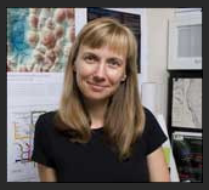
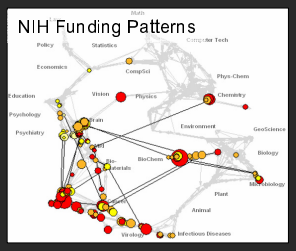
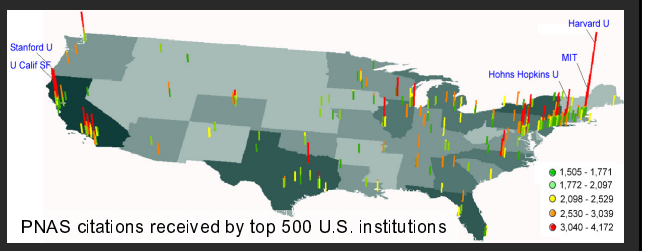
Mapping the Structure and Evolution of Science Locally and Globally



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

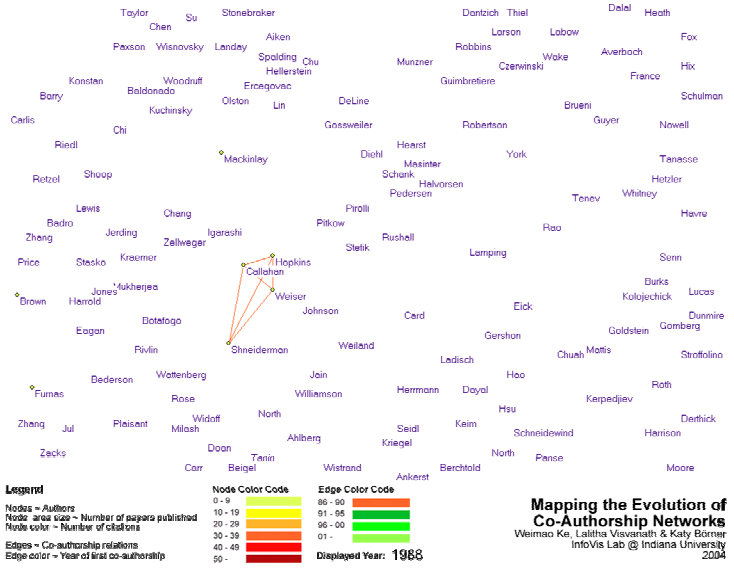
This Talk has Three Parts:

1. **Mapping Science: Why?**
2. Mapping Science: How?

3. Mapping Science: Applications

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

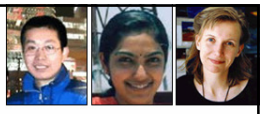
Ke, Visvanath & Börner, (2004) Won 1st price at the IEEE InfoVis Contest.

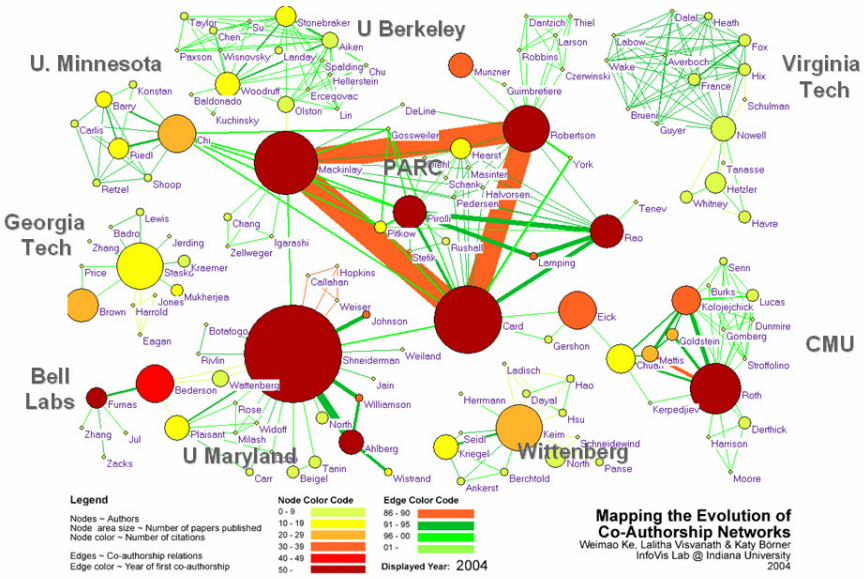


4

Mapping the Evolution of Co-Authorship Networks

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





5

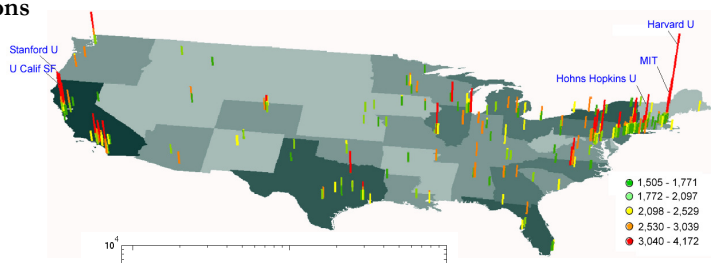
Spatio-Temporal Information Production and Consumption of Major U.S. Research Institutions

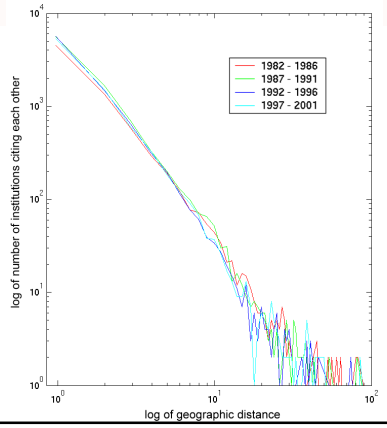
Börner & Penumarthy. (2005)

Does Internet lead to more global citation patterns, i.e., more citation links between papers produced at geographically distant research instructions?

Analysis of top 500 most highly cited U.S. institutions.

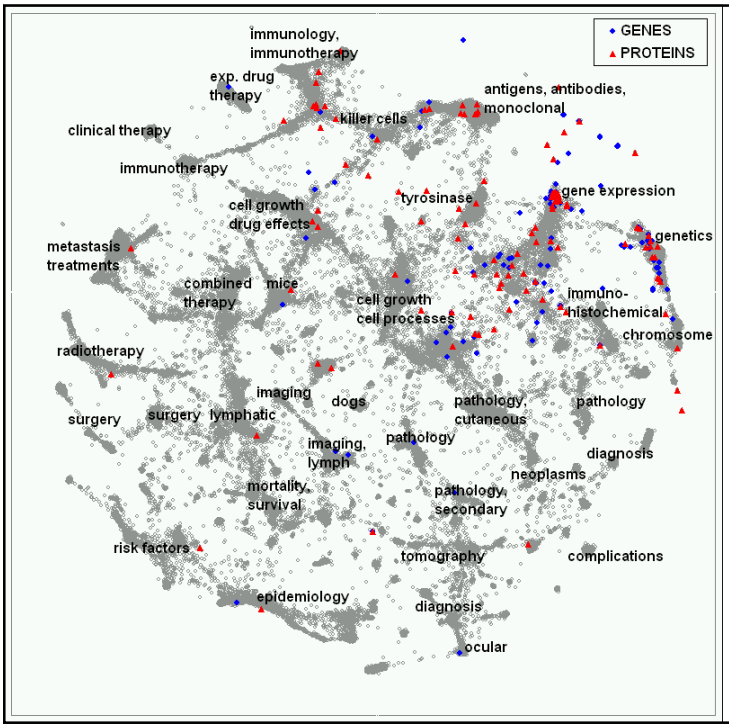
Each institution is assumed to produce and consume information.





