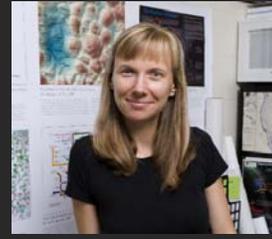


Communicating the Structure and Evolution of Science

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



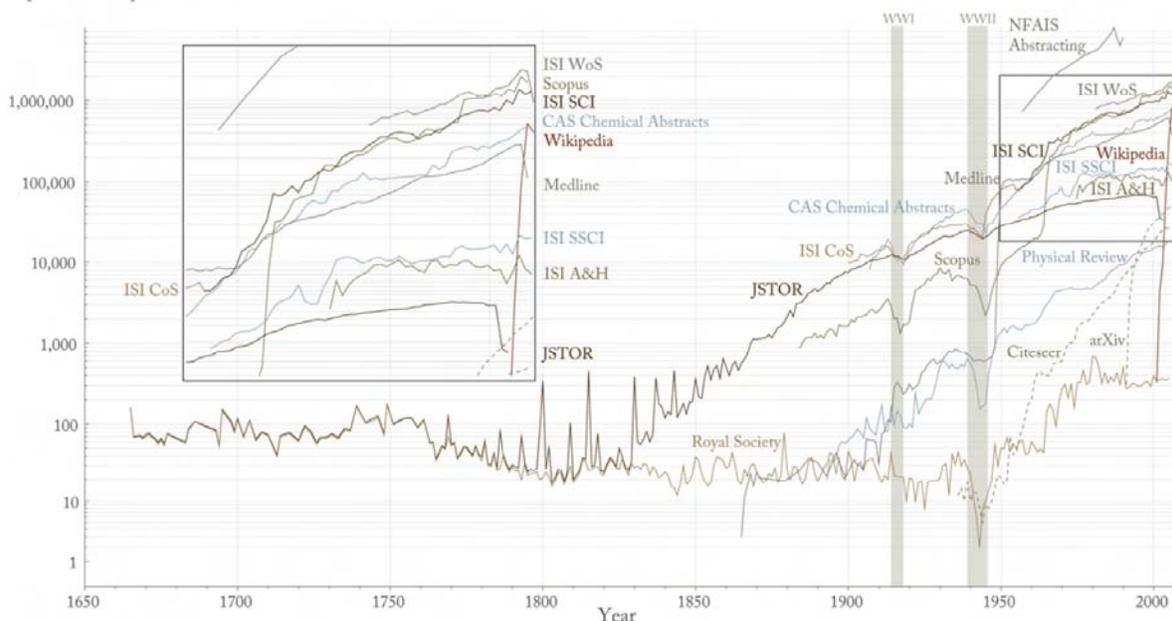
*SAP Media X 2009 Winter Quarter Seminar Series
 Y2E2 Room 292A, Jerry Yang and Akiko Yamazaki Environment & Energy Building
 Stanford University*

Wednesday, January 21, 2009



Growth of Scientific Knowledge, 1665 to 2006

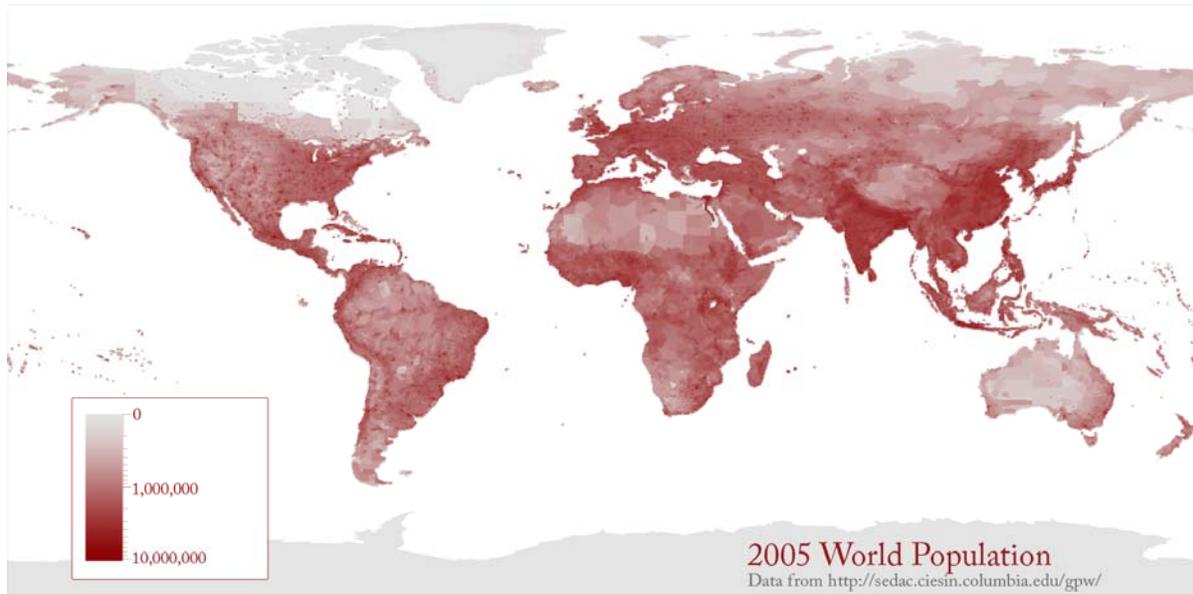
Papers & Wikipedia Entries



*'Atlas of Science: Guiding the Navigation and Management of Scholarly Knowledge', Part E: The Rise of Science and Technology.
 Chart showing the number of papers/wikipedia entries for different databases and publication years.
 Contact Katy Börner <katy@indiana.edu> or Elisha Hardy <ehardy@indiana.edu> for details.*

2005 World Population

The population map uses a quarter degree box resolution. Boxes with zero people are given in white. Darker shades of red indicate higher population counts per box using a logarithmic interpolation. The highest density boxes appear in Mumbai, with 11,687,850 people in the quarter degree block, Calcutta (10,816,010), and Shanghai (8,628,088).



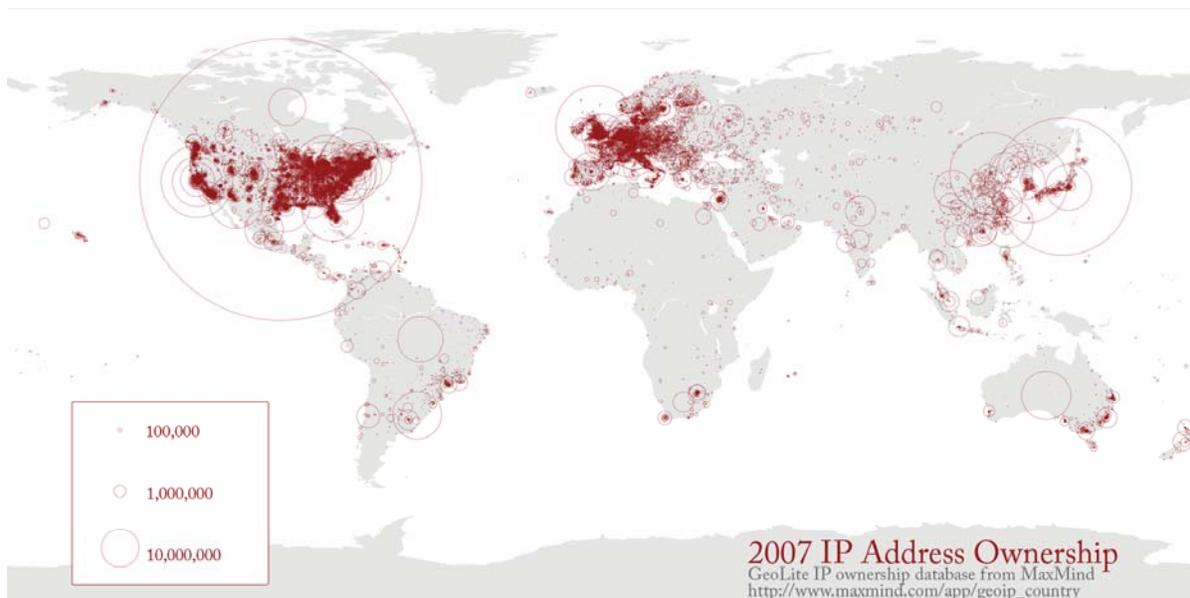
2000 Night on Earth

This image shows city lights at night. It was composed from hundreds of pictures made by orbiting satellites. The seaboard of Europe, the eastern United States, and Japan are particularly well lit. Many cities exist near rivers or oceans so that goods can be exchanged cheaply by boat. The central parts of South America, Africa, and Australia are rather dark despite their high population density, see map to the left.



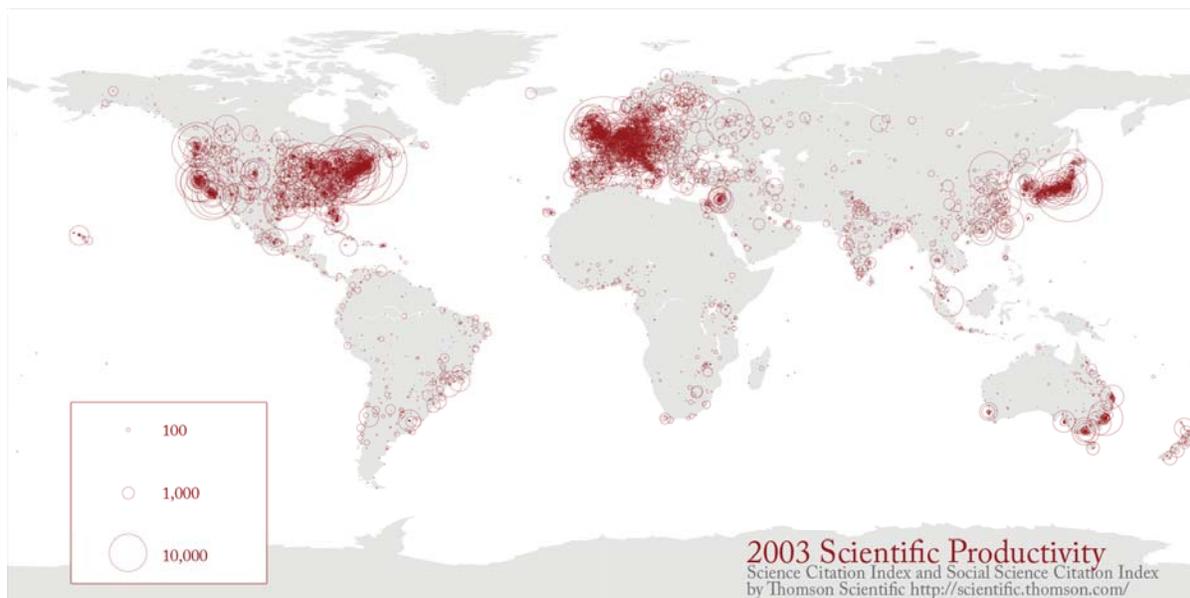
2007 IP Address Ownership

This map shows IP address ownership by location. Each owner is represented by a circle and the area size of the circle corresponds to the number of IP addresses owned. The largest circle denotes MIT's holdings of an entire class A subnet, which equates to 16,581,375 IP addresses. The countries that own the most IP addresses are US (560 million), Japan (130 million), Great Britain (47 million).

