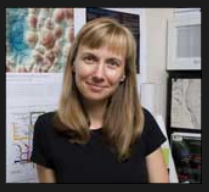
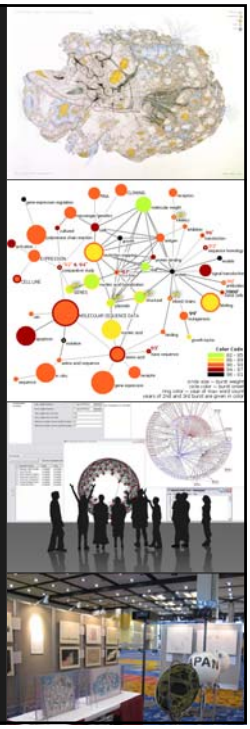
Mapping the Evolution of Science



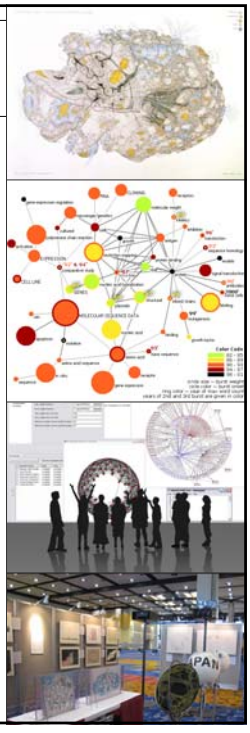
Dr. Katy Börner
Cyberinfrastructure for Network Science Center, Director
Information Visualization Laboratory, Director
School of Library and Information Science
Indiana University, Bloomington, IN
katy@indiana.edu

Heraeus Seminar on the Evolution of Physics, Bad Honnef, Germany
January 22, 2008



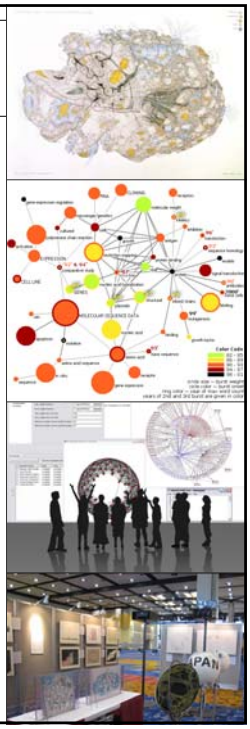
Overview

- Concepts -> Conceptualization
- Models -> Descriptive Models
  -> Process Models
- CI Design -> IVC, NWB, SDB
- Communication -> "Mapping Science" Exhibit



Overview

- **Concepts** **-> Conceptualization**
- Models -> Descriptive Models
  -> Process Models
- CI Design -> IVC, NWB, SDB
- Communication -> "Mapping Science" Exhibit



Börner: "Mapping the Evolution of Science"

Conceptualizing Science

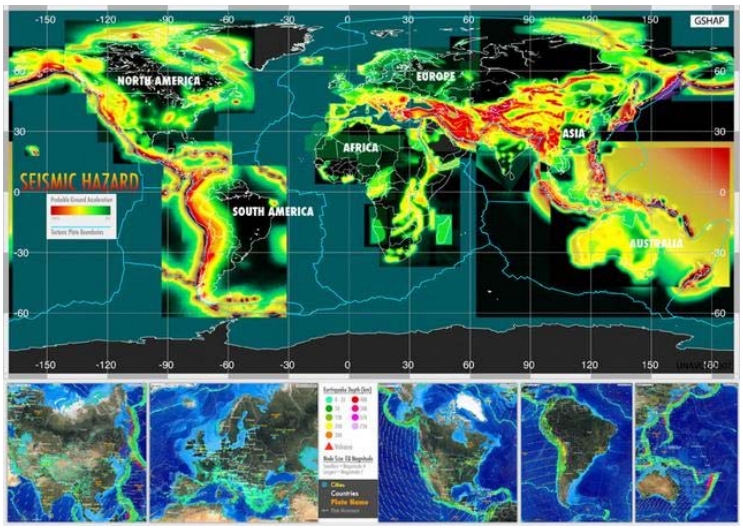
Hypothetical Model of the Evolution and Structure of Science, by Daniel Zeller
On display in 3rd iteration of exhibit

Authors are mortal. Papers are immortal.
Densely knit communities. The importance of weak links.
Cumulative structure of science. Good and bad years. Monsters = 'the unknown' or voids.
Impact of funding on science (yellow).

Tectonic Movements and Earthquake Hazard Predictions

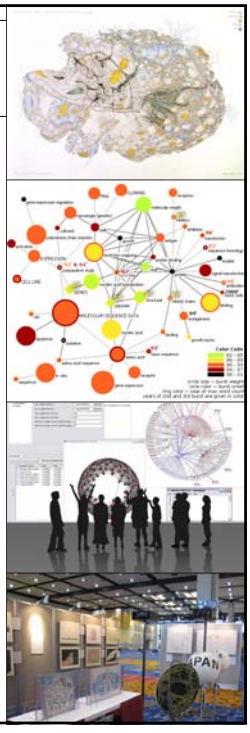
Michael W. Hamburger, Lou Estey, Chuck Meertens (Data & Visualization), Elisha Hardy (Graphic Design)

On display in 3rd iteration of exhibit.



Overview

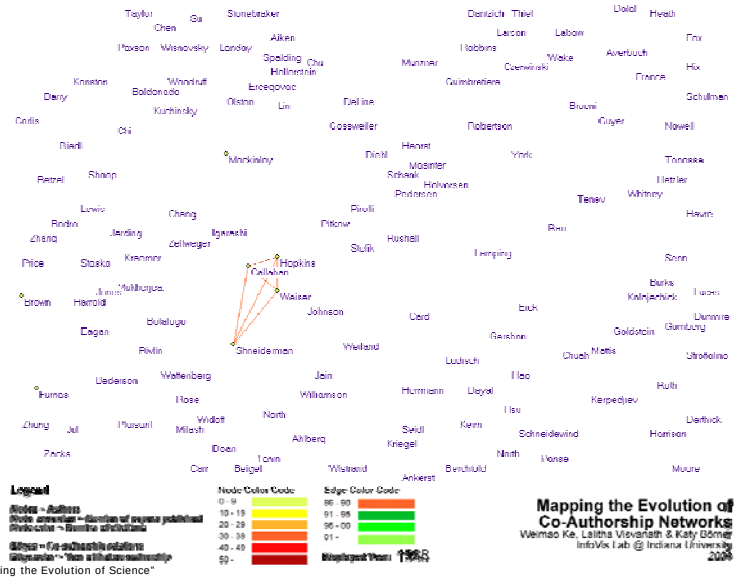
- Concepts -> Conceptualization
- **Models** **-> Descriptive Models**
  -> Process Models
- CI Design -> IVC, NWB, SDB
- Communication -> "Mapping Science" Exhibit



Börner: "Mapping the Evolution of Science"