$\gamma_{82\text{-}86} = 1.94$ (R^2=91.5%)
$\gamma_{87\text{-}91} = 2.11$ (R^2=93.5%)
$\gamma_{92\text{-}96} = 2.01$ (R^2=90.8%)
$\gamma_{97\text{-}01} = 2.01$ (R^2=90.7%)

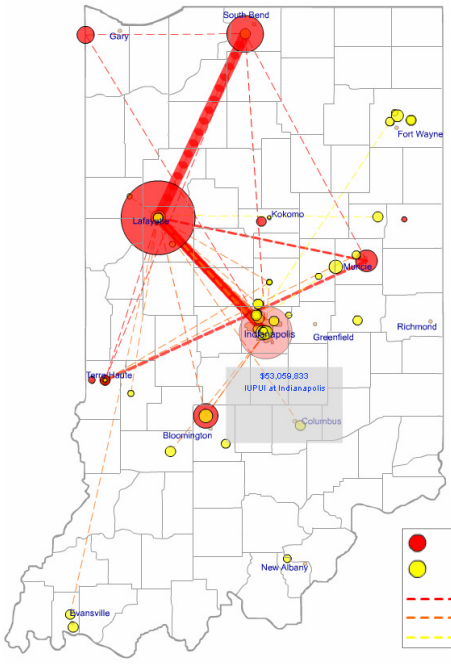
6



Mapping Medline Papers, Genes, and Proteins Related to Melanoma Research

Boyack, Mane & Börner. (2004) IV Conference, pp. 965-971.

7



Mapping Indiana's Intellectual Space

(Ke, Börner & Mei, 2005)

Identify

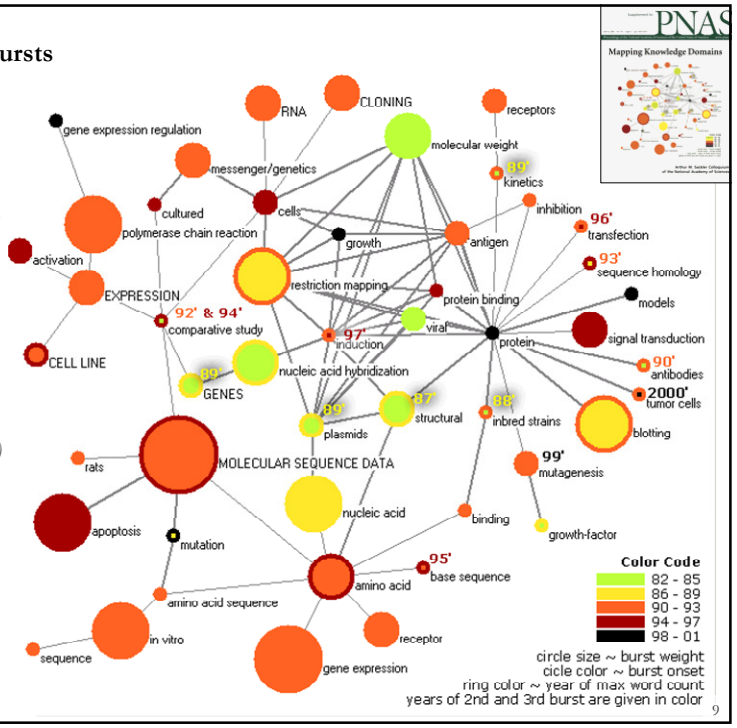
- Pockets of innovation
- Pathways from ideas to products
- Interplay of industry and academia

8

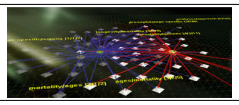
Mapping Topic Bursts

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.



9



Mapping 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.

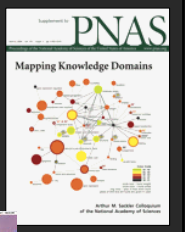
For Society

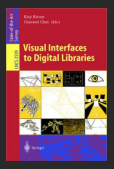
- Dramatically improved access to scientific knowledge and expertise.

10

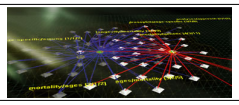
This Talk has Three Parts:

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





- *Shiffrin, Richard M. and Börner, Katy (Eds.) (2004). Mapping Knowledge Domains. Proceedings of the National Academy of Sciences of the United States of America, 101(Suppl_1).*
- *Börner, Katy, Chen, Chaomei, and Boyack, Kevin. (2003). Visualizing Knowledge Domains. In Blaise Cronin (Ed.), Annual Review of Information Science & Technology, Volume 37, Medford, NJ: Information Today, Inc./American Society for Information Science and Technology, chapter 5, pp. 179-255.*
- *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. BROADENING By citation By terms	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 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

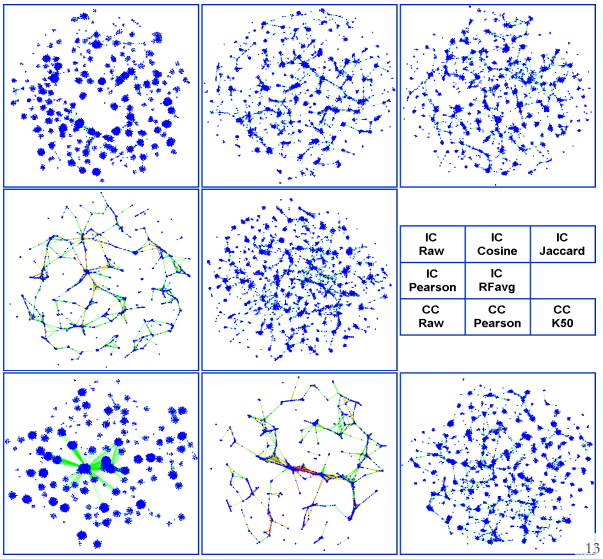
Börner, Chen & Boyack.. (2003) Visualizing Knowledge Domains. In Blaise Cronin (Ed.), Annual Review of Information Science & Technology, Volume 37, Medford, NJ: Information Today, Inc./American Society for Information Science and Technology, chapter 5, pp. 179-255.

12

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.

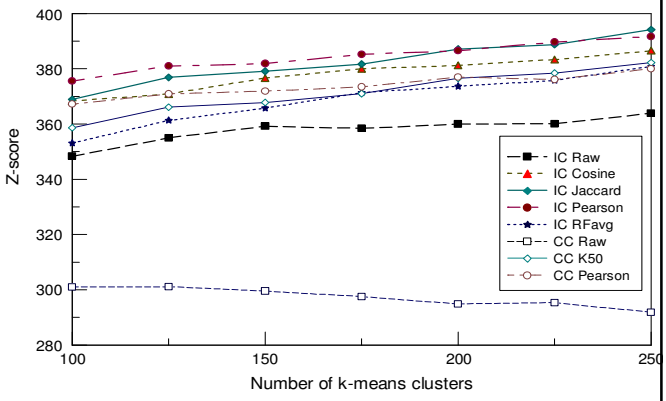


IC Raw	IC Cosine	IC Jaccard
IC Pearson	IC RFavg	
CC Raw	CC Pearson	CC K50

13

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.

