

Scholarly Data, Network Science, and (Google) Maps – Modeling, Mapping, and Exploring the Dynamic Landscape of Science

Dr. Katy Börner

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

*Talk at Borders in Bloomington, IN
2007.05.10*

The image shows a screenshot of the 'Places & Spaces: Mapping Science' website and several photographs of the exhibit. The website header features the logo 'places & spaces &' and a navigation menu with links for Home, Browse Maps, Compare & Contrast Maps, Schedule, and Connect. The main content area includes an 'Exhibit Purpose and Goals' section, a 'The Places & Spaces: Mapping Science' section with a 'ORDER DVD HERE!' button, and a section titled '"Places & Spaces: Mapping Science" on display at the New York Hall of Science, Dec. 9, 2006 - Feb. 25, 2007.' Below this is a note about the exhibit's display at the NYPL Science, Industry, and Business Library (Madison/34th), New York, from April 3rd to August 31st, 2006. To the right of the website screenshot is a photograph of five people standing in front of the exhibit. Below the website screenshot are two more photographs: one showing the exhibit's entrance with a sign that reads 'Places & Spaces Cartography of the Physical and Abstract' and another showing two people interacting with a large digital display.

Places & Spaces: Mapping Science
An exhibition created to demonstrate the power of maps to understand, navigate, and manage not only physical places, but also abstract information spaces.

Home Browse Maps Compare & Contrast Maps Schedule Connect

Home

Exhibit Purpose and Goals

The Places & Spaces: Mapping Science exhibit has been created to demonstrate the power of maps. An initial theme of this exhibit is to compare and contrast first maps of our entire planet with the first maps of all of science as we know it.

Check out the [schedule of physical showings](#) and come see with your own eyes the extent to which maps can be employed to help make sense of the flood of information we are confronted with and how domain maps can be used to locate complex and beautiful information.

"Places & Spaces: Mapping Science" on display at the New York Hall of Science, Dec. 9, 2006 - Feb. 25, 2007.

Places & Spaces at the NYPL Science, Industry, and Business Library (Madison/34th), New York, April 3rd - August 31st, 2006.

Places & Spaces: Mapping Science
a science exhibit that introduces people to maps of sciences, their makers and users.
Scimaps.org

Exhibit Curators:
Dr. Katy Börner & Julie Smith

The Power of Maps

Four Early Maps of Our World
VERSUS
Six Early Maps of Science



(1st Iteration of Places & Spaces Exhibit - 2005)

The Power of Reference Systems

Four Existing Reference Systems
VERSUS
Six Potential Reference Systems of Science



(2nd Iteration of Places & Spaces Exhibit - 2006)



"Places & Spaces: Mapping Science"
 on display at the New York Hall of Science
(5 mins from LaGuardia Airport)
 December 9th, 2006 – February 25th, 2007.

TOPIC MAP: HOW SCIENTIFIC PARADIGMS RELATE

GEOGRAPHIC MAP: WHERE SCIENCE GETS DONE

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

Nanotechnology

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

<p>All Topics</p> <p><i>Sweep through all 776 scientific paradigms</i></p>	<p>Nanotechnology</p> <p><i>Science on the tiny scale of molecules</i></p>	<p>Francis H. C. CRICK</p> <p><i>Co-discovered DNA's double helix</i></p>	<p>Albert EINSTEIN</p> <p><i>Revitalized physics with Relativity theories</i></p>	<p>Michael E. FISHER</p> <p><i>Models critical phase transitions of matter</i></p>	<p>Susan T. FISKE</p> <p><i>Connects perception and stereotypes</i></p>
<p>Sustainability</p> <p><i>The science behind our long-term hopes</i></p>	<p>Biology & Chemistry</p> <p><i>The interface between these two vital fields</i></p>	<p>Joshua LEDERBERG</p> <p><i>Pioneer in bacterial genetic mechanisms</i></p>	<p>Derek J. de Solla PRICE</p> <p><i>Known as the "Father of Scientometrics"</i></p>	<p>Richard N. ZARE</p> <p><i>Lives laser chemistry in molecular dynamics</i></p>	<p>About this display</p> <p><i>People & organizations that helped create it</i></p>

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

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

TOPIC MAP: HOW SCIENTIFIC PARADIGMS RELATE

GEOGRAPHIC MAP: WHERE SCIENCE GETS DONE

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

Nanotechnology

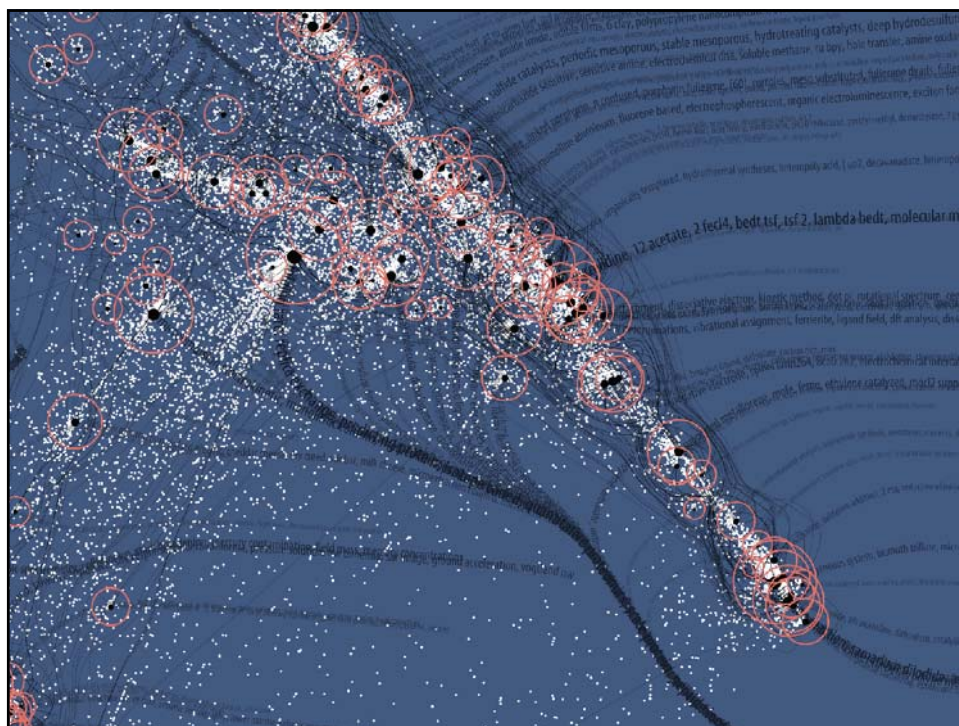
This overlay shows the distribution of nanotechnology within the paradigms of science. The majority of current work in nanotechnology takes place in physics, chemistry, and materials science, at the upper right portion of the map. However, an increasing amount of nanotechnology is being applied in the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

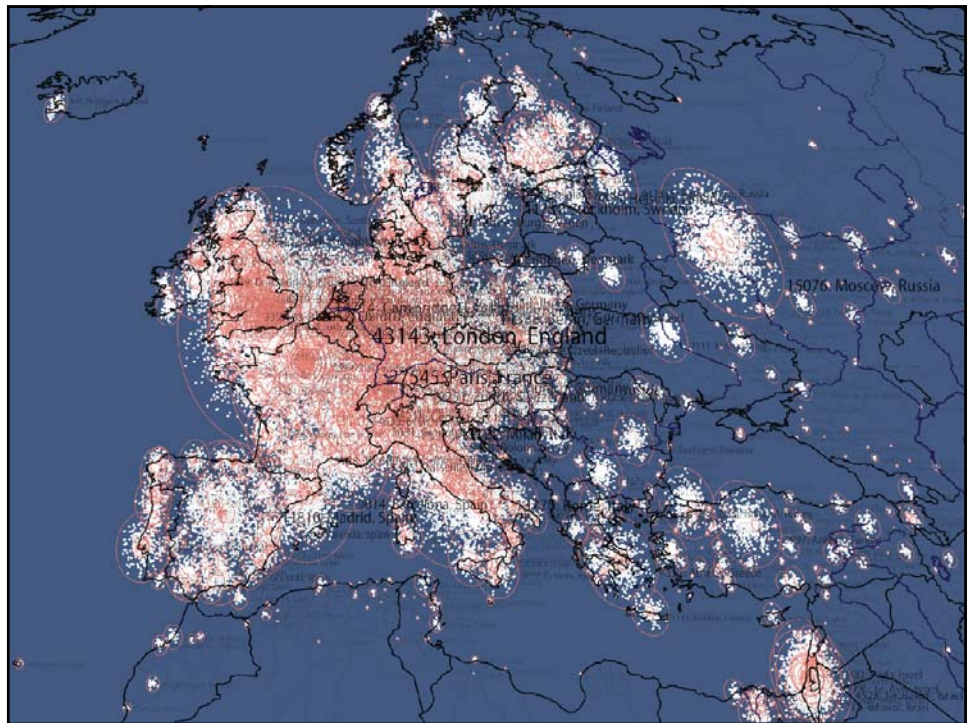
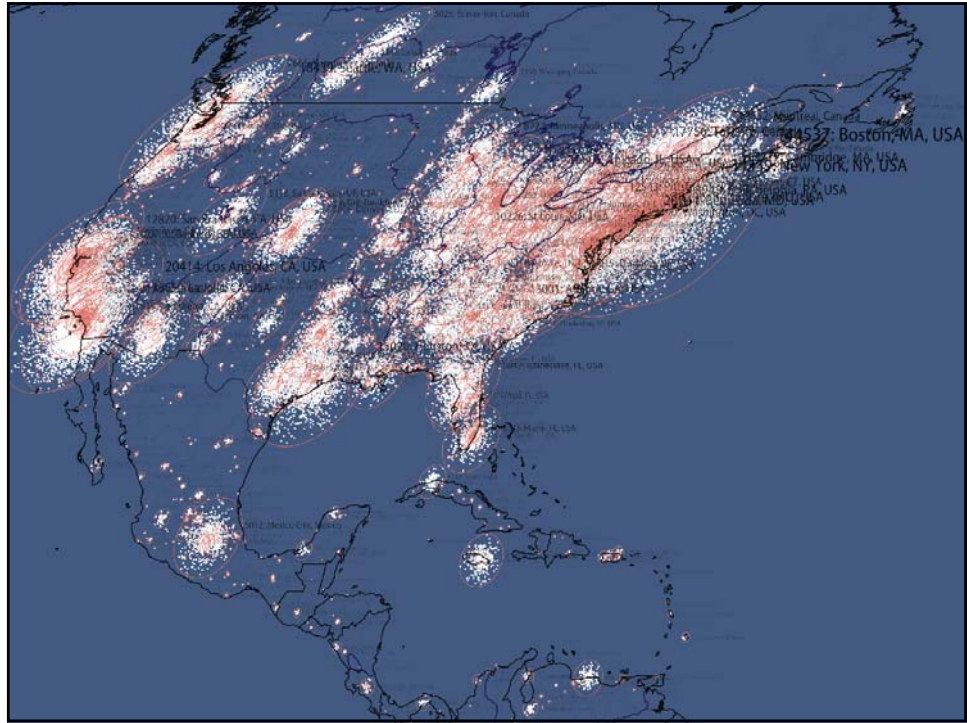
All Topics	Nanotechnology	Francis H. C. CRICK	Albert EINSTEIN	Michael E. FISHER	Susan T. FISKE
<i>Sweep through all 776 scientific paradigms</i>	<i>Science on the tiny scale of molecules</i>	<i>Co-discovered DNA's double helix</i>	<i>Revitalized physics with Relativity theories</i>	<i>Models critical phase transitions of matter</i>	<i>Connects perception and stereotypes</i>
Sustainability	Biology & Chemistry	Joshua LEDERBERG	Derek J. de Solla PRICE	Richard N. ZARE	About this display
<i>The science behind our long-term hopes</i>	<i>The interface between these two vital fields</i>	<i>Pioneer in bacterial genetic mechanisms</i>	<i>Known as the "Father of Scientometrics"</i>	<i>Uses laser chemistry in molecular dynamics</i>	<i>People & organizations that helped create it</i>

