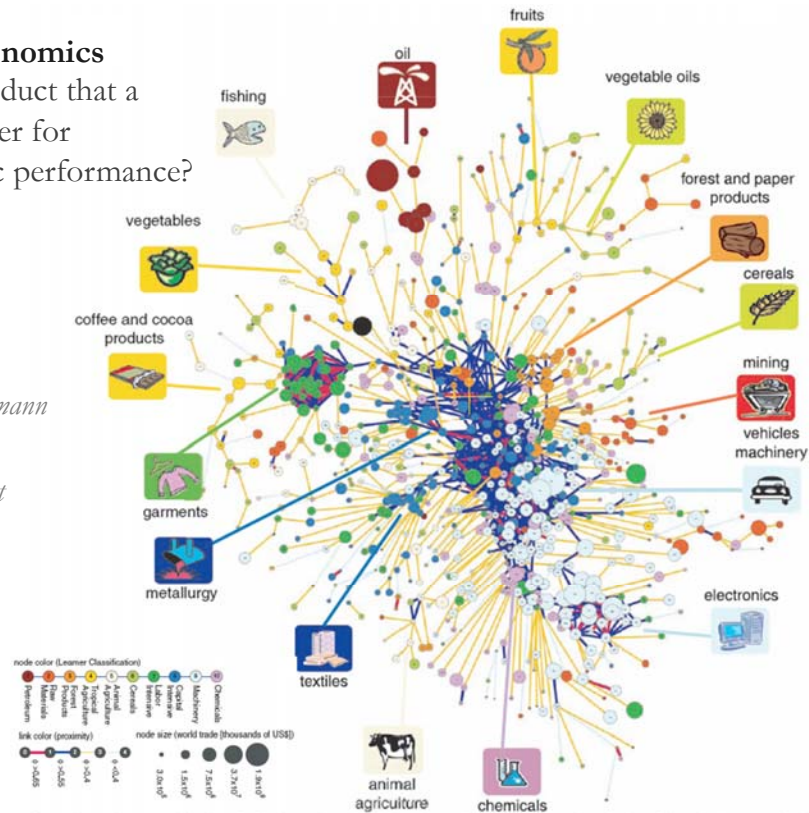


Fig. 1. The product space. (A) Hierarchically clustered proximity matrix representing the 775 SITC-4 product classes exported in the 1998–2000 period. (B) Network representation of the product space. Links are color coded with their proximity value. The sizes of the nodes are proportional to world trade,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

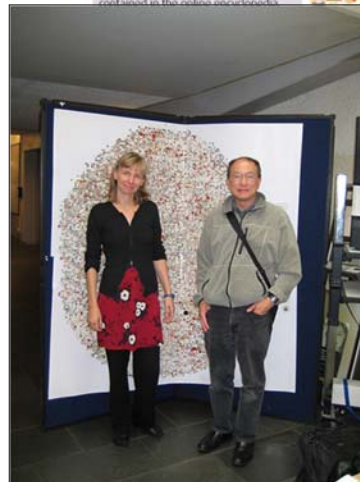


Second sight

Image: Bruce W. Hest and Todd M. Holloway

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.



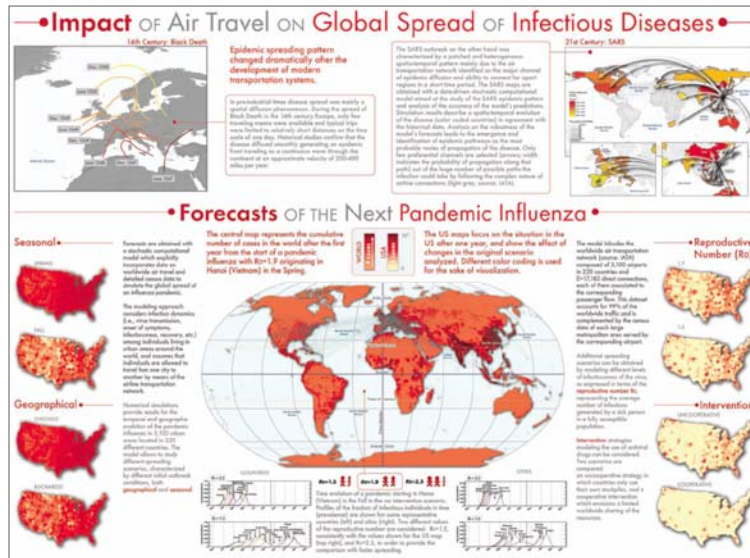
Computational Epidemics

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).

