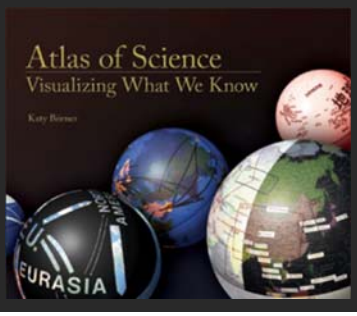
Identifying Overlap, Gaps and Emerging Research Areas Using Multi-Level Science Maps

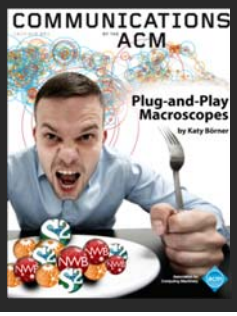
Katy Börner
Cyberinfrastructure for Network Science Center, Director
Information Visualization Laboratory, Director
School of Library and Information Science
Indiana University, Bloomington, IN, USA and
Royal Netherlands Academy of Arts and Sciences (KNAW), NL
katy@indiana.edu

With special thanks to the members at the Cyberinfrastructure for Network Science Center; the Sci2, NWB, and EpiC teams; and the VIVO Collaboration

Portfolio Analysis Symposium
NIH Natcher Auditorium, Bethesda, MD

Tuesday July 24, 2012 • 8:45-9:30am

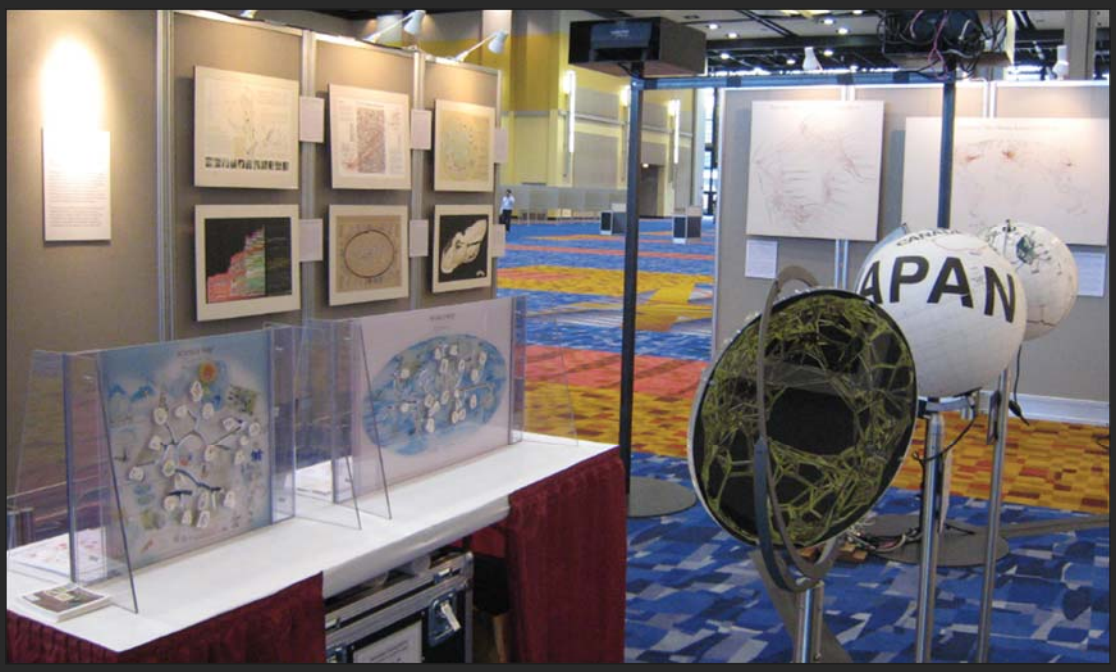




Science Maps

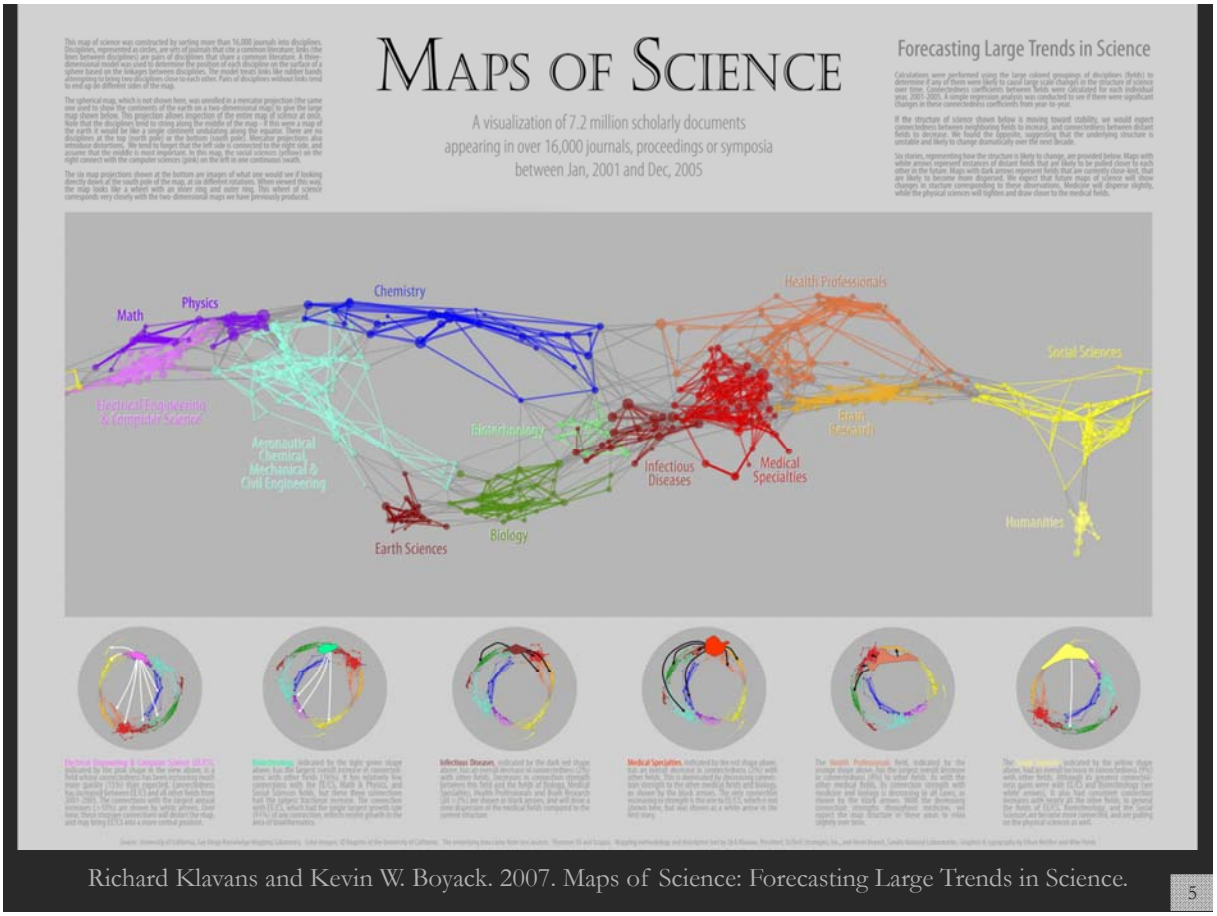
- Different Datasets
- Different Analyses and Visualizations
- Different User Groups

Places & Spaces: Mapping Science Exhibit (http://scimaps.org)



After eight years, there now exist 80 out of 100 maps.

Mapping Science Exhibit at MEDIA X, Wallenberg Hall, Stanford University, 2009
http://mediax.stanford.edu, http://scaleindependentthought.typepad.com/photos/scimaps



Richard Klavans and Kevin W. Boyack. 2007. Maps of Science: Forecasting Large Trends in Science.

