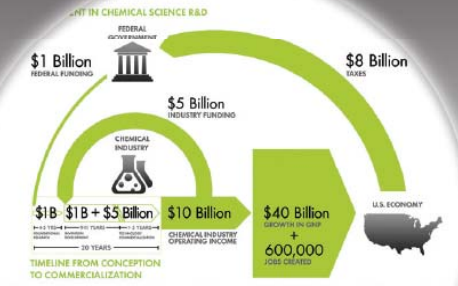
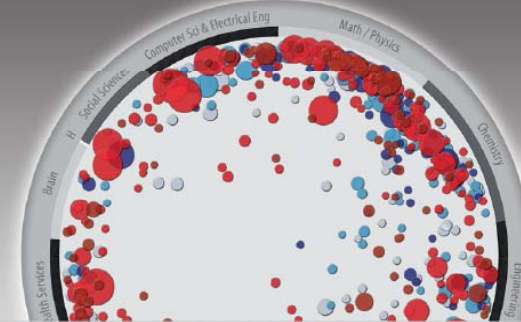


Mapping Innovation for Congress

CREATING AN **Innovation Dashboard** TO INFORM POLICY MAKERS

A Seminar and Interactive Session | Thursday, December 10, 2009 | Rayburn HOB 2103 | 3:00 p.m. - 4:30 p.m.



U.S. VULNERABILITIES IN SCIENCE

National Strengths are more accurately assessed using a reference-based classification system.

Chemical Research & Development Powers the U.S. Innovation Engine

Macroeconomic Implications of Public and Private R&D Investments in Chemical Science

Have you ever wondered what an "Innovation Dashboard" of your state or congressional district would reveal? Where would the "Pockets of Innovation" lie and what pathways do those innovations take to impact education, health, or profit? Where are the jobs or the innovative education programs? How would you use such a dashboard in your daily work?

This seminar series brings together R&D researchers with a strong background in data analysis and visualization and selected Members of Congress and Congressional Staff. The main goal is to collectively identify features and functionality of an "Innovation Dashboard" for Congress.

AGENDA

- 3:00 p.m. Welcome & Introductions
- 3:10 p.m. Overview of R&D Analyses and Maps
- 3:45 p.m. Open Discussion Forum
- 4:30 p.m. Adjourn

Take ASTRA's Mapping Innovation for Congress Survey—please visit <https://www.surveymonkey.com/s/CX7PLB>

The first session in this series features Dr. Katy Börner of Indiana University. Dr. Börner will describe different approaches to R&D data analysis and visualization, discuss sample maps from the international Mapping Science exhibit (<http://scimaps.org>) and insights derived from them, as well as near-term developments in the mapping of R&D data. This presentation is followed by a structured but open discussion on desirable features and functionality for "State-Oriented R&D Fact Sheets" and "Innovation Dashboards" and how they would be used to understand, improve, and communicate State and Congressional District innovation metrics.

Please RSVP to the rhoege@comcast.net by December 7, 2009 and confirm your attendance as seating is limited.

Acknowledgments: This work is supported in part by ASTRA, the CyberInfrastructure for Network Science Center and the School of Library and Information Science at Indiana University, the National Science Foundation under Grant Nos. 0810730, 0810731, and 0810732, and the James S. McDonnell Foundation.



ASTRA's Mapping Innovation for Congress Baseline Survey 121009

Would it be useful to have innovation information that compares different States?

		Response Percent
Yes	<input type="checkbox"/>	89.5%
No	<input type="checkbox"/>	5.3%
Not Sure	<input type="checkbox"/>	5.3%

ASTRA's Mapping Innovation for Congress Baseline Survey 121009

Do you (and Congress generally) have adequate data to make decisions on innovation policy?