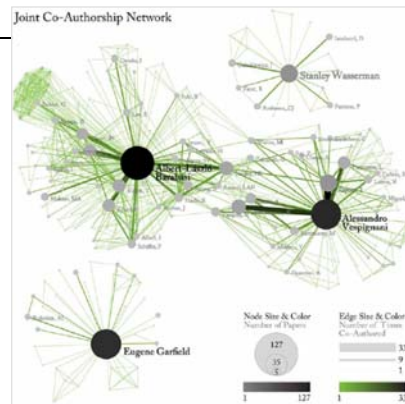
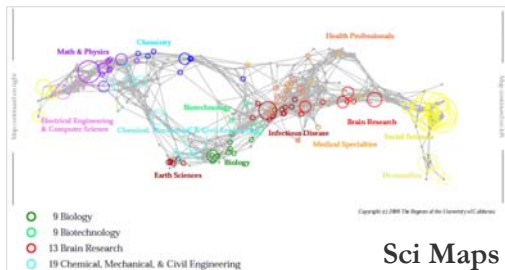
Reaction-diffusion processes and metapopulation models in heterogeneous networks, V. Colizza, R. Pastor-Satorras, A. Vespignani, *Nature Physics* 3, 276-282 (2007).

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).

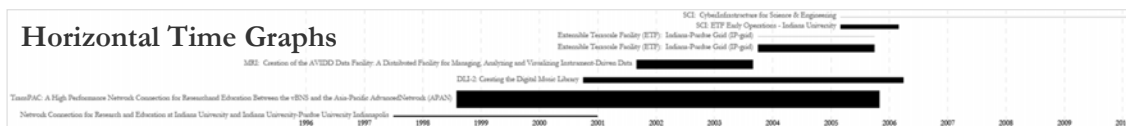


Sci2 Tool – “Open Code for S&T Assessment”

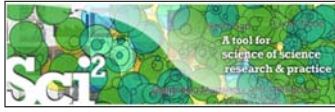
OSGi/CIShell powered tool with NWB plugins and many new scientometrics and visualizations plugins.



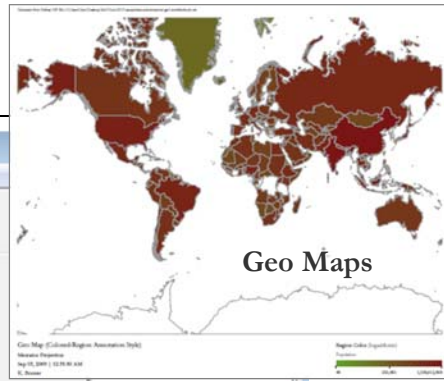
Horizontal Time Graphs



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>).

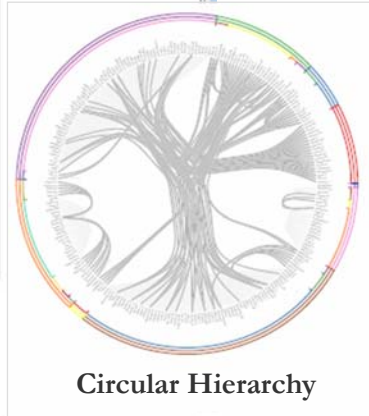
The primary investigators are Katy Börner, In SciTech Strategies Inc. The Sci² tool was developed by J. Duhon, Patrick A. Phillips, Chintan Tank, a Cyberinfrastructure Shell (<http://cishell.org>) for Network Science Center (<http://cns.slis.indiana.edu>). Many algorithm plugins were derived from the 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	100%
<input checked="" type="checkbox"/>	Load and Clean ISI File	09/03/2009	00:15:05 AM	100%