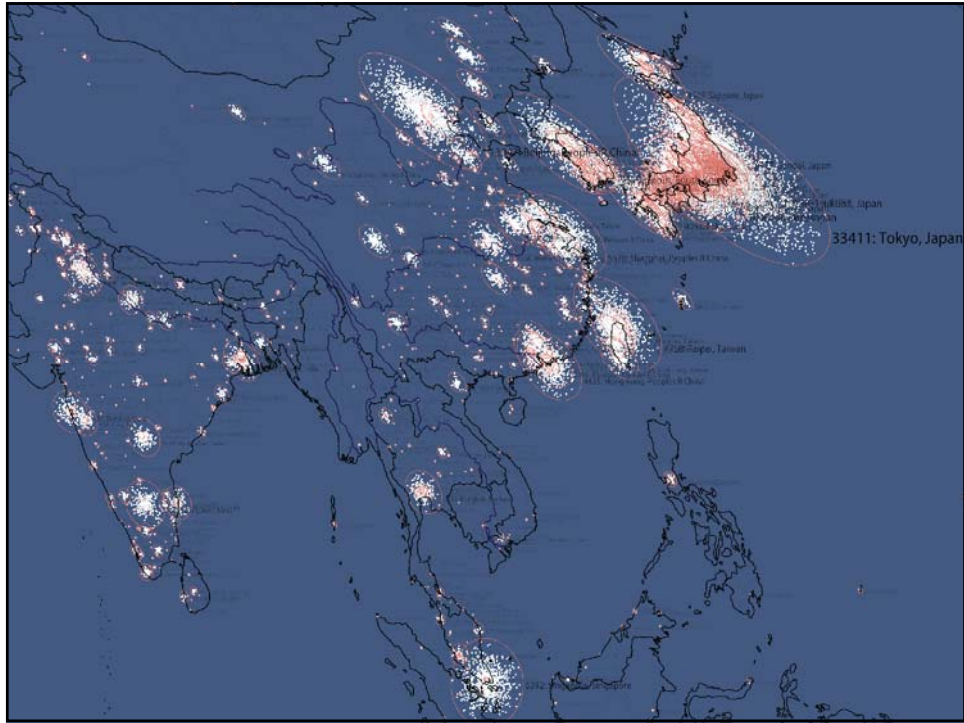
We sweep slowly through adjoining related topics, lighting up the places in the world that study each topic. You may select a subset of the topics that deal with these three interesting subjects by touching it.

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



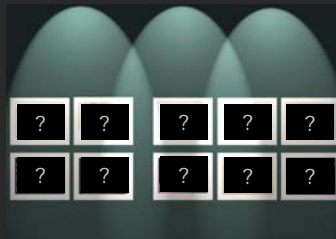






The Power of Forecasts

Four Existing Forecasts
VERSUS
Six Potential Science 'Weather' Forecasts



(3rd Iteration of Places & Spaces Exhibit - 2007)

10 Iterations in 10 Years

part I:

2005 The power off maps

2006 The power off reference systems

2007 The power off forecasts

Here four relevant maps from any domain are compared with six powerful maps/reference systems/forecasts.

part II:

2008 Science maps for industry

2009 Science maps for science policy makers

2010 Science maps for researchers

2011 Science maps as visual interfaces to digital libraries

2012 Science maps for kids

2013 Science maps as daily science forecasts

Here four relevant maps from that specific domain, e.g., kids, are compared with six science maps for kids.

part III:

2014 How to lie with maps

Examples how to do a bad/misleading job.

Behind the Scene:

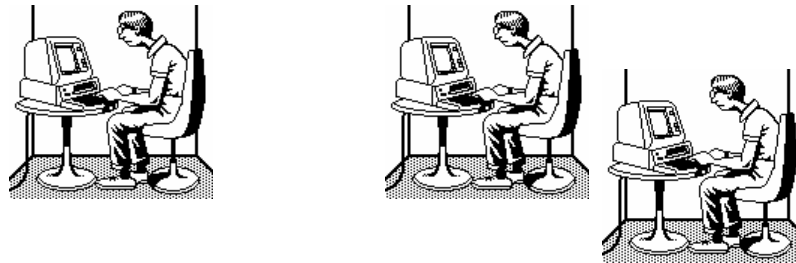
- **General Process**
- **People**
- **Scholarly Data Integration**
- **Network Science &
(Google) Maps**

How to Make a Science Map

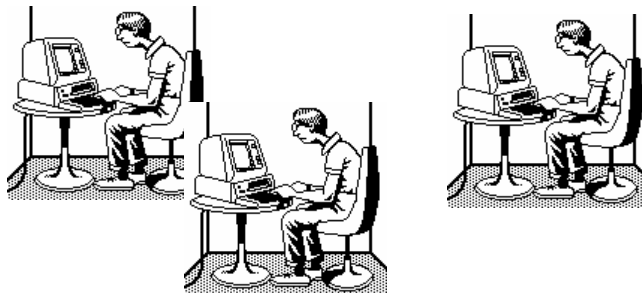
DATA EXTRACTION	UNIT OF ANALYSIS	MEASURES	LAYOUT (often one code does both similarity and ordination steps)		DISPLAY
			SIMILARITY	ORDINATION	
SEARCHES ISI INSPEC Eng Index Medline ResearchIndex Patents etc.	COMMON CHOICES Journal Document Author Term	COUNTS/FREQUENCIES Attributes (e.g. terms) Author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

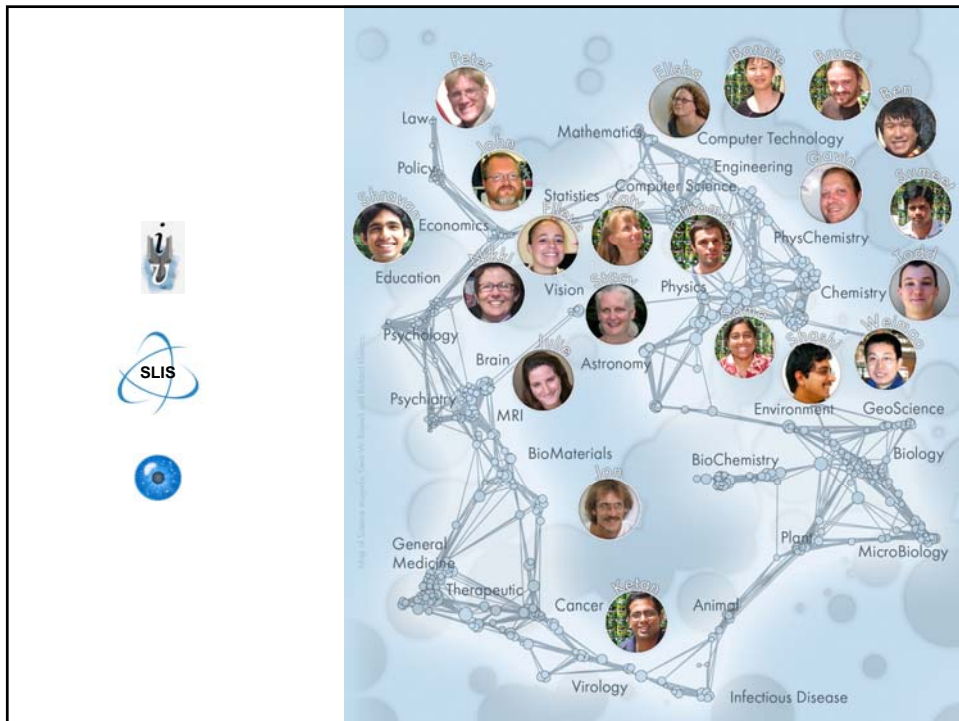
Shiffrin, Richard M. and Börner, Katy (Eds.) (2004). *Mapping Knowledge Domains. Proceedings of the National Academy of Sciences of the United States of America*, 101(Suppl_1).

Börner, Katy, Chen, Chaomei, and Boyack, Kevin. (2003). *Visualizing Knowledge Domains. In Blaise Cronin (Ed.), Annual Review of Information Science & Technology, Volume 37, Medford, NJ: Information Today, Inc./ American Society for Information Science and Technology, chapter 5, pp. 179-255.*