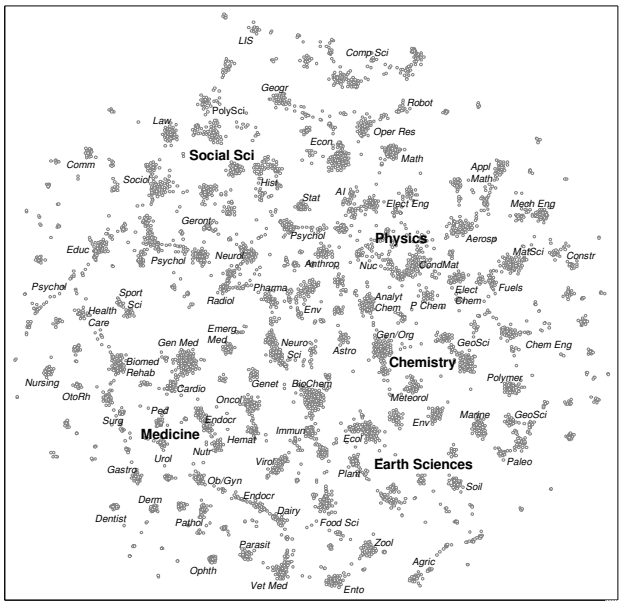


Boyack, Kevin W., Klavans, R. and Börner, Katy. (2005). Mapping the Backbone of Science. Scientometrics. 64(3), 351-374.

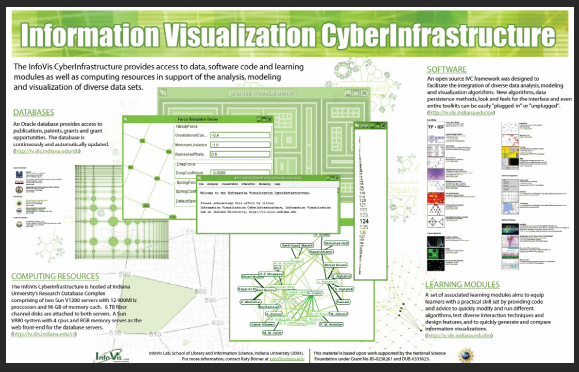
14

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



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

CIShell
cyberinfrastructure shell

Cyberinfrastructure Shell
http://cishell.org

NetworkWorkbench
A Workbench for Network Scientists

HPNAP
High Performance Network Applications Program

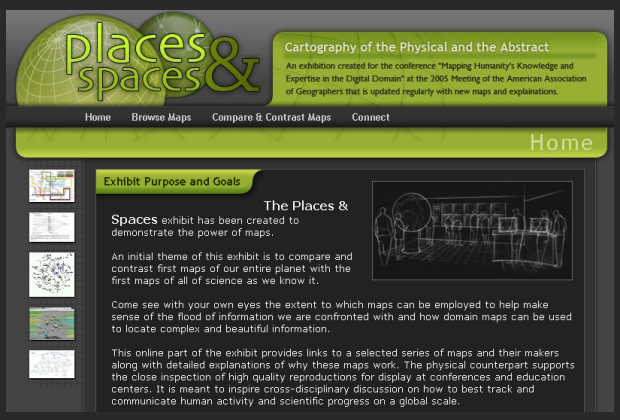
SEI: Network Workbench: A Large-Scale Network Analysis, Modeling and Visualization Toolkit for Biomedical, Social Science and Physics Research. NSF IIS-0513650 award (Katy Börner, Albert-Laszlo Barabasi, Santiago Schnell, Alessandro Vespignani & Stanley Wasserman, Eric Wernert (Senior Personnel), $1,120,926)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

EO

This Talk has Three Parts:

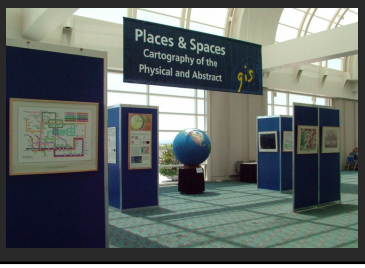
1. Mapping Science: Why?
2. Mapping Science: How?

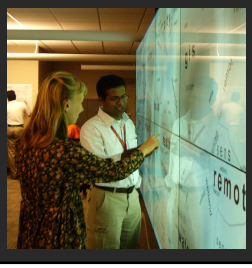
3. **Mapping Science: Applications**

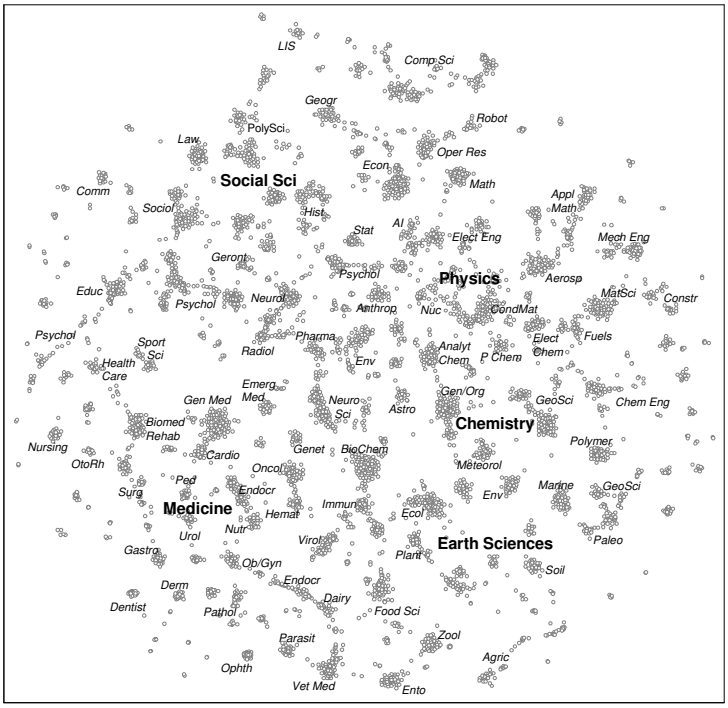


Places & Spaces: Mapping Science is a science exhibit that introduces people to maps of sciences, their makers and users.