		Response Percent
Most of the Time	<input type="checkbox"/>	5.0%
Some of the Time	<input type="checkbox"/>	25.0%
Usually Not	<input type="checkbox"/>	55.0%
Not Sure	<input type="checkbox"/>	15.0%

ASTRA's Mapping Innovation for Congress Baseline Survey 121009

What types of data would you like to have more readily available to you? (Please choose all that apply)

		Response Percent
R&D Funding		76.2%
Agency Budget Data		52.4%
Academic Degree Production Data		42.9%
STEM Education Data		52.4%
General Economic Statistics that Affect		
Cybersecurity		14.3%
State and Congressional Migration, Visas, Foreign Talent in U.S.		23.8%
Internet & Broadband		23.8%
International Venture Capital		38.1%
Comparisons between		
Income and Unemployment Information		52.4%
New Business Start-ups		52.4%
Job Creation		47.6%

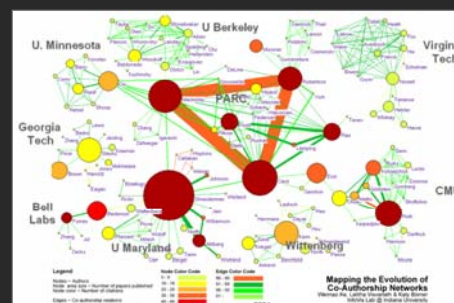
Science and Technology (S&T) Studies

Can be conducted at different levels:

- local (individual),
- meso (local, e.g., one institute, one funding agency), or
- global level (all of science or world wide).

Using

- Statistical Analysis/Profiling
- Temporal Analysis (When)
- Geospatial Analysis (Where)
- Topical Analysis (What)
- Network Analysis (With Whom?)
- Modeling (Why)





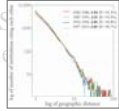
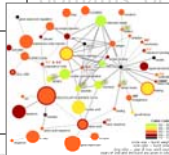



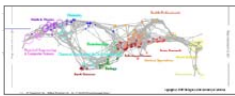



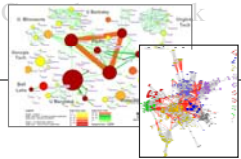

Type of Analysis vs. Level of Analysis

	<i>Micro/Individual</i> (1-100 records)	<i>Meso/Local</i> (101–10,000 records)	<i>Macro/Global</i> (10,000 < records)
<i>Statistical Analysis/Profiling</i>	Individual person and their expertise profiles	Larger labs, centers, universities, research domains, or states	All of NSF, all of USA, all of science.
<i>Temporal Analysis (When)</i>	Funding portfolio of one individual	Mapping topic bursts in 20-years of PNAS	113 Years of Physics Research
<i>Geospatial Analysis (Where)</i>	Career trajectory of one individual	Mapping a states intellectual landscape	PNAS Publications
<i>Topical Analysis (What)</i>	Base knowledge from which one grant draws.	Knowledge flows in Chemistry research	VxOrd/Topic maps of NIH funding
<i>Network Analysis (With Whom?)</i>	NSF Co-PI network of one individual	Co-author network	NSF's core competency

5



Type of Analysis vs. Level of Analysis

	<i>Micro/Individual</i> (1-100 records)	<i>Meso/Local</i> (101–10,000 records)	<i>Macro/Global</i> (10,000 < records)
<i>Statistical Analysis/Profiling</i>	Individual person and their expertise profiles	Larger labs, centers, universities, research domains, or states	All of NSF, all of USA, all of science. 
<i>Temporal Analysis (When)</i>	Funding portfolio of one individual	Mapping topic bursts in 20-years of PNAS 	113 Years of Physics Research 
<i>Geospatial Analysis (Where)</i>	Career trajectory of one individual	Mapping a states intellectual landscape 	PNAS Publications 
<i>Topical Analysis (What)</i>	Base knowledge from which one grant draws. 	Knowledge flows in Chemistry research 	VxOrd/Topic maps of NIH funding 
<i>Network Analysis (With Whom?)</i>	NSF Co-PI network of one individual 	Co-author network 	NSF's core competency 

6

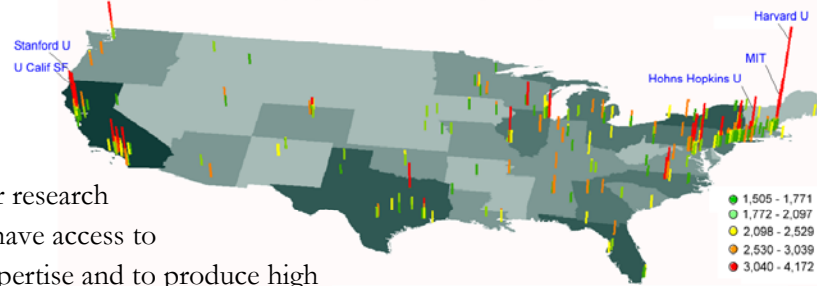
Spatio-Temporal Information Production and Consumption of Major U.S. Research Institutions