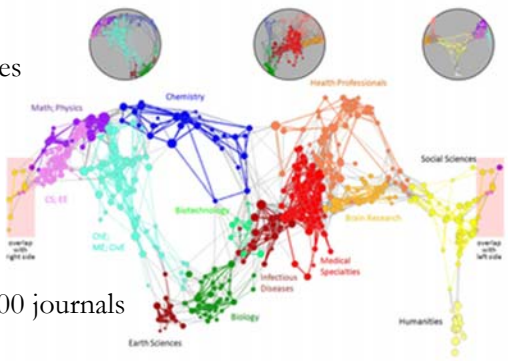
The UCSD Map of Science and Classification System

2007 Map:

Data: WoS and Scopus for 2001–2005, 7.2 million papers, >16,000 separate journals, proceedings, series

Similarity Metric: Combination of bibliographic coupling and keyword vectors

Number of Disciplines: 13; ***Subdisciplines:*** 554



2010 Map:

Data: WoS and Scopus for 2001–2010; about 25,000 journals

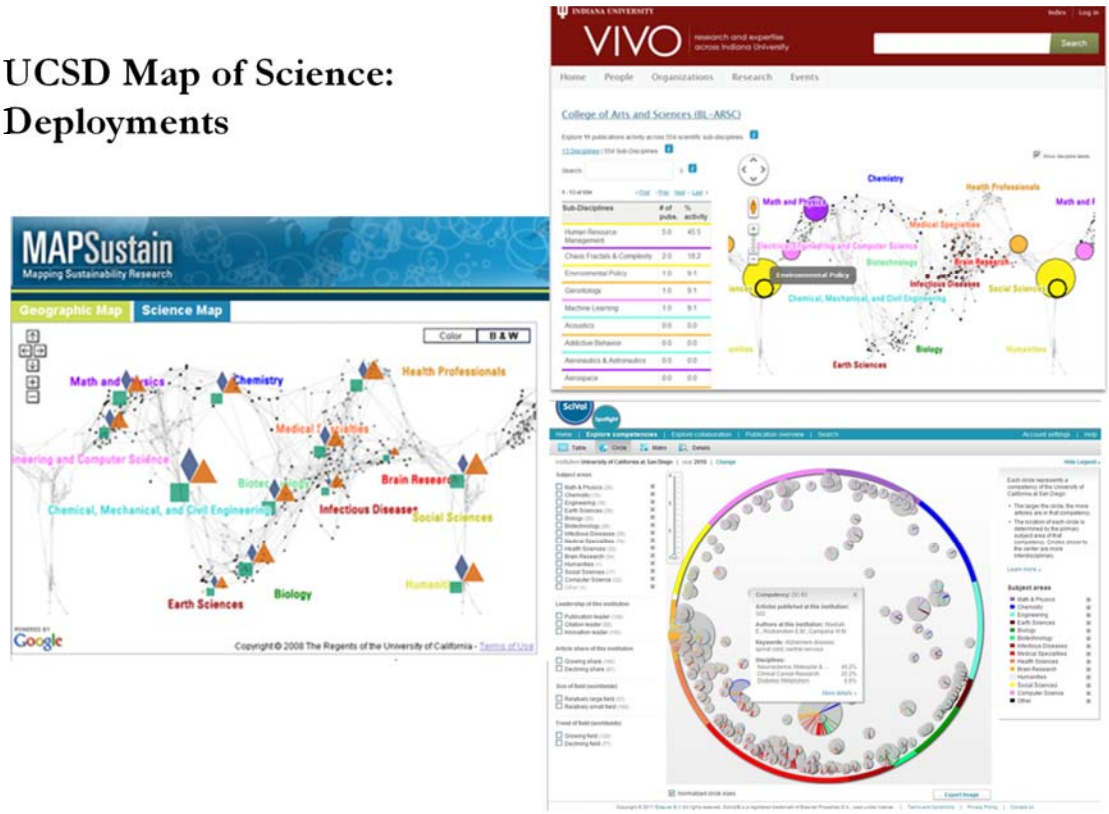
Number of Disciplines: 13; ***Subdisciplines:*** 554

Map Design and Usage:

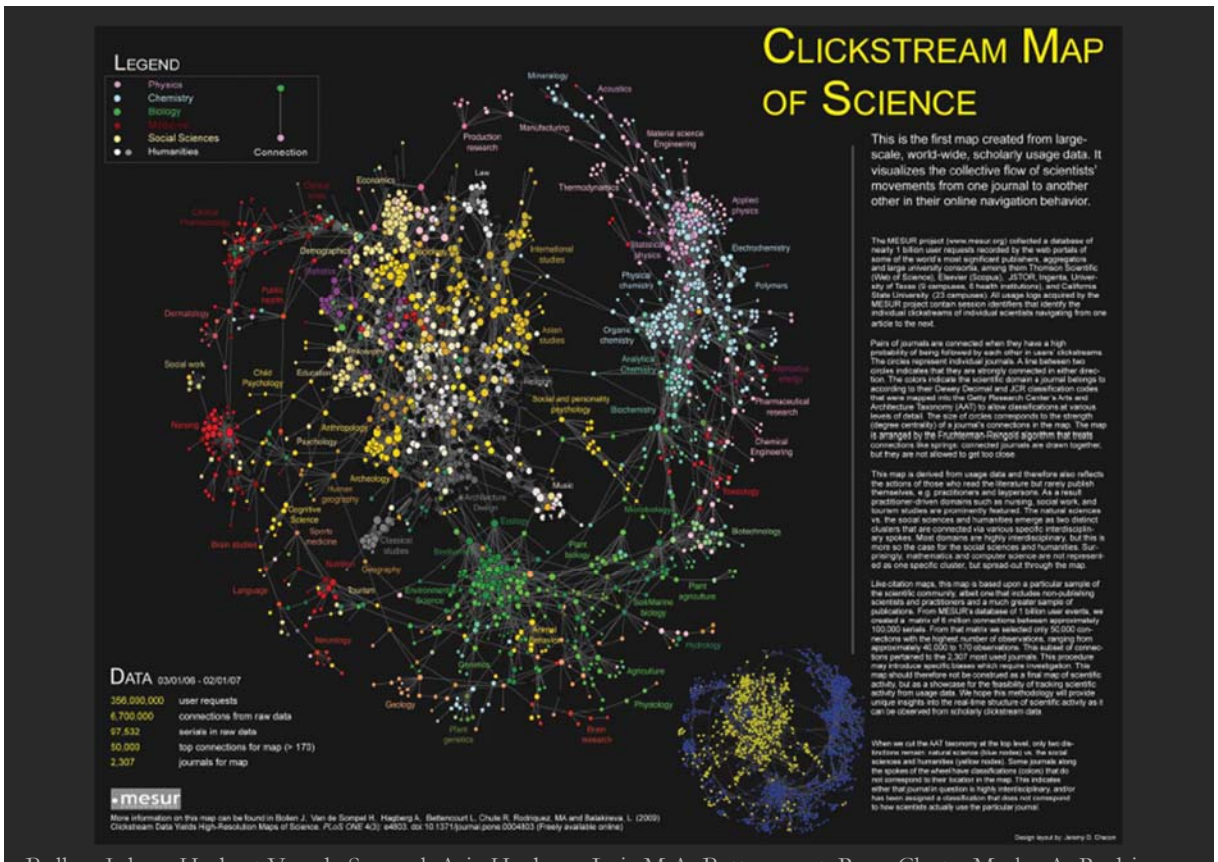
Map places 554 subdisciplines on the surface of a sphere—those with papers that cite the same base knowledge are placed in closer proximity. The spheric layout is then flattened using a Mercator projection. Each node is labeled and has an extensive list of journal names and key phrases as metadata, which can be used to "science locate" journal publications as well as nonjournal data such as patents or grants.

Börner, Katy, Richard Klavans, et al. (2012) Design and Update of a Classification System: The UCSD Map of Science. PLoS ONE 7(7): e39464. doi:10.1371/journal.pone.0039464