Mapping the Evolution of Co-Authorship Networks

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

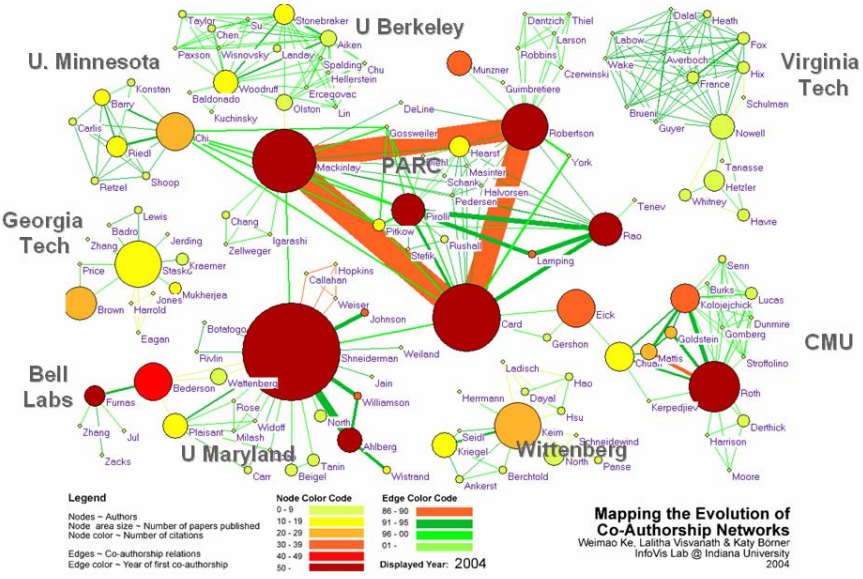


Börner: "Mapping the Evolution of Science"

7

Mapping the Evolution of Co-Authorship Networks

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



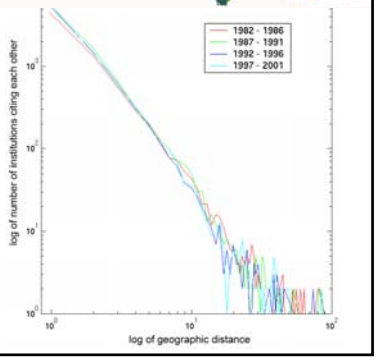
8

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

Börner, Katy, Penumarthy, 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 instructions?



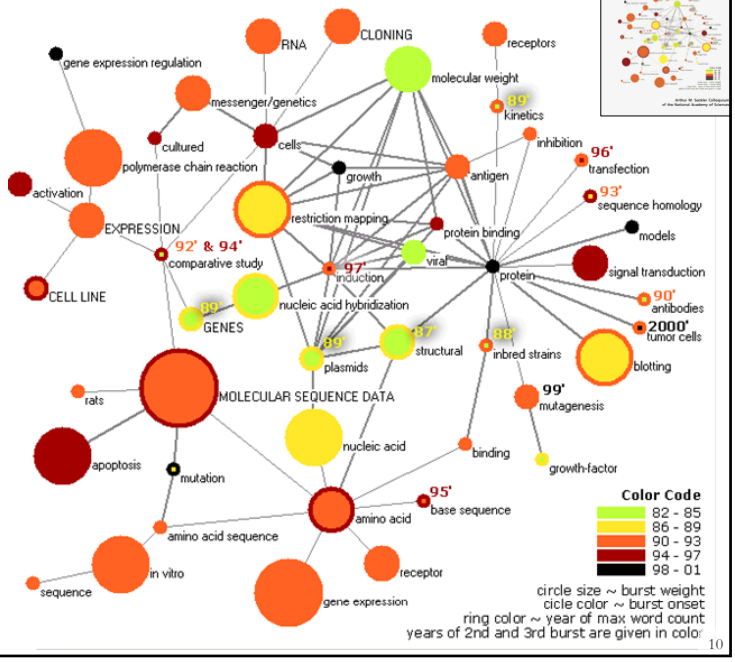
Contributions:

- 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.

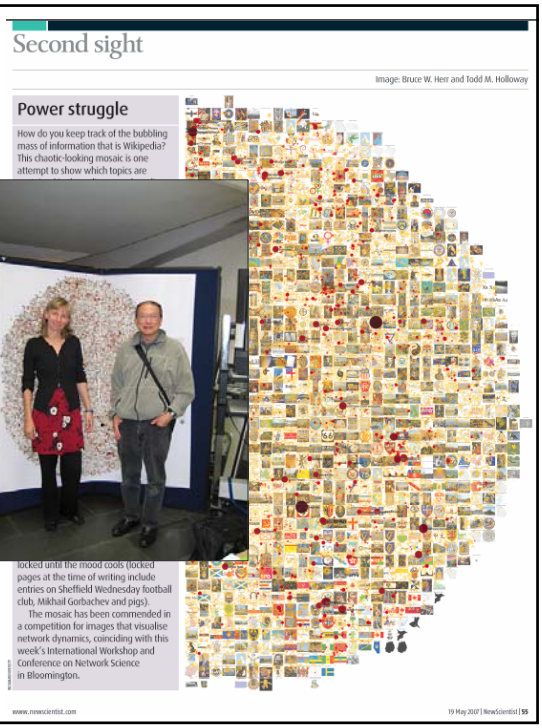


10

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



Second sight

Image: Bruce W. Herr and Todd M. Holloway

Power struggle

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

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.

www.newscientist.com

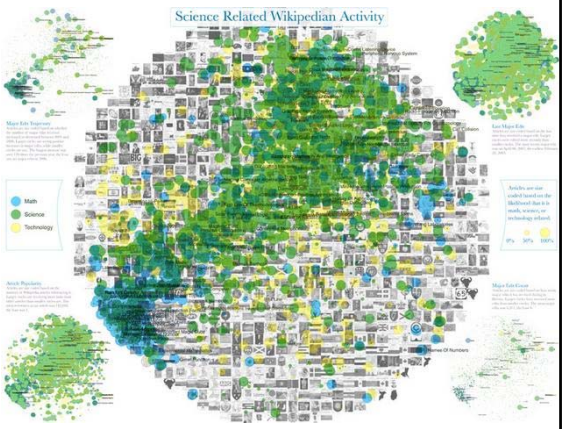
19 May 2007 | NewScientist | 55

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.

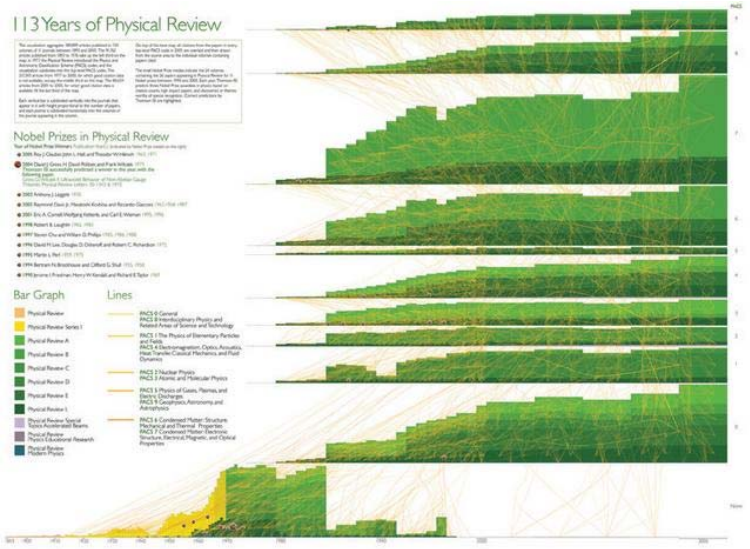
Corners show articles size coded according to
-article edit activity (top left),
- number of major edits (top right),
- number of bursts in edit activity (bottom, right)
- indegree (bottom left).