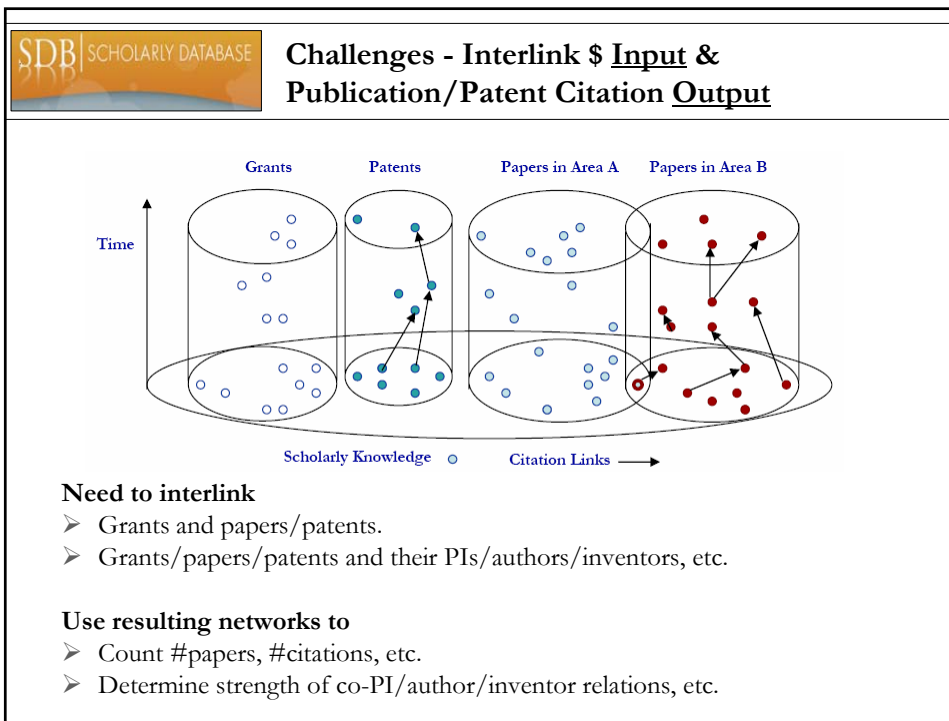
The Cyberinfrastructure Developers

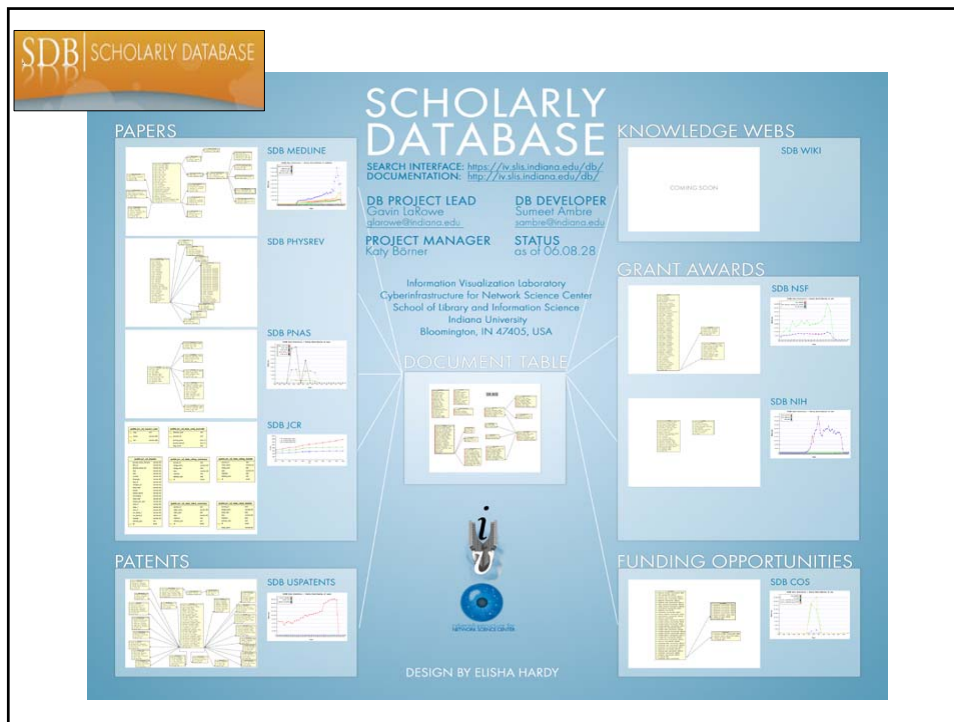




Behind the Scene:

- People
- **Scholarly Data Integration**
- Network Science & (Google) Maps





Scholarly Database: Web Interface

Search across publications, patents, grants.
Download records and/or (evolving) co-author, paper-citation networks.

SDB SCHOLARLY DATABASE

Home Search Admin Logout

Select Database

COS
 NIH
 NSF
 USPAT
 MEDLINE
 PHYSREV
 PNAS

Author(s) Last Name Middle Name First Name
 Title: e.g. Classifying DNA
 Journal: e.g. Journal of Biological Sciences

Publication Range
 From: to (default Year range is 1945-2005)

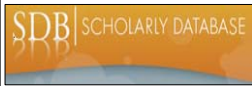
SDB SCHOLARLY DATABASE

Home Search Admin Logout

NIH (336 Matching Records)

1. JAMES, ERIC (2001) GLUCOCORTICOID RECEPTOR-MEDIATED CATABACT.
 [DE SCHEPFIN(Applied's Abstract)] Calancts are a stress sink in those undergoing steroid therapy, restricting the efficacy of these compounds. Steroid-induced catabolism is a protein substrate, frequently outside the central axonal axis and cell.
 2. JAMES, GARTH (2001) THE USE OF BIOPOLYMER TO COUNTER BIOTERRORISM.
 [DE SCHEPFIN (excludes from Applied's Abstract)] It is possible that terrorists will contaminate public drinking water supplies with biological agents, such as bacteria, viruses, or toxins, because greater every day. This web search of information feeds...
 3. JAMES, JUSTIN (2001) Free quantity of antibodies autoantibodies.
 [DE SCHEPFIN (provided by applied)] Systemic sclerosis (scleroderma) is a debilitating, multi-system disease of unknown etiology, which is characterized by a broad spectrum of disease manifestations with varying organ involvement. Treatment's effectiveness...
 4. JAMES, LAURA (2001) NOVEL THERAPIES FOR ACETABROMPHEN TOXICITY.
 [DE SCHEPFIN (adapted from the application)] The long-term goal of this award is to develop therapies, based on more therapeutic data, that can be applied to the treatment of the acetabromphen (APAP) overdose patient. All therapeutic data, APAP is methyl...
 5. JAMES, LAURA (2001) NOVEL THERAPIES FOR ACETABROMPHEN TOXICITY.
 [DE SCHEPFIN (adapted from the application)] The long-term goal of this award is to develop therapies...

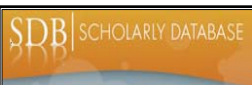
<< Prev 1 2 3 4 5 6 7 8 9 10 Next >>



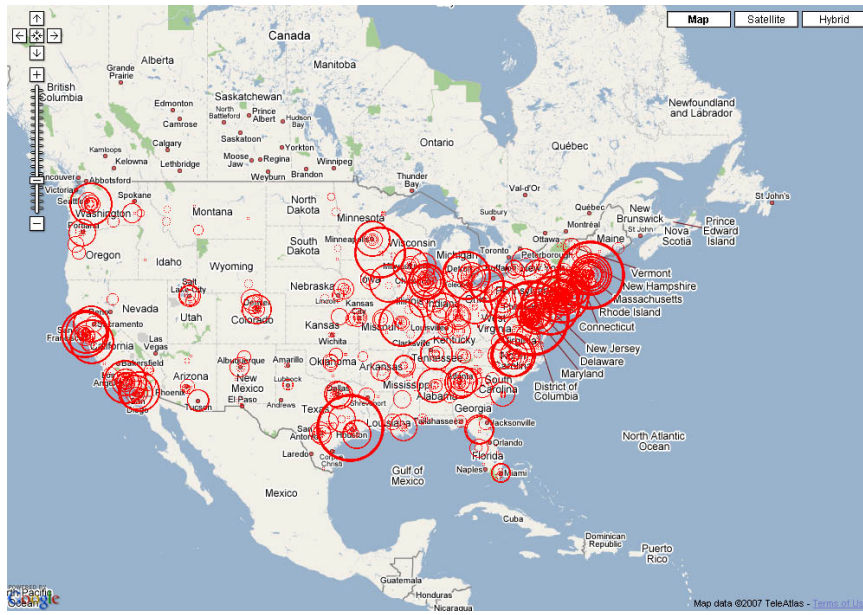
Scholarly Database: # Records & Years Covered

Datasets available via the Scholarly Database (* future feature)

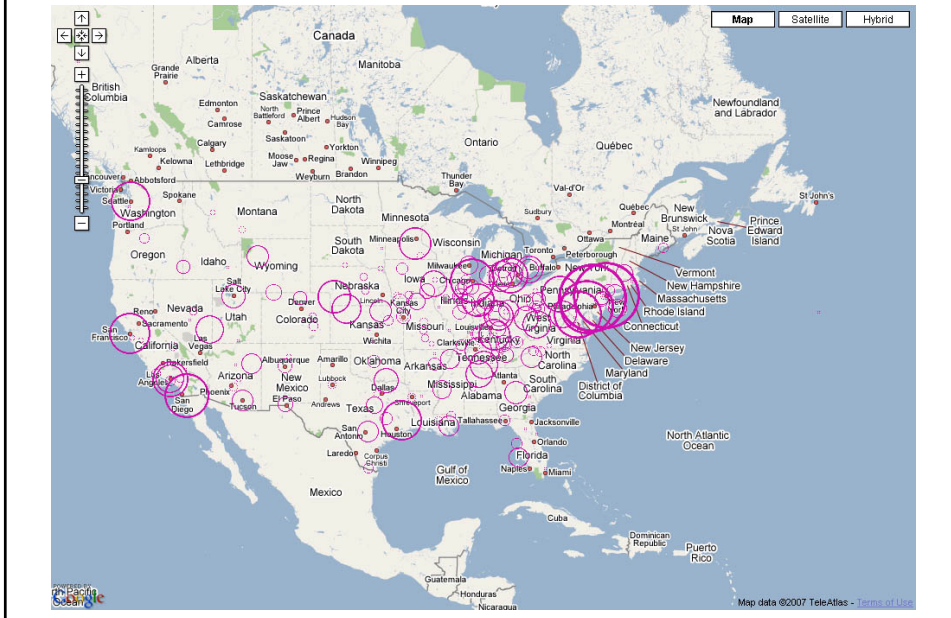
Dataset	# Records	Years Covered	Updated	Restricted Access
Medline	13,149,741	1965-2005	Yes	
PhysRev	398,005	1893-2006		Yes
PNAS	16,167	1997-2002		Yes
JCR	59,078	1974, 1979, 1984, 1989 1994-2004		Yes
USPTO	3,179,930	1976-2004	Yes*	
NSF	174,835	1985-2003	Yes*	
NIH	1,043,804	1972-2002	Yes*	
Total	18,021,560	1893-2006	4	3



