

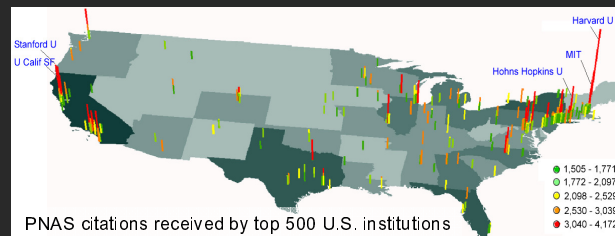
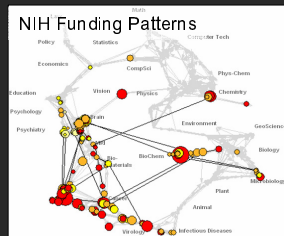
Mapping the Structure and Evolution of Science Locally and Globally

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



*New Horizons in Internet Site Development
Banbury Center, Cold Spring Harbor Laboratory, New York, October 24, 2006*



This Talk has Three Parts:

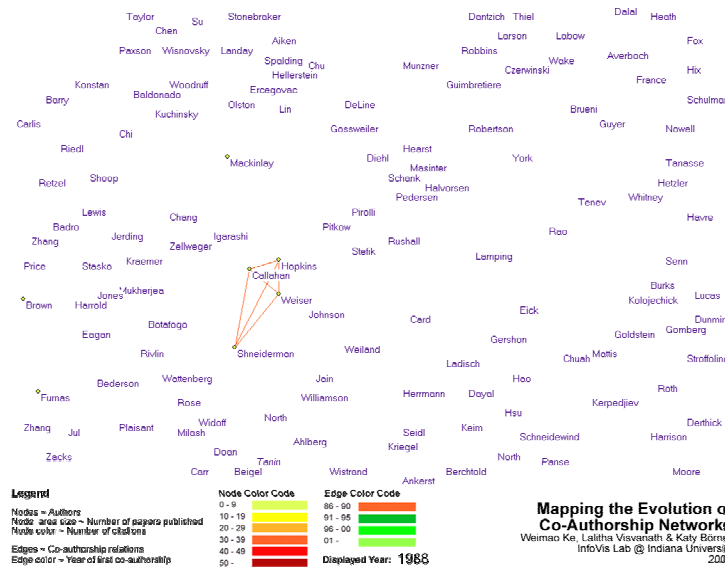
1. Mapping Science: Why?
2. Mapping Science: How?
3. Mapping Science: Applications