113 Years of Physical Review

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

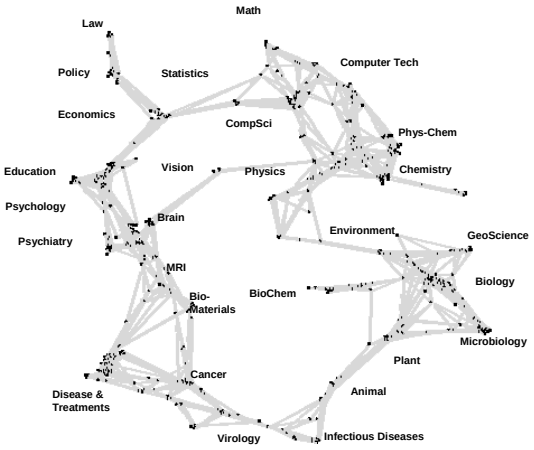
Bruce W. Herr II and Russell Duhon (Data Mining & Visualization), Elisha F. Hardy (Graphic Design), Shashikant Penumarthy (Data Preparation) and Katy Börner (Concept)



'Base Map' of Science

Kevin W. Boyack & Richard Klavans,

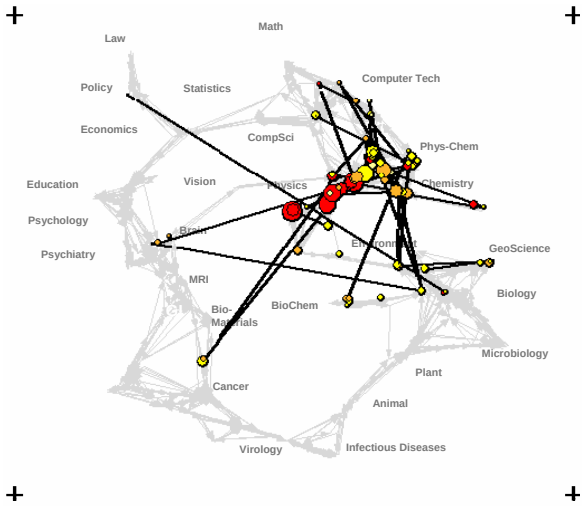
- Uses combined SCI/SSCI from 2002
  - 1.07M papers, 24.5M references, 7,300 journals
  - Bibliographic coupling of papers, aggregated to journals
- Initial ordination and clustering of journals gave 671 clusters
- Coupling counts were reaggregated at the journal cluster level to calculate the
  - (x,y)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



Science map applications: Identifying core competency

Kevin W. Boyack & Richard Klavans,

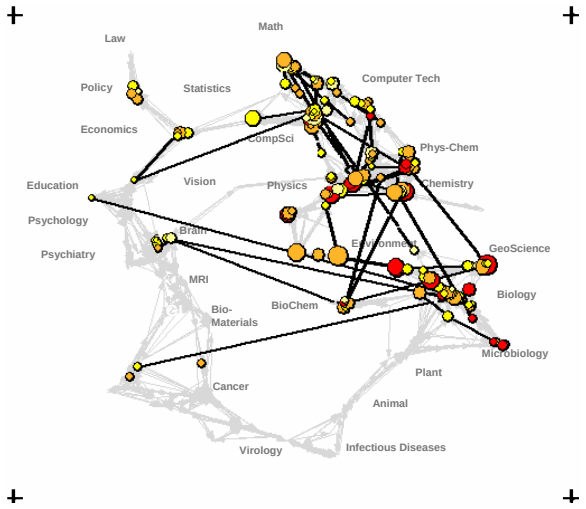
Funding patterns of the US Department of Energy (DOE)



Science map applications: Identifying core competency

Kevin W. Boyack & Richard Klavans,

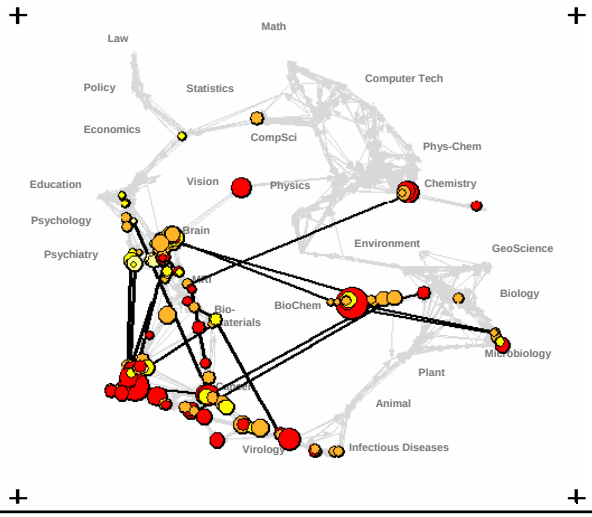
Funding Patterns of the National Science Foundation (NSF)

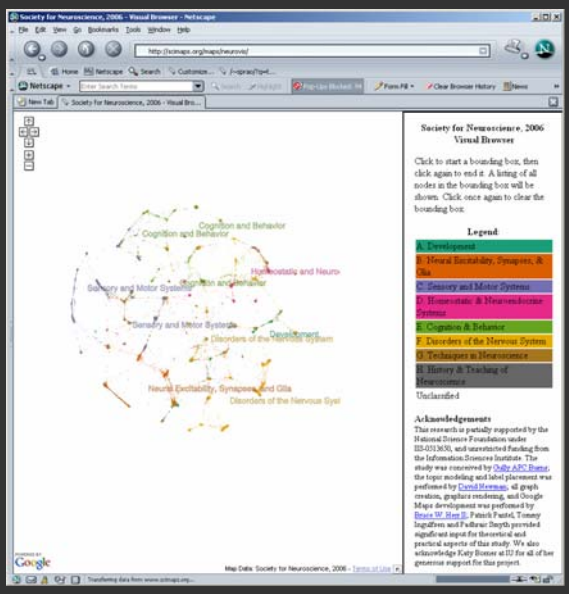


Science map applications: Identifying core competency

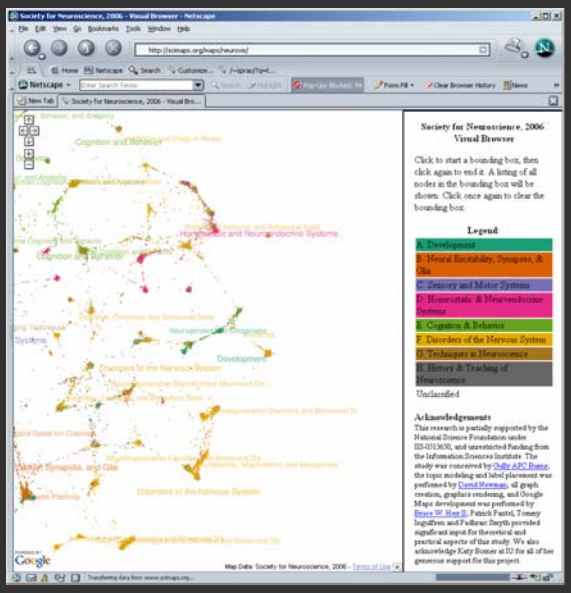
Kevin W. Boyack & Richard Klavans,

Funding Patterns of the National Institutes of Health (NIH)