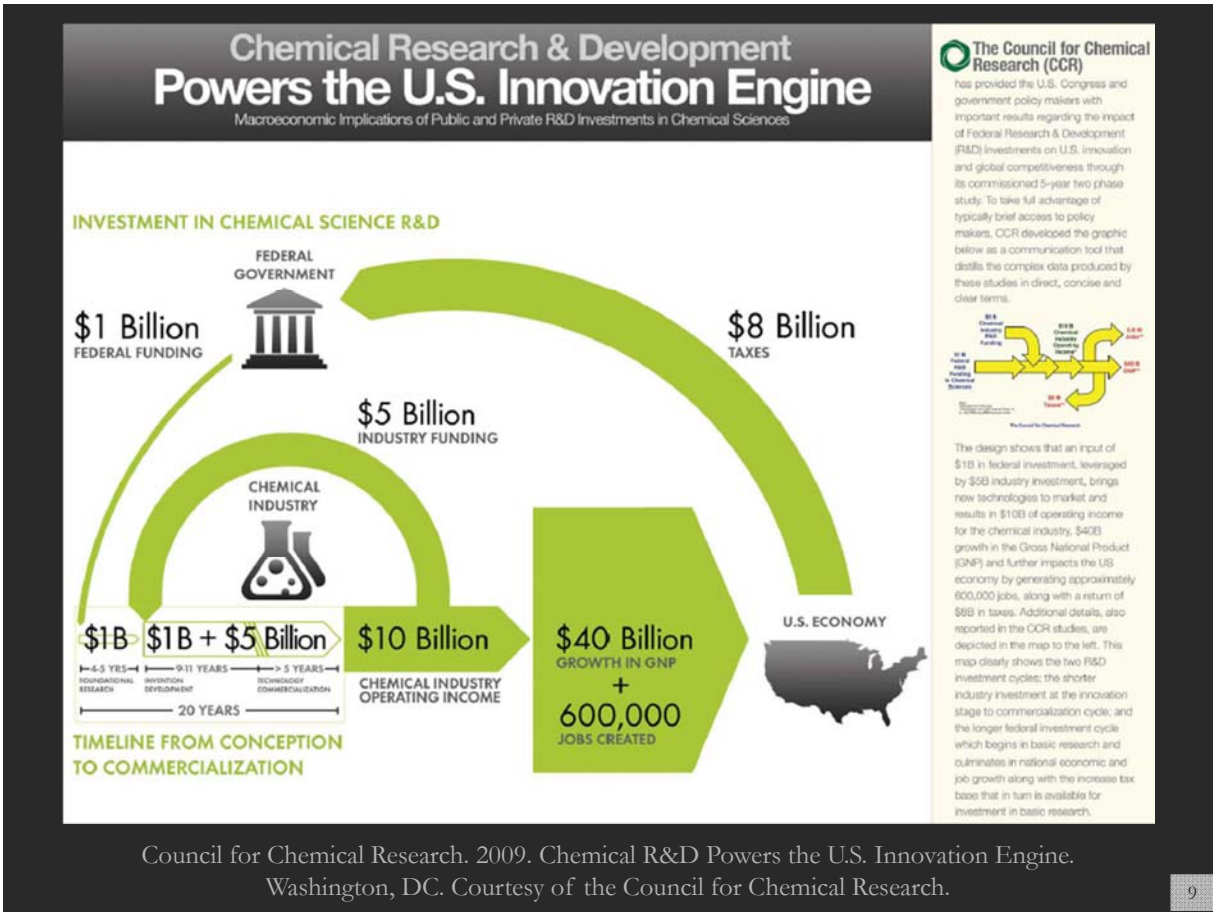
UCSD Map of Science: Deployments



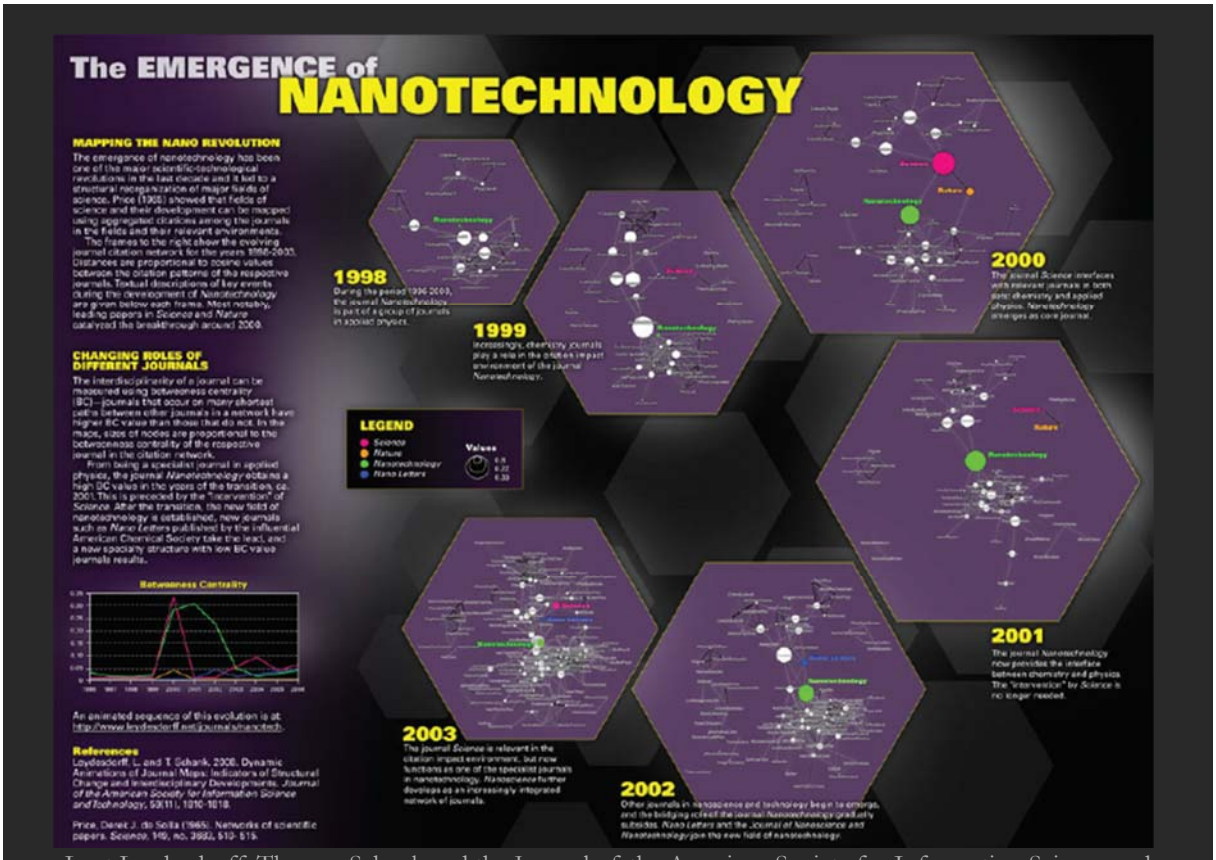
Börner, Katy, Richard Klavans, et al. (2012) Design and Update of a Classification System: The UCSD Map of Science. PLoS ONE 7(7): e39464. doi:10.1371/journal.pone.0039464