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>

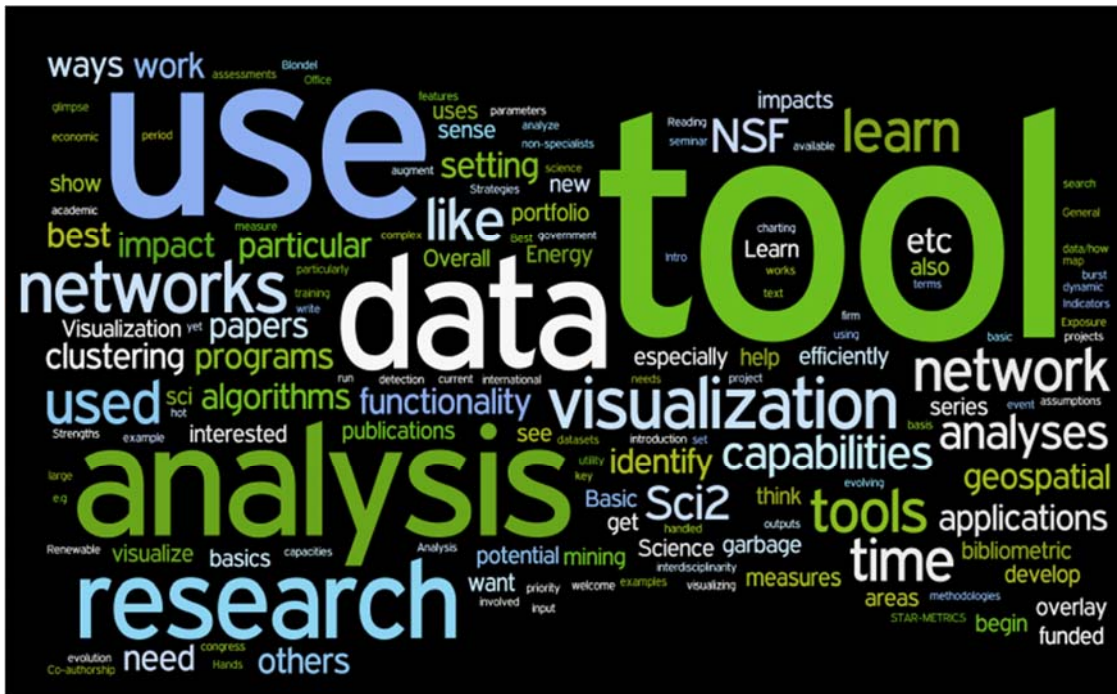
<http://sci2.wiki.cns.in.edu>

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

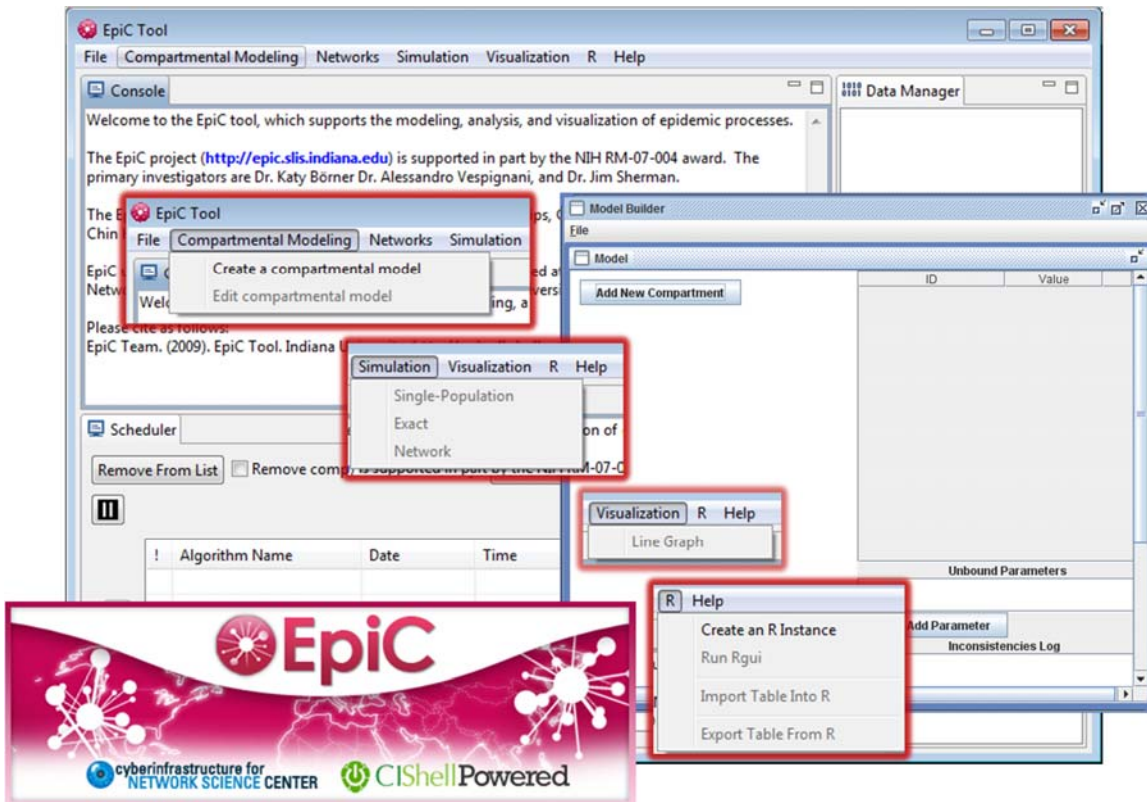
NIH Office of Extramural Research and Katy Börner (2010) [Network Visualizations Using SPIRES Data and the Sci2 Tool](#). Office of Extramural Research at NIH.



Wordle.net of “Interest to Learn” response by users from more than 40 countries



61

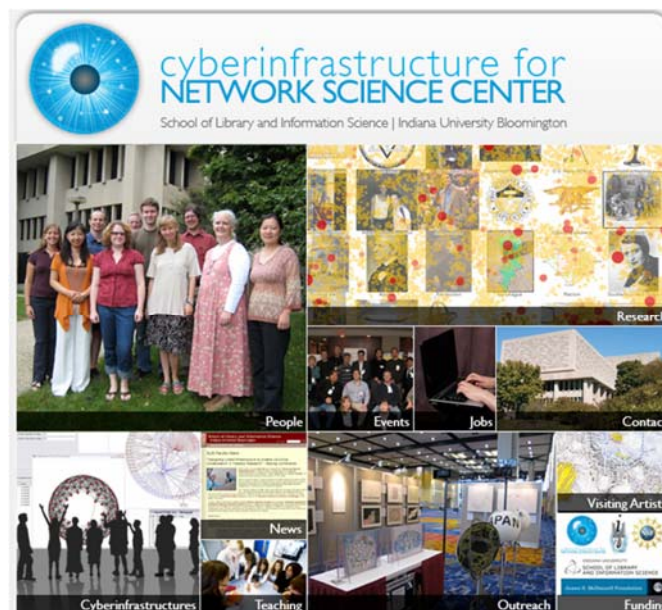


62

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

- | | |
|--------|---|
| USA | <ul style="list-style-type: none"> ➤ <i>Cytoscape</i> (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). ➤ <i>MAEviz</i> (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 | <ul style="list-style-type: none"> ➤ <i>Taverna Workbench</i> (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. ➤ <i>TEXTrend</i> (http://textrend.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. ➤ <i>DynaNets</i> (http://www.dynanets.org) Coordinated by Peter M.A. Sloot at the University of Amsterdam, The Netherlands develops algorithms to study evolving networks. ➤ <i>SISOB</i> (http://sisob.lcc.uma.es) An Observatory for Science in Society Based in Social Models. <p>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.</p> |

63



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>

64