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*



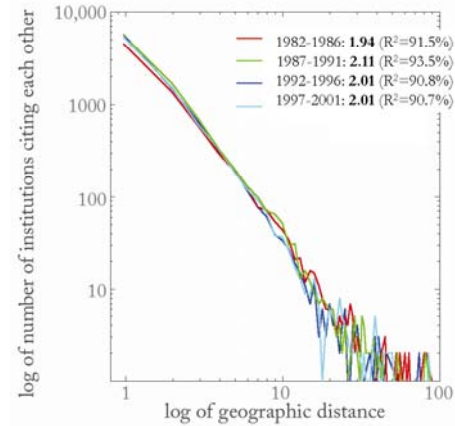
Research questions:

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?



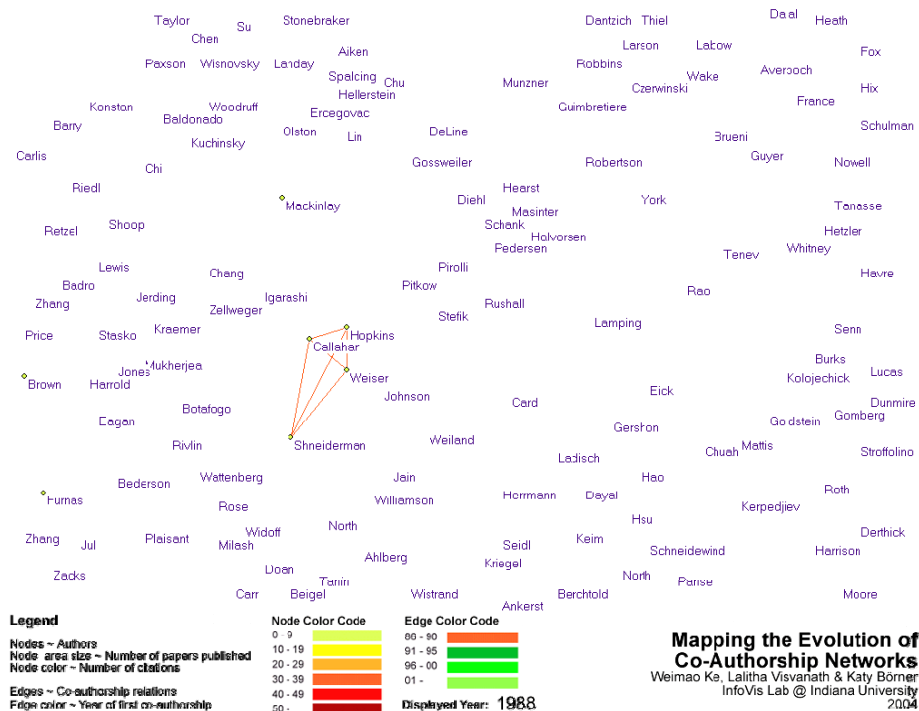
Contributions:

- 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 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
Weimao Ke, Lalitha Visvanath & Katy Börner
InfoVis Lab @ Indiana University
2004

Interactive World and Science Map of S&T Jobs

Angela Zoss, Michael Conover, Katy Börner (in preparation).

Visualization of Job Postings

Map of Science | Geographic

Map of Science | Geographic

Map of Science | Geographic

Postdoc at Harvard Medical School
[Link to Post](#)

Map of Science

Scientific domains are highly interconnected. The boundaries between different domains are often fuzzy. One way of thinking about the relationships between domains is to conceptualize all scientific domains as existing within a large **network of research**.

Creating a network of scientific research can be accomplished by looking at scientific journals and their articles. The UCSD Map of Science used here is the product of a large study by researchers at the University of California San Diego using 7.2 million papers and over 16,000 separate journals, proceedings, and series from Thomson Scientific and Scopus over the five year period from 2001 to 2005. The researchers used citations between the papers and journals to **cluster journals** into small groups of highly related journals.

Those clusters are represented by 554 individual nodes in the network. The links between the clusters show that some clusters are related to other clusters but are not as tightly connected as the journals that make up each cluster. Then the clusters are labeled both by the content area shared by the journals in the cluster and by the overarching scientific domain for that cluster (represented by one of 13 colors).