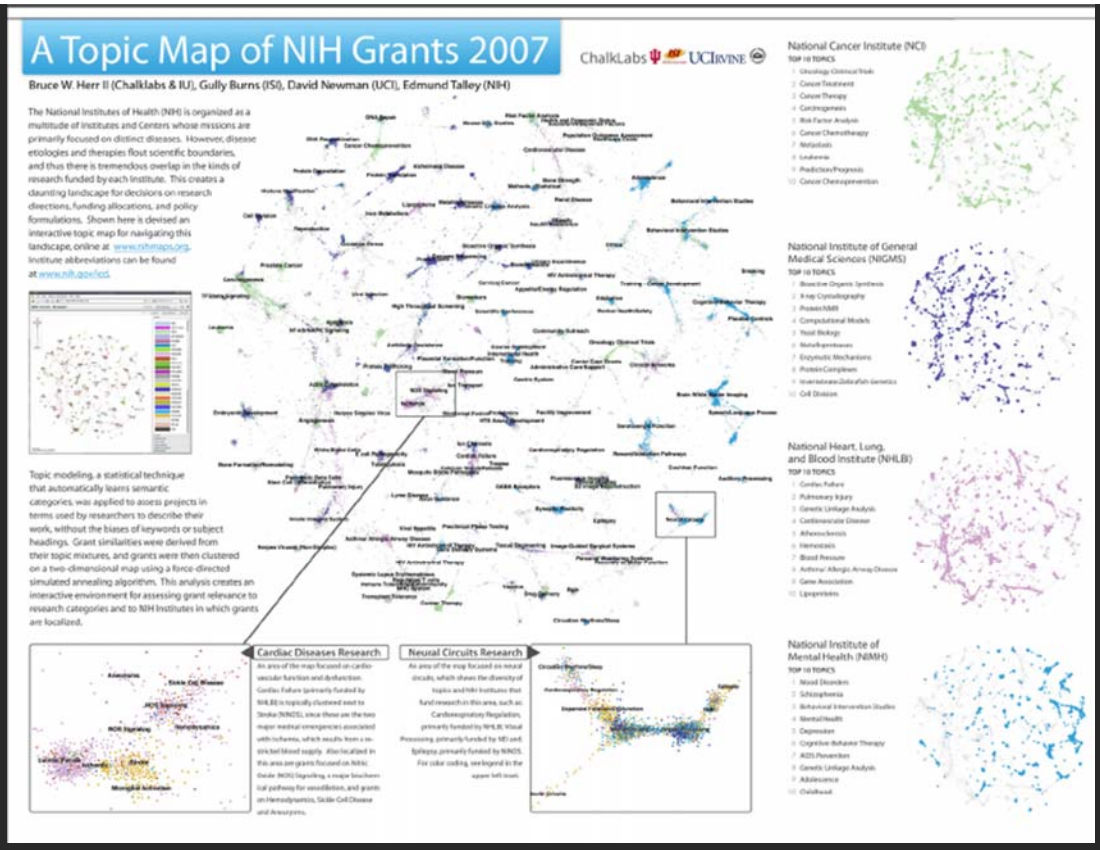
Bollen, Johan, Herbert Van de Sompel, Aric Hagberg, Luis M.A. Bettencourt, Ryan Chute, Marko A. Rodriquez, Lyudmila Balakireva. 2008. A Clickstream Map of Science.



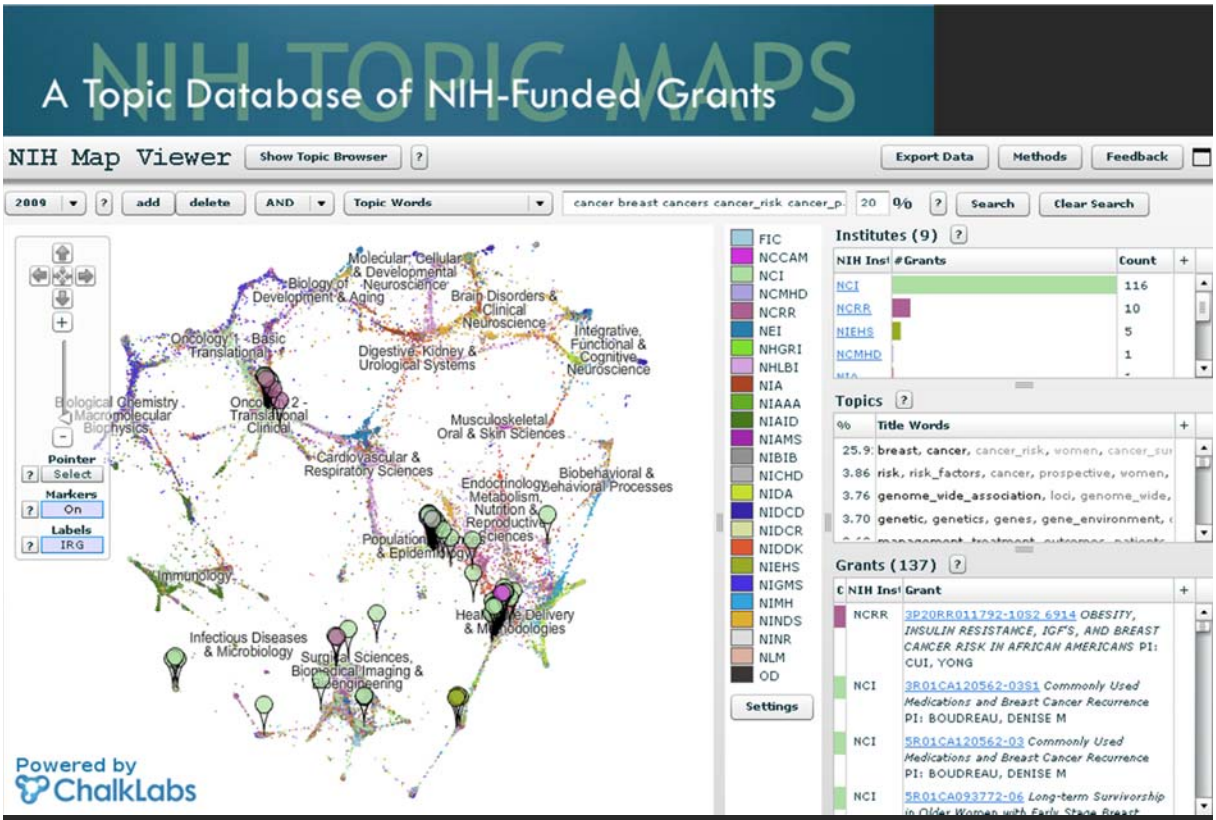
Council for Chemical Research. 2009. Chemical R&D Powers the U.S. Innovation Engine. Washington, DC. Courtesy of the Council for Chemical Research.