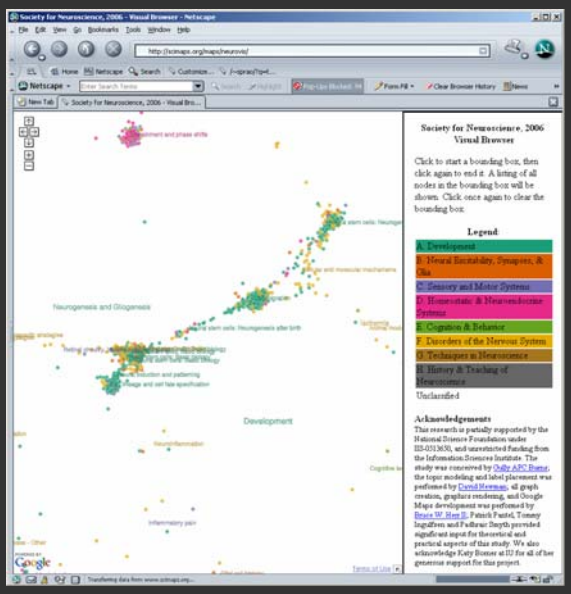




Bruce W. Herr II, Gully Burns (USC), David Newman (UCI), Society for Neuroscience, 2006 Visual Browser, 2007, http://scimaps.org/maps/neurovis/

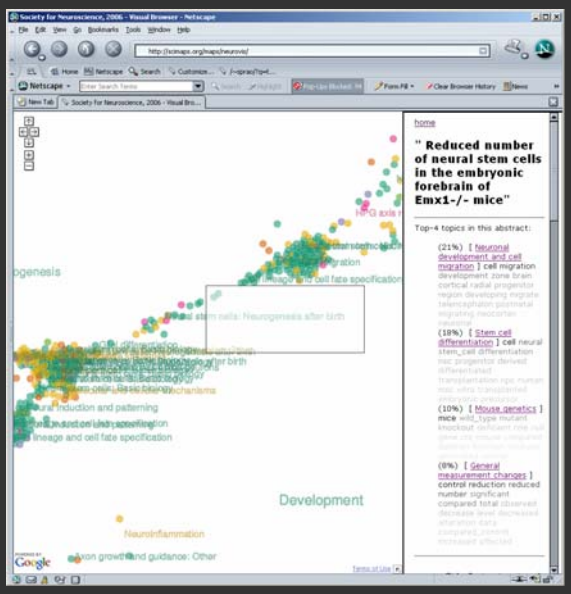


Bruce W. Herr II, Gully Burns (USC), David Newman (UCI), Society for Neuroscience, 2006 Visual Browser, 2007, http://scimaps.org/maps/neurovis/

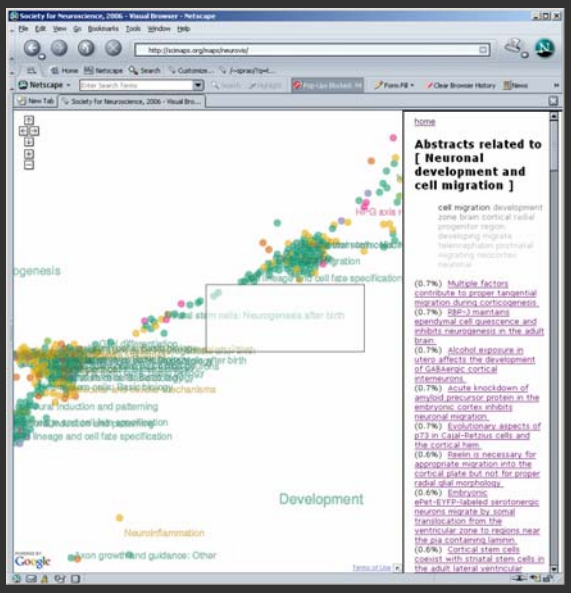


Bruce W. Herr II, Gully Burns (USC), David Newman (UCI), Society for Neuroscience, 2006 Visual Browser, 2007, http://scimaps.org/maps/neurovis/

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



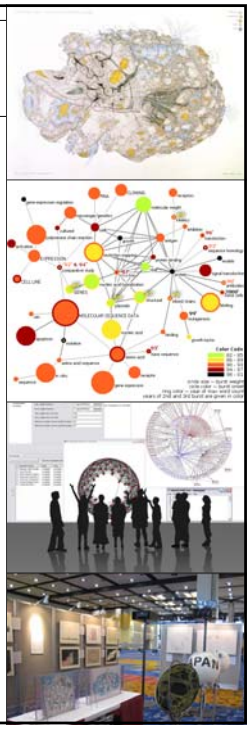
Bruce W. Herr II, Gully Burns (USC), David Newman (UCI), Society for Neuroscience, 2006 Visual Browser, 2007, http://scimaps.org/maps/neurovis/



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

Overview

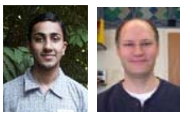
- Concepts -> Conceptualization
- **Models** -> Descriptive Models
  -> Process Models
- CI Design -> IVC, NWB, SDB
- Communication -> "Mapping Science" Exhibit



Börner: "Mapping the Evolution of Science"

The TARL Model (Topics, Aging, and Recursive Linking)

Börner, Katy, Maru, Jeegar & Goldstone, Robert. (2004). The Simultaneous Evolution of Author and Paper Networks. Proceedings of the National Academy of Sciences of the United States of America,. Vol. 101(Suppl. 1), 5266-5273.



Basic Assumptions

- Co-author and paper-citation networks co-evolve.
- Authors come and go. Papers are forever.
- Only authors that are 'alive' are able to co-author.
- All existing (but no future) papers can be cited.

Unique Features

- Author and paper networks grow simultaneously.
- Preferential attachment is modeled as an emergent property of the elementary, local networking activity of authors reading and citing papers, but also the references listed in papers.
- The number of topics is linearly correlated with the clustering coefficient of the resulting network and can be determined from the cluster coefficient observed in real world networks.
- The model incorporates aging, i.e., a bias for authors to cite recent papers and hence papers are not only clustered by topic, but also in time.