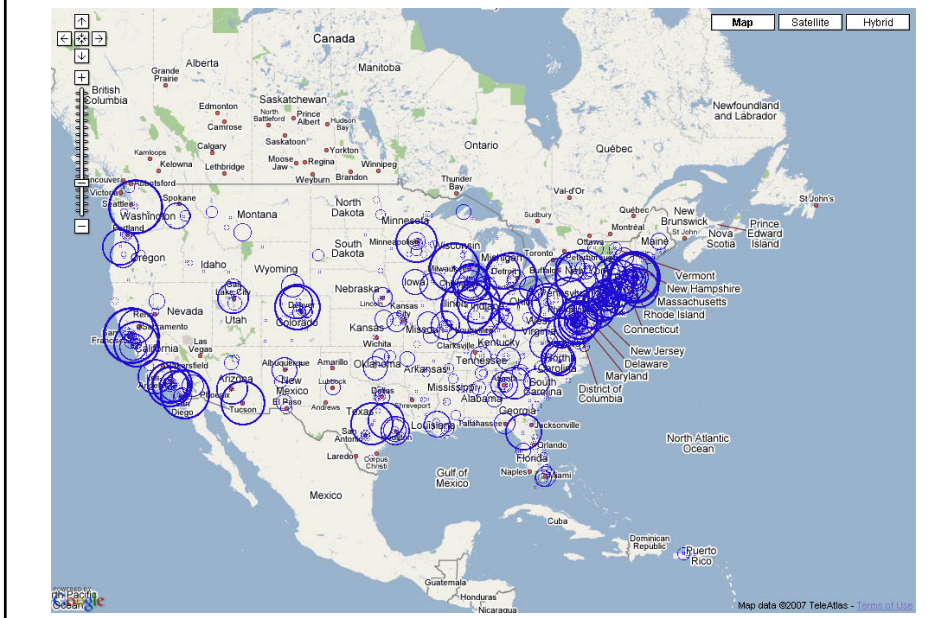
Medline Publications

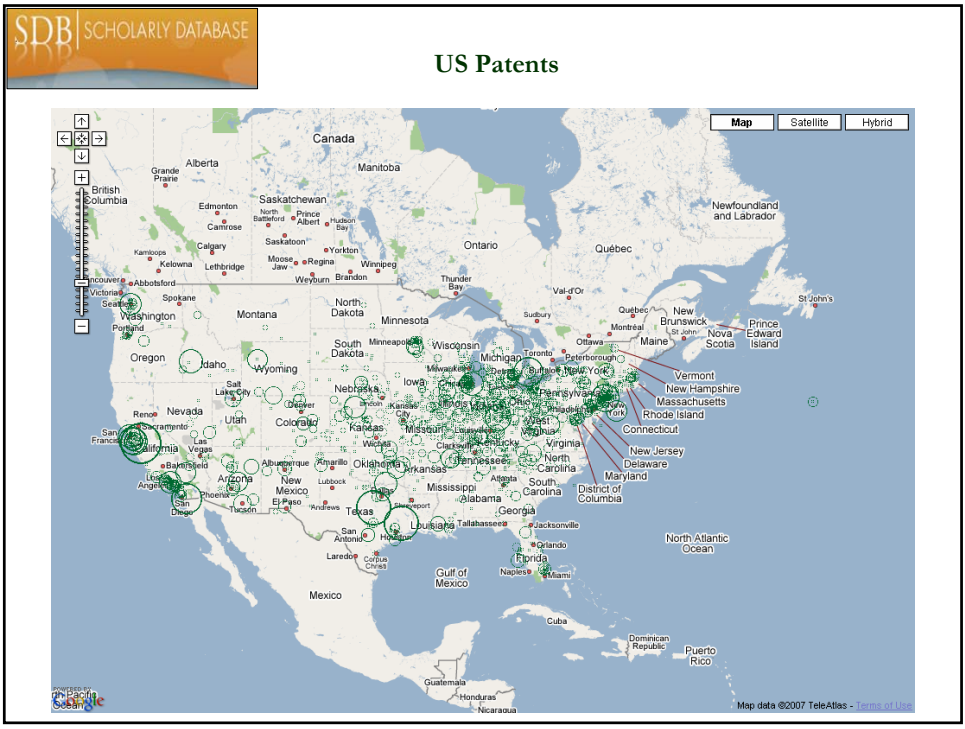


NIH



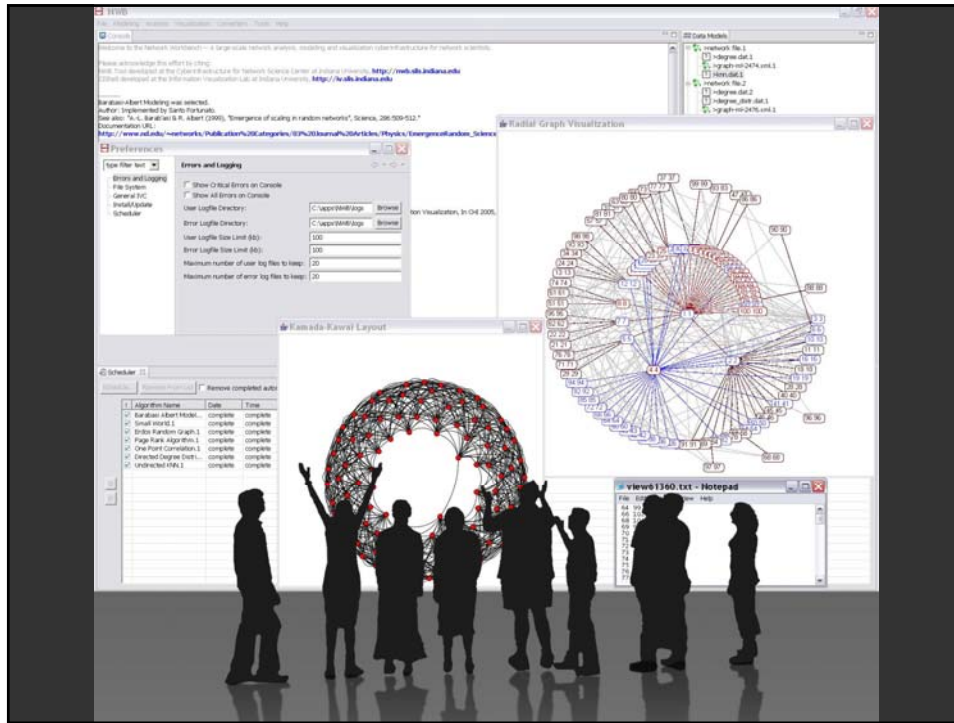
NSF Funding





Behind the Scene:

- People
- Scholarly Data Integration
- Network Science & (Google) Maps



Building Market Places not Cathedrals



- Requires the design & implementation of ‘software glue’ that can interlink datasets and algorithms written in different languages using different data formats.
- The smaller the glue or ‘CI Shell’, the more likely it can be maintained.
- Dataset and algorithm ‘plugins’ are provided by application holders/ community.
- Applications resemble custom ‘fillings’.



Cyberinfrastructure Shell (CIShell)

<http://cishell.org>

CIShell is an 'empty shell' that supports

- Easy integration of new datasets and algorithms by algorithm developers and
- Easy usage of algorithms by algorithm users.

Its plug-and-play architecture supports the integration and utilization of diverse

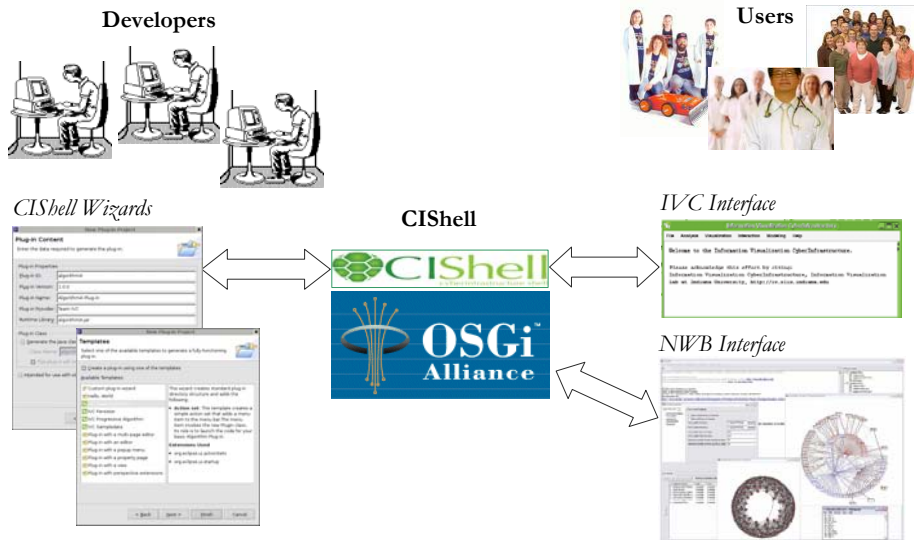
- Datasets, e.g., stored in files, databases, streaming data.
- Algorithms, e.g., data processing, analysis, modeling, visualization.
- Interfaces, e.g., remote services, scripting engines, peer-to-peer clients.
- Services, e.g., workflow support, scheduler.

Hence, it can be used for custom UI/Toolkit development.

Network Workbench: A Large-Scale Network Analysis, Modeling and Visualization Toolkit for Biomedical, Social Science and Physics Research. NSF IIS-0513650 award (Katy Börner, Albert-László Barabási, Santiago Schnell, Alessandro Vespignani & Stanley Wasserman, Eric Werner (Senior Personnel), \$1,120,926) Sept. 05 - Aug. 08.
<http://nwb.slis.indiana.edu>

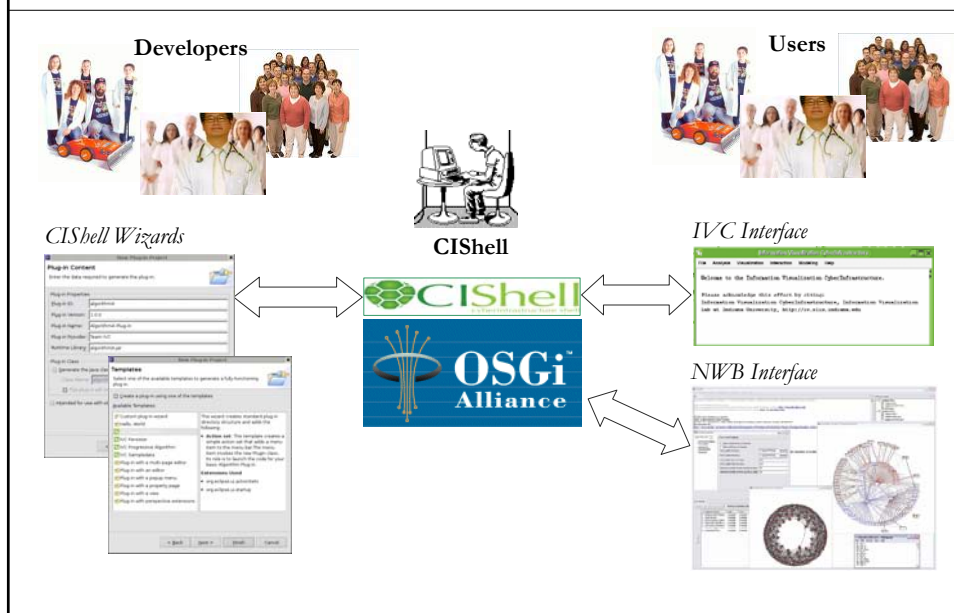


CIShell – Needs of Algorithm Developers & Users





CIShell – Needs of Algorithm Developers & Users



CIShell – Technical Details

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

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

- A standardized, component oriented, computing environment for networked services.
- Successfully used in the industry from high-end servers to embedded mobile devices since 7 years.
- Alliance members include IBM (Eclipse), Sun, Intel, Oracle, Motorola, NEC and many others.
- Widely adopted in open source realm, especially since Eclipse 3.0 that uses OSGi R4 for its plugin model.

Advantages of Using OSGi

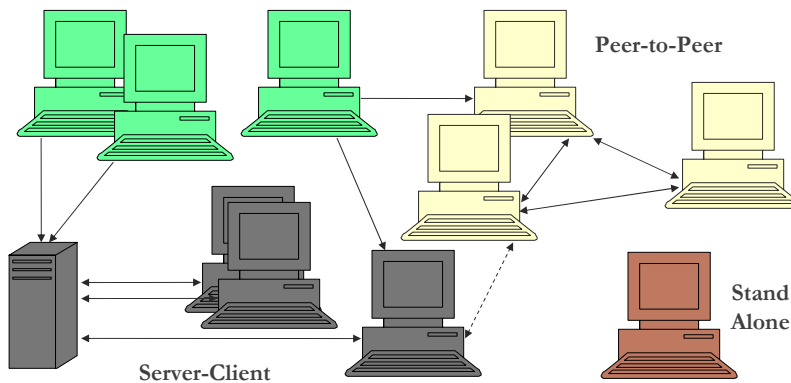
- Any CIShell algorithm is a service that can be used in any OSGi-framework based system.
- Using OSGi, running CIShells/tools can be connected via RPC/RMI supporting peer-to-peer sharing of data, algorithms, and computing power.

Ideally, CIShell becomes a standard for creating OSGi Services for algorithms. Developed Tools/CI, e.g., IVC & NWB, provide a reference GUI for underlying services.