Loet Leydesdorff, Thomas Schank and the Journal of the American Society for Information Science and Technology. 2010. The Emergence of Nanoscience & Technology.

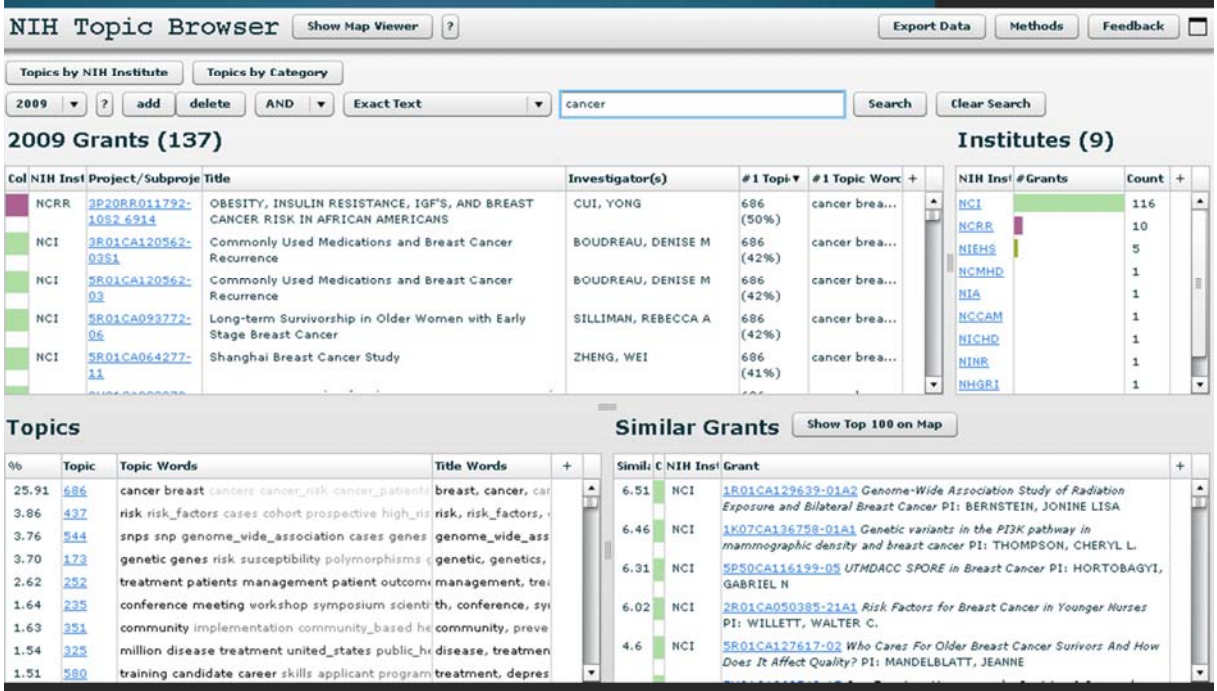


Bruce W. Herr II, Gully Burns, David Newman, Edmund Talley. 2007. A Topic Map of NIH Grants 2007