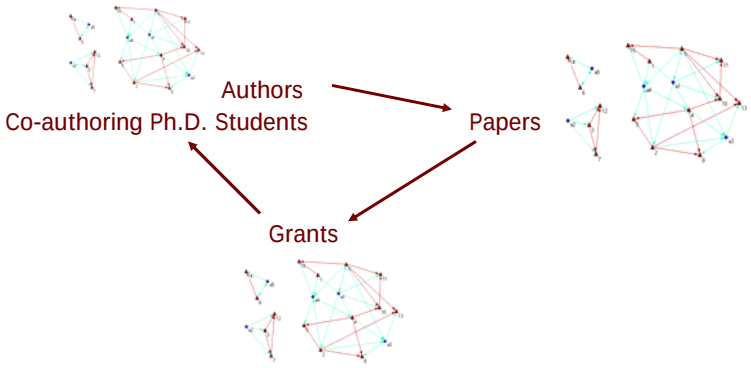
pseudo code

```
// Initialization
generate #_papers papers and assign a random topic to each paper;
generate #_authors authors and assign a random topic to each author;
randomly assign #_co-authors+1 authors to papers of the same topic;
// Simulation
for each year do {
    add #_new_authors new authors, deactivate authors older than #_author_age;
    for each topic do {
        randomly partition set of authors into author_groups of size #_co-authors+1;
        for each author_group do {
            for each new_paper to be produced, do {
                generate new_paper;
                randomly select #_read_papers from existing papers;
                get all references of read_papers up to #_reference_path_length;
                for each new_paper_reference do {
                    select a time_slice from (start year to curr_year-1) with probability given in aging
                    randomly select a paper published or cited in this time_slice; as a new_paper_ref
                    add the new_paper_reference to new_paper;
                }
            }
        }
    }
    add all new papers to the set of existing papers;
    add new links to author and paper information;
}
```

The TARL Model (Topics, Aging, and Recursive Linking)

Modeling Network Ecologies

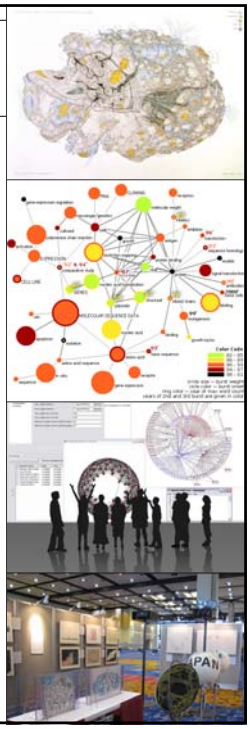
Most real world networks exist within a delicate ecology of networks.



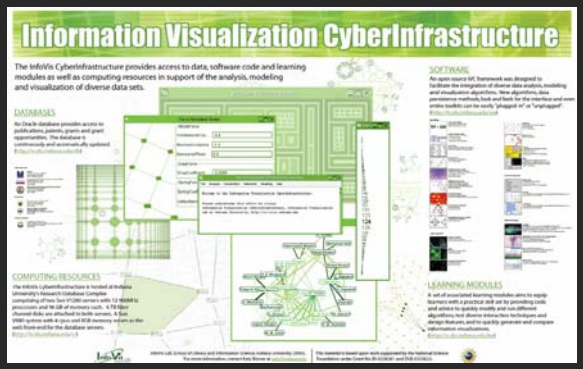
To fully understand, e.g., the 'rich get richer effect' or the 'diffusion of knowledge', different networks need to be considered simultaneously and the interplay of network structure and network dynamics has to be studied.

Overview

- Concepts -> Conceptualization
- Models -> Descriptive Models
  -> Process Models
- **CI Design** **-> IVC, NWB, SDB**
- Communication -> "Mapping Science" Exhibit



Börner: "Mapping the Evolution of Science"



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

SDB | SCHOLARLY DATABASE

Scholarly Database
http://sdb.slis.indiana.edu

CIShell
cyberinfrastructure shell

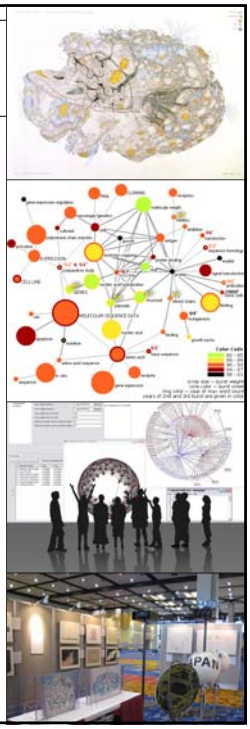
NetworkWorkbench
A Workbench for Network Scientists

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-Laszlo Barabasi, Santiago Schnell, Alessandro Vespignani & Stanley Wasserman, Eric Wernert (Senior Personnel), $1,120,926) Sept. 05 - Aug. 08.
http://nwb.slis.indiana.edu

HPNAP
High Performance Network Applications Program

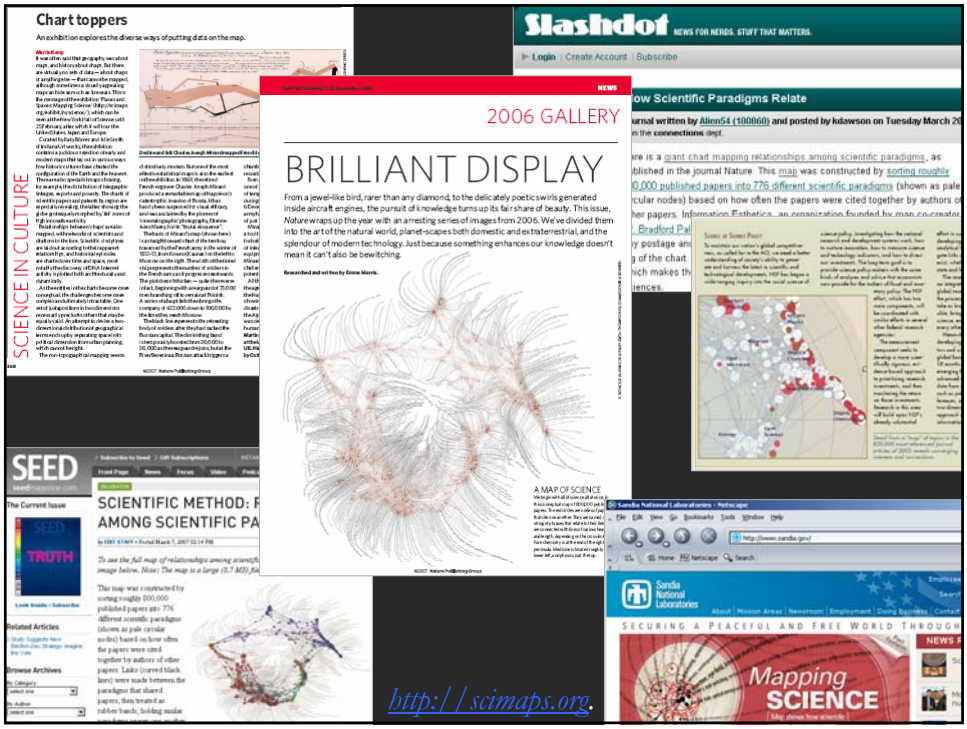
Overview

- Concepts -> Conceptualization
- Models -> Descriptive Models
  -> Process Models
- CI Design -> IVC, NWB, SDB
- **Communication** **-> "Mapping Science" Exhibit**