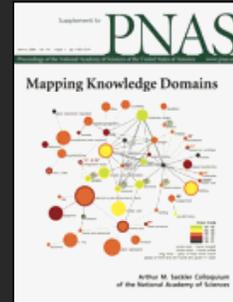
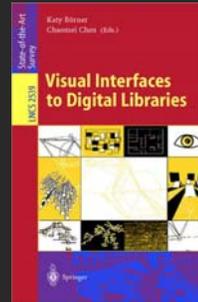


2003 Scientific Productivity

Shown is where science is performed today. Each circle indicates a geographic location at which scholarly papers are published. The larger the circle the more papers are produced. Boston, MA, London, England, and New York, NY are the top three paper production areas. Note the strong resemblance with the Night on Earth and the IP Ownership maps and the striking differences to the world population map.



Computational Scientometrics: Studying Science by Scientific Means



- Börner, Katy, Chen, Chaomei, and Boyack, Kevin. (2003). *Visualizing Knowledge Domains*. In Blaise Cronin (Ed.), *Annual Review of Information Science & Technology*, Medford, NJ: Information Today, Inc./American Society for Information Science and Technology, Volume 37, Chapter 5, pp. 179-255. <http://ivl.slis.indiana.edu/km/pub/2003-borner-arist.pdf>
- Shiffrin, Richard M. and Börner, Katy (Eds.) (2004). *Mapping Knowledge Domains*. *Proceedings of the National Academy of Sciences of the United States of America*, 101(Suppl_1). http://www.pnas.org/content/vol101/suppl_1/
- Börner, Katy, Sanyal, Soma and Vespignani, Alessandro (2007). *Network Science*. In Blaise Cronin (Ed.), *Annual Review of Information Science & Technology*, Medford, NJ, Volume 41, Chapter 12, pp. 537-607. <http://ivl.slis.indiana.edu/km/pub/2007-borner-arist.pdf>
- *Places & Spaces: Mapping Science* exhibit, see also <http://scimaps.org>.

Chart toppers

An exhibition explores the diverse ways of putting data on the map.

Map of the World
How often do you see the world on a map? Not often enough. But there are several ways of doing it—about 200 different ways. This exhibit explores the diverse ways of putting data on the map. It features a variety of maps, including a world map, a map of the United States, and a map of the world's oceans. The exhibit is designed to be interactive, allowing visitors to explore the maps and learn more about the data they represent.

BRILLIANT DISPLAY

From a jewel-like bird, rarer than any diamond, to the delicately poetic swirls generated inside aircraft engines, the pursuit of knowledge turns up its fair share of beauty. This issue, Nature wraps up the year with an arresting series of images from 2006. We've divided them into the art of the natural world, planet-scapes both domestic and extraterrestrial, and the splendor of modern technology. Just because something enhances our knowledge doesn't mean it can't also be bewitching.

Researched and written by Emma Miskin.

How Scientific Paradigms Relate

Journal written by Allen54 (180860) and posted by kdawson on Tuesday March 20, 2007 in the connections dept.

There is a giant chart mapping relationships among scientific paradigms, as published in the journal Nature. This map was constructed by sorting roughly 800,000 published papers into 776 different scientific paradigms (shown as pale circular nodes) based on how often the papers were cited together by authors of other papers. Information Esthetics, an organization founded by map co-creator Bradford Paltridge, is the organization that makes this chart. The chart makes these connections.

The map was constructed by sorting roughly 800,000 published papers into 776 different scientific paradigms (shown as pale circular nodes) based on how often the papers were cited together by authors of other papers. Lines (called black lines) were made between the paradigms that shared papers, then treated as rubber bands, holding similar paradigms near one another.

<http://scimaps.org>

“Science of Science” Opportunities

Advantages for Funding Agencies

- Supports monitoring of (long-term) money flow and research developments, evaluation of funding strategies for different programs, decisions on project durations, funding patterns.
- Staff resources can be used for scientific program development, to identify areas for future development, and the stimulation of new research areas.

Advantages for Researchers

- Easy access to research results, relevant funding programs and their success rates, potential collaborators, competitors, related projects/publications (**research push**).
- More time for research and teaching.

Advantages for Industry

- Fast and easy access to major results, experts, etc.
- Can influence the direction of research by entering information on needed technologies (**industry-pull**).

Advantages for Publishers

- Unique interface to their data.
- Publicly funded development of databases and their interlinkage.

For Society

- Dramatically improved access to scientific knowledge and expertise.

Process of Analyzing and Mapping Knowledge Domains

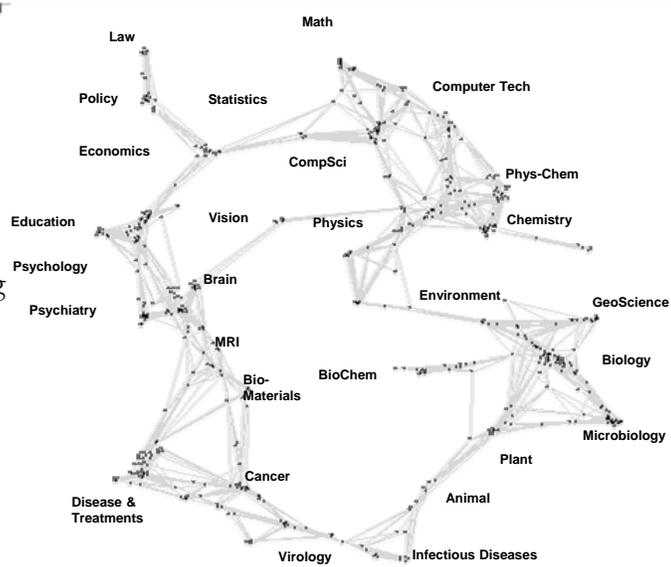
DATA EXTRACTION	UNIT OF ANALYSIS	MEASURES	LAYOUT (often one code does both similarity and ordination steps)		DISPLAY
			SIMILARITY	ORDINATION	
SEARCHES ISI INSPEC Eng Index Medline ResearchIndex Patents etc.	COMMON CHOICES Journal Document Author Term	COUNTS/FREQUENCIES Attributes (e.g. terms) Author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

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.

Latest 'Base Map' of Science

Kevin W. Boyack, Katy Börner, & Richard Klavans (2007). *Mapping the Structure and Evolution of Chemistry Research*. 11th International Conference on Scientometrics and Informetrics. pp. 112-123.

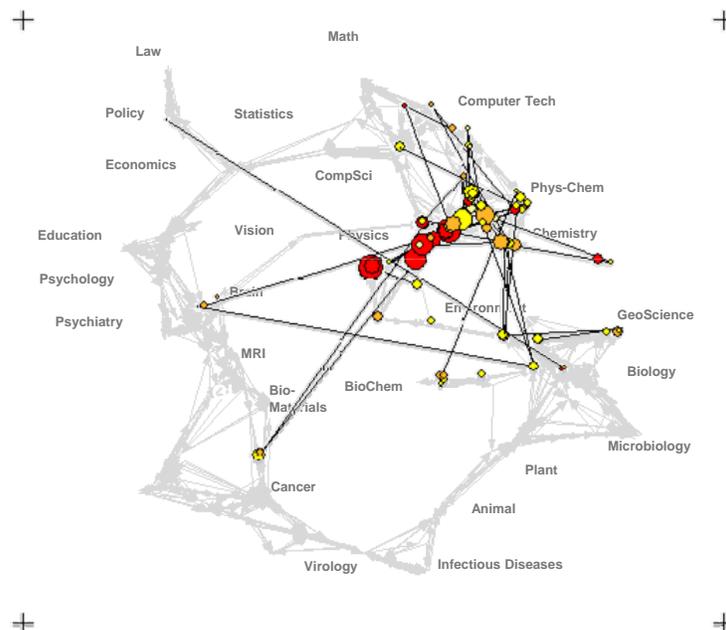
- Uses combined SCI/SSCI from 2002
 - 1.07M papers, 24.5M references, 7,300 journals
 - Bibliographic coupling of papers, aggregated to journals
- Initial ordination and clustering of journals gave 671 clusters
- Coupling counts were reaggregated at the journal cluster level to calculate the
 - (x,y)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



Science map applications: Identifying core competency

Kevin W. Boyack, Katy Börner, & Richard Klavans (2007).

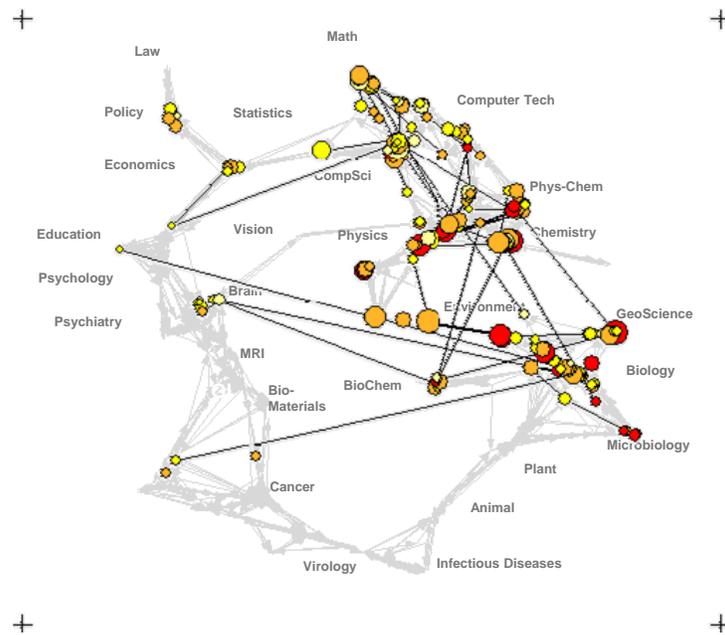
Funding patterns of the US Department of Energy (DOE)



Science map applications: Identifying core competency

Kevin W. Boyack, Katy Börner, & Richard Klavans (2007).

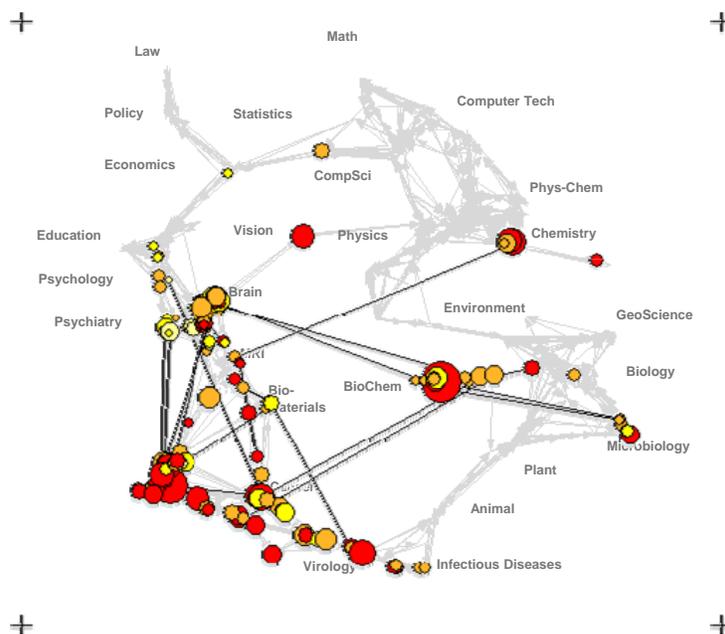
Funding Patterns of the National Science Foundation (NSF)



Science map applications: Identifying core competency

Kevin W. Boyack, Katy Börner, & Richard Klavans (2007).

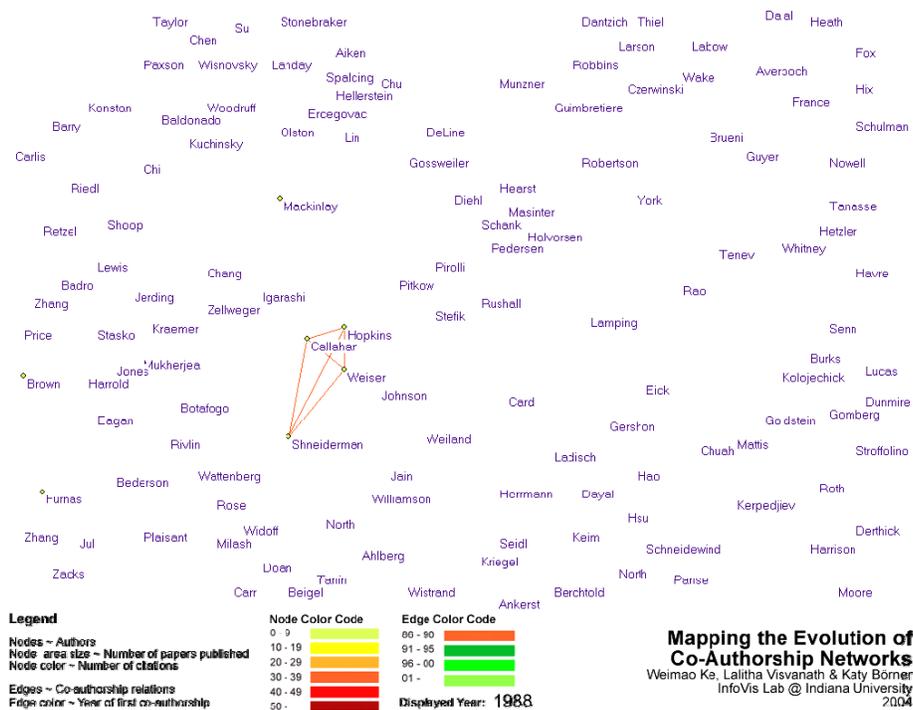
Funding Patterns of the National Institutes of Health (NIH)



Sample Science Studies

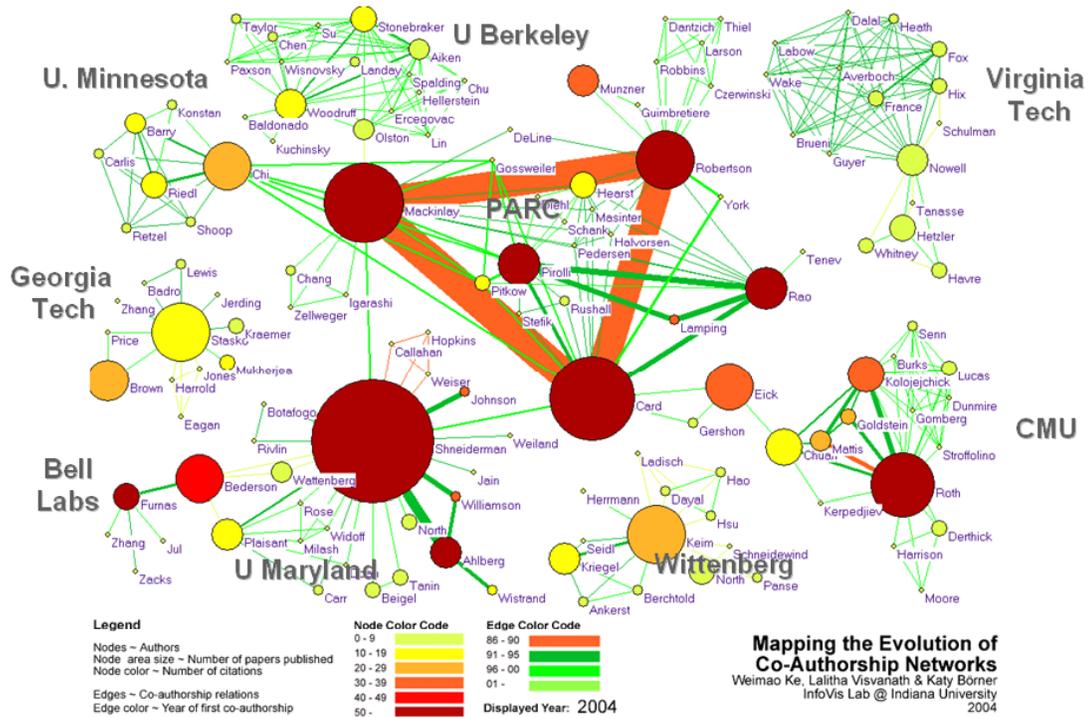
Mapping the Evolution of Co-Authorship Networks

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



Mapping the Evolution of Co-Authorship Networks

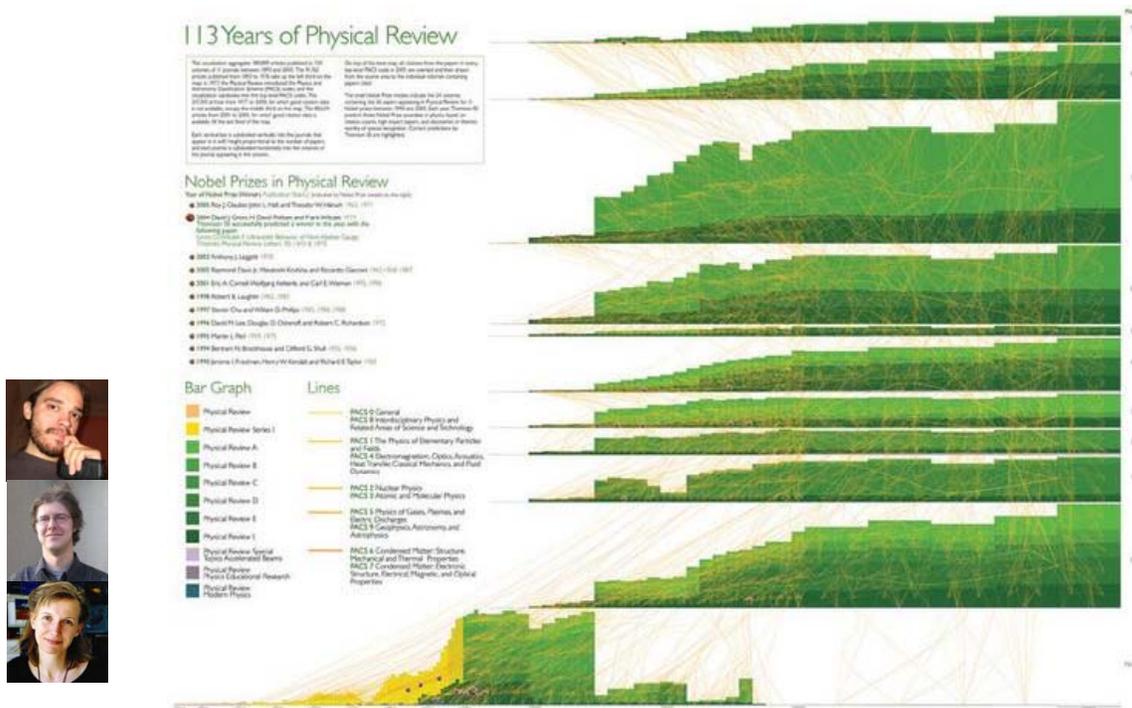
Ke, Viswanath & Börner, (2004) Won 1st price at the IEEE InfoVis Contest



113 Years of Physical Review

http://scimaps.org/dev/map_detail.php?map_id=171

Bruce W. Herr II and Russell Dabon (Data Mining & Visualization), Elisha F. Hardy (Graphic Design), Shashikant Penumarthy (Data Preparation) and Katy Börner (Concept)



Spatio-Temporal Information Production and Consumption of Major U.S. Research Institutions

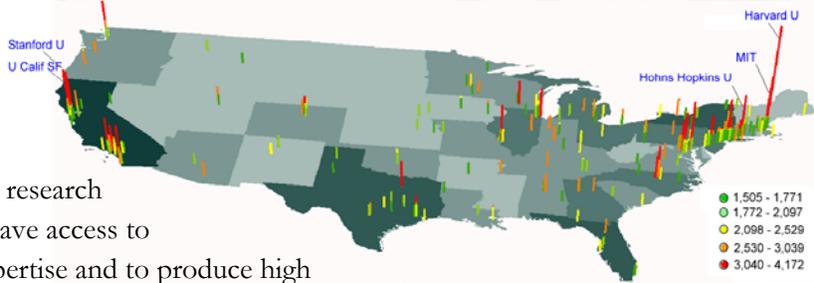
Research Institutions

Börner, Katy, Penumarty, Shashikant, Meiss, Mark and Ke, Weimao. (2006) *Mapping the Diffusion of Scholarly Knowledge Among Major U.S. Research Institutions. Scientometrics. 68(3), pp. 415-426.*



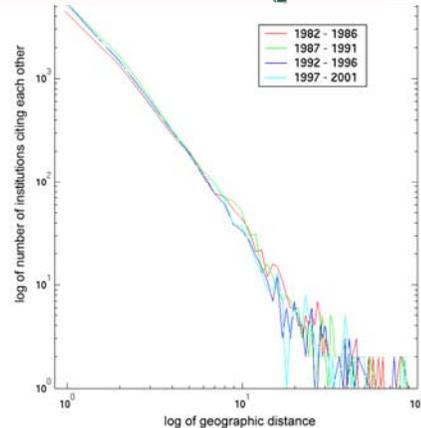
Research questions:

1. Does space still matter in the Internet age?
2. Does one still have to study and work at major research institutions in order to have access to high quality data and expertise and to produce high quality research?
3. Does the Internet lead to more global citation patterns, i.e., more citation links between papers produced at geographically distant research institutions?



Contributions:

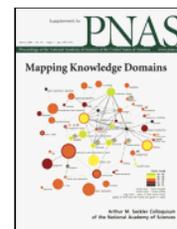
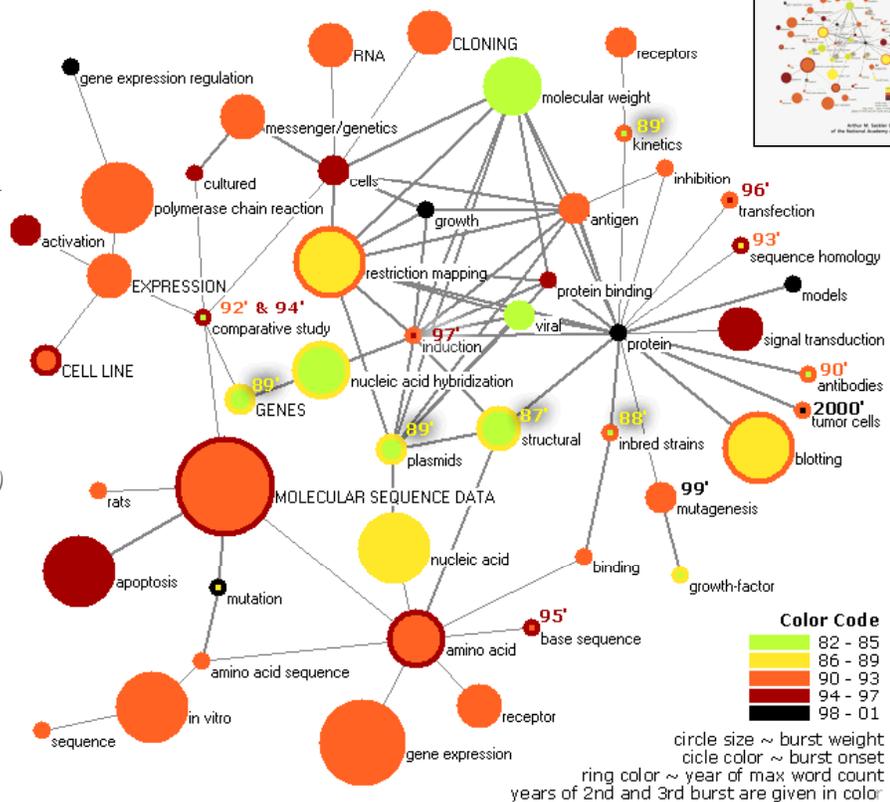
- Answer to Qs 1 + 2 is YES.
- Answer to Qs 3 is NO.
- Novel approach to analyzing the dual role of institutions as information producers and consumers and to study and visualize the diffusion of information among them.



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.*

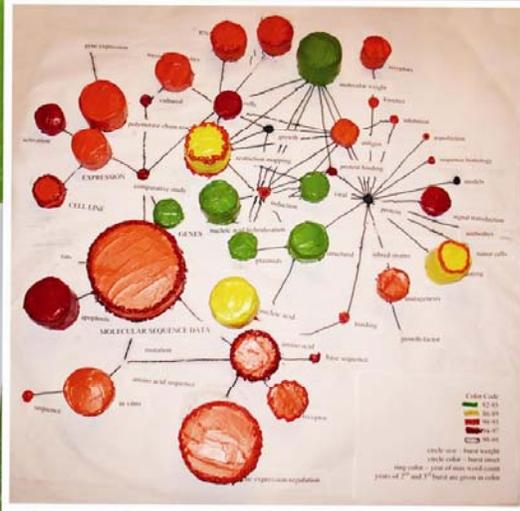


Merry Christmas and Happy New Year! 2008

Jon Burgoyne Katy Börner
Russell J. Duhon

Shravan Rajagopal
Heng (Michael) Zhang
Bruce W. Herr II
Julie M. Smith
Chung Yang (Kenneth) Lee

Kristin F. Reed
Stacy Kowalevzk
Micah Linnemcier
Bryan J. Hook
Nianli Ma
Carol Walter
Rengpeng Hu
Richard Pinapati
Todd Holloway
Peter A. Hook
Gonzalez Jr Micah Lin



Weixia (Bonnie) Huang
Elisha F. Hardy Fileve Palmer

Cake created by Kristin Reed and Lydia Nichols. They insisted on having a legend!

<http://ella.slis.indiana.edu/~katy>

<http://scimaps.org>

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

Examining the Evolution and Distribution of Patent Classifications

1 Patents Granted Over the Last 20 Years

Class	Title	Patents
200	Chemistry of Carbon Compounds	13,348
432	Electric Wiring and Body Training Compositions	12,527
214	Frictional	11,461
436	Surfaces	10,844
676	Work Methods or Microencapsulated Articles	10,491
75	Manufacturing and Testing	10,071
222	Internal Combustion Engines	9,462
340	Combinatorial Chemistry	8,283
224	Electrical Components and Circuits	7,700
27	Motor Vehicle	6,884
Total		35,059

Top Classes 1998-2002

1310	Drug Bio/Wiring and Body Training Compositions	28,791
314	Drug Bio/Wiring and Body Training Compositions	28,776
436	Frictional	21,774
432	Chemistry of Carbon Compounds	19,744
424	Drug Bio/Wiring and Body Training Compositions	19,207
676	Work Methods or Microencapsulated Articles	13,514
227	Acidic Solid State Devices Inc., Transistors, Cathodes	13,204
385	Information Processing System Organization	12,995
393	Computer Graphics Processing, Operator Interface Processing and Selective Visual Display Systems	12,919
227	Optical Systems and Devices	12,131
283	Static Information Storage and Retrieval	12,022
Total		140,000

In the United States, each patent gets assigned to one out of more than 450 classes covering broad application domains. An examination of the size and growth of patent classes provides insight about patenting trends.

Treemaps, a space-filling technique developed in the HCI Lab at the University of Maryland, are used to communicate major results. Treemaps represent a tree structure as nested rectangles with each rectangle representing a node. A rectangular area is first allocated to hold the representation of the tree, and this area is then subdivided into a set of rectangles that represent the top level of the tree. This process continues recursively on the resulting rectangles to represent each lower level of the tree. The parent-child relationship is indicated by enclosing the child rectangle by its parent rectangle. Typically, the size of each rectangle corresponds to the size of the node. Additional information about a node, e.g., its age or value, can be represented by the color of the respective rectangle.

2 Fast Growth Domains 1983 - 1987 / 1998 - 2002

Slow Growth Domains 1983 - 1987 / 1998 - 2002

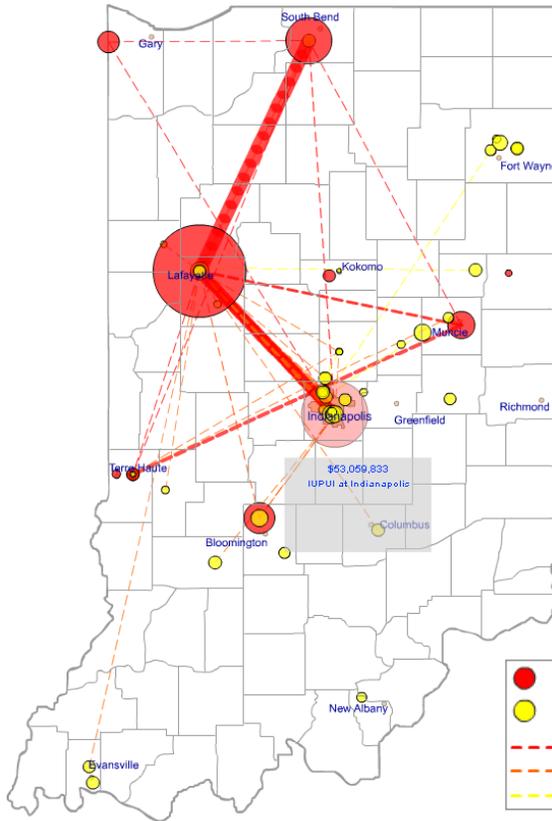
Shown is a comparison of the patent class space for 1983 to 1987 and 1998 to 2002. There is a predominance of growth in the 1998 to 2002 patent space, which correlates to the increase in patent grants during this period. By comparing the growth in categories, one can distinguish between domains that have been receiving a larger amount of patent grants.

3 Apple Computer

Depicted above is how Apple Computers' portfolio has changed in yearly increments from 1980 to 2002.

Jerome Lemelson

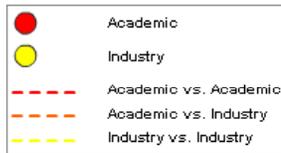
Lemelson's patent holdings below show a more even distribution over multiple classes. No class dominates over a majority of the years for granted patents; instead they are distributed more broadly over the intellectual space.



Mapping Indiana's Intellectual Space

Identify

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



Wikipedian Activity

Studying large scale social networks such as Wikipedia

Vizzards 2007 Entry

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

Rendered as Google Map:

<http://scimaps.org/maps/wikipedia>

Jan 8th, 2008 Data Version on Gigapan:

<http://gigapan.org/vienGigapan.php?id=5042>



Second sight

Image: Bruce W. Herr and Todd M. Holloway

Power struggle



To spot where arguments are taking place, Herr suggests. If rival contributors are repeatedly changing each other's entries, for example, a page could be locked until the mood cools (locked 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.



Science Related Wikipedian Activity

http://scimaps.org/dev/map_detail.php?map_id=165

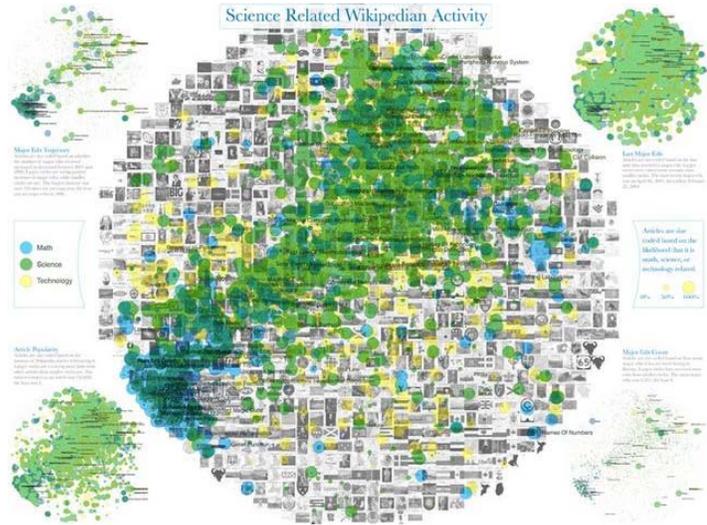
Same base map.

Overlaid are 3,599 math (blue), 6,474 science (green), and 3,164 technology relevant articles (yellow).

All other articles are given in grey.

Corners show articles size coded according to

- article edit activity (top left),
- number of major edits (top right),
- number of bursts in edit activity (bottom, right)
- indegree (bottom left).



**Society for Neuroscience, 2006
Visual Browser**

Click to start a bounding box, then click again to end it. A listing of all nodes in the bounding box will be shown. Click once again to clear the bounding box.

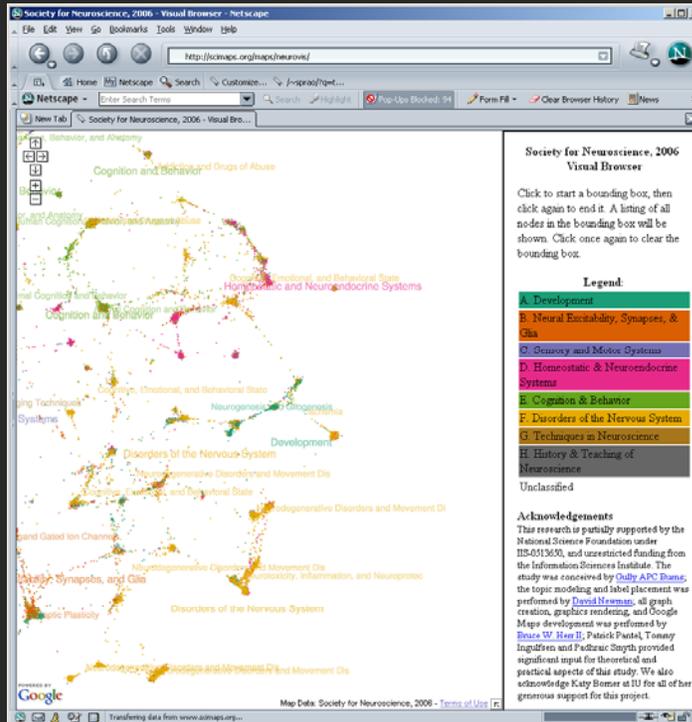
Legend

- A. Development
- B. Neural Excitability, Synapses, & Glia
- C. Sensory and Motor Systems
- D. Homeostatic & Neuroendocrine Systems
- E. Cognition & Behavior
- F. Disorders of the Nervous System
- G. Techniques in Neuroscience
- H. History & Teaching of Neuroscience
- Unclassified

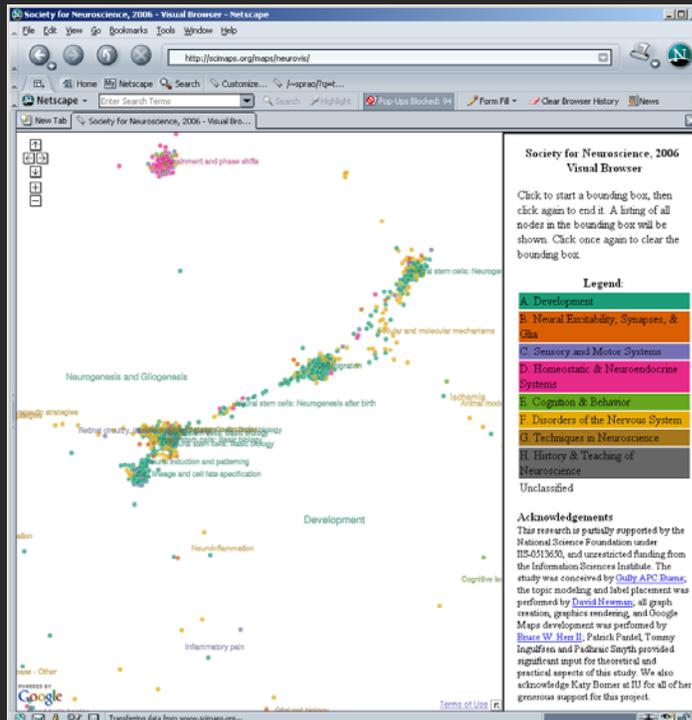
Acknowledgements
This research is partially supported by the National Science Foundation under IIS-0513650, and unrestricted funding from the Information Sciences Institute. The study was conceived by Gully APC Burns, the topic modeling and label placement was performed by David Newman, all graph creation, graphics rendering, and Google Maps development was performed by Bruce W. Herr II, Patrick Paudel, Tommy Ingulfsen and Padraic Smyth provided significant input for the theoretical and practical aspects of this study. We also acknowledge Katy Bonner at UI for all of her generous support for this project.

Map Data: Society for Neuroscience, 2006 - Terms of Use

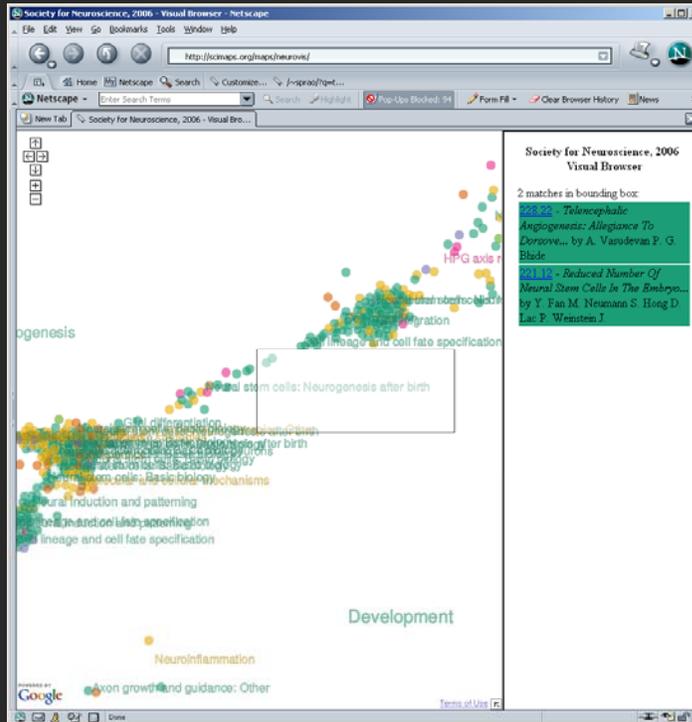
Bruce W. Herr II, Gully Burns (USC), David Newman (UCI), Society for Neuroscience, 2006
Visual Browser, 2007, <http://scimaps.org/maps/neurovis/>



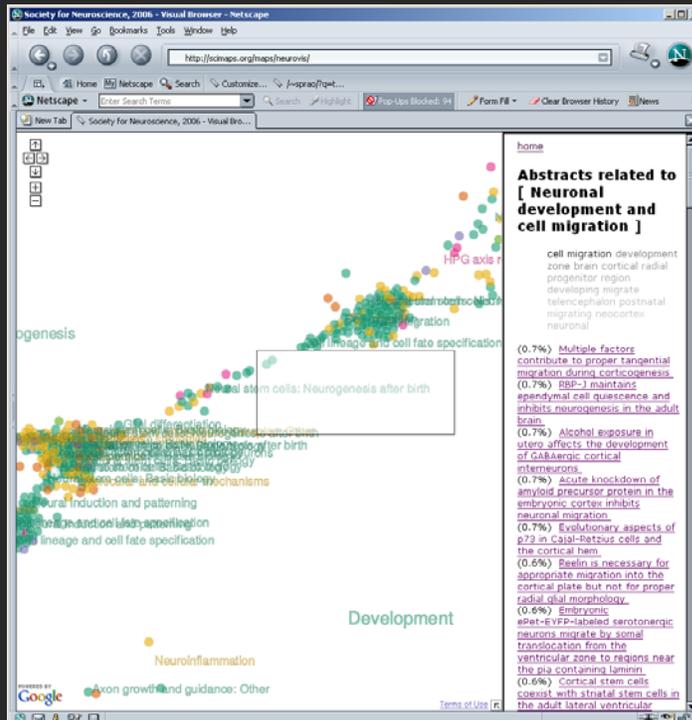
Bruce W. Herr II, Gully Burns (USC), David Newman (UCI), Society for Neuroscience, 2006
Visual Browser, 2007, <http://scimaps.org/maps/neurovis/>



Bruce W. Herr II, Gully Burns (USC), David Newman (UCI), Society for Neuroscience, 2006
Visual Browser, 2007, <http://scimaps.org/maps/neurovis/>

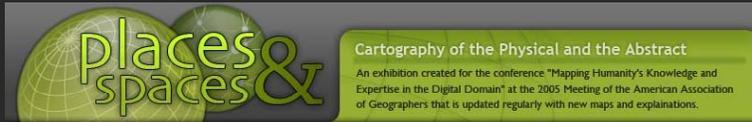


Bruce W. Herr II, Gully Burns (USC), David Newman (UCI), Society for Neuroscience, 2006
 Visual Browser, 2007, <http://scimaps.org/maps/neurovis/>



Bruce W. Herr II, Gully Burns (USC), David Newman (UCI), Society for Neuroscience, 2006
 Visual Browser, 2007, <http://scimaps.org/maps/neurovis/>

Mapping Science Exhibit



Home Browse Maps Compare & Contrast Maps Connect

Home



Exhibit Purpose and Goals

The Places & Spaces 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.

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.

This online part of the exhibit provides links to a selected series of maps and their makers along with detailed explanations of why these maps work. The physical counterpart supports the close inspection of high quality reproductions for display at conferences and education centers. It is meant to inspire cross-disciplinary discussion on how to best track and communicate human activity and scientific progress on a global scale.



Places & Spaces: Mapping Science

a science exhibit that introduces people to maps of sciences, their makers and users.

<http://scimaps.org>

Exhibit Curators: Dr. Katy Börner & Elisha F. Hardy



Mapping Science Exhibit – 10 Iterations in 10 years

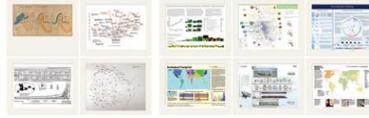
<http://scimaps.org/>



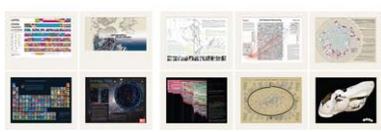
The Power of Maps (2005)



Science Maps for Economic Decision Makers (2008)



The Power of Reference Systems (2006)



Science Maps for Science Policy Makers (2009)

Science Maps for Scholars (2010)

Science Maps as Visual Interfaces to Digital Libraries (2011)

Science Maps for Kids (2012)

Science Forecasts (2013)

The Power of Forecasts (2007)



How to Lie with Science Maps (2014)



Exhibit has been shown in 49 venues on four continents. Also at

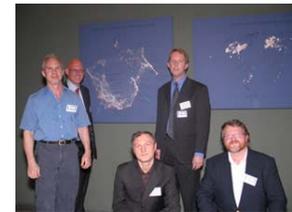
- NSF, 10th Floor, 4201 Wilson Boulevard, Arlington, VA.
- Chinese Academy of Sciences, China, May 17-Nov. 15, 2008.
- University of Alberta, Edmonton, Canada, Nov 10-Jan 31, 2009
- Center of Advanced European Studies and Research, Bonn, Germany, Dec. 11-19, 2008.



Illuminated Diagram Display

W. Bradford Paley, Kevin W. Boyack, Richard Kalvans, and Katy Börner (2007)

Mapping, Illuminating, and Interacting with Science. SIGGRAPH 2007.



Questions:

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

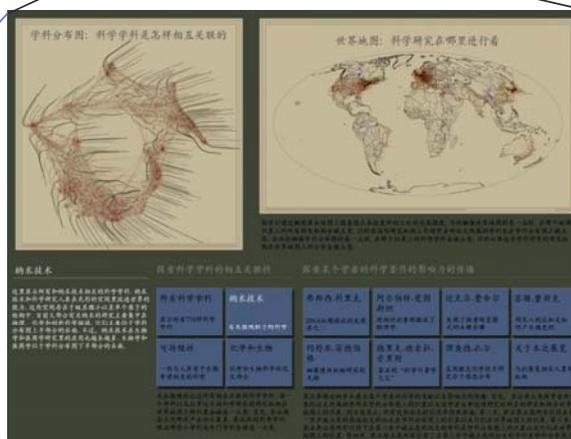


Large-scale, high resolution prints illuminated via projector or screen.

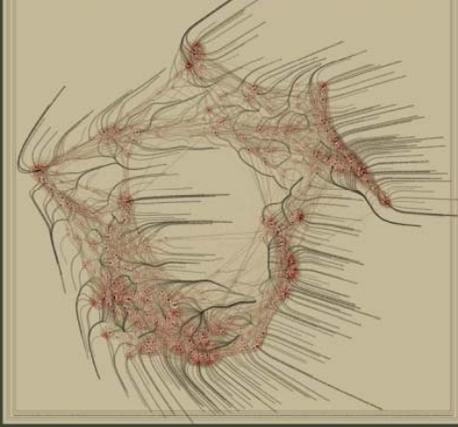
Interactive touch panel.

Contributions:

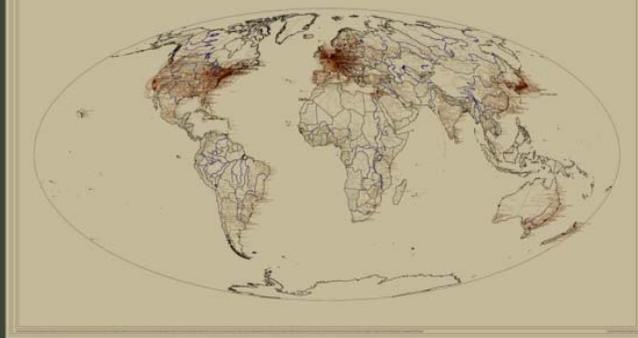
- Interactive, high resolution interface to access and make sense of data about scholarly activity.



学科分布图：科学学科是怎样相互关联的



世界地图：科学研究在哪里进行着



你可以通过触摸屏在地图上随意指点来改变所到之处的光亮强度。当你触摸世界地图的某一点时，在那个地理位置上的所有研究机构会被点亮。同时在这些研究机构工作的学者的论文所属的学科会在学科分布图上被点亮。而当你触摸学科分布图的某一点时，在那个位置上的科学学科会被点亮，同时从事这些学科研究的研究机构在世界地图上的分布会被点亮。

纳米技术

这里显示所有和纳米技术相关的科学学科。纳米技术和科学研究人在无形的空间里改造世界的的能力。这些空间存在于极其微小以至单个原子的结构中。目前大部分有关纳米的研究主要集中在物理、化学和材料科学领域。它们主要位于学科分布图上半部分的右面。不过，纳米技术在生物学和医药学研究里的应用也越来越多。生物学和医药学位于学科分布图下半部分的右面。



探索科学学科的相互关联性

所有科学学科 显示所有776种科学学科	纳米技术 有关微观粒子的科学
可持续性 一些与人类寄予长期希望相关的科学	化学和生物 化学和生物科学的交叉部分

光标缓慢的扫过所有相互关联的科学学科，每一个学科以及从事这方面科学研究的研究机构在世界地图上的位置会被逐一点亮。首先，显示屏会点亮那些产出论文最多、最活跃的科学学科，然后那些小学科或冷门学科会被逐一点亮。

探索某个学者的科学著作的影响力的传播

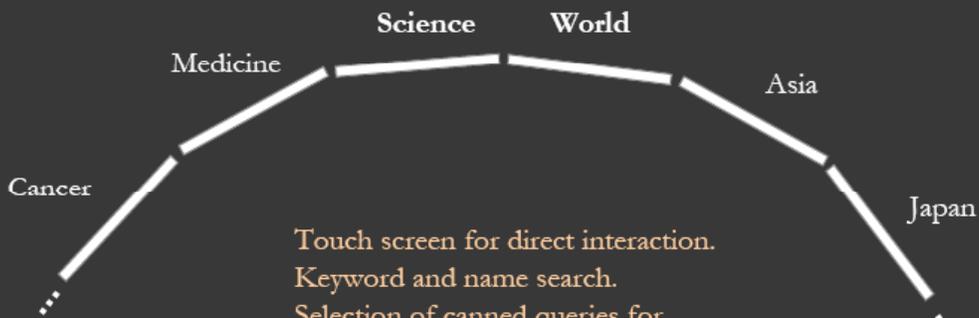
弗郎西·科里克 DNA双螺旋结构的发现者之一	阿尔伯特·爱因斯坦 用相对论重新激活了物理学	迈克尔·费舍尔 发现了物质转变模式的关键步骤	苏珊·费斯克 研究人的认知是如何产生偏见的
约舒亚·雷德伯格 细菌遗传机制研究先驱	德里克·德索拉·普里斯 著名的“科学计量学之父”	理查德·扎尔 采用激光化学技术研究分子动态分布	关于本次展览 与此展览相关人员和机构

显示屏通过四步来展示某个学者对科学的贡献以及影响力的传播。首先，显示屏点亮该学者所发表的论文所属的学科在学科分布图上的位置以及该学者从事这项研究所在的研究机构在世界地图上的位置。到目前为止，所有这些论文的引用率仍然很高。第二步，显示屏点亮所有引用在第一步中被点亮的原始论文的论文在学科分布图上的位置以及它们在世界地图上的位置。第三步，显示屏点亮所有引用了在第二步中被点亮的论文的论文在学科分布图上的位置以及它们在世界地图上的位置。第四步，显示屏点亮所有引用了在第三步中被点亮的论文的论文在学科分布图上的位置以及它们在世界地图上的位置。

Re-implementation of Illuminated Diagram Software

by Advanced Visualization Lab, Indiana University

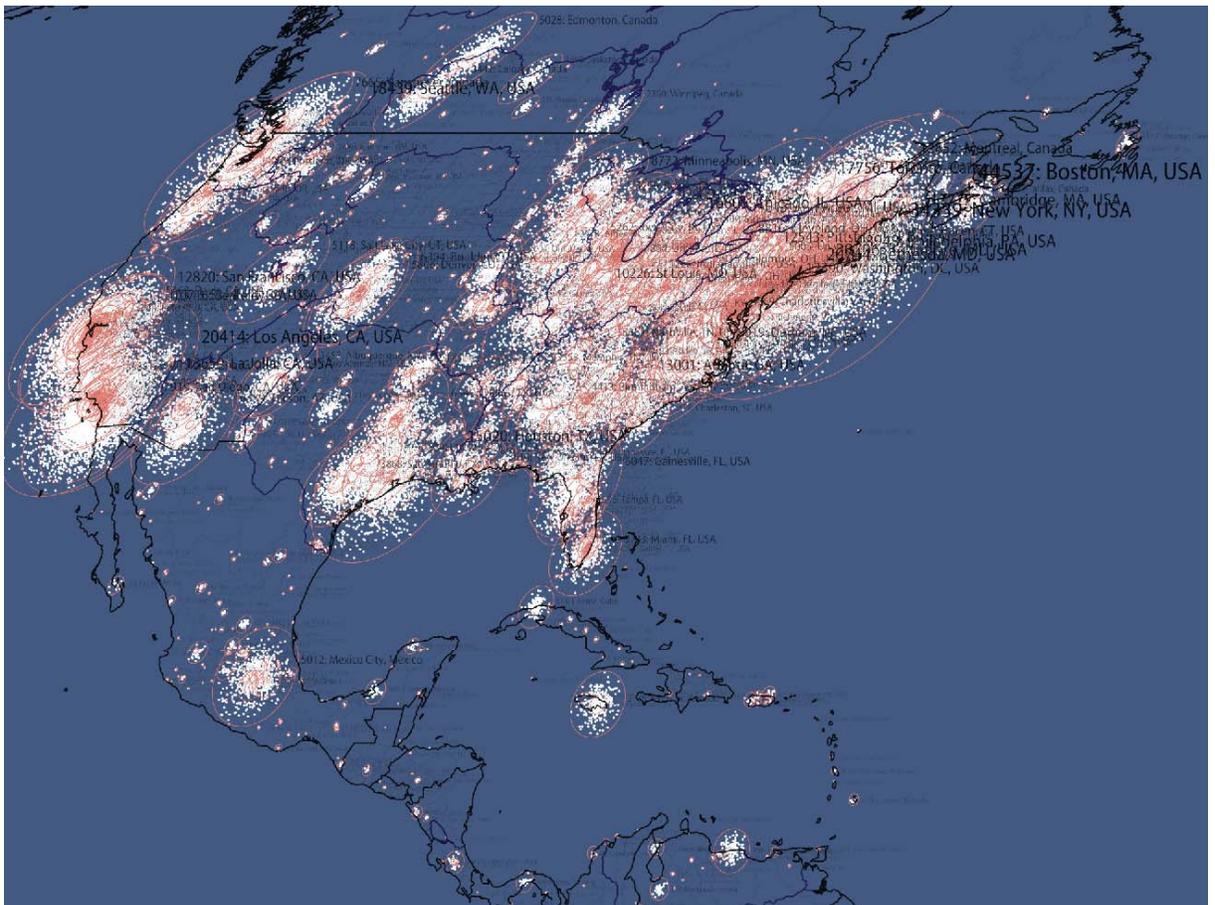
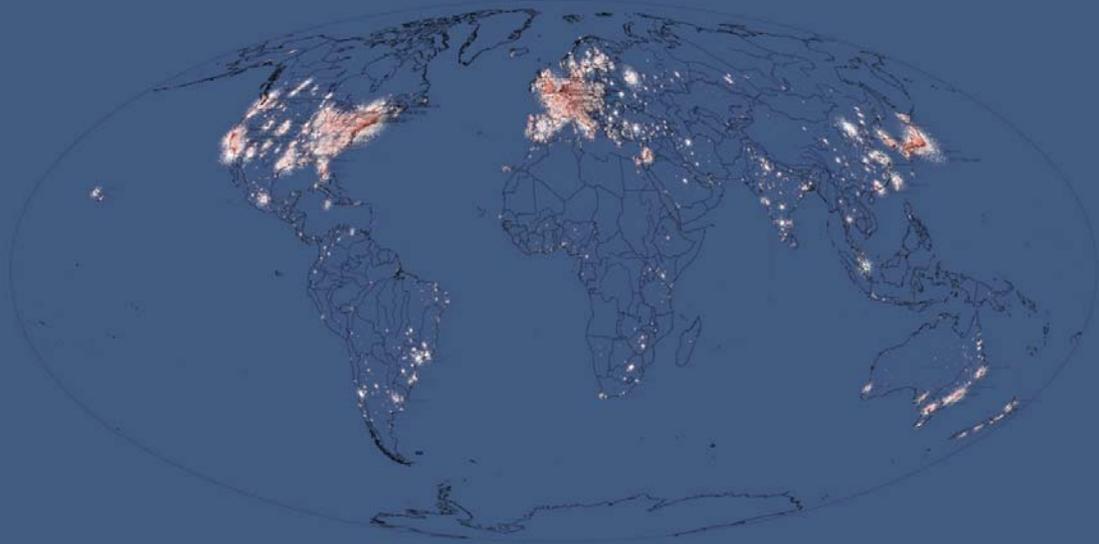
Drives unlimited number of ID screens.

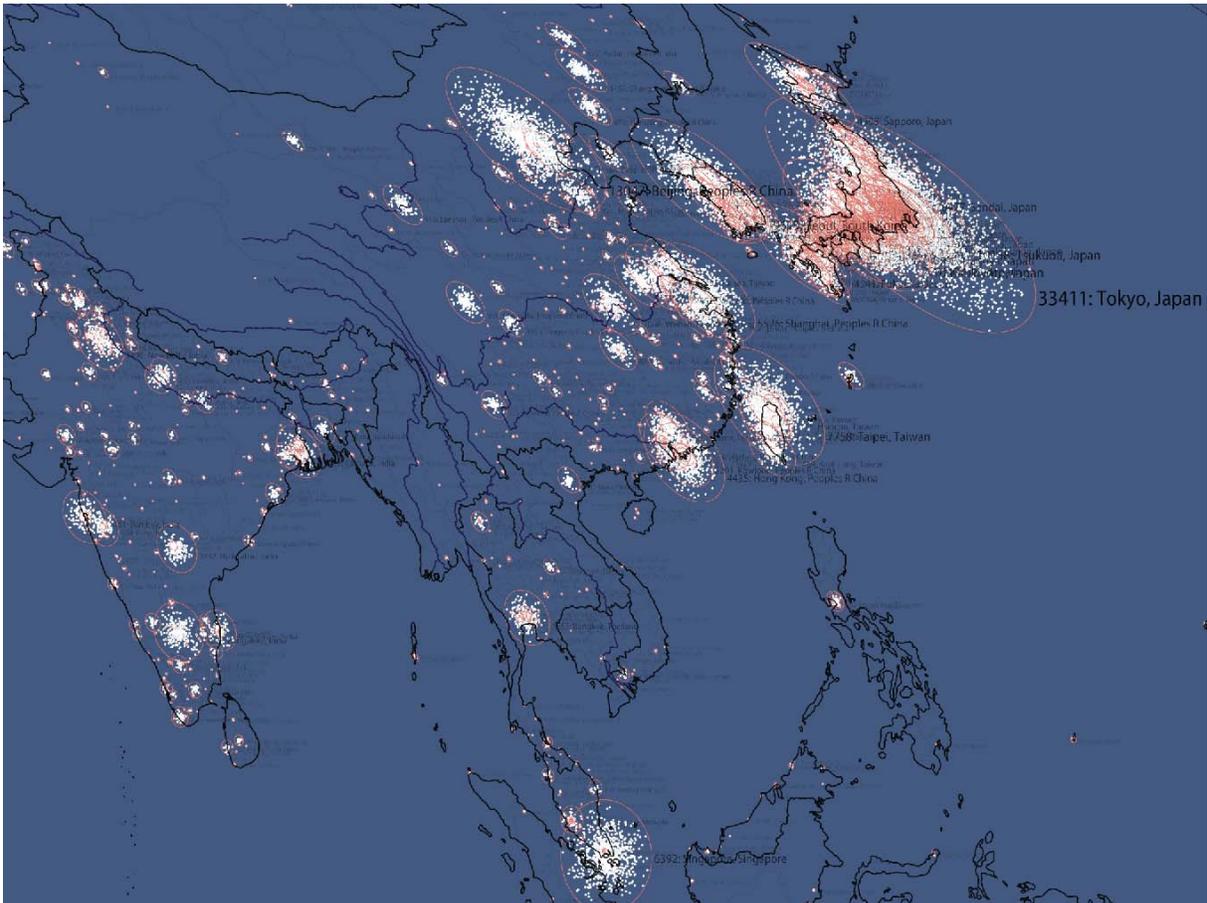
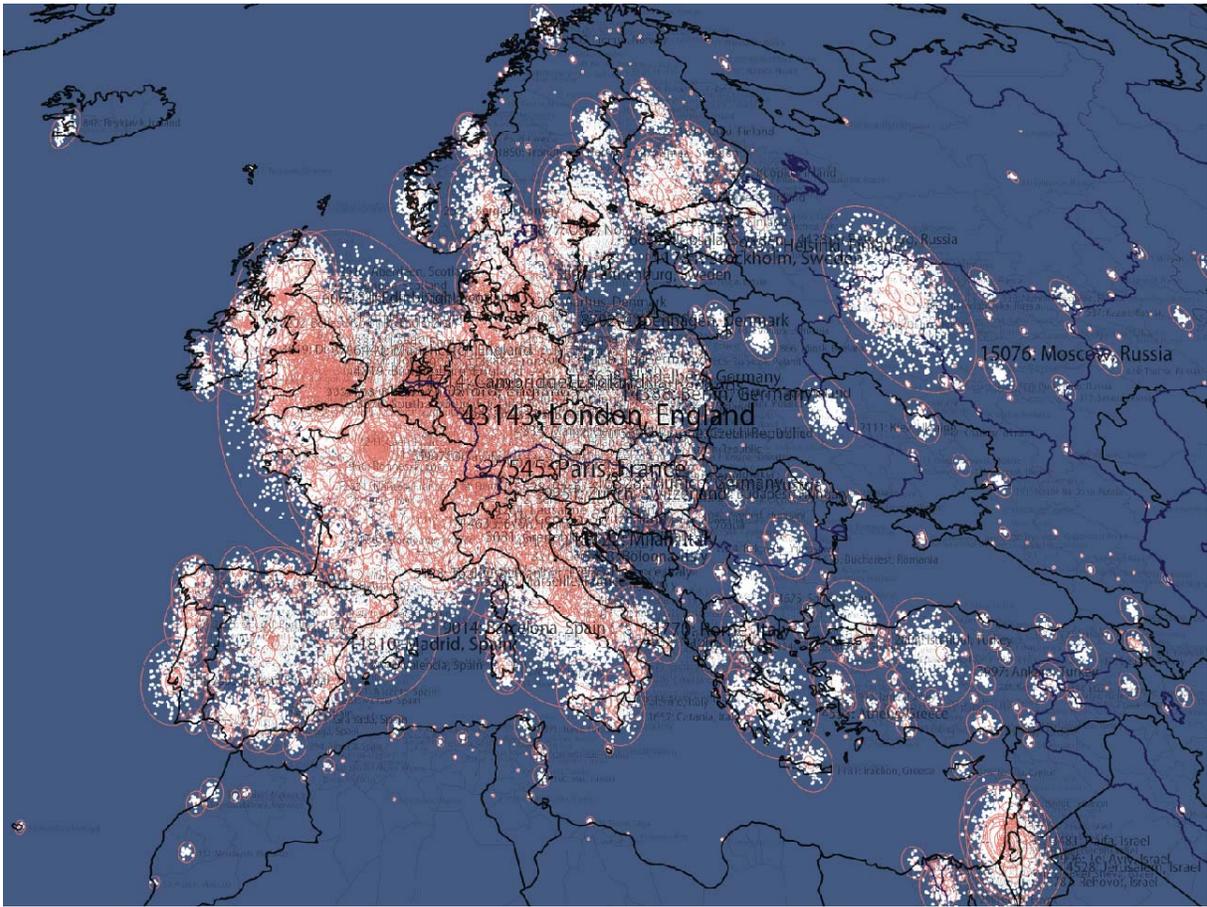


- Touch screen for direct interaction.
 Keyword and name search.
 Selection of canned queries for
- interdisciplinary research areas
 - famous people
 - activity patterns, e.g., bursts, trends, etc.

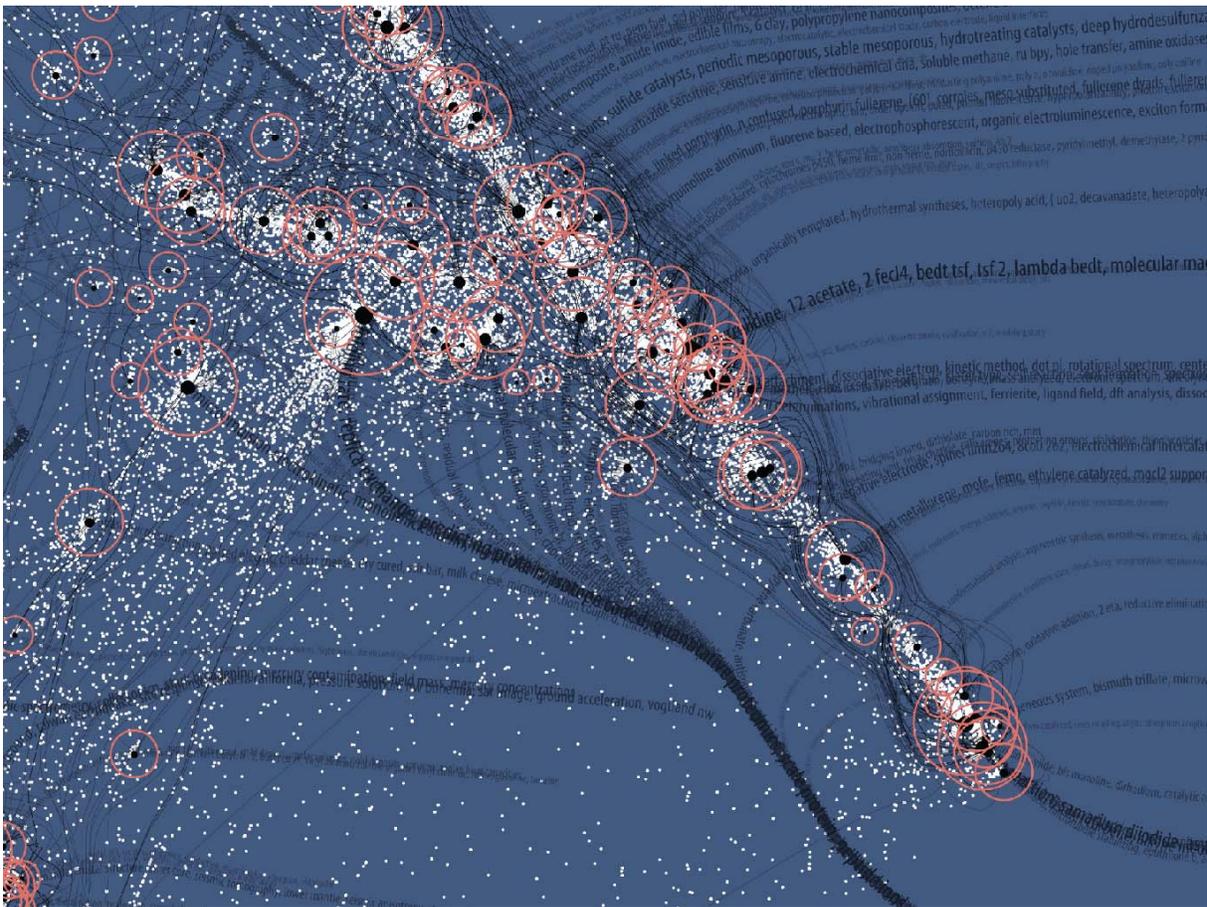


GEOGRAPHIC MAP: WHERE SCIENCE GETS DONE





TOPIC MAP: HOW SCIENTIFIC PARADIGMS RELATE





Inventors



Hands-On Science Maps for Kids, by Filipe Palmer (Painting), Julia Smith (Data Acquisition), Eksha Hardy and Kitty Elmer (Graphic Design), BLOOMINGTON, IN, 2006. Courtesy of Indiana University. Learn more at www.sciemaps.org. This map plots the locations of where scientific papers were published; each light green dot represents 50 or fewer papers; they are scattered around the exact location for visibility, within a labeled green circle whose size is proportional to the number of papers published in that place. The base map is part of an "illuminated diagram" display which used a computer and two projectors projecting spots of light on the panel to highlight different kinds of scientific research on a sliding map of scientific paradigms and the areas in the world where such science was performed. Base map research by Kevin Baksh and Dik Kikstra, cartography by John Deacon, data from Thompson ISI graphics and typography by its Bradford Philp. Copyright © 2006 by Bradford Philp, all rights reserved.



Science of Science Cyberinfrastructure



Science of Science Cyberinfrastructure — P O R T A L —

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

Introduction
E. O. Wilson writes in *Consilience: The Unity of Knowledge* (1998): "Features that distinguish science from pseudoscience are repeatability, economy, mensuration, heuristics, and consilience."
Please see Börner's [recent presentation](#) at the *A Deeper Look at the Visualization of Scientific Discovery* NSF Workshop for a general introduction of the needs and the resources provided here.

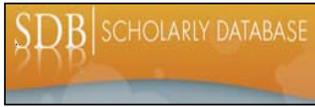
Needs Analysis
As part of the "IIS: Towards a Macroscopic for Science Policy Decision Making" NSF SBE-0738111 award, interviews with science policy makers are conducted to identify what science of science research results and tools might be most desirable and effective. So far, 30 formal, one-hour interviews have been conducted with science policy makers at university campus level, program officer level, and division director level for governmental, state, and private foundations. Data compilation will start in October 2008 and resulting report can be ordered by sending a request to Mark Price (maaprice@indiana.edu).

Conceptualization of Science
A science of science requires a theoretically grounded and practically useful conceptualization of the structure and evolution of science. A special journal issue entitled "*Science of Science: Conceptualizations and Models of Science*" edited by [Katy Börner](#), Indiana University & [Andrea Schamhorst](#), Royal Netherlands Academy of Arts and Sciences invites contributions on this topic. It will be published in the *Journal of Informetrics* 3(1) in January 2009.

Scholarly Database
The [Scholarly Database \(SDB\)](#) at Indiana University aims to serve researchers and practitioners interested in the analysis, modeling, and visualization of large-scale scholarly datasets. The database currently provides access to over 20 million papers, patents and grants. Resulting datasets can be downloaded in bulk. Register for free access at <https://sdb.slis.indiana.edu/>.

Cyberinfrastructures
The Scientometrics filling of the [Network Workbench \(NWB\) Tool](#) provides a unique distributed, shared resources environment for large-scale network analysis, modeling, and visualization. Thomson Scientific/ISI, Scopus and Google Scholar data, EndNote and Bibtext files, or NSF awards can be read and diverse networks can be extracted and studied. Download [User Manual with focus on Scientometrics](#).

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

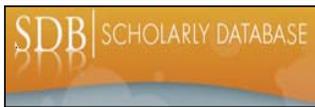


Scholarly Database: Web Interface

Search across publications, patents, grants.

Download records and/or (evolving) co-author, paper-citation networks.

Register for free access at <http://sdb.slis.indiana.edu>



Scholarly Database: # Records & Years Covered

Datasets available via the Scholarly Database

Dataset	#Records	Years Coverage	updated	Restricted Access
Medline	16,053,495	1898-2008	Yes	
PhysRev	398,005	1893-2006		Yes
PNAS	16,167	1997-2002		Yes
JCR	59,078	1974,1979,1984,1989,1994-2004		Yes
USPTO	3,710,952	1976-2007	Yes	
NSF	174,835	1985-2003	Yes	
NIH	1,043,804	1972-2002	Yes	
Total	21,456,336	1893-2008	4	3

Aim for comprehensive temporal, geospatial, and topic coverage.

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

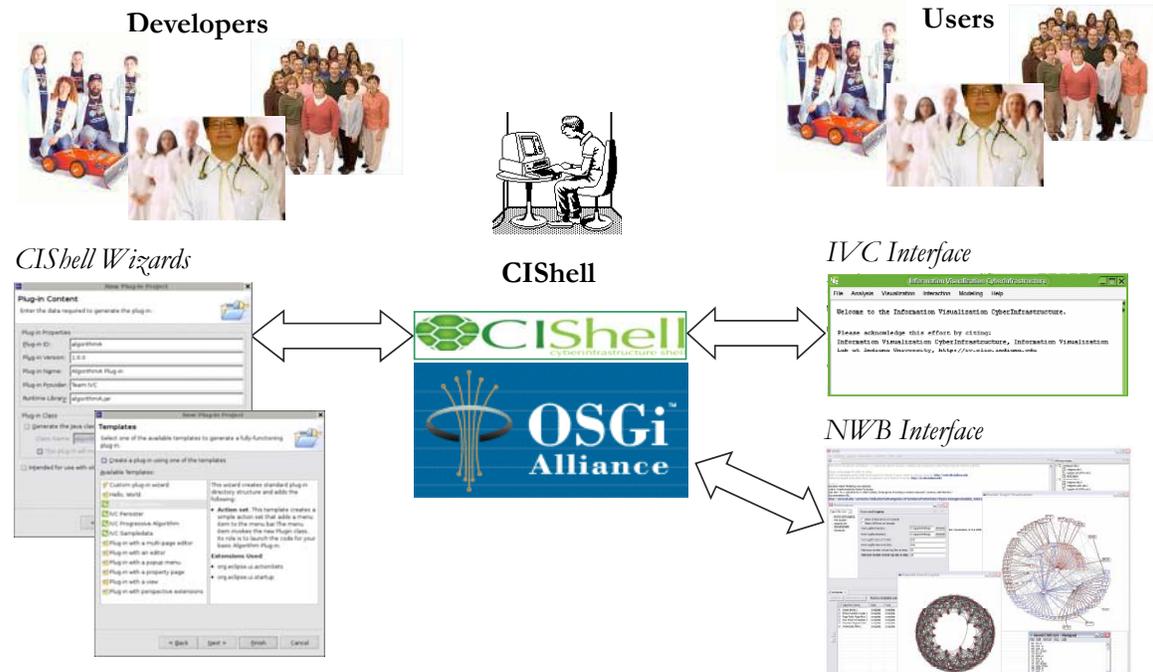


Software Team: Lead: Micah Linnemeier
 Members: Patrick Phillips, Russell Duhon, Tim Kelley & Ann McCranie
 Previous Developers: Weixia (Bonnie) Huang, Bruce Herr, Heng Zhang, Duygu Balcan, Bryan Hook, Ben Markines, Santo Fortunato, Felix Terkhorn, Ramya Sabbineni, Vivek S. Thakre & Cesar Hidalgo



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

Amount: \$1,120,926, NSF IIS-0513650 award
Duration: Sept. 2005 - Aug. 2009
Website: <http://nwb.slis.indiana.edu>



Preprocessing [Edit](#)

Remove Nodes

[Extract Top Nodes](#)
[Extract Nodes Above or Below Val](#)
[Delete High Degree Nodes](#)
[Delete Random Nodes](#)
[Delete Isolates](#)

Remove Edges

[Extract Top Edges](#)
[Extract Edges Above or Below Val](#)
[Remove Self Loops](#)
[Trim By Degree²](#)
[Pathfinder Network Scaling](#)

Sampling

[Snowball Sampling \(n nodes\)](#)
[Node Sampling](#)
[Edge Sampling](#)

Transformations

[Symmetrize](#)
[Dichotomize](#)
[Multipartite Joining](#)

Modeling [Edit](#)

General

[Random Graph](#)
[Watts-Strogatz Small World](#)
[Barabási-Albert Scale-Free](#)

Structured

[CAN](#)
[Chord](#)

Unstructured

[Hypergrid](#)
[PRU](#)

Other

[TARL](#)
[Discrete Network Dynamics](#)

Analysis [Edit](#)

General Purpose

[Network Analysis Toolkit²](#)

Unweighted & Undirected

[Based on degree/](#)
[Node Degree](#)
[Node Distribution](#)
[Based on clustering](#)
[k-Nearest Neighbor](#)
[Watts Strogatz Clustering Coefficient](#)
[Watts Strogatz Clustering Coefficient](#)

Based on path

[Diameter](#)
[Average Shortest Path](#)
[Shortest Path Distribution](#)
[Node Betweenness Centrality](#)

Based on components

[Connected Components](#)
[Weak Component Clustering](#)

K-Core

[Extract K-Core²](#)
[Annotate K-Core²](#)

Unweighted & Directed

Based on degree

[Node Indegree](#)
[Node Outdegree](#)
[Indegree Distribution](#)
[Outdegree Distribution](#)

Based on local graph structure

[k-Nearest Neighbor](#)
[Single Node In-Out Degree Correla](#)

Unnamed Category?

[Page Rank](#)
[Based on local graph structure](#)

[Dyad Reciprocity²](#)
[Arc Reciprocity²](#)

[Adjacency Transitivity²](#)

Based on components

[Weak Component Clustering](#)
[Extract Attractors²](#)

Visualization [Edit](#)

Tools

[GUESS](#)
[GnuPlot²](#)

Predefined Positions Layout

[DrL \(VxOrd\)](#)
[Pre-defined Positions \(prefuse beta\)²](#)

Move

[Circular](#)

Tree Layouts

[Radial Tree \(prefuse alpha\)](#)
[Radial Tree with Annotations \(prefuse beta\)²](#)
[Tree Map](#)
[Tree View](#)
[Balloon Graph \(prefuse alpha\)²](#)

Network Layouts

[Force Directed with Annotation \(prefuse beta\)](#)
[Kamada-Kawai \(JUNG\)](#)
[Fruchterman-Reingold \(JUNG\)](#)
[Fruchterman-Reingold with Annotation \(prefuse beta\)](#)
[Spring \(JUNG\)](#)
[Small World \(prefuse alpha\)](#)

Other Layouts

[Parallel Coordinates \(demo\)²](#)
[LaNet \(k-Core Decomposition\)](#)

Scientometrics [Edit](#)

Extract Network From Table

[Extract Co-Authorship Network](#)
[Extract Co-Occurrence Network From Table²](#)
[Extract Directed Network From Table²](#)

Extract Network From Another Network

[Extract Bibliographic Coupling Similarity Network](#)
[Extract Co-Citation Similarity Network²](#)

Cleaning

[Remove ISI Duplicate Records](#)
[Detect Duplicate Nodes](#)
[Remove Rows With Multitudinous Fields²](#)

erner: Mapping the Structure and Dynamics of Science >>



EpiC will Build on and Extend NWB

INDIANA UNIVERSITY

IU News Room

Sunday, May 4, 2008

IU News from all eight campuses

Browse by Topic

Search GO

[For Journalists](#) | [Archives](#) | [Site Index](#) | [Contact Us](#) | [Public Affairs](#)

Browse by Topic

- [Arts & Humanities](#)
- [Athletics](#)
- [Business](#)
- [Education](#)
- [General](#)
- [Health & Medicine](#)
- [Law](#)
- [Public & Environmental Affairs](#)
- [Science](#)
- [Social Science](#)
- [Technology](#)

Newsroom Home > Indiana University Media Relations > News Release

[E-mail this page](#)
[Print this page](#)

Last modified: Tuesday, April 8, 2008

\$1.2 million NIH project will help track and predict epidemics

FOR IMMEDIATE RELEASE
April 8, 2008

BLOOMINGTON, Ind. -- The National Institutes of Health has given \$1.2 million to Indiana University researchers to build the ultimate international epidemic research tool.

Media Contacts

- Neal Moore
ngmoore@indiana.edu
317-278-9208
- David Bricker
brickerd@indiana.edu
812-856-9035

News by Topic

- [General News](#)
- [Graduate Studies](#)
- [Life Sciences](#)
- [Science](#)
- [Technology](#)

[More Topics >>](#)

Multimedia News



cyberinfrastructure for NETWORK SCIENCE CENTER

School of Library and Information Science | Indiana University Bloomington



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