Maps of science like this one can be used to understand many different data sets and how they can be represented by topic. Here we are looking at the topics that appear in job postings from large in-

Search for Jobs

Search for Jobs

Search for Jobs

Search

Search

Search

Biotechnology

Copyright © 2008 The Regents of the University of California - [Terms of Use](#)

13

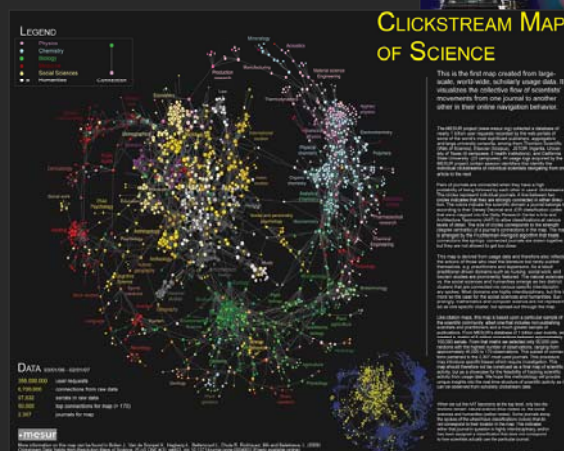
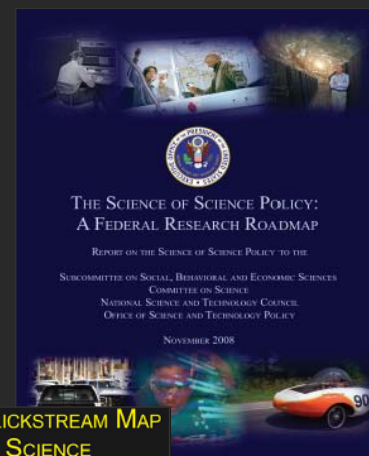
Science and Technology (S&T) Studies

Results are communicated using

- Top-N lists
- Profiles
- Graphs and charts
- Geographic map overlays
- Science maps

By means of

- Alerts
- White papers
- Reports
- Science forecasts

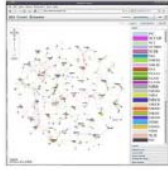


A Topic Map of NIH Grants 2007

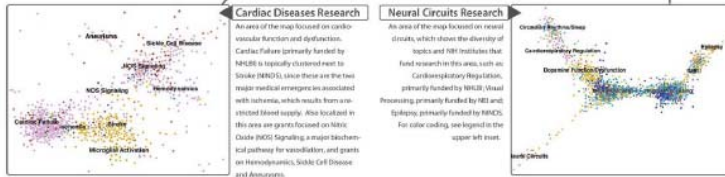
ChalkLabs UCI UCIRVINE

Bruce W. Herr II (Chalklabs & IU), Gully Burns (IS), David Newman (UCI), Edmund Talley (NIH)

The National Institutes of Health (NIH) is organized as a multitude of Institutes and Centers whose missions are primarily focused on distinct diseases. However, disease etiologies and therapies blur scientific boundaries, and thus there is tremendous overlap in the kinds of research funded by each institute. This creates a daunting landscape for decisions on research directions, funding allocations, and policy formulations. Shown here is devised an interactive topic map for navigating this landscape online at www.nih.gov/td.

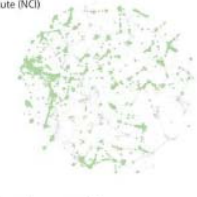


Topic modeling, a statistical technique that automatically learns semantic categories, was applied to mission projects in terms used by researchers to describe their work, without the biases of keywords or subject headings. Grant similarities were derived from their topic mixtures, and grants were then clustered on a two-dimensional map using a force-directed simulated annealing algorithm. This analysis creates an interactive environment for assessing grant relevance to research categories and to NIH Institutes in which grants are localized.