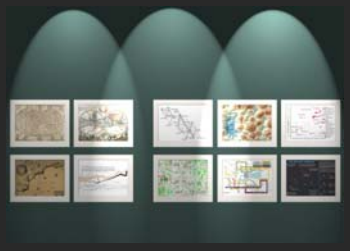
Börner: "Mapping the Evolution of Science"

Places & Spaces: Mapping Science *exhibit, see also http://scimaps.org.*



The Power of Maps

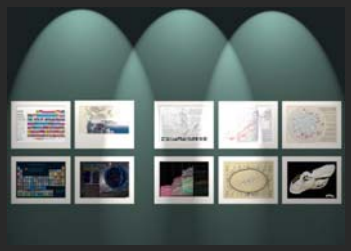
Four Early Maps of Our World
VERSUS
Six Early Maps of Science



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

The Power of Reference Systems

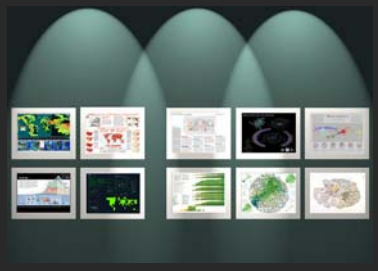
Four Existing Reference Systems
VERSUS
Six Potential Reference Systems of Science



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

The Power of Forecasts

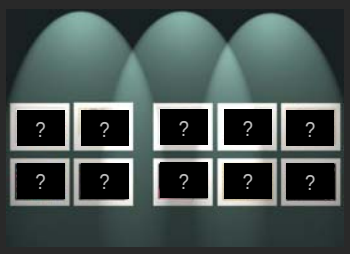
Four Existing Forecasts
VERSUS
Six Potential Science 'Weather' Forecasts



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

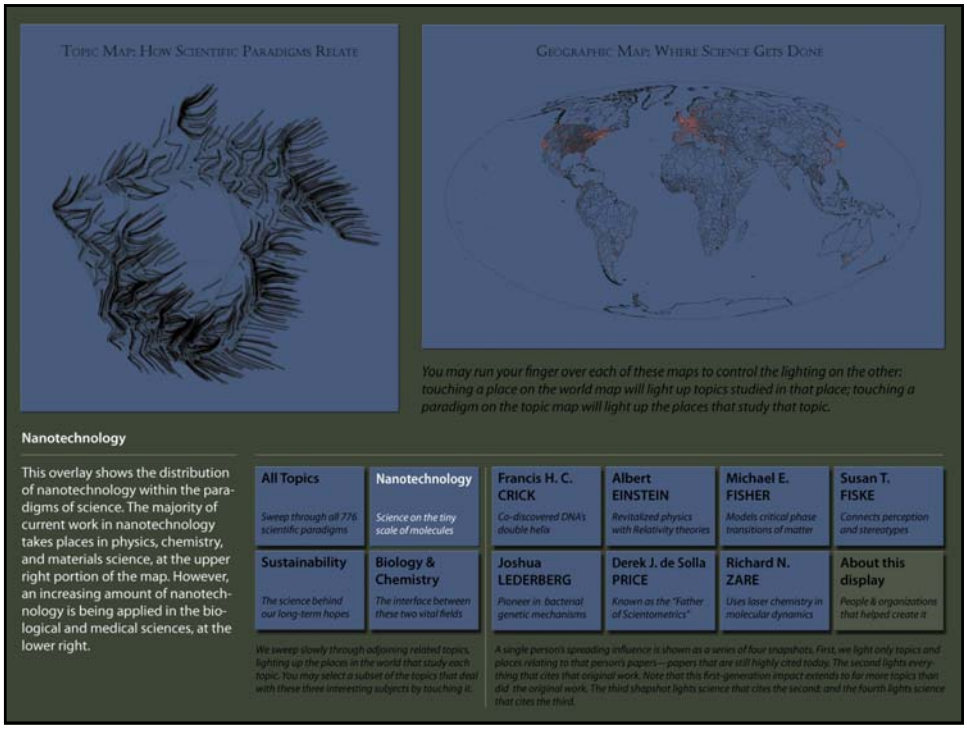
Science Maps for Economic Decision Making

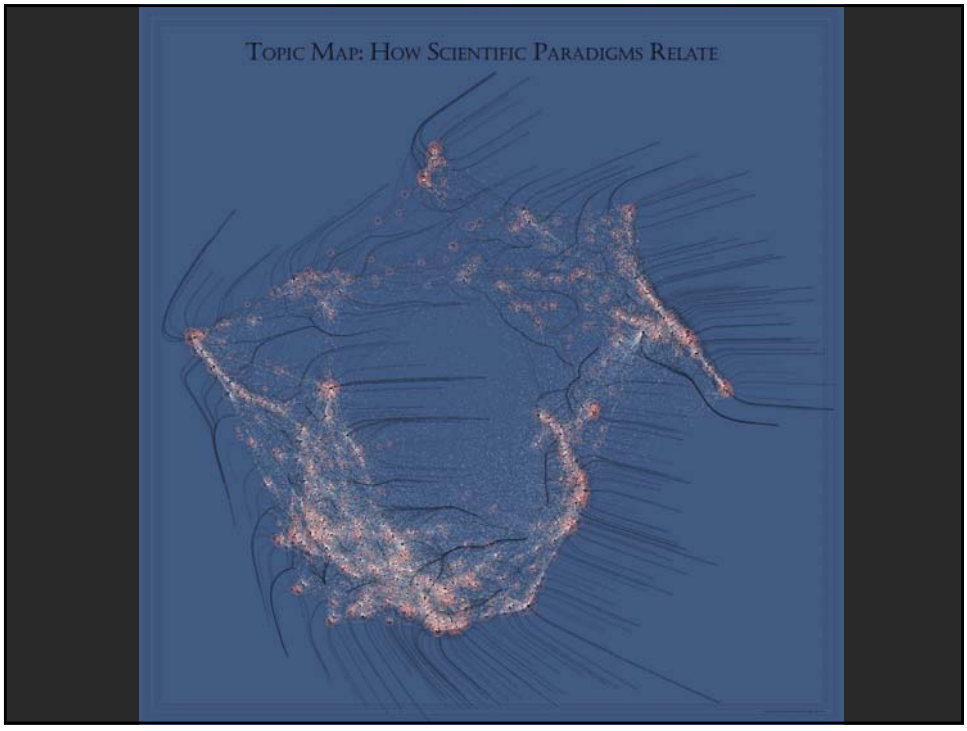
Four Existing Maps
VERSUS
Six Science Maps