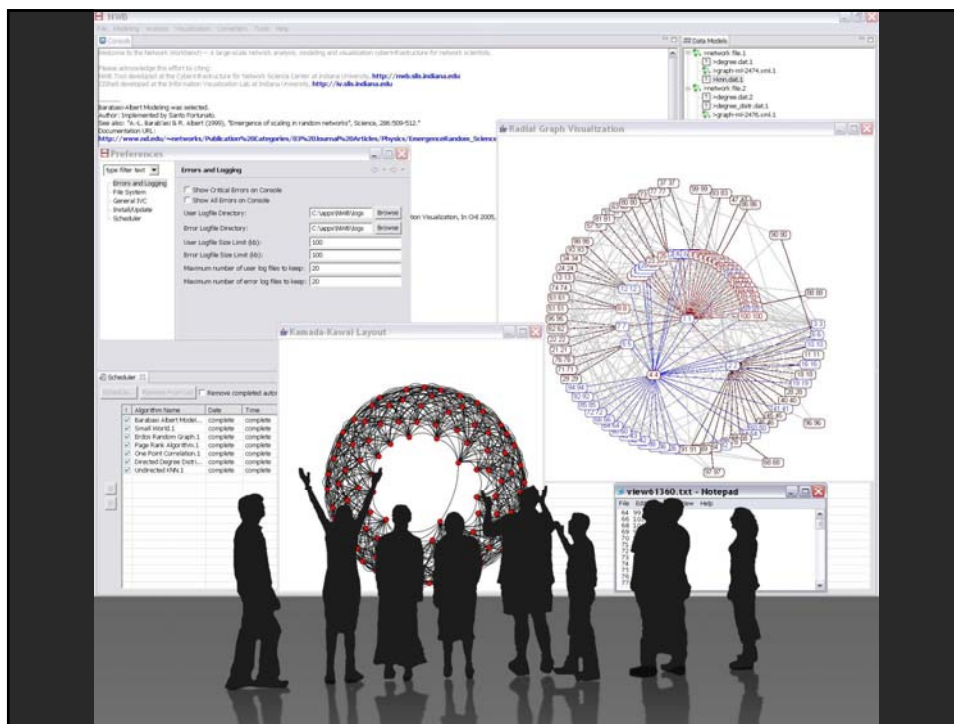
CIShell layer cake.



Data-Algorithm Repositories



CIShell applications can be deployed as distributed data and algorithm repositories, stand alone applications, peer-to-peer architectures, and server-client architectures.



Network Workbench (NWB)

Investigators: Katy Börner, Albert-Laszlo Barabasi, Santiago Schnell, Alessandro Vespignani & Stanley Wasserman, Eric Wernert



Software Team: Lead: Weixia (Bonnie) Huang
Developers: Bruce Herr, Ben Markines, Santo Fortunato, Cesar Hidalgo, Ramya Sabbineni, Vivek S. Thakre, & Russell Duhon



Goal: Develop a large-scale network analysis, modeling and visualization toolkit for biomedical, social science and physics research.

Amount: \$1,120,926 NSF IIS-0513650 award.

Duration: Sept. 2005 - Aug. 2008

Website: <http://nwb.slis.indiana.edu>





NWB Advisory Board

- James Hendler (Semantic Web)
- Jason Leigh (CI)
- Neo Martinez (Biology)
- Michael Macy, Cornell University (Sociology)
- Ulrik Brandes (Graph Theory)
- Mark Gerstein, Yale University (Bioinformatics)
- Stephen North (AT&T)
- Tom Snijders, University of Groningen (Social Science)



NWB CI Deliverables

Glue:

- CIShell Core programmer team lead by Bonnie Huang

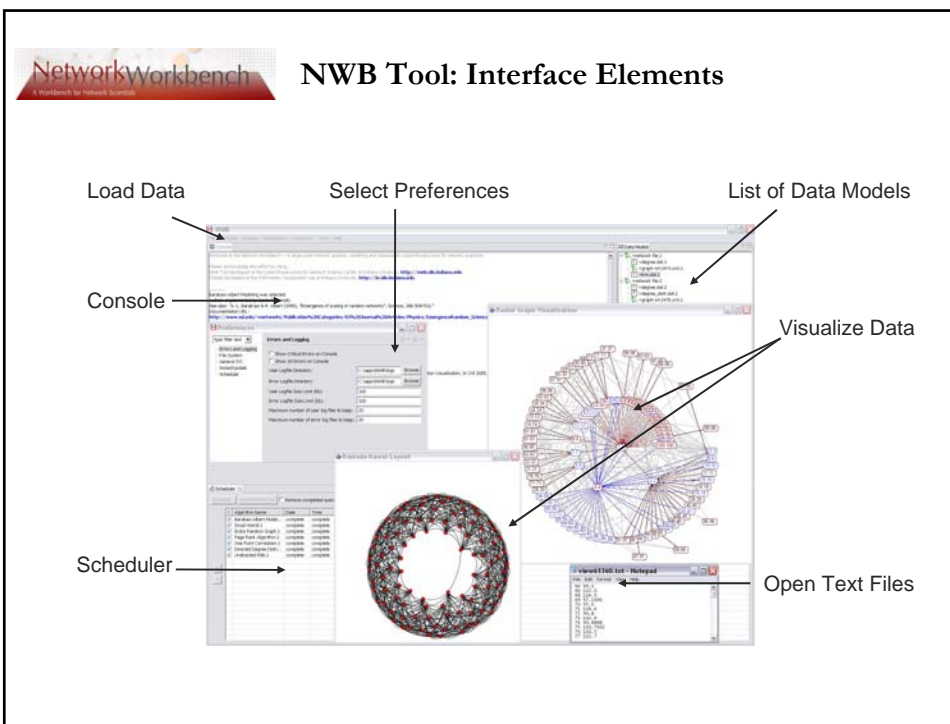
Tools, Services & Portals:

- NWB Tool Lead by Alex Vespignani with input from other PIs
- SciMaps Service Online Lead by Katy Borner
- Bio Portal Lead by Laszlo Barabasi & Santiago Schnell

All three are prototypical instantiations of CIShell serving as reference implementations.

Documentation/Registry/Market Place:

- NWB Community Wiki Lead by Katy Borner





NWB Tool 0.2.0: List of Algorithms

Category	Algorithm	Language	Analysis Algorithm	Language
Preprocessing	Directory Hierarchy Reader	JAVA	Attack Tolerance	JAVA
Modeling	Erdős-Rényi Random	FORTRAN	Error Tolerance	JAVA
	Barabási-Albert Scale-Free	FORTRAN	Betweenness Centrality	JAVA
	Watts-Strogatz Small World	FORTRAN	Site Betweenness	FORTRAN
	Chord	JAVA	Average Shortest Path	FORTRAN
	CAN	JAVA	Connected Components	FORTRAN
	Hypergrid	JAVA	Diameter	FORTRAN
	PRU	JAVA	Page Rank	FORTRAN
Visualization	Tree Map	JAVA	Shortest Path Distribution	FORTRAN
	Tree Viz	JAVA	Watts-Strogatz Clustering Coefficient	FORTRAN
	Radial Tree / Graph	JAVA	Watts-Strogatz Clustering Coefficient Versus Degree	FORTRAN
	Kamada-Kawai	JAVA	Directed k-Nearest Neighbor	FORTRAN
	Force Directed	JAVA	Undirected k-Nearest Neighbor	FORTRAN
	Spring	JAVA	Indegree Distribution	FORTRAN
	Fruchterman-Reingold	JAVA	Outdegree Distribution	FORTRAN
	Circular	JAVA	Node Indegree	FORTRAN
	Parallel Coordinates (demo)	JAVA	Node Outdegree	FORTRAN
Tool	XMGrace		One-point Degree Correlations	FORTRAN
			Undirected Degree Distribution	FORTRAN
			Node Degree	FORTRAN
			k Random-Walk Search	JAVA
			Random Breadth First Search	JAVA
			CAN Search	JAVA

The screenshot shows the NetworkWorkbench website interface. The main content area is titled "Master List of Algorithms" and includes a green icon indicating availability in the nwb 0.2.0 release. Below this, there is a "Load Data" section with a list of data formats: IXI², NWB, Pajek (.net), GraphML (.xml), and XGMML. There are also sections for "Databases", "Streaming Data", "Sample Data", and "Sampling". A diagram on the right side of the page shows a network graph with nodes labeled "Jung", "text/NWB", "NWBModel", "A Prefuse", "B Prefuse", and "Pajek", connected by arrows. The browser window title is "NWB Community Wiki: Home Page browse - Netscape" and the address bar shows "https://nwb.slis.indiana.edu/community/?n=Algorithms.HomePage".

<https://nwb.slis.indiana.edu/community>

A Potential Future: Science Maps in Action

KIDS first ...



All maps of science are on sale
via <http://scimaps.org>

TOPIC MAP: HOW SCIENTIFIC PARADIGMS RELATE

GEOGRAPHIC MAP: WHERE SCIENCE GETS DONE

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

Nanotechnology

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

All Topics	Nanotechnology	Francis H. C. CRICK	Albert EINSTEIN	Michael E. FISHER	Susan T. FISKE
<i>Sweep through all 776 scientific paradigms</i>	<i>Science on the tiny scale of molecules</i>	<i>Co-discovered DNA's double helix</i>	<i>Revitalized physics with Relativity theories</i>	<i>Models critical phase transitions of matter</i>	<i>Connects perception and stereotypes</i>
Sustainability	Biology & Chemistry	Joshua LEDERBERG	Derek J. de Solla PRICE	Richard N. ZARE	About this display
<i>The science behind our long-term hopes</i>	<i>The interface between these two vital fields</i>	<i>Pioneer in bacterial genetic mechanisms</i>	<i>Known as the "Father of Scientometrics"</i>	<i>Uses laser chemistry in molecular dynamics</i>	<i>People & organizations that helped create it</i>

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


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

Inventors & Inventions


*Science Puzzle Map for Kids by Fileve Palmer, Julie Smith, Elisha Hardy and Katy Börner, Indiana University, 2006.
 (Base map taken from Illuminated Diagram display by Kevin Boyack, Richard Klavans, and W. Bradford Paley.)*








My Science Story
By _____



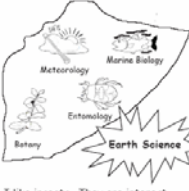
There are seven main fields of science. They are...



social science, mathematics, physics, chemistry, earth science, medicine, and psychology. I like to study earth science.

Color earth science green.

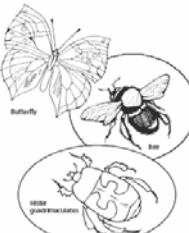
Earth scientists study the weather, plants and trees, marine life, insects, and much more.



I like insects. They are interesting to look at and study.


Color in the insect.

For more information about the map of science for kids or this activity, please contact Kara Bomer (kbomer@indiana.edu) or Tavis Holary (tholary@indiana.edu) at the School of Library and Information Science, Indiana University. These materials were compiled by Tavis Holary in 2006.




Butterfly
Bee
Scorpion

There are many types of insects in the world. Bees, butterflies, and beetles are just a few.



I want to be an **entomologist** when I grow up. Then I can study insects all the time.

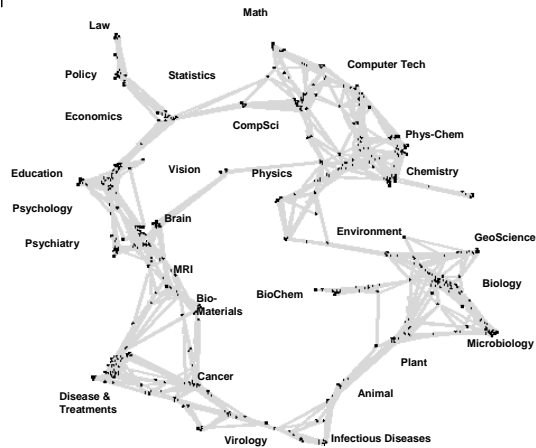
... my **SPONSORS** next ...