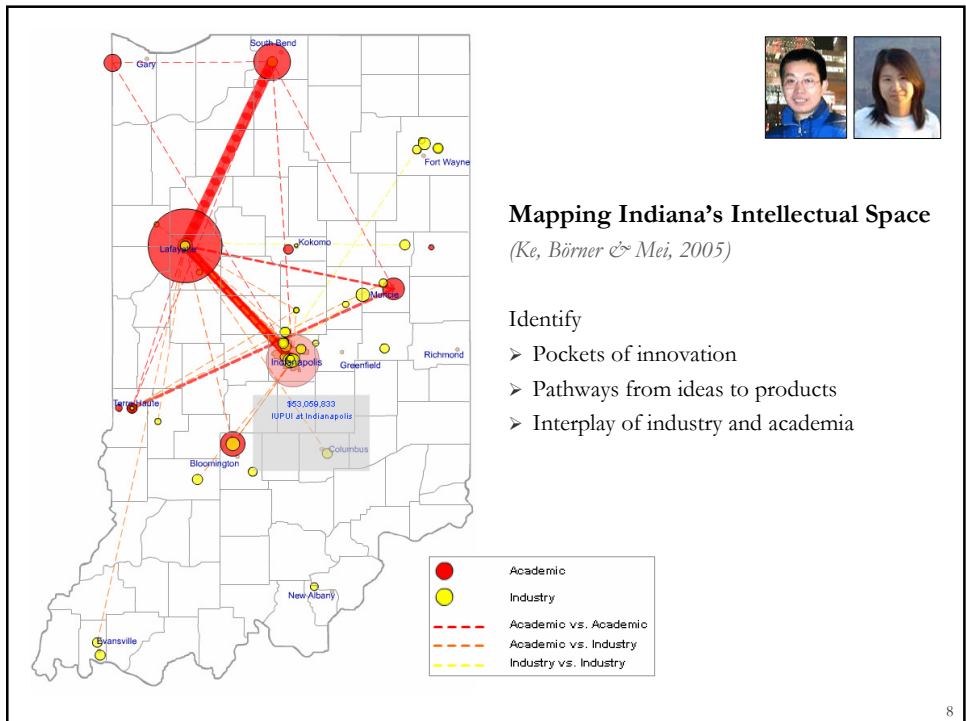
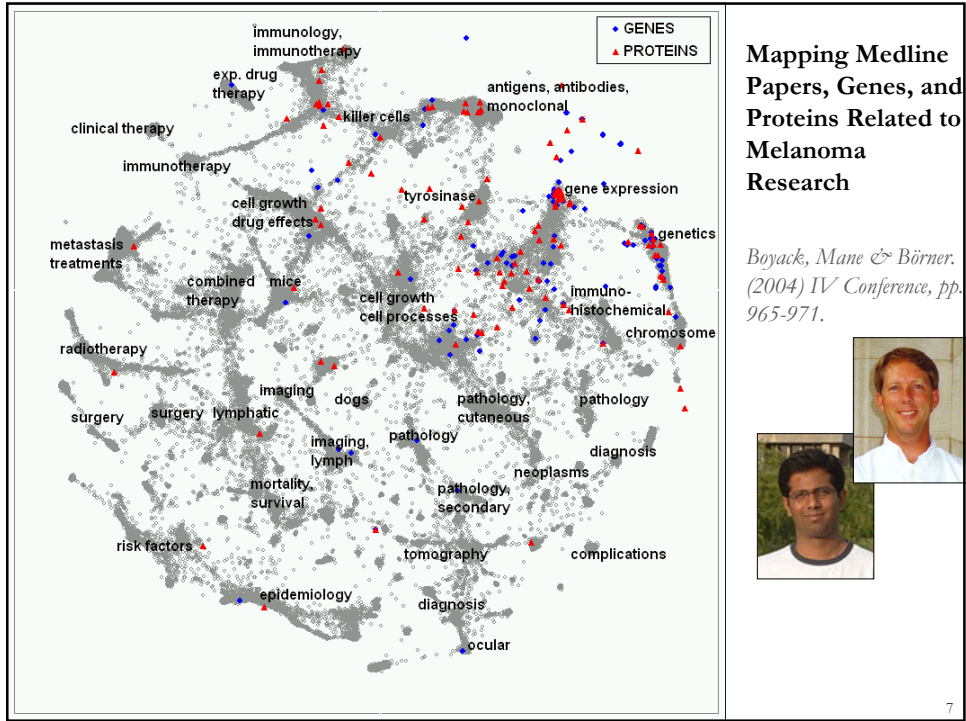
This Talk has Three Parts:

1. Mapping Science: Why?
2. Mapping Science: How?
3. Mapping Science: Applications

Mapping the Evolution of Co-Authorship Networks in Information Visualization, 1988 - 2004

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



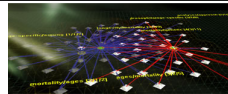


This Talk has Three Parts:

1. Mapping Science: Why?
2. Mapping Science: How?
3. Mapping Science: Applications



- Shiff, 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.*
- Börner, Katy, Sanyal, Soma and Vespignani, Alessandro (in press). *Network Science. In Blaise Cronin (Ed.), Annual Review of Information Science & Technology, Information Today, Inc./ American Society for Information Science and Technology, Medford, NJ.*



Process of Mapping Knowledge Domains

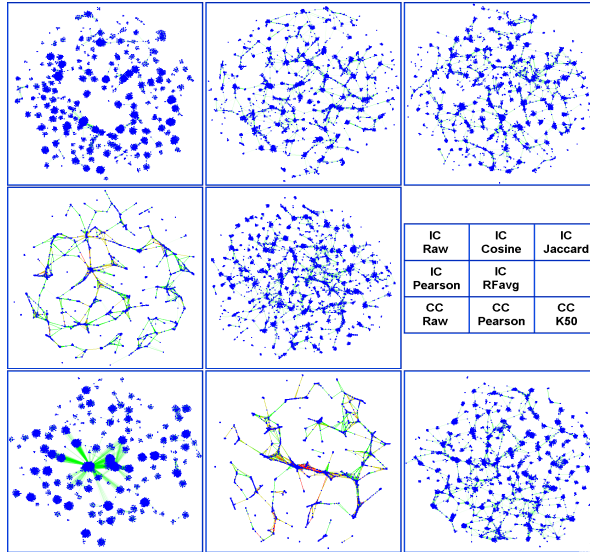
DATA EXTRACTION	UNIT OF ANALYSIS	MEASURES	LAYOUT (often one code does both similarity and ordination steps)		DISPLAY
			SIMILARITY	ORDINATION	
SEARCHES ISI INSPEC Eng Index Medline ResearchIndex Patents etc.	COMMON CHOICES Journal Document Author Term	COUNTS/FREQUENCIES Attributes (e.g. terms) Author 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) 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, Topics Pathfinder networks (PFNet) Self-organizing maps (SOM) includes SOM, ET-maps, etc. CLUSTER ANALYSIS SCALAR Triangulation Force-directed placement (FDP)	INTERACTION Browse Pan Zoom Filter Query Detail on demand ANALYSIS
BROADENING By citation By terms					