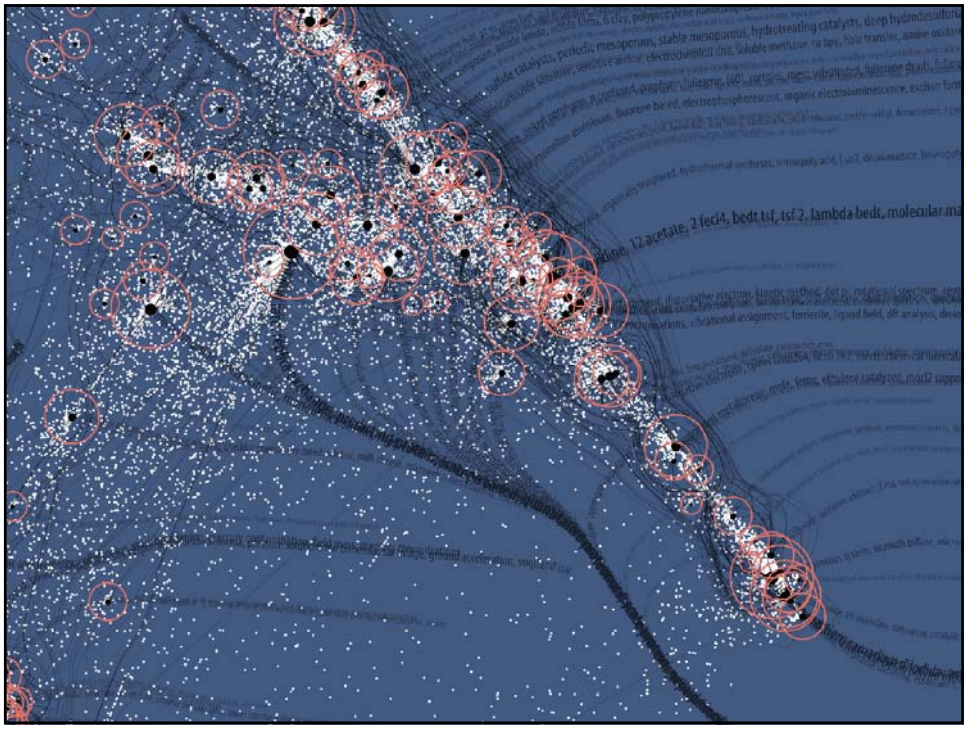


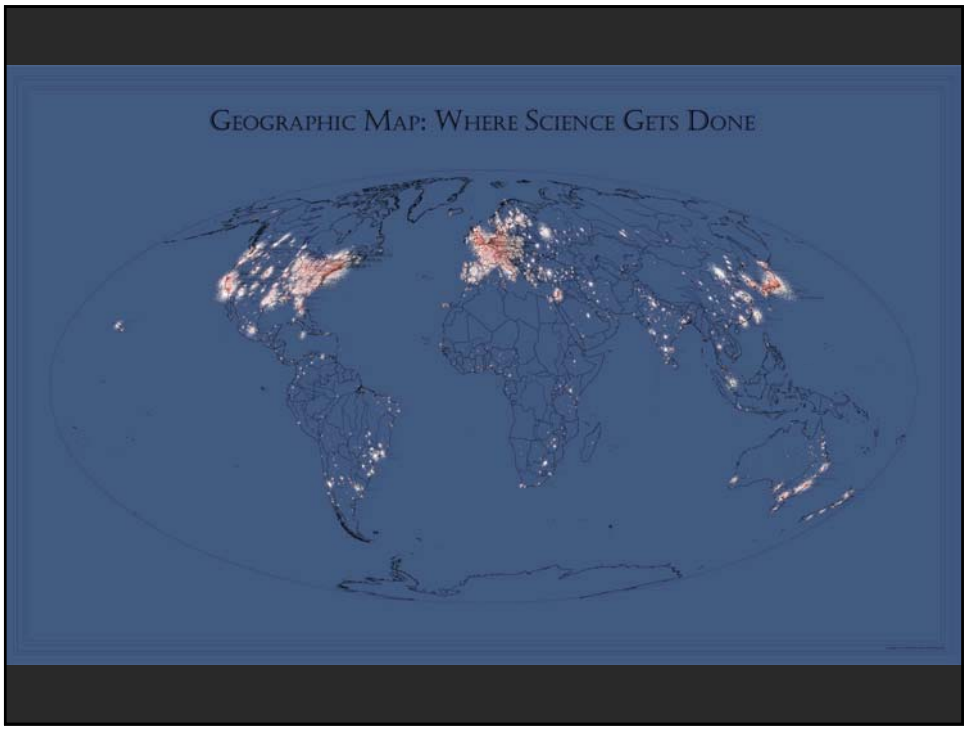
(4th Iteration of Places & Spaces Exhibit - 2008)

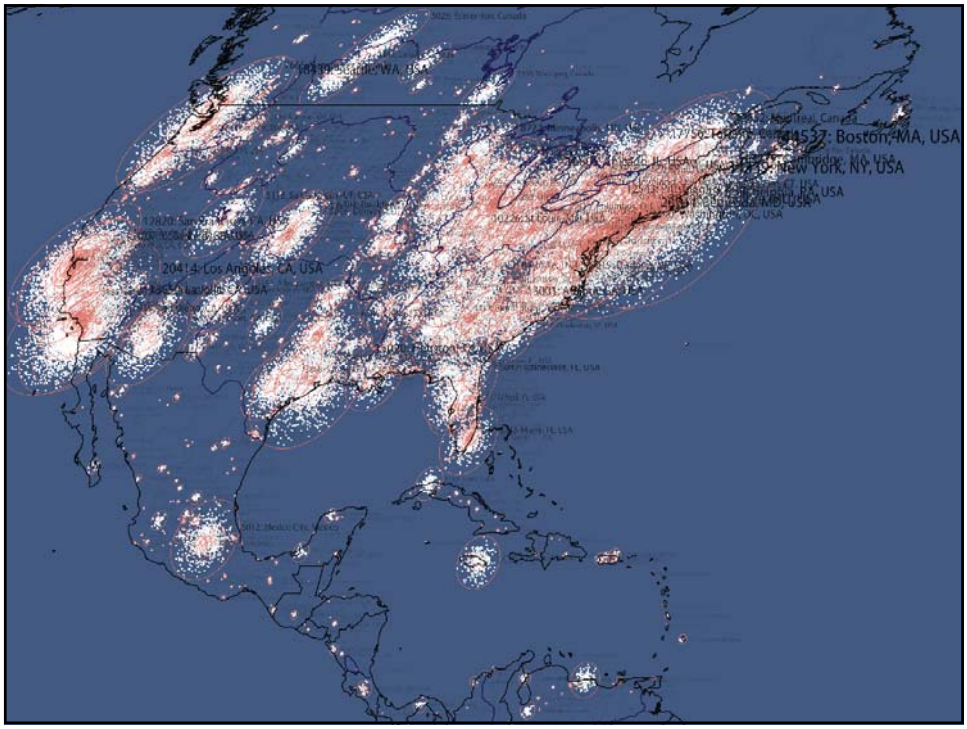
"Places & Spaces: Mapping Science"
on display at on display at the American Museum of Science and Energy, Oak Ridge, TN,
September 7, 2007- January 7, 2008.