Latest 'Base Map' of Science

Kevin W. Boyack & Richard Klavans, unpublished work.

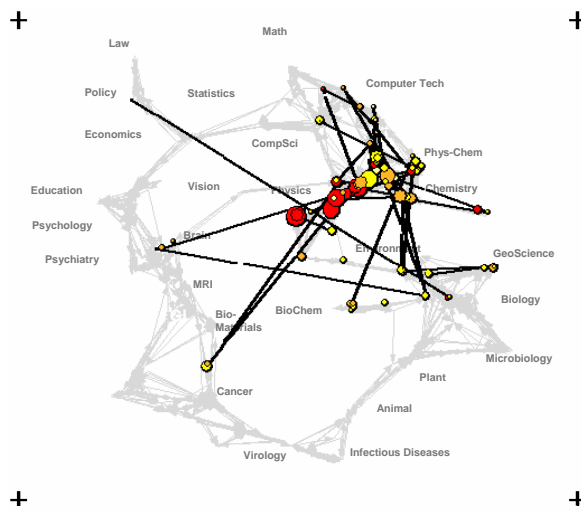
- Uses combined SCI/SSCI from 2002
 - 1.07M papers, 24.5M references, 7,300 journals
 - Bibliographic coupling of papers, aggregated to journals
- Initial ordination and clustering of journals gave 671 clusters
- Coupling counts were reaggregated at the journal cluster level to calculate the
 - (x,y) positions for each journal cluster
 - by association, (x,y) positions for each journal



Science map applications: Identifying core competency

Kevin W. Boyack & Richard Klavans, unpublished work.

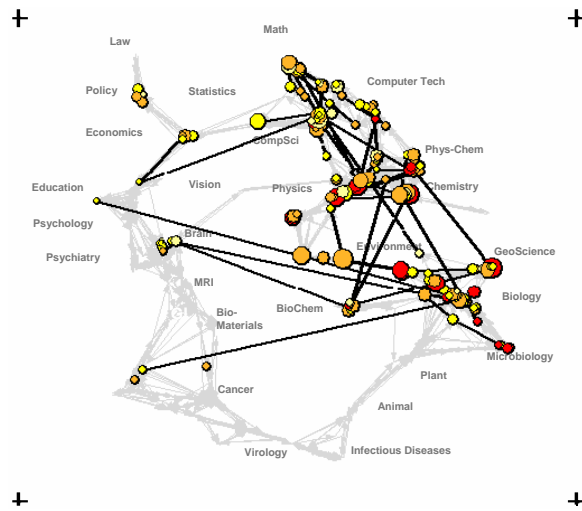
Funding patterns of the US Department of Energy (DOE)



Science map applications: Identifying core competency

Kevin W. Boyack & Richard Klavans, unpublished work.

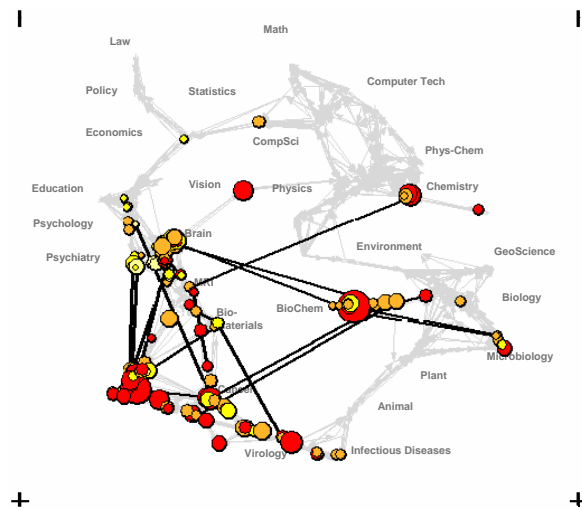
Funding Patterns of the National Science Foundation (NSF)



Science map applications: Identifying core competency

Kevin W. Boyack & Richard Klavans, unpublished work.

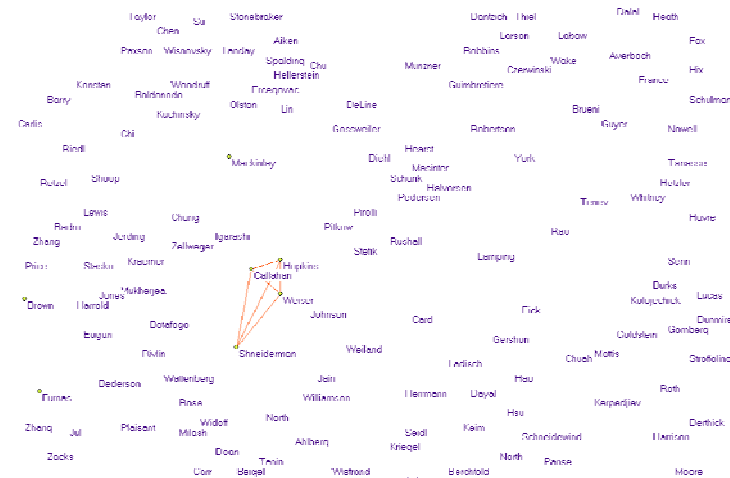
Funding Patterns of the National Institutes of Health (NIH)



... then SCIENTISTS ...

Mapping the Evolution of Co-Authorship Networks

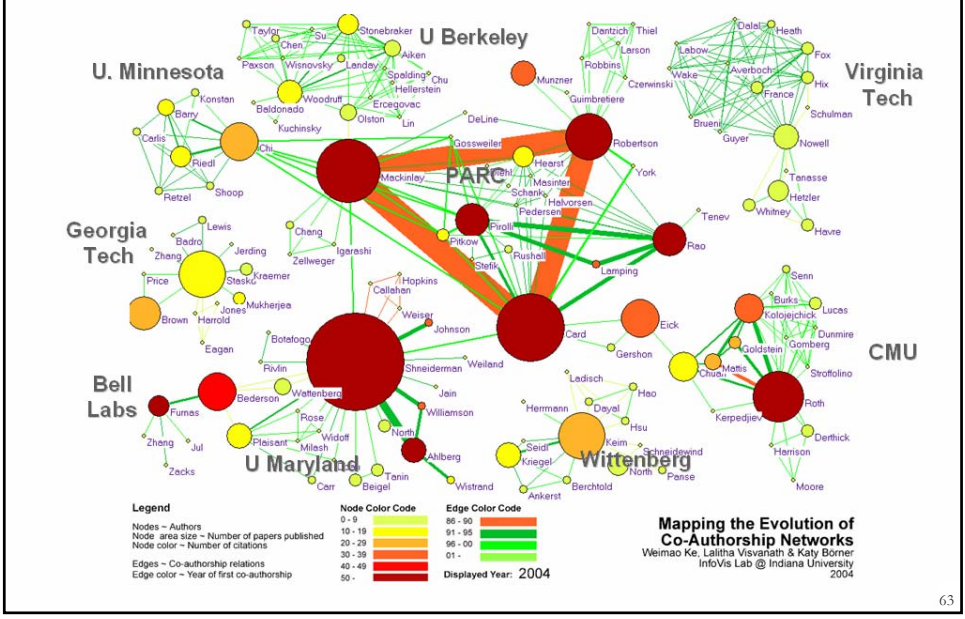
Ke, Viswanath & Börner, (2004) Won 1st price at the IEEE InfoVis Contest.



Mapping the Evolution of Co-Authorship Networks
 Wernoo Ke, Latha Viswanath & Katy Börner
 InfoVis Lab @ Indiana University
 2004

Mapping the Evolution of Co-Authorship Networks

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



63



Lab/Center Management System

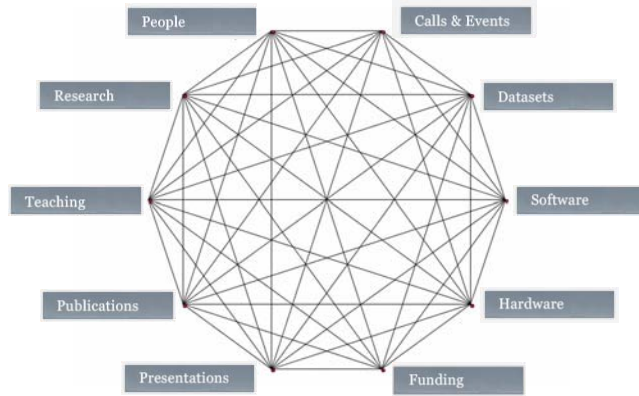
- People
- Research
- Teaching
- Publications
- Presentations