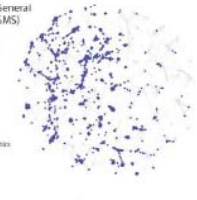
National Cancer Institute (NCI)

- TOP 10 TOPICS
- 1 Oncology/Clinical Trials
 - 2 Cancer Treatment
 - 3 Cancer Therapy
 - 4 Carcinogenesis
 - 5 Risk Factor Analysis
 - 6 Cancer Chemotherapy
 - 7 Metastasis
 - 8 Leukemia
 - 9 Prevention/Prognosis
 - 10 Cancer Chemoprevention



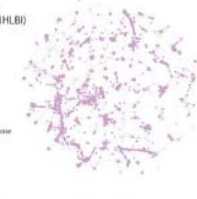
National Institute of General Medical Sciences (NIGMS)

- TOP 10 TOPICS
- 1 Bioactive Organic Synthesis
 - 2 X-ray Crystallography
 - 3 Protein-NMR
 - 4 Computational Models
 - 5 Yeast Biology
 - 6 Metabolism
 - 7 Enzymatic Mechanisms
 - 8 Protein Complexes
 - 9 Invertebrate/Zebrafish Genetics
 - 10 Cell Division



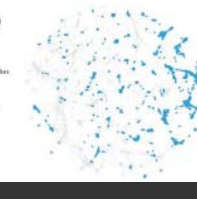
National Heart, Lung, and Blood Institute (NHLBI)

- TOP 10 TOPICS
- 1 Cardiac Failure
 - 2 Pulmonary Hypertension
 - 3 Genetic Linkage Analysis
 - 4 Cardiovascular Disease
 - 5 Atherosclerosis
 - 6 Hemostasis
 - 7 Blood Pressure
 - 8 Asthma/Allergic Airway Disease
 - 9 Gene Association
 - 10 Lipoproteins



National Institute of Mental Health (NIMH)

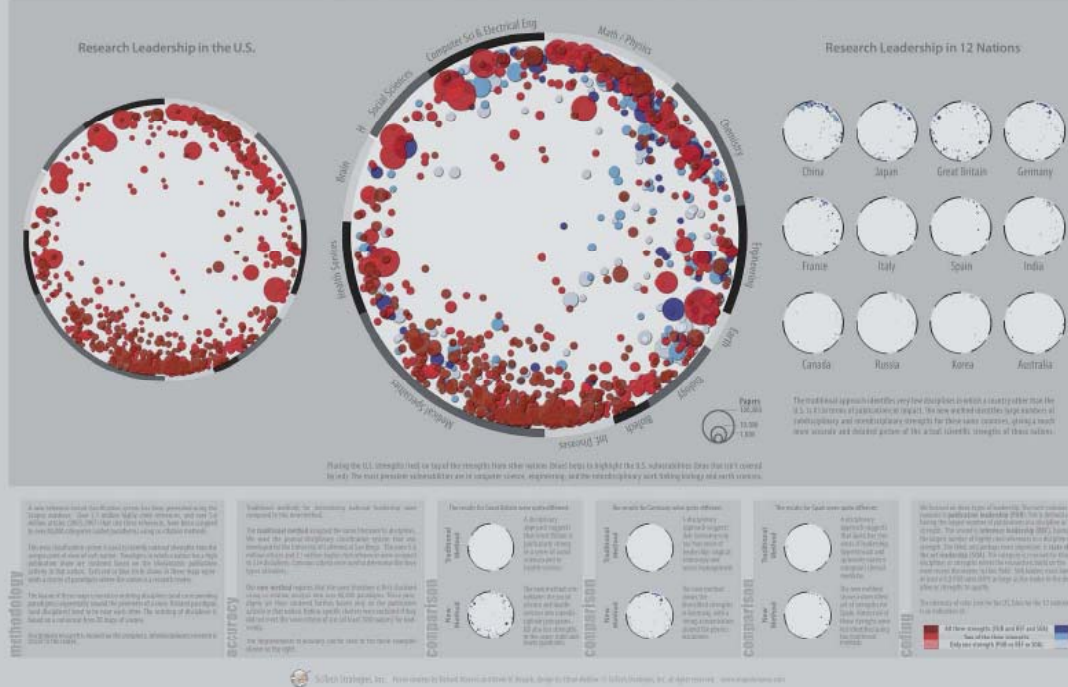
- TOP 10 TOPICS
- 1 Mood Disorders
 - 2 Schizophrenia
 - 3 Behavioral/Neuroscience Studies
 - 4 Mental Health
 - 5 Depression
 - 6 Cognitive Behavior Therapy
 - 7 ADD/Attention
 - 8 Genetic Linkage Analysis
 - 9 Adolescence
 - 10 Childhood