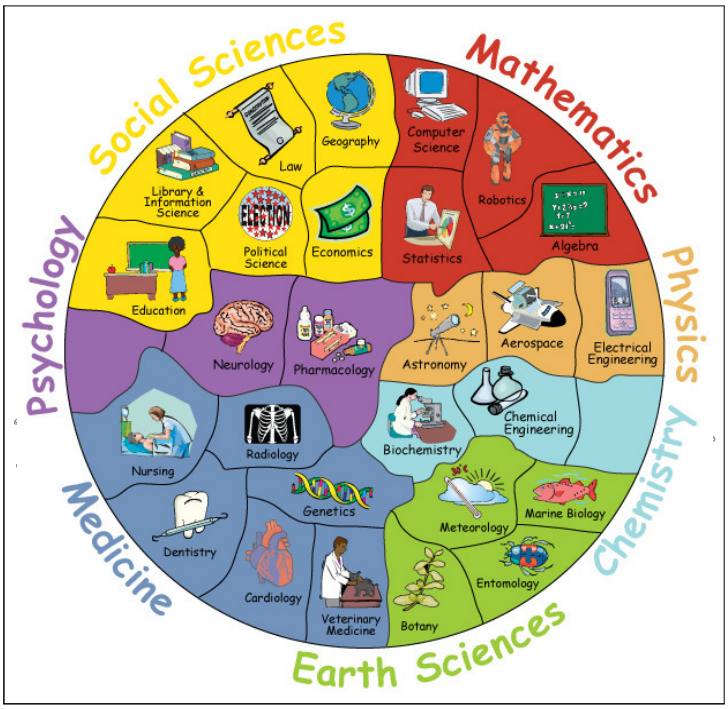
http://scimaps.org/exhibit



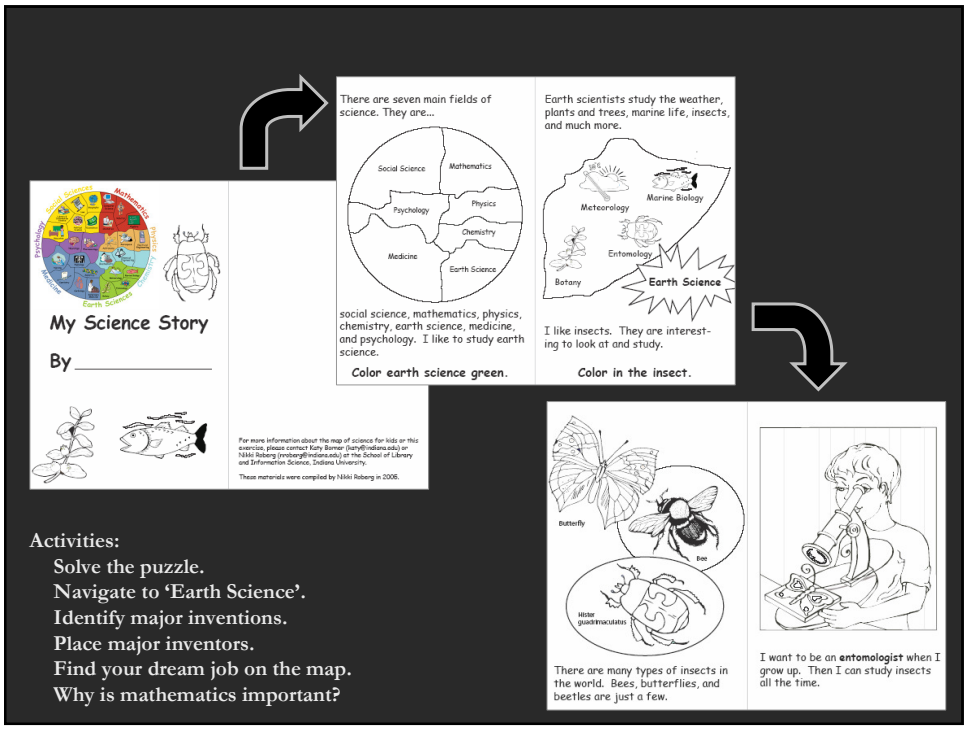



LIS
Comp Sci
Geogr
Robot
PolySci
Law
Oper Res
Econ
Social Sci
Math
Comm
Appl Math
Sociol
Hist
Stat
AI
Elect Eng
Mech Eng
Geront
Psychol
Physics
Aerosp
Educ
Psychol
Neurol
Anthrop
Nuc
CondMat
MatSci
Constr
Psychol
Sport Sci
Pharma
Fuels
Radiol
Analyt Chem
Elect Chem
Health Care
Env
P Chem
Emerg Med
Gen/Org
GeoSci
Chem Eng
Gen Med
Neuro Sci
Astro
Biomed
Chemistry
Rehab
Polymer
Nursing
Genet
BioChem
Cardio
Meteorol
OtoRh
Oncol
Marine
GeoSci
Ped
Endocr
Env
Surg
Medicine
Immun
Ecol
Hemat
Nutr
Paleo
Urol
Virol
Earth Sciences
Gastro
Plant
Ob/Gyn
Soil
Endocr
Derm
Dairy
Dentist
Pathol
Food Sci
Parasit
Zool
Agric
Ophth
Vet Med
Ento

19


Social Sciences
Mathematics
Physics
Chemistry
Earth Sciences
Medicine
Psychology
Geography
Law
Library & Information Science
Political Science
Economics
Education
Computer Science
Robotics
Statistics
Algebra
Neurology
Pharmacology
Astronomy
Aerospace
Electrical Engineering
Chemical Engineering
Biochemistry
Radiology
Nursing
Genetics
Meteorology
Marine Biology
Dentistry
Entomology
Cardiology
Veterinary Medicine
Botany

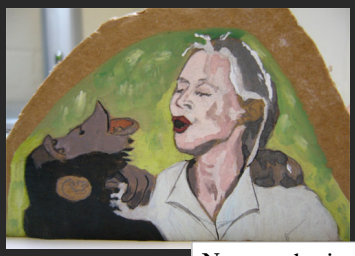
20


My Science Story
By _____
There are seven main fields of science. They are...
Social Science
Mathematics
Psychology
Physics
Chemistry
Medicine
Earth Science
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.
Meteorology
Marine Biology
Entomology
Botany
Earth Science
I like insects. They are interesting to look at and study.
Color in the insect.
Butterfly
Bee
Hister quadrimaculatus
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?

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

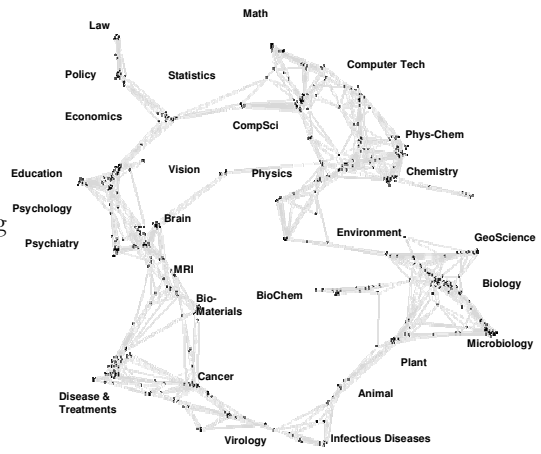


New puzzle pieces

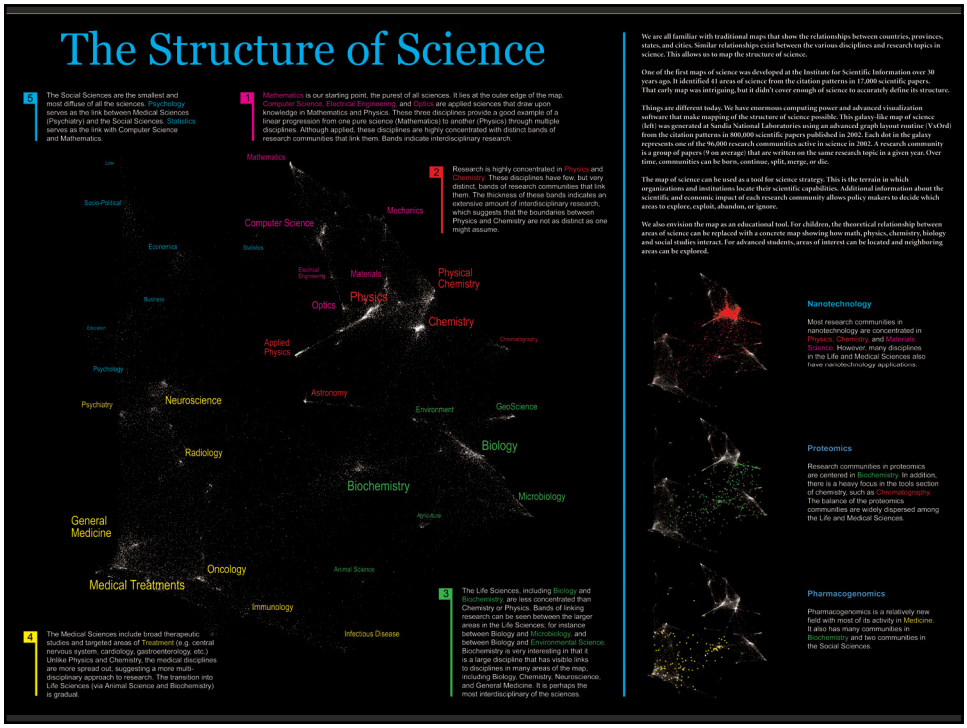
Latest 'Base Map' of Science

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



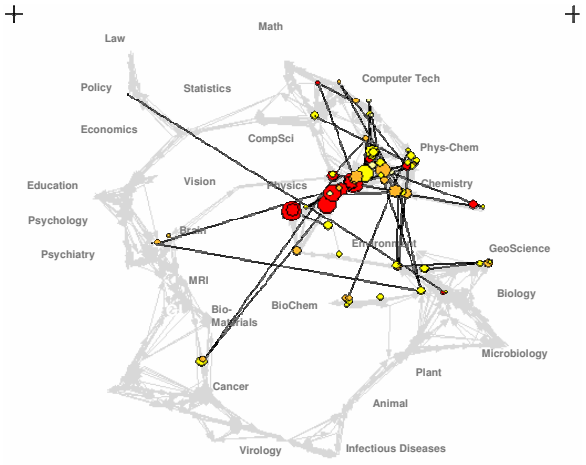
24



Science map applications: Identifying core competency

Kevin W. Boyack & Richard Klavans, unpublished work.