- Calls & Events
- Datasets
- Software
- Hardware
- Funding

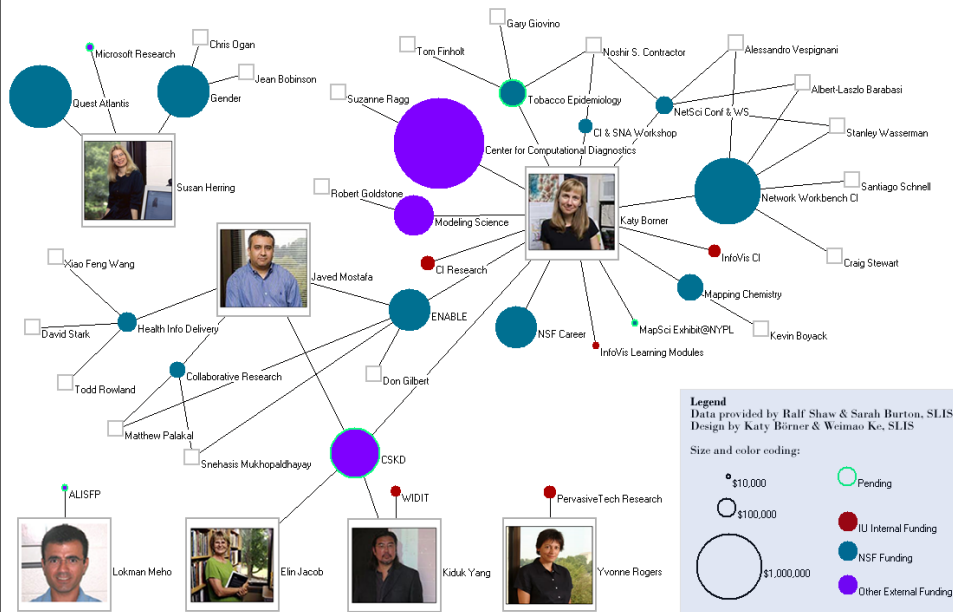
<https://inl.slis.indiana.edu>



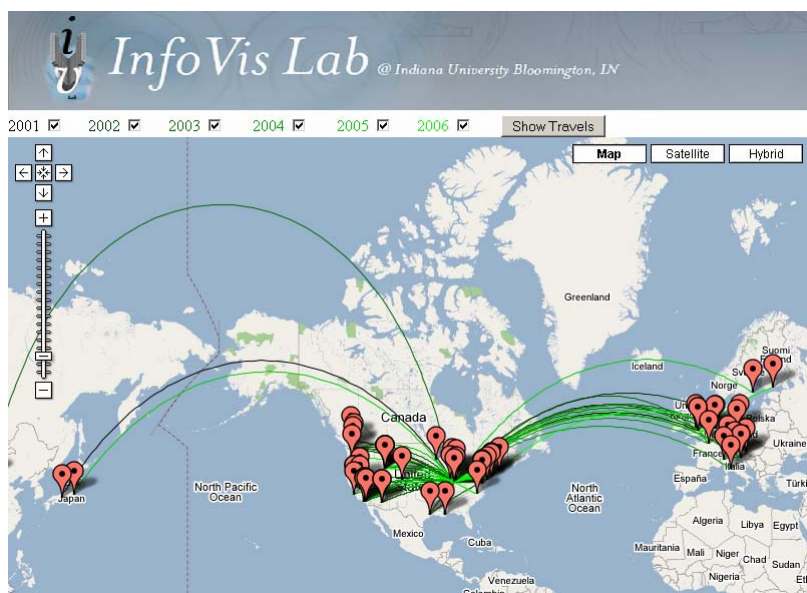
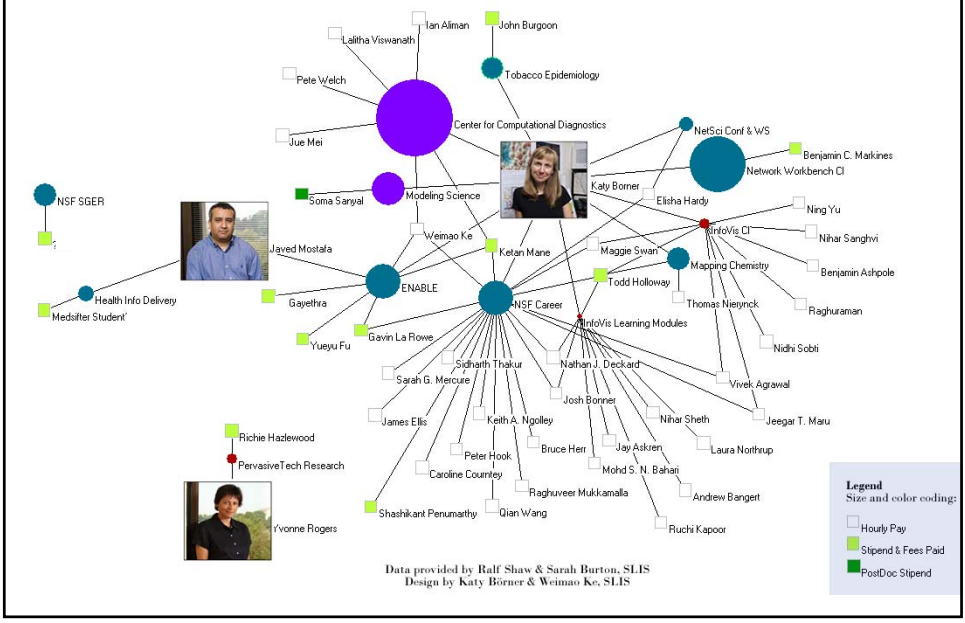
Data Entities and Interlinkages

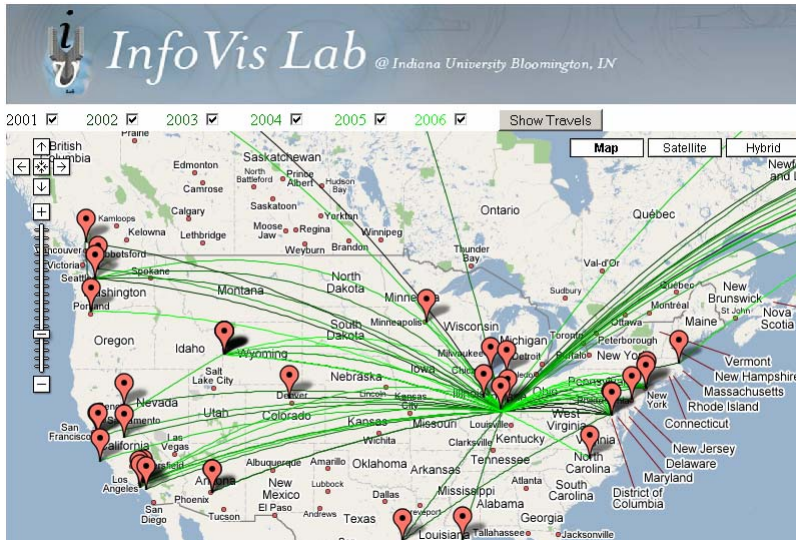


Awarded & Pending Funding of SLIS Faculty Members in January 2006

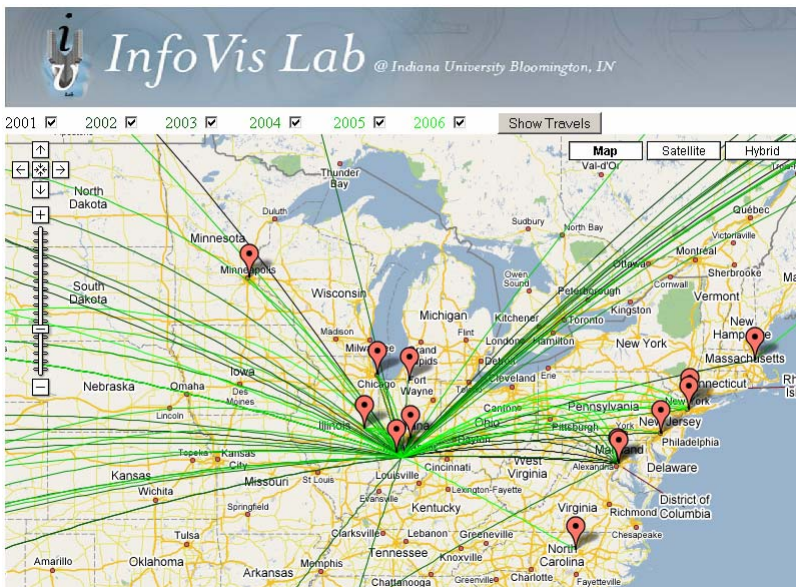


Student Support by SLIS Faculty Members in January 2006

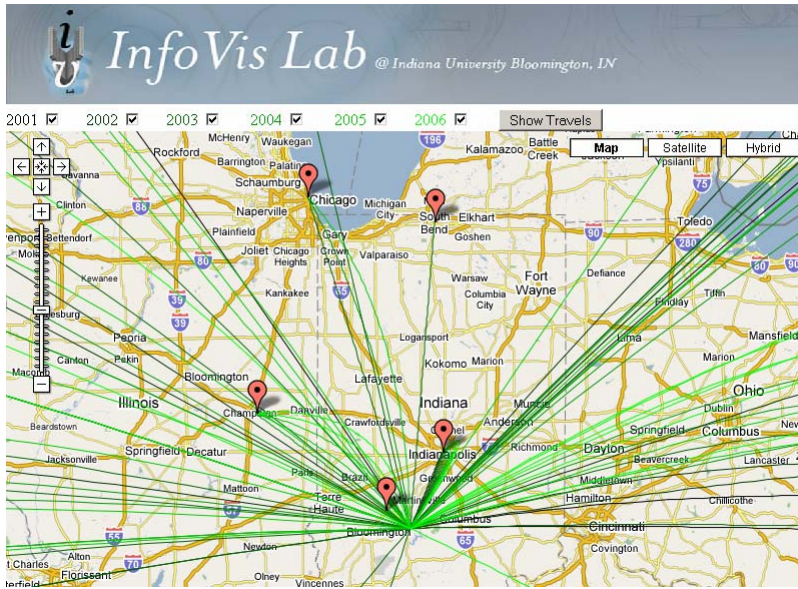




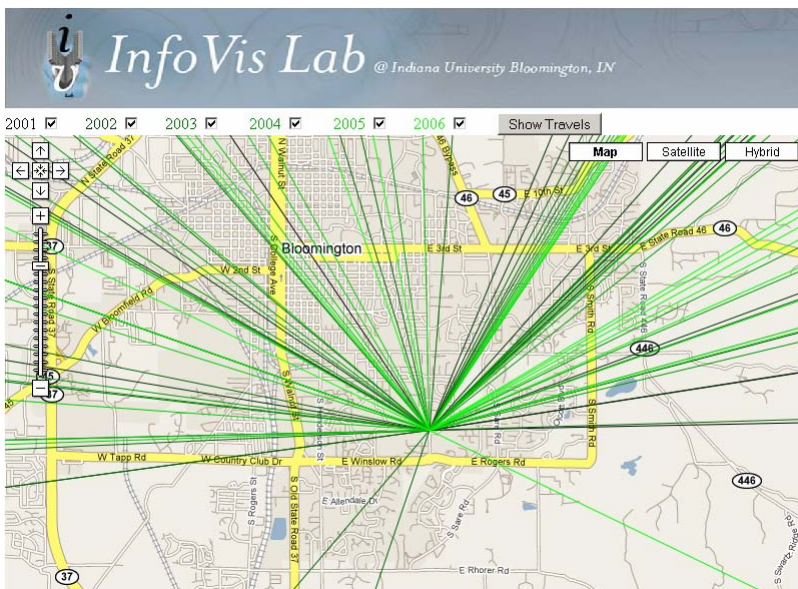
By Thomas Neiryck, 2007.



By Thomas Neiryck, 2007.