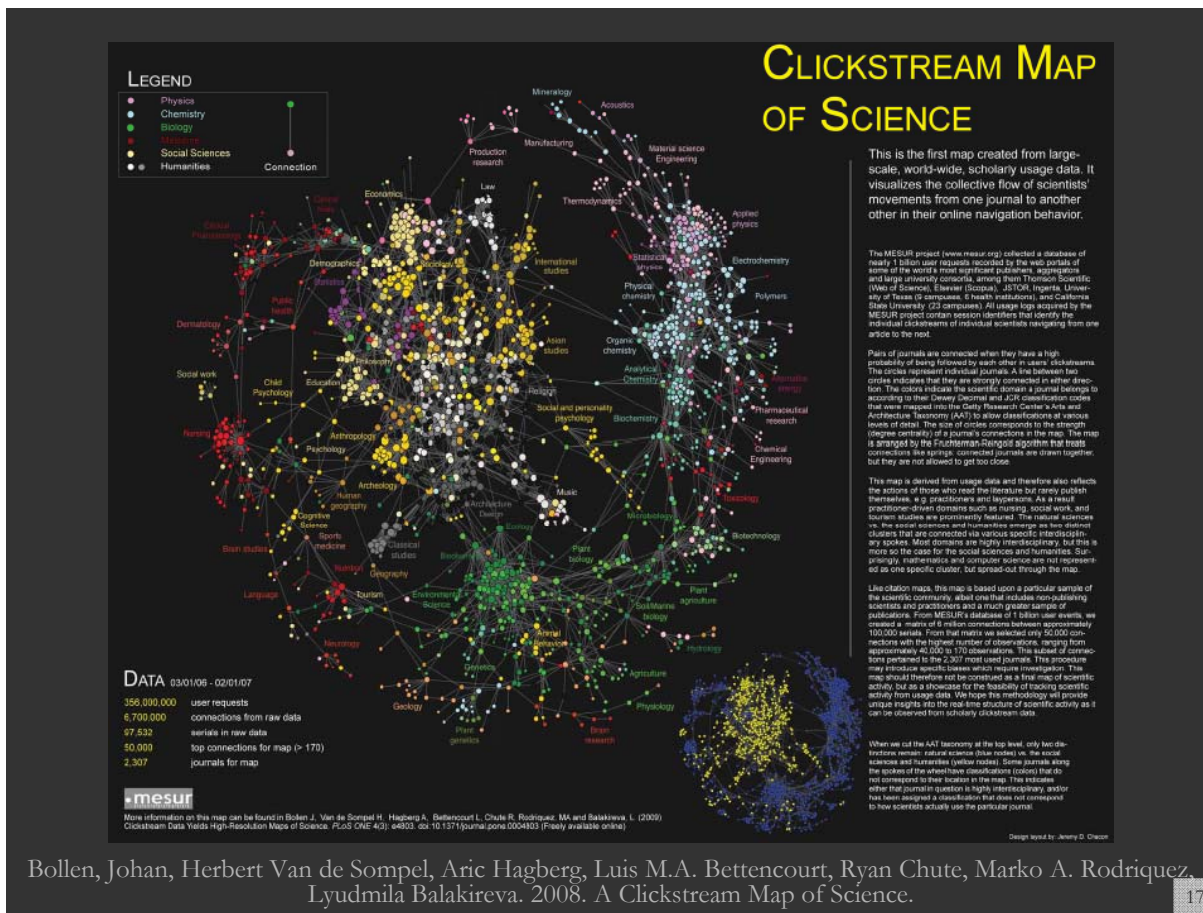
Herr II, Bruce W., Gully Burns, David Newman, Edmund Talley. 2007. A Topic Map of NIH Grants 2007. Bloomington, IN.