Funding patterns of the US Department of Energy (DOE)

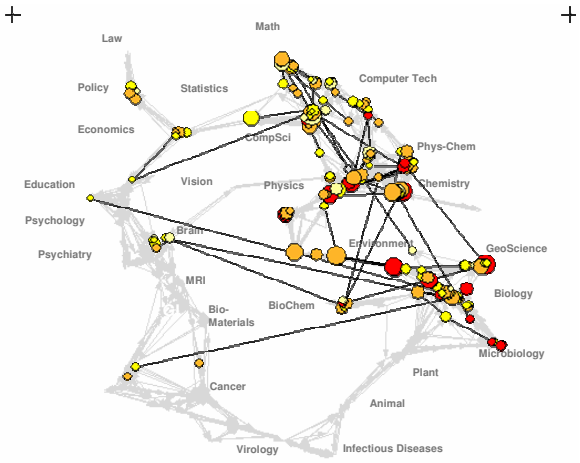


26

Science map applications: Identifying core competency

Kevin W. Boyack & Richard Klavans, unpublished work.

Funding Patterns of the National Science Foundation (NSF)

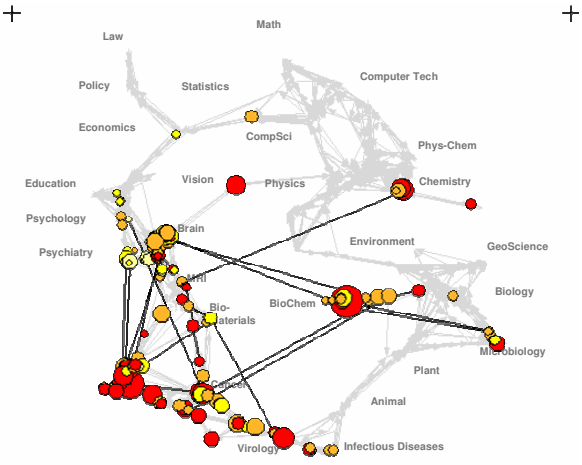


27

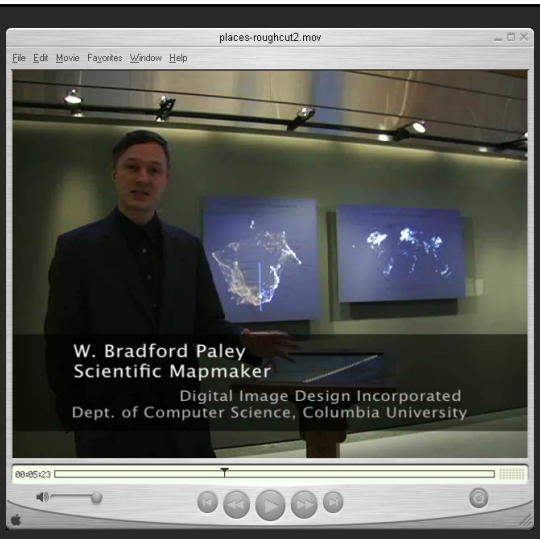
Science map applications: Identifying core competency

Kevin W. Boyack & Richard Klavans, unpublished work.

Funding Patterns of the National Institutes of Health (NIH)