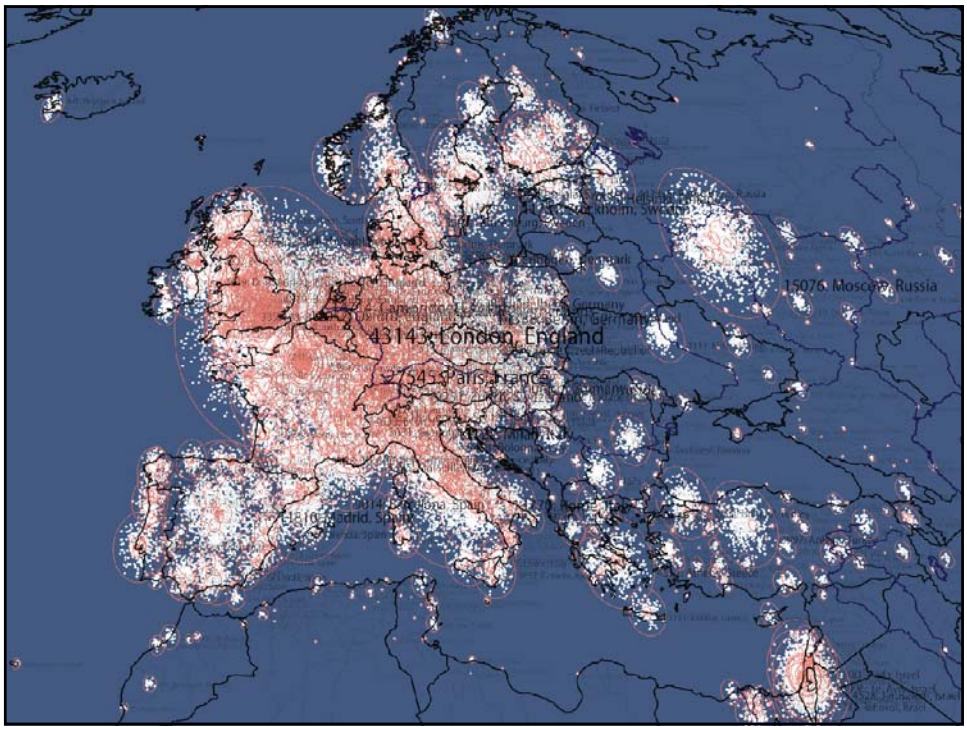

Topic Map: How Scientific Paradigms Relate
Geographic Map: Where Science Gets Done
You may run your finger over each of these maps to control the lighting on the other: touching a place on the world map will light up topics studied in that place; touching a paradigm on the topic map will light up the 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
Sweep through all 776 scientific paradigms
Nanotechnology
Science on the tiny scale of molecules
Francis H. C. CRICK
Co-discovered DNA's double helix
Albert EINSTEIN
Revitalized physics with Relativity theories
Michael E. FISHER
Models critical phase transitions of matter
Susan T. FISKE
Connects perception and stereotypes
Sustainability
The science behind our long-term hopes
Biology & Chemistry
The interface between these two vital fields
Joshua LEDERBERG
Pioneer in bacterial genetic mechanisms
Derek J. de Solla PRICE
Known as the "Father of Scientometrics"
Richard N. ZARE
Uses laser chemistry in molecular dynamics
About this display
People & organizations that helped create it
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 shapshot lights science that cites the second; and the fourth lights science that cites the third.

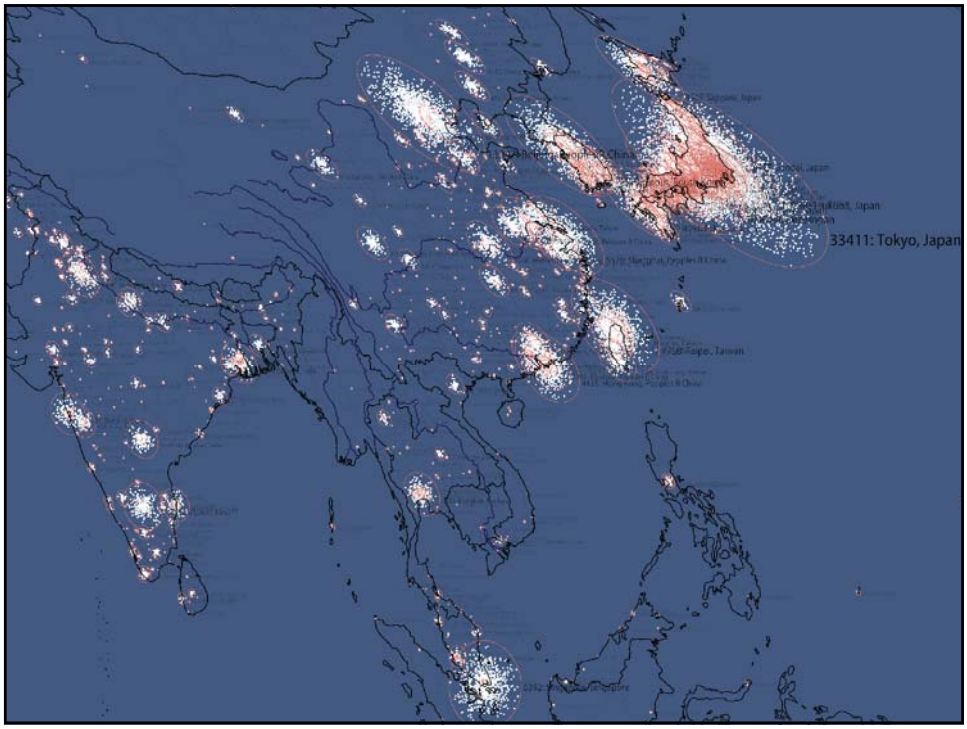

Topic Map: How Scientific Paradigms Relate




Geographic Map: Where Science Gets Done


20414 Los Angeles, CA, USA
New York, NY, USA


15076 Moscow, Russia
43143 London, England
27543 Paris, France


33411: Tokyo, Japan

54
SOCCER

References

Shiffrin, Richard M. and Börner, Katy. (April 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, Chen, Chaomei, and Boyack, Kevin. (2003). **Visualizing Knowledge Domains**. In Blaise Cronin (Ed.), Annual Review of Information Science & Technology, Information Today, Inc./American Society for Information Science and Technology, Medford, NJ, volume 37, chapter 5, pp. 179-255. http://ivl.slis.indiana.edu/km/pub/2003-borner-arist.pdf

Börner, Katy, Sanyal, Soma and Vespignani, Alessandro (2007). **Network Science**. In Blaise Cronin (Ed.), Annual Review of Information Science & Technology, Information Today, Inc./American Society for Information Science and Technology, Medford, NJ, volume 41, chapter 12, pp. 537-607. http://ivl.slis.indiana.edu/km/pub/2007-borner-arist.pdf

Boyack, Kevin W, Börner, Katy & Klavans, Richard. (2007). Mapping the Structure and Evolution of Chemistry Research. Proceedings of the 11th International Conference on Scientometrics and Informetrics (ISSI 2007), Madrid, Spain, June 25-27, pp. 112-123.

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

These and more papers are linked from http://ivl.slis.indiana.edu/publications/

The End.