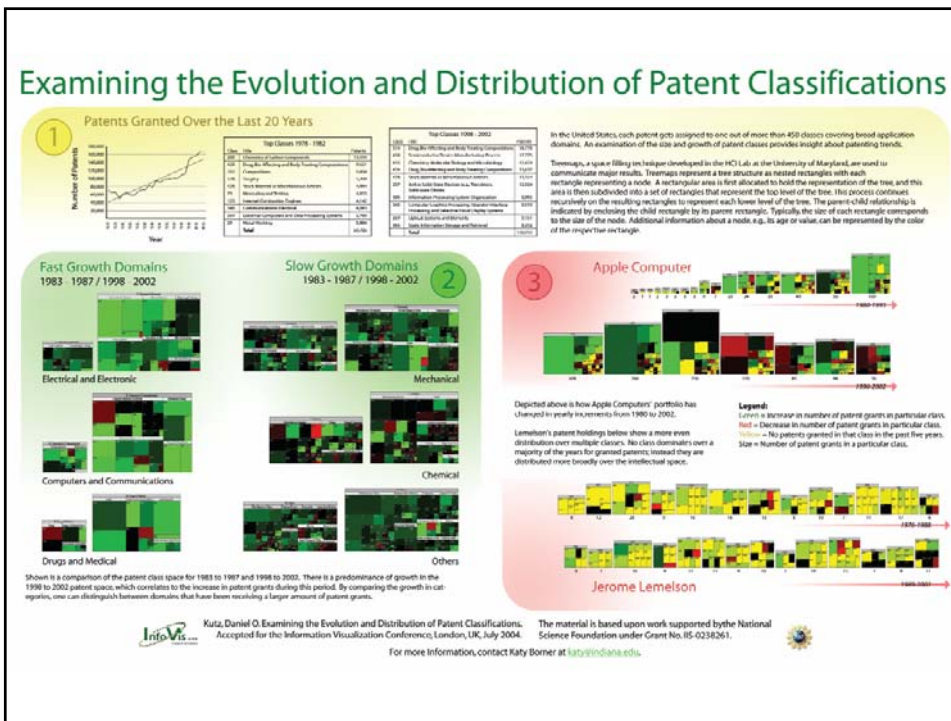


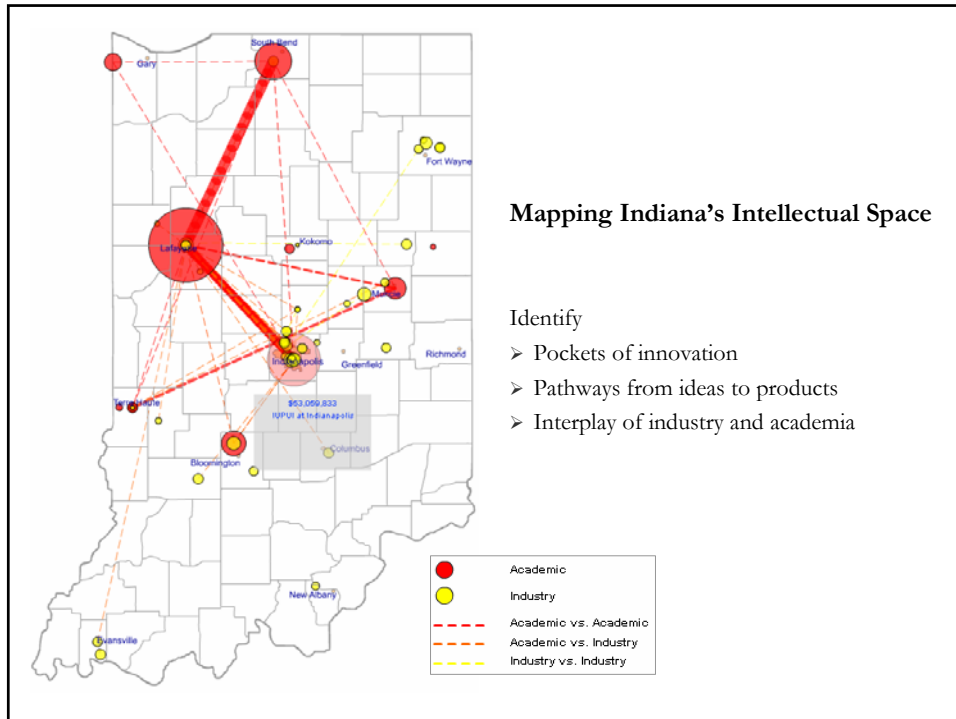
By Thomas Neiryck, 2007.



By Thomas Neiryck, 2007.

... and INDUSTRY too.



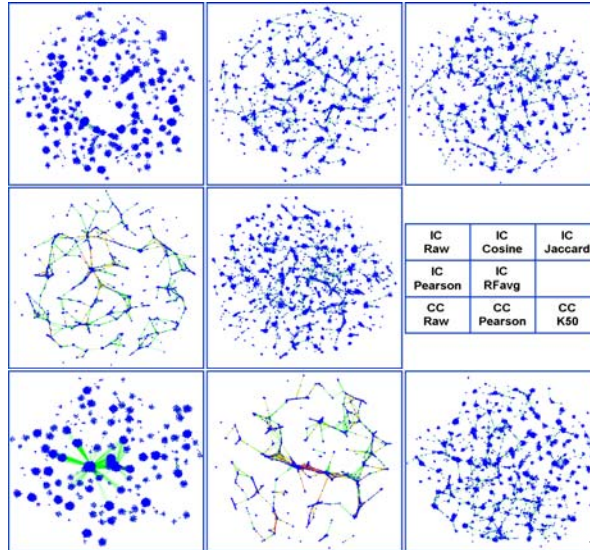


The End.

Comparison of Similarity Metrics



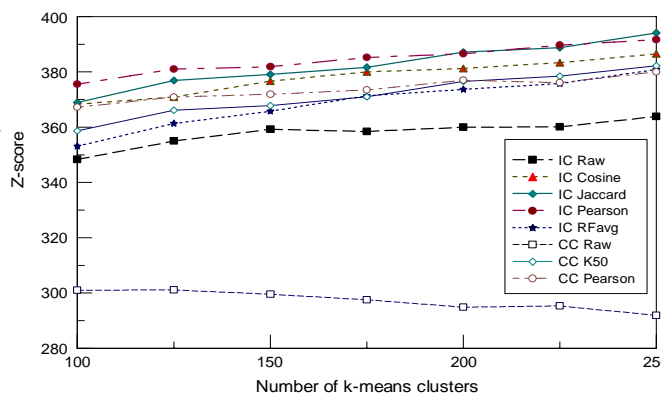
- ISI file year 2000, SCI and SSCI: 7,121 journals.
- Different similarity metrics
 - Inter-citation (raw counts, cosine, modified cosine, Jaccard, RF, Pearson)
 - Co-citation (raw counts, cosine, modified cosine, Pearson)
- Maps were compared based on
 - regional accuracy,
 - the scalability of the similarity algorithm, and
 - the readability of the layouts.



Boyack, Kevin W., Klavans, R. and Börner, Katy. (2005). *Mapping the Backbone of Science*. *Scientometrics*. 64(3), 351-374.

Selecting the similarity measure with the best regional accuracy

- For each similarity measure, the VxOrd layout was subjected to k-means clustering using different numbers of clusters.
- Resulting cluster/category memberships were compared to actual category memberships using entropy/mutual information method by Gibbons & Roth, 2002.
- Increasing Z-score indicates increasing distance from a random solution.
- Most similarity measures are within several percent of each other.



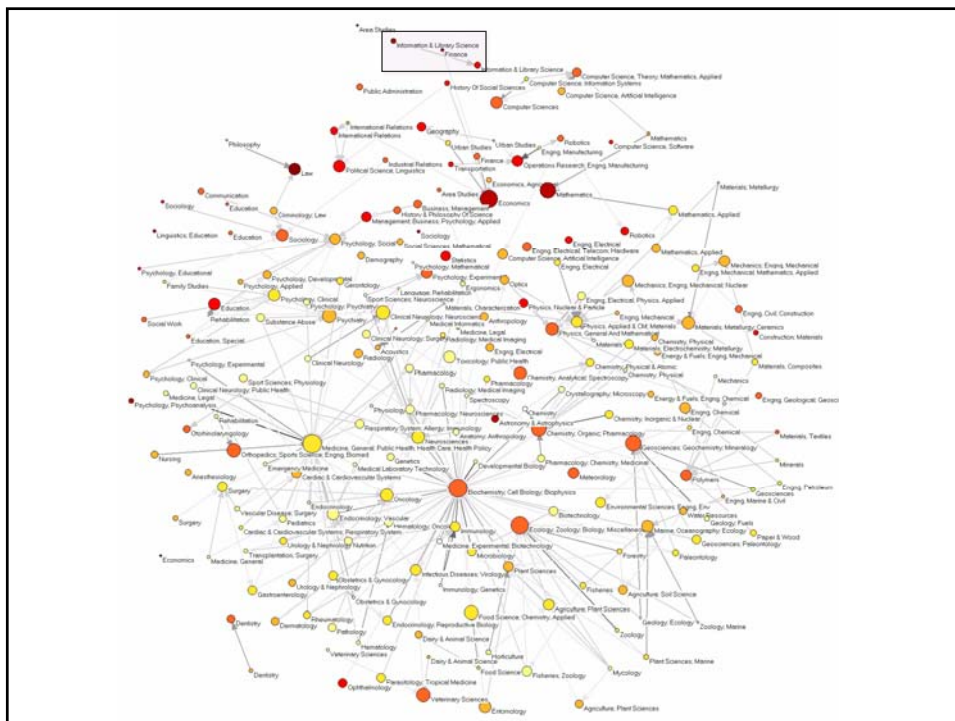
Boyack, Kevin W., Klavans, R. and Börner, Katy. (2005). *Mapping the Backbone of Science*. *Scientometrics*. 64(3), 351-374.

Structural Map: Studying Disciplinary Diffusion

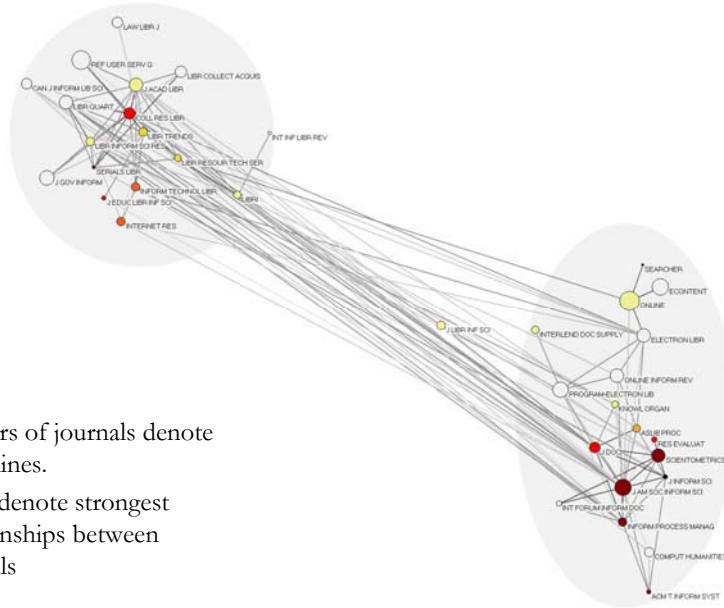
- The 212 nodes represent clusters of journals for different disciplines.
- Nodes are labeled with their dominant ISI category name.
- Circle sizes (area) denote the number of journals in each cluster.
- Circle color depicts the independence of each cluster, with darker colors depicting greater independence.
- Lines denote strongest relationships between disciplines (citing cluster gives more than 7.5% of its total citations to the cited cluster).



81



Zoom Into Structural Map



- Clusters of journals denote disciplines.
- Lines denote strongest relationships between journals

83

Information Visualization CyberInfrastructure

The InfoVis CyberInfrastructure provides access to data, software code and learning modules as well as computing resources in support of the analysis, modeling and visualization of diverse data sets.

DATABASES

An Oracle database provides access to publications, papers, grants and grant opportunities. The database is continuously and automatically updated. It includes:

- Publications
- Grants
- Grant Opportunities
- Publications
- Grants
- Grant Opportunities

COMPUTING RESOURCES

The InfoVis CyberInfrastructure is hosted at Indiana University's Research Database Complex, consisting of three Sun T3000 servers with 12 GB RAM, 100 MB cache and 100 GB of disk space. A Sun 1000 server acts as the central CPU server, which acts as the database server.

SOFTWARE

An open source framework was designed to facilitate the integration of diverse data analysis, modeling and visualization algorithms. The algorithms, data and visualization methods form and work for the InfoVis CyberInfrastructure can be easily "plugged in" or "unplugged" from the framework.

LEARNING MODULES

A set of associated learning modules aims to equip learners with a practical skill set and the underlying methods and software to quickly identify and use different algorithms and diverse visualization techniques and data visualization tools for the generation and comparison of network visualizations.

CAREER: Visualizing Knowledge Domains. NSF IIS-0238261 award

(Katy Börner, \$440,000) Sept. 03-Aug. 08.

<http://iiv.slis.indiana.edu/>

NetworkWorkbench

A Workbench for Network Scientists



SEI: Network Workbench: A Large-Scale Network Analysis, Modeling and Visualization Toolkit for Biomedical, Social Science and Physics Research. NSF IIS-0513650 award (Katy Börner, Albert-László Barabási, Santiago Schnell, Alessandro Vespignani & Stanley Wasserman, Craig Stewart (Senior Personnel), \$1,120,926) Sept. 05 - Aug. 08. <http://nwb.slis.indiana.edu>