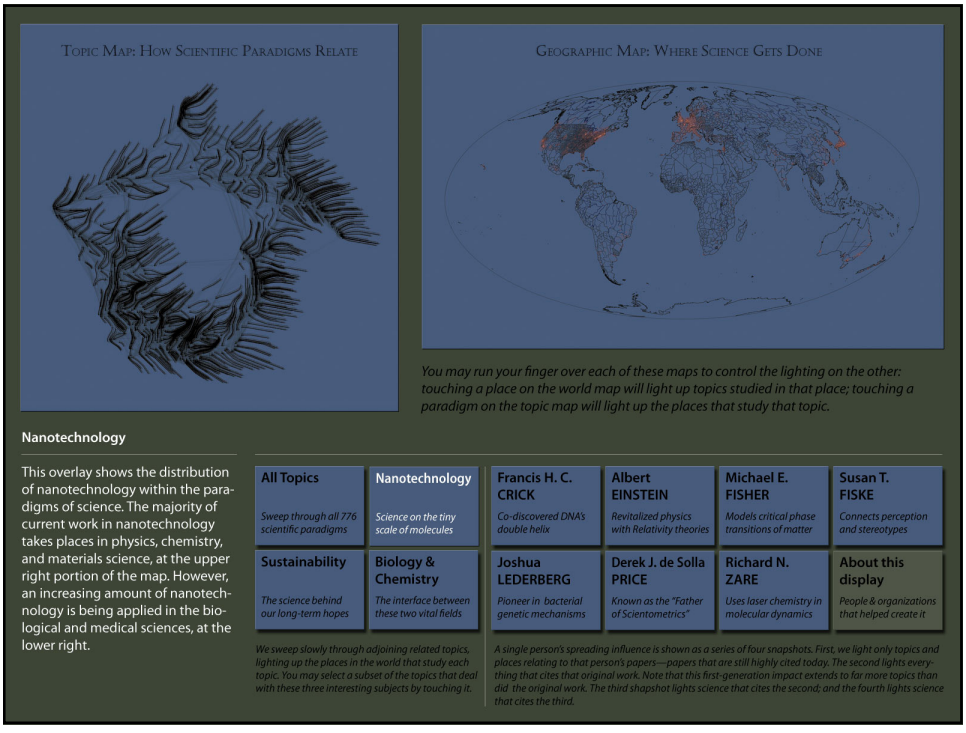


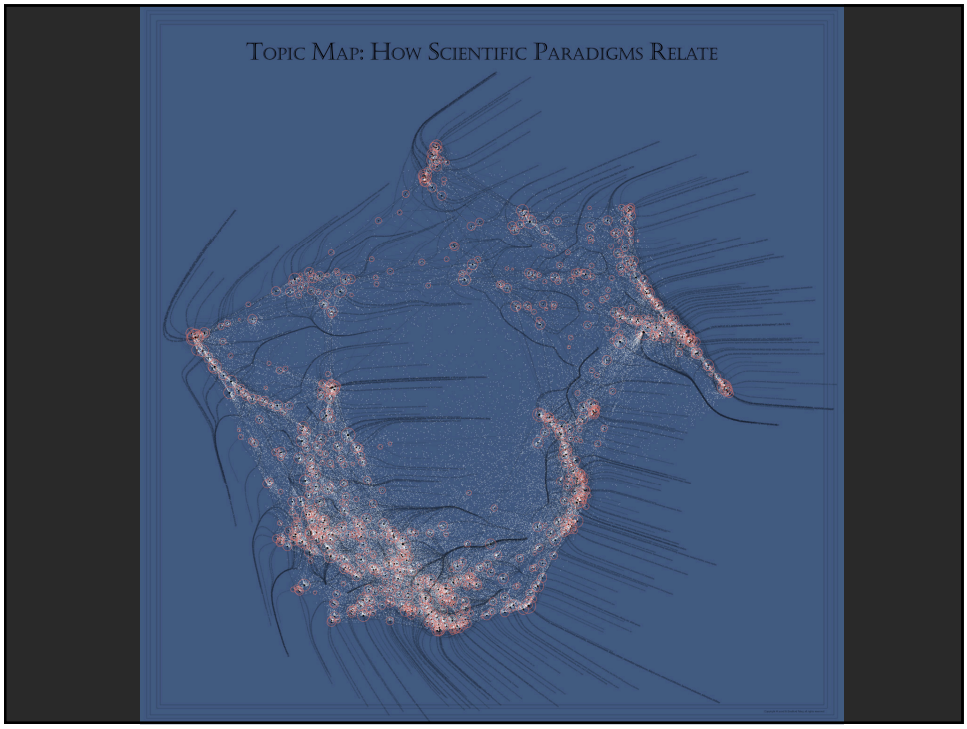
28

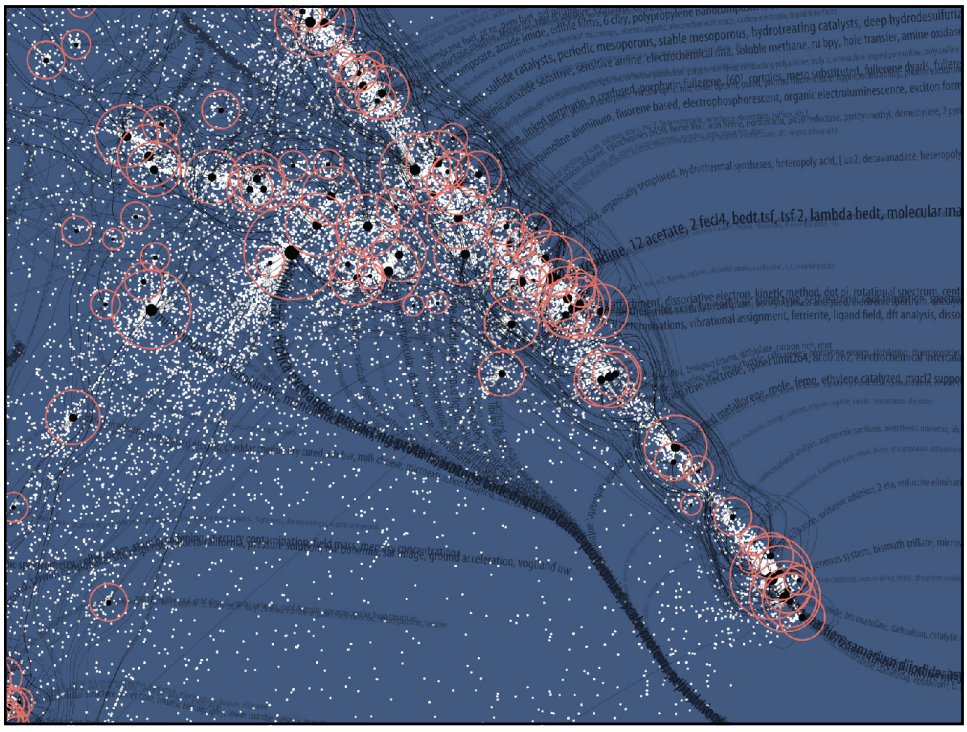


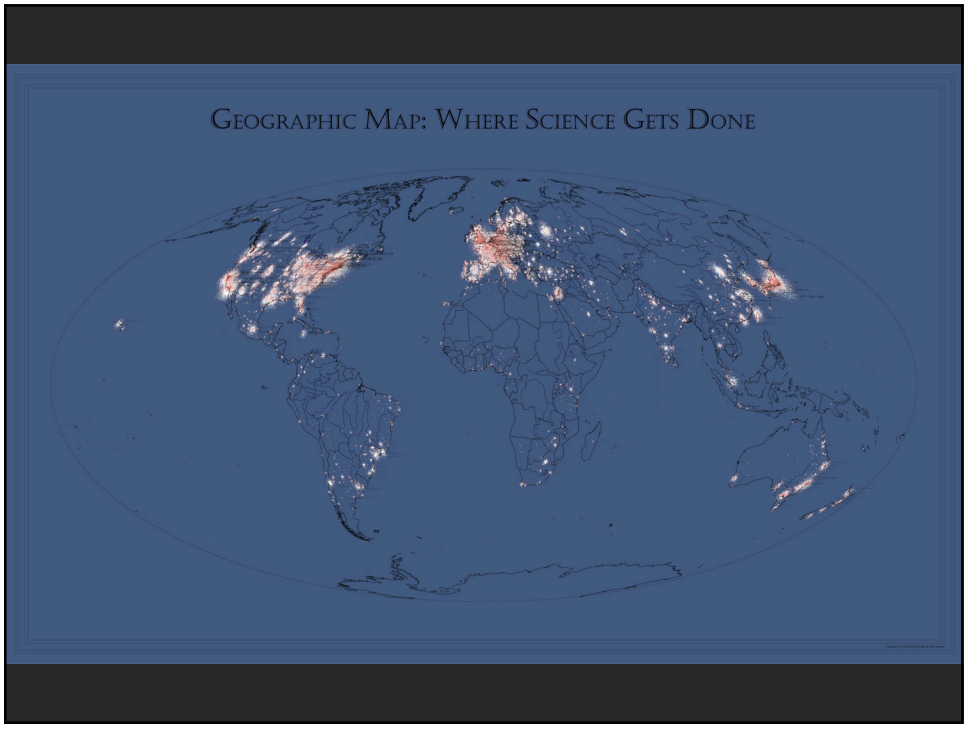
"Places & Spaces: Mapping Science" Illuminated Diagram Display

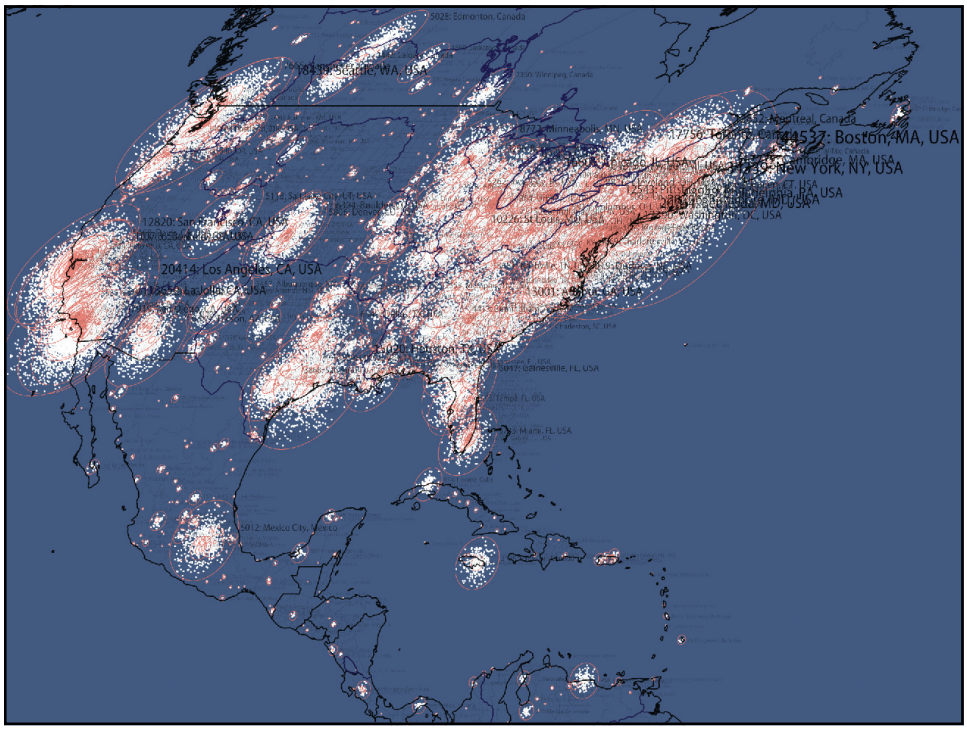
Was on display at the NYPL Science, Industry, and Business Library
Madison/34th, New York City, April 3rd - August 31st, 2006.