U.S. VULNERABILITIES IN SCIENCE

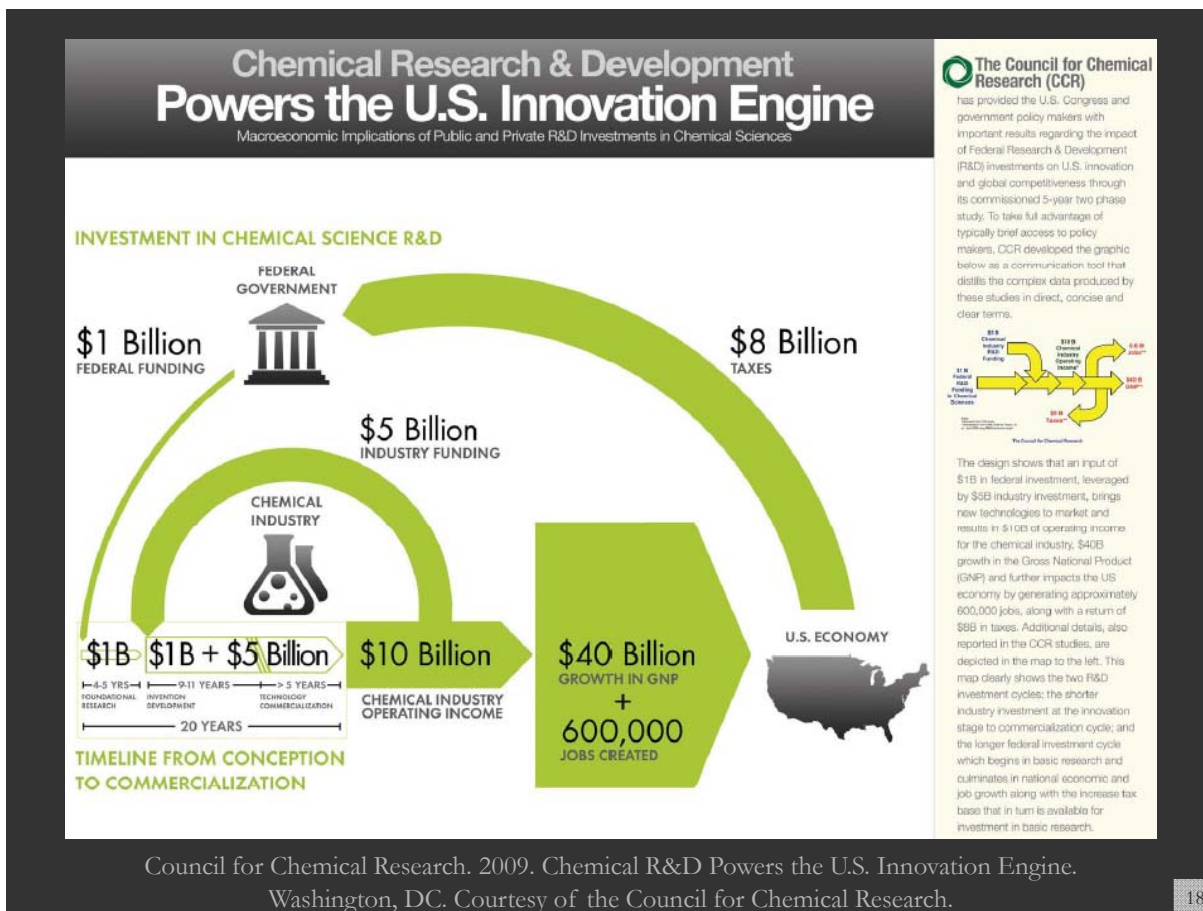
National Strengths are more accurately assessed using a reference-based classification system.



U.S. Vulnerabilities in Science - Kevin W. Boyack, Richard Klavans - 2009



Bollen, Johan, Herbert Van de Sompel, Aric Hagberg, Luis M.A. Bettencourt, Ryan Chute, Marko A. Rodriguez, Lyudmila Balakireva. 2008. A Clickstream Map of Science. 17



Council for Chemical Research. 2009. Chemical R&D Powers the U.S. Innovation Engine. Washington, DC. Courtesy of the Council for Chemical Research. 18

Illuminated Diagram Display

W. Bradford Paley, Kevin W. Boyack, Richard Kalvans, and Katy Börner (2007)
Mapping, Illuminating, and Interacting with Science. SIGGRAPH 2007.



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

Questions:

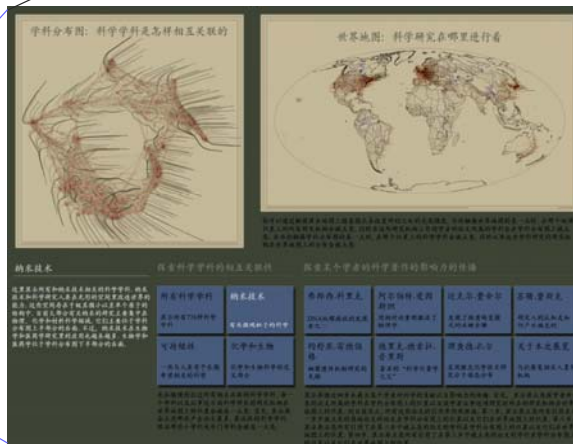
- Who is doing research on what topic and where?
- What is the 'footprint' of interdisciplinary research fields?
- What impact have scientists?



Interactive touch panel.

Contributions:

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



TOPIC MAP: HOW SCIENTIFIC PARADIGMS RELATE

GEOGRAPHIC MAP: WHERE SCIENCE GETS DONE

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

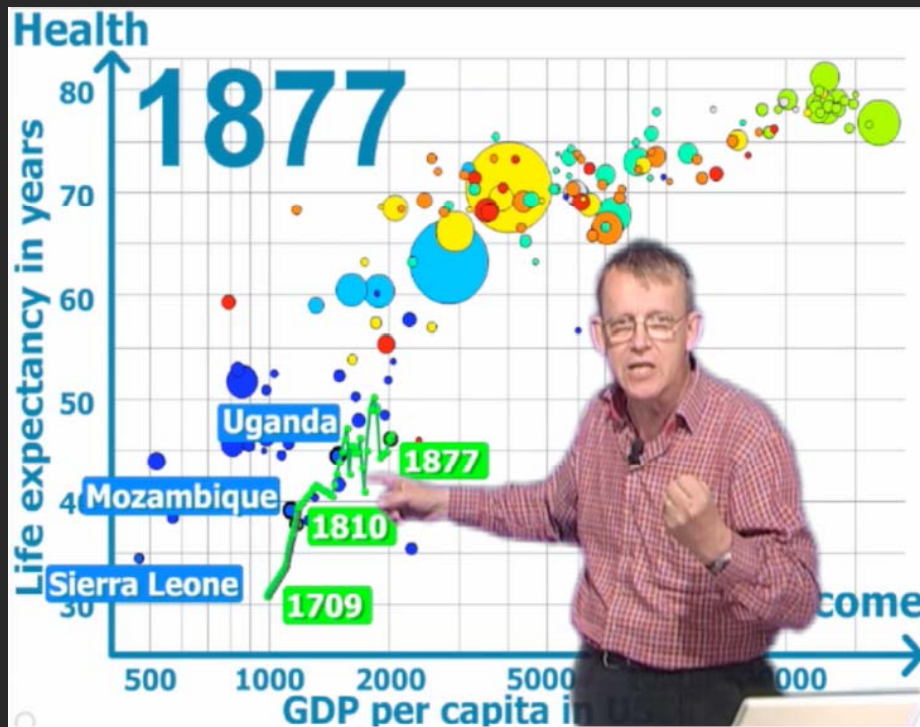
Nanotechnology

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

<p>All Topics</p> <p>Sweep through all 376 scientific paradigms</p>	<p>Nanotechnology</p> <p>Science on the tiny scale of molecules</p>	<p>Francis H. C. CRICK</p> <p>Co-discovered DNA's double helix</p>	<p>Albert EINSTEIN</p> <p>Revitalized physics with Relativity theories</p>	<p>Michael E. FISHER</p> <p>Models critical phase transitions of matter</p>	<p>Susan T. FISKE</p> <p>Connects perception and stereotypes</p>
<p>Sustainability</p> <p>The science behind our long-term hopes</p>	<p>Biology & Chemistry</p> <p>The interface between these two vital fields</p>	<p>Joshua LEDERBERG</p> <p>Pioneer in bacterial genetic mechanisms</p>	<p>Derek J. de Solla PRICE</p> <p>Known as the "Father of Scientometrics"</p>	<p>Richard N. ZARE</p> <p>Uses laser chemistry in molecular dynamics</p>	<p>About this display</p> <p>People & organizations that helped create it</p>

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

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



http://www.ted.com/talks/bans_rosling_shows_the_best_stats_you_ve_ever_seen.html

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 "TSL: 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 Scharnhorst, 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>