Börner, Chen & Boyack.. (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.*

Comparison of Similarity Metrics



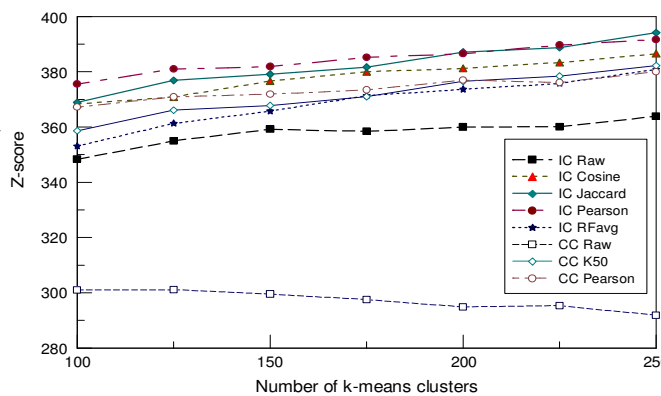
- ISI file year 2000, SCI and SSCI: 7,121 journals.
- Different similarity metrics
 - Inter-citation (raw counts, cosine, modified cosine, Jaccard, RF, Pearson)
 - Co-citation (raw counts, cosine, modified cosine, Pearson)
- Maps were compared based on
 - regional accuracy,
 - the scalability of the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

A 'Backbone' Map of Science & Social Science

- The map is comprised of 7,121 journals from year 2000.
- Each dot is one journal
- An *IC-Jaccard* similarity measure was used.
- Journals group by discipline.
- Groups are labeled by hand.
- Large font size labels identify major areas of science.
- Small labels denote the disciplinary topics of nearby large clusters of journals.



15

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, patent grants and grant opportunities. The database is updated weekly and automatically updated. <http://slis.indiana.edu/infvis>

COMPUTING RESOURCES
The InfoVis CyberInfrastructure is located at Indiana University Research Database Complex consisting of two Sun T2000 servers with 12 400GB disks and 16 GB of memory each. 400GB shared disks are attached to both servers. A Sun V9000 system with 4 Sun 400GB disks is located in the InfoVis CyberInfrastructure building and is used for the database servers. <http://slis.indiana.edu/infvis>

SOFTWARE
An open source PC framework was designed to facilitate the integration of diverse data analysis, modeling and visualization algorithms. These algorithms, data performance metrics, links and tools for the framework and more user tools can be easily "plugged in" or "unplugged". <http://slis.indiana.edu/infvis>

LEARNING MODULES
A set of associated learning modules aims to equip learners with practical skills and to provide code and advice to quickly modify and run different algorithms and diverse research techniques and infrastructure visualizations. <http://slis.indiana.edu/infvis>



Cyberinfrastructure Shell
<http://cisbell.org>

CAREER: Visualizing Knowledge Domains. NSF IIS-0238261 award
(Katy Börner, \$451,000) Sept. 03-Aug. 08.
<http://in.slis.indiana.edu/>



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, Eric Wernert (Senior Personnel), \$1,120,926) Sept. 05 - Aug. 08. <http://lww.slis.indiana.edu>

This Talk has Three Parts:

1. Mapping Science: Why?
2. Mapping Science: How?
3. Mapping Science: Applications

The screenshot shows the website for 'places & spaces & Cartography of the Physical and the Abstract'. The website has a green and grey color scheme. The main navigation bar includes 'Home', 'Browse Maps', 'Compare & Contrast Maps', and 'Connect'. The 'Home' page features a section titled 'Exhibit Purpose and Goals' with the following text:

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.

Below the text is a small image of a globe and a wireframe model of a building. To the right of the website screenshot is a photograph of five people standing in front of the exhibit. Below this is a large text block:

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

<http://scimaps.org/exhibit>

At the bottom of the screenshot are three photographs: the first shows the exhibit booth with a sign that reads 'Places & Spaces Cartography of the Physical and Abstract'; the second shows a woman pointing at a large interactive map on a wall; the third shows a man and a woman standing in front of the exhibit.



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.

Butterfly

bee

Walter Guadagnoli

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.

Activities:

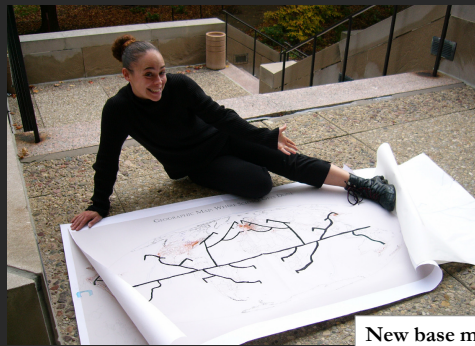
- Solve the puzzle.
- Navigate to 'Earth Science'.
- Identify major inventions.
- Place major inventors.
- Find your dream job on the map.
- Why is mathematics important?

For more information about the map of science for kids or this exercise, please contact Katy Borner (katy@indiana.edu) or Nikki Roberg (nroberg@indiana.edu) at the School of Library and Information Science, Indiana University. These materials were compiled by Nikki Roberg in 2006.

Next P&S exhibit in **New York Hall of Science** opens Dec. 9th, 2006. It features a **science map for kids** that resembles a slide-n-seek puzzle



but uses a geospatial and science base map and new puzzle pieces of inventions and inventors.



New base maps

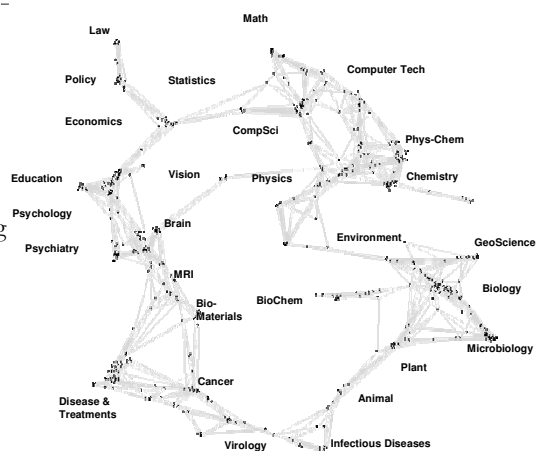


New puzzle pieces

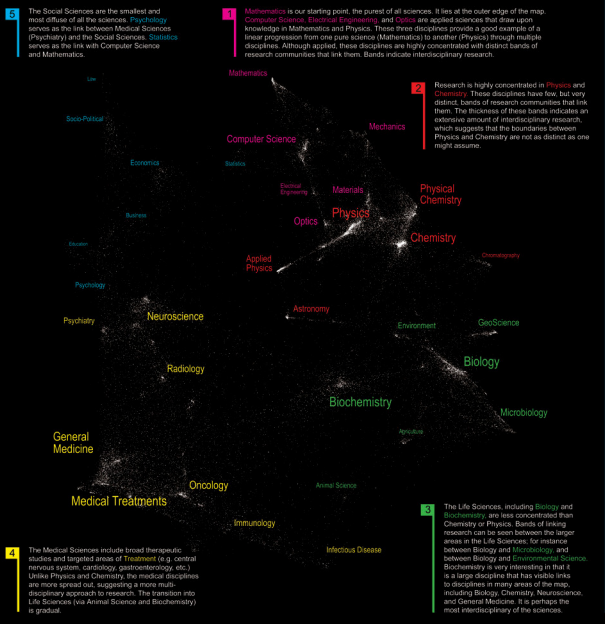
Latest 'Base Map' of Science

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



The Structure of Science



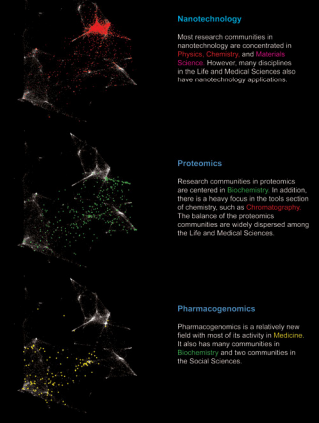
We are all familiar with traditional maps that show the relationships between countries, provinces, states, and cities. Similar relationships exist between the various disciplines and research topics in science. This allows us to map the structure of science.

One of the first maps of science was developed at the Institute for Scientific Information over 30 years ago. It identified 43 areas of science from the citation patterns in 17,000 scientific papers. That early map was intriguing, but it didn't cover enough of science to accurately define its structure.

Things are different today. We have enormous computing power and advanced visualization software that make mapping of the structure of science possible. This galaxy-like map of science (left) was generated at Science National Laboratories using an advanced graph layout routine (VxHub) from the citation patterns in 800,000 scientific papers published in 2002. Each dot in the galaxy represents one of the 36,000 research communities active in science in 2002. A research community is a group of papers (9 on average) that are written on the same research topic in a given year. Over time, communities can be born, continue, split, merge, or die.

The map of science can be used as a tool for science strategy. This is the terrain in which organizations and institutions locate their scientific capabilities. Additional information about the scientific and economic impact of each research community allows policy makers to decide which areas to explore, exploit, defend, or ignore.

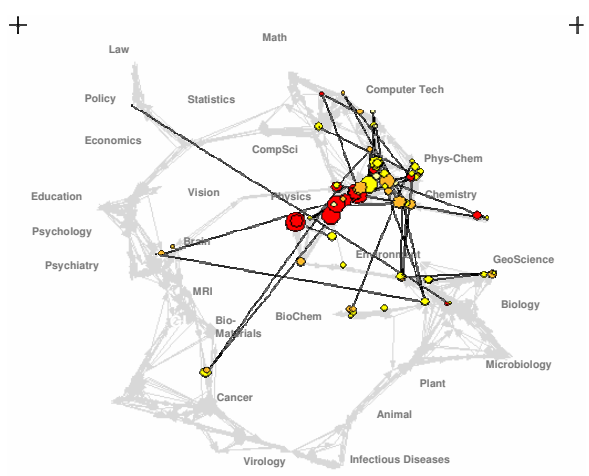
We also envision the map as an educational tool. For children, the theoretical relationship between areas of science can be replaced with a concrete map showing how math, physics, chemistry, biology and social studies interact. For advanced students, areas of interest can be located and neighboring areas can be explored.



Science map applications: Identifying core competency

Kevin W. Boyack & Richard Klavans, unpublished work.

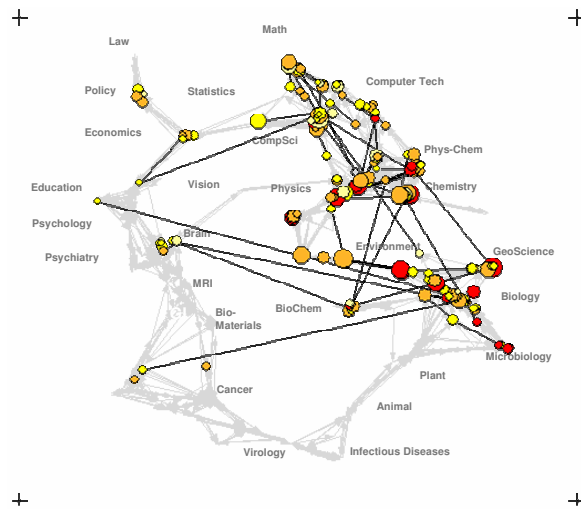
Funding patterns of the US Department of Energy (DOE)



Science map applications: Identifying core competency

Kevin W. Boyack & Richard Klavans, unpublished work.

Funding Patterns of the National Science Foundation (NSF)

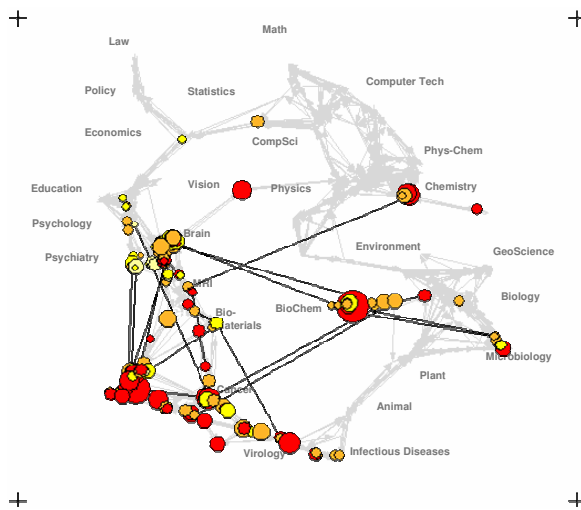


27

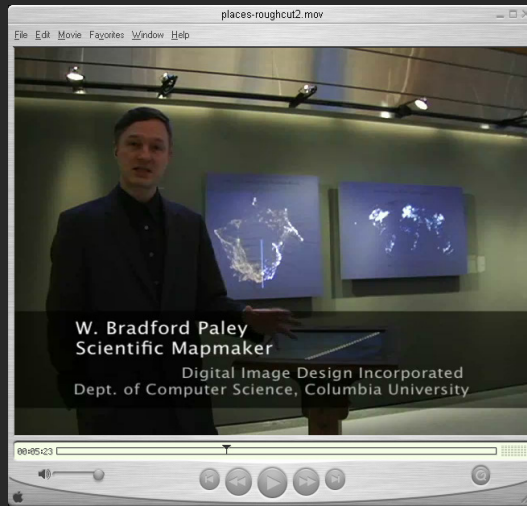
Science map applications: Identifying core competency

Kevin W. Boyack & Richard Klavans, unpublished work.

Funding Patterns of the National Institutes of Health (NIH)



28



W. Bradford Paley
Scientific Mapmaker
Digital Image Design Incorporated
Dept. of Computer Science, Columbia University

"Places & Spaces: Mapping Science" Illuminated Diagram Display

Was on display at the NYPL Science, Industry, and Business Library
Madison/34th, New York City, April 3rd - August 31st, 2006.



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.

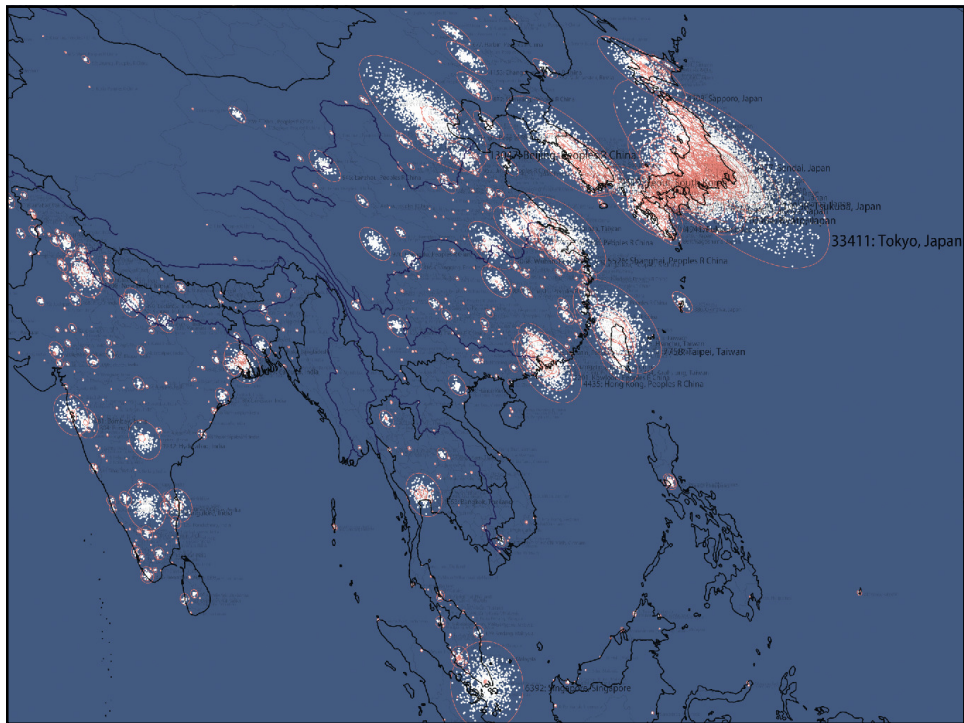
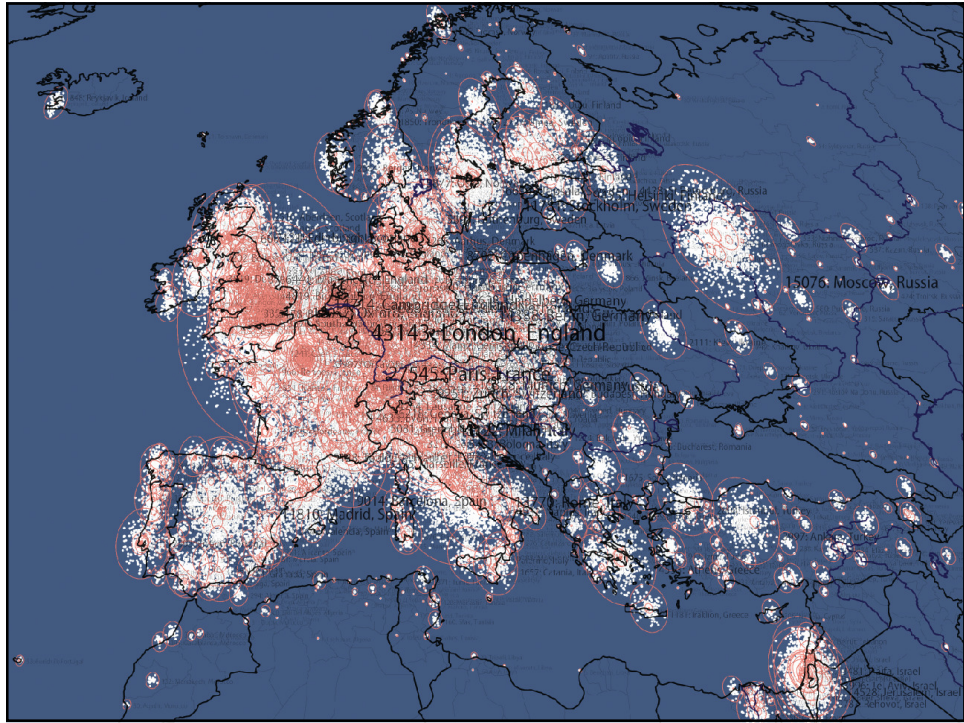
Nanotechnology

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.

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

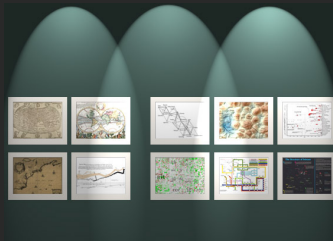
We sweep slowly through adjoining related topics, lighting up the places in the world that study each topic. You may select a subset of the topics that deal with these three interesting subjects by touching it.

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. 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.



The Power of Maps

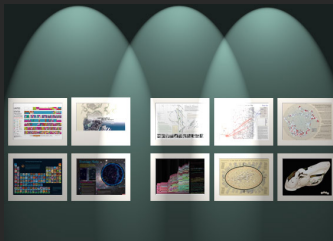
Four Early Maps of Our World
VERSUS
Six Early Maps of Science



(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)



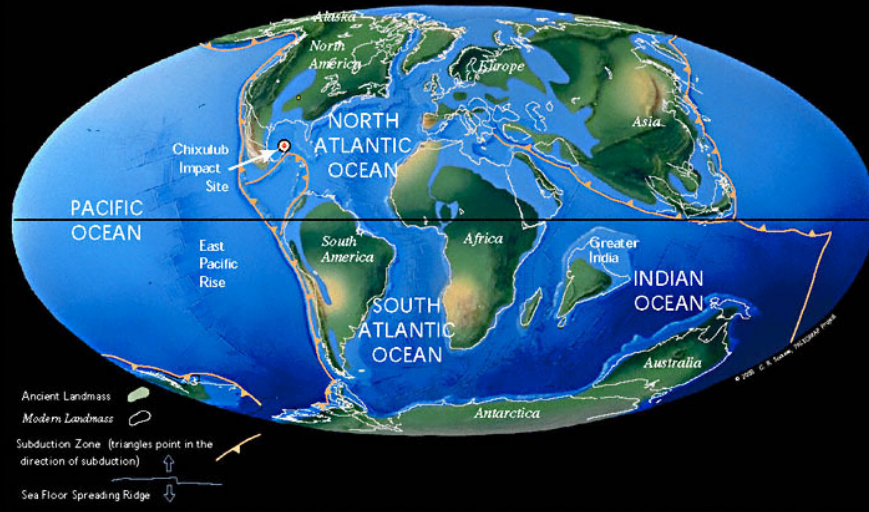
The Power of Forecasts

Four Existing Forecasts
VERSUS
Six Potential Science 'Weather' Forecasts



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

K/T Boundary 66 Ma



<http://www.scolese.com/>

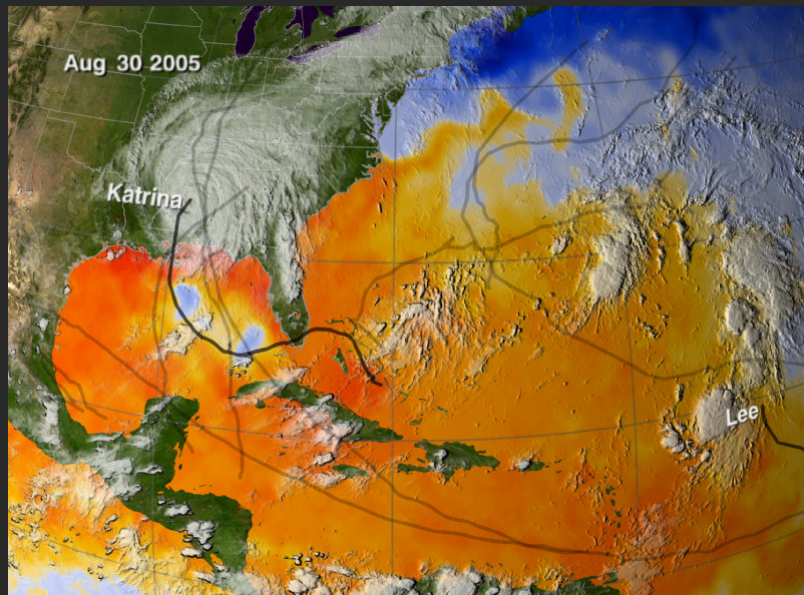
Warnings & Forecasts Graphical Forecasts National Maps Radar Rivers Air Quality Satellite Climate

To view local information, select area of interest and click on the image below.

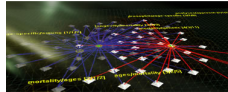


Weather Forecast for Tuesday, April 18, 2006
DOC/NOAA/NWS/NCEP/Hydro-meteorological Prediction Center
Prepared by Fries based on HPC, SPC, and TPC forecasts.

<http://www.weather.gov>



Named Storms, available online at <http://svs.gsfc.nasa.gov/vis/a000000/a003200/a003279>



References

- Börner, Katy. Mapping All of Science: How to Collect, Organize and Make Sense of Mankind's Scholarly Knowledge and Expertise. Submitted to *Environment and Planning B*, Special Issue on *Mapping Humanity's Knowledge and Expertise in the Digital Domain*.
- Börner, Katy, Penumathy, Shashikant, Meiss, Mark and Ke, Weimao. Mapping the Diffusion of Scholarly Knowledge Among Major U.S. Research Institutions. Accepted for *Scientometrics*. Dedicated issue on the *10th International Conference of the International Society for Scientometrics and Informetrics* held in Stockholm.
- Holloway, Todd, Božicevic, Miran and Börner, Katy. Analyzing and Visualizing the Semantic Coverage of Wikipedia and Its Authors. Submitted to *Complexity*, Special issue on *Understanding Complex Systems*. Also available as [cs.IR/0512085](https://arxiv.org/abs/0512085).
- Katy Börner. (2006) Semantic Association Networks: Using Semantic Web Technology to Improve Scholarly Knowledge and Expertise Management. In Vladimir Geroimenko & Chaomei Chen (eds.) *Visualizing the Semantic Web*, Springer Verlag, 2nd Edition, chapter 11, pp. 183-198.
- Boyack, Kevin W., Klavans, R. and Börner, Katy. (2005). Mapping the Backbone of Science. *Scientometrics*, 64(3), 351-374.
- Hook, Peter A. and Börner, Katy. (2005) Educational Knowledge Domain Visualizations: Tools to Navigate, Understand, and Internalize the Structure of Scholarly Knowledge and Expertise. In Amanda Spink and Charles Cole (eds.) *New Directions in Cognitive Information Retrieval*. Springer-Verlag, Netherlands, chapter 5, pp. 187-208.
- Börner, Katy, Dall'Asta, Luca, Ke, Weimao and Vespignani, Alessandro. (April 2005) Studying the Emerging Global Brain: Analyzing and Visualizing the Impact of Co-Authorship Teams. *Complexity*, special issue on *Understanding Complex Systems*, 10(4): pp. 58 - 67. Also available as [cond-mat/0502147](https://arxiv.org/abs/cond-mat/0502147).
- Ord, Terry J., Martins, Emilia P., Thakur, Sidharth, Mane, Ketan K., and Börner, Katy. (2005) Trends in animal behaviour research (1968-2002): Ethoinformatics and mining library databases. *Animal Behaviour*, 69, 1399-1413. [Supplementary Material](#).
- Mane, Ketan K. and Börner, Katy. (2004). [Mapping Topics and Topic Bursts in PNAS](#). *Proceedings of the National Academy of Sciences of the United States of America*, 101(Suppl. 1):5287-5290. Also available as [cond-mat/0402380](https://arxiv.org/abs/cond-mat/0402380).
- Börner, Katy, Maru, Jeeagar and Goldstone, Robert. (2004). [The Simultaneous Evolution of Author and Paper Networks](#). *Proceedings of the National Academy of Sciences of the United States of America*, 101(Suppl.1):5266-5273. Also available as [cond-mat/0311459](https://arxiv.org/abs/cond-mat/0311459).



All Maps of Science Are for Sale

http://vw.indiana.edu/places/spaces/ordermaps/

Home My Netscape Search

Netscape - Enter Search Terms Search Highlight Pop-Ups Blocked: 45

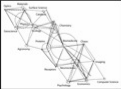
New Tab Places & Spaces > Order Maps

PHOTO PAPER		WATERCOLOR PAPER	
QUANTITY	PRICE (per map)	QUANTITY	PRICE (per map)
1 - 4	\$50 \$60	1 - 4	\$95
5 - 9	\$45 \$55	5 - 9	\$90
10+	\$40 \$50	10+	\$85

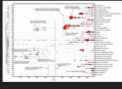
Shipping and handling within US is \$12. Watercolor paper used in NYPL exhibit.

* Shipping and handling within US is about \$12.
All maps are printed by [SpectralImaging](#)

Six Early Maps of Science (1st Iteration of Places & Spaces Exhibit)



Henry Small, 1996 Map of Sciences: A Network Representation of the 43 Fourth Level Clusters Based on Data from the 1996 Science Citation Index, 1999



Steven Morris, Timeline of 60 years of anthrax research literature, 2005

20 x 26" on Photo Paper

20 x 26" on Photo Paper