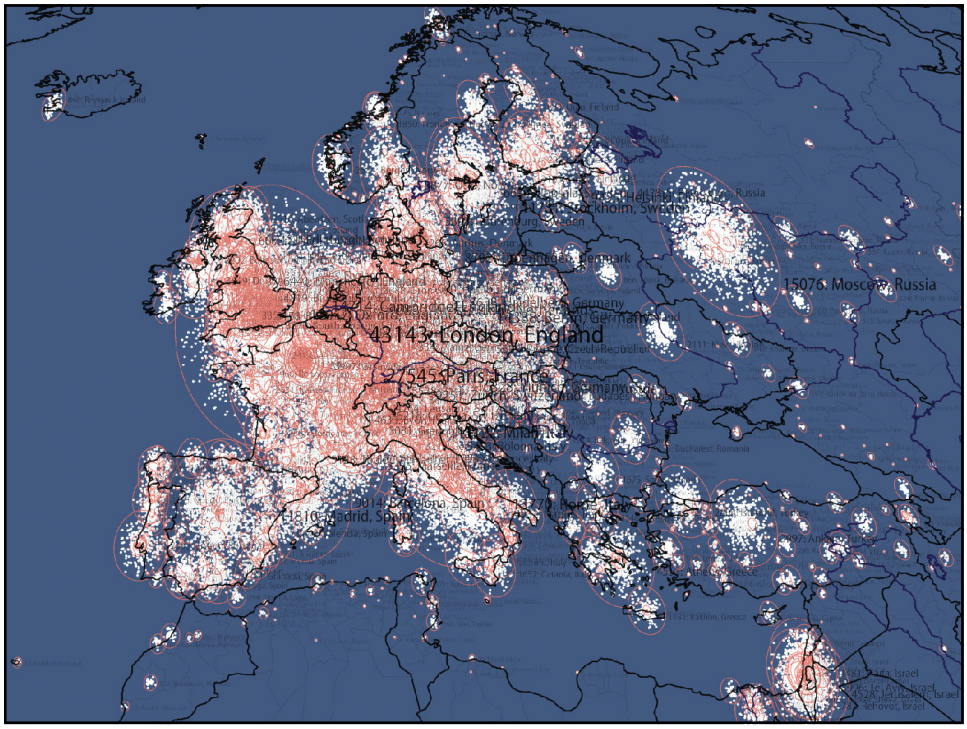


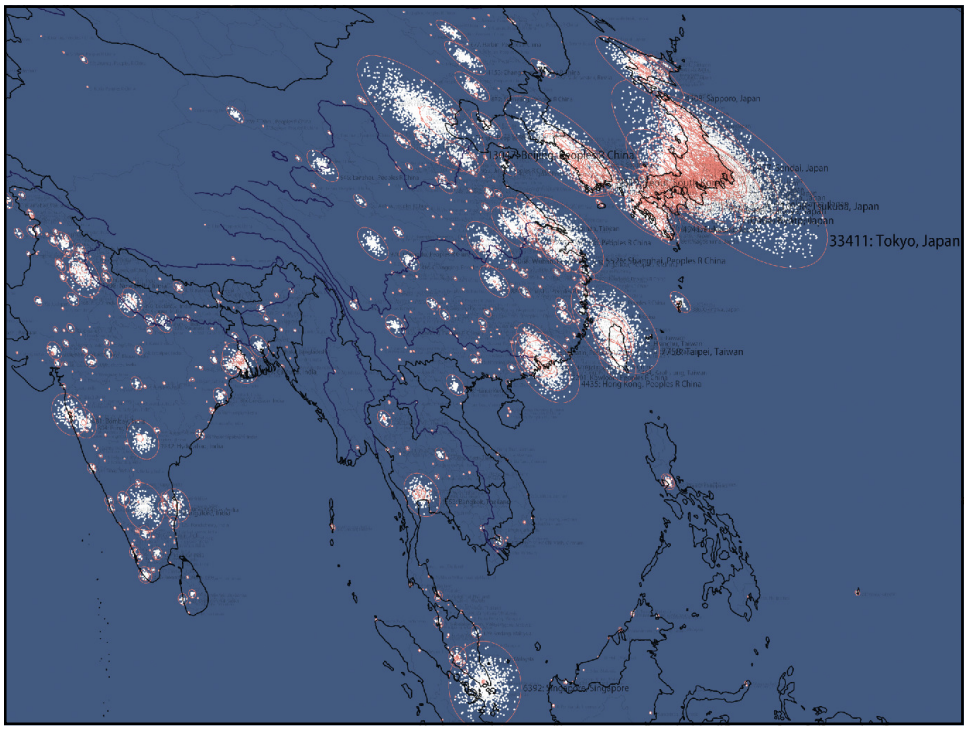

Topic Map: How Scientific Paradigms Relate




Geographic Map: Where Science Gets Done

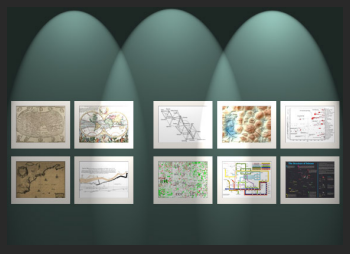

44537: Boston, MA, USA
New York, NY, USA
20414-Los Angeles, CA, USA


15076: Moscow, Russia
43143: London, England
27545: Paris, France


33411: Tokyo, Japan

The Power of Maps

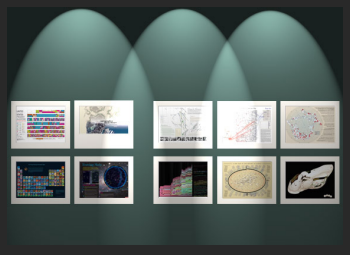
Four Early Maps of Our World
VERSUS
Six Early Maps of Science



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

The Power of Reference Systems

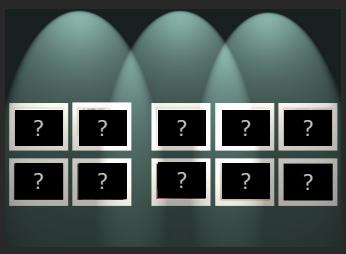
Four Existing Reference Systems
VERSUS
Six Potential Reference Systems of Science



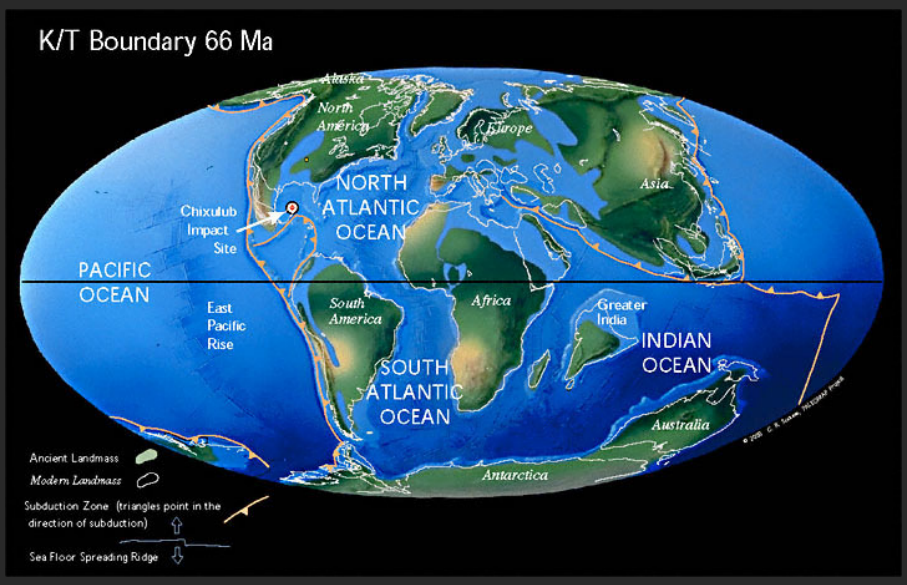
(2nd Iteration of Places & Spaces Exhibit - 2006)

The Power of Forecasts

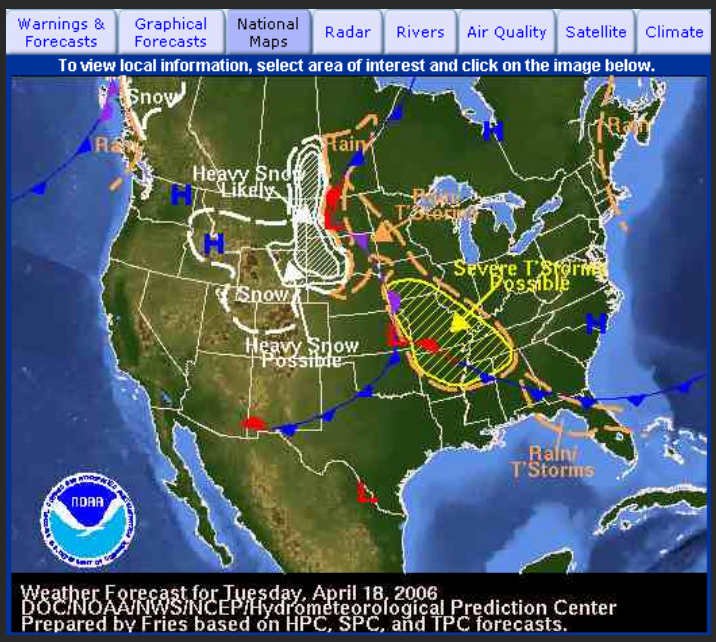
Four Existing Forecasts
VERSUS
Six Potential Science 'Weather' Forecasts



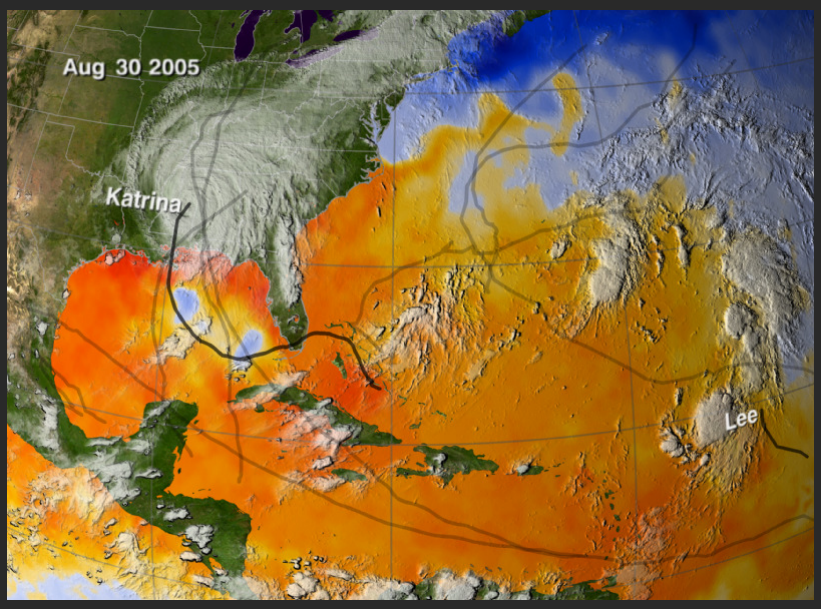
(3rd Iteration of Places & Spaces Exhibit - 2007)



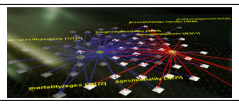
http://www.scotese.com/



http://www.weather.gov



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

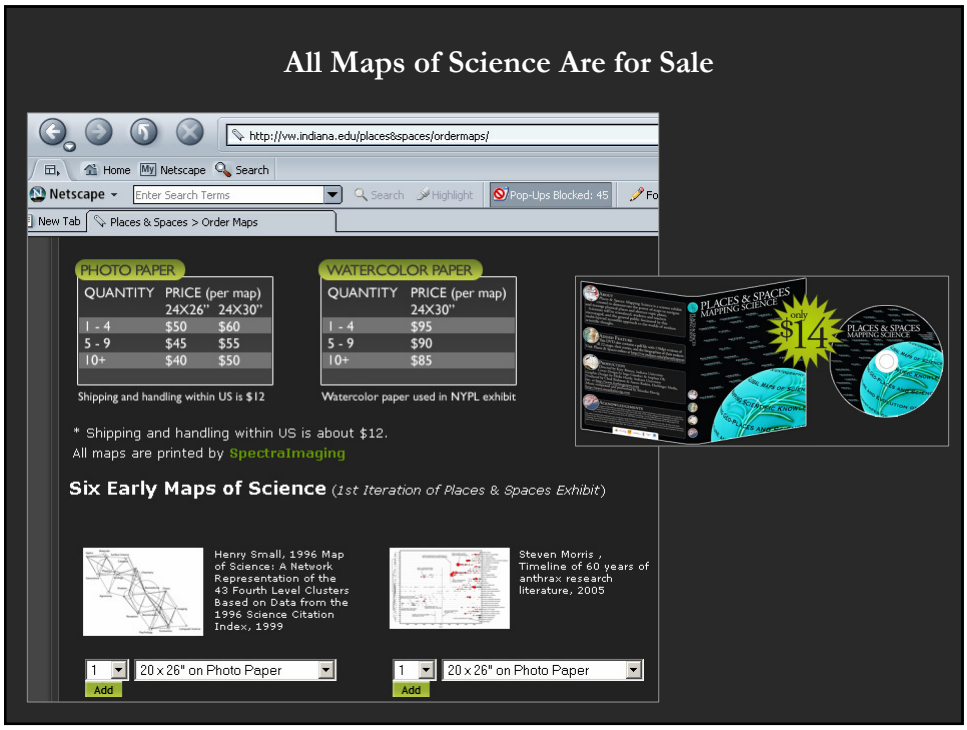


References

- 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, Penumarthy, 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.
- 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.
- Ord, Terry J., Martins, Emília 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.
- Börner, Katy, Maru, Jeegar 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.

44


Interested to get your own science map?
Contact the map makers!
katy@indiana.edu


All Maps of Science Are for Sale
http://vw.indiana.edu/places&spaces/ordermaps/
Places & Spaces > Order Maps
PHOTO PAPER
QUANTITY PRICE (per map) 24X26" 24X30"
1 - 4 $50 $60
5 - 9 $45 $55
10+ $40 $50
Shipping and handling within US is $12
WATERCOLOR PAPER
QUANTITY PRICE (per map) 24X30"
1 - 4 $95
5 - 9 $90
10+ $85
Watercolor paper used in NYPL exhibit
only $14
* Shipping and handling within US is about $12.
All maps are printed by SpectraImaging
Six Early Maps of Science (1st Iteration of Places & Spaces Exhibit)
Henry Small, 1996 Map of Science: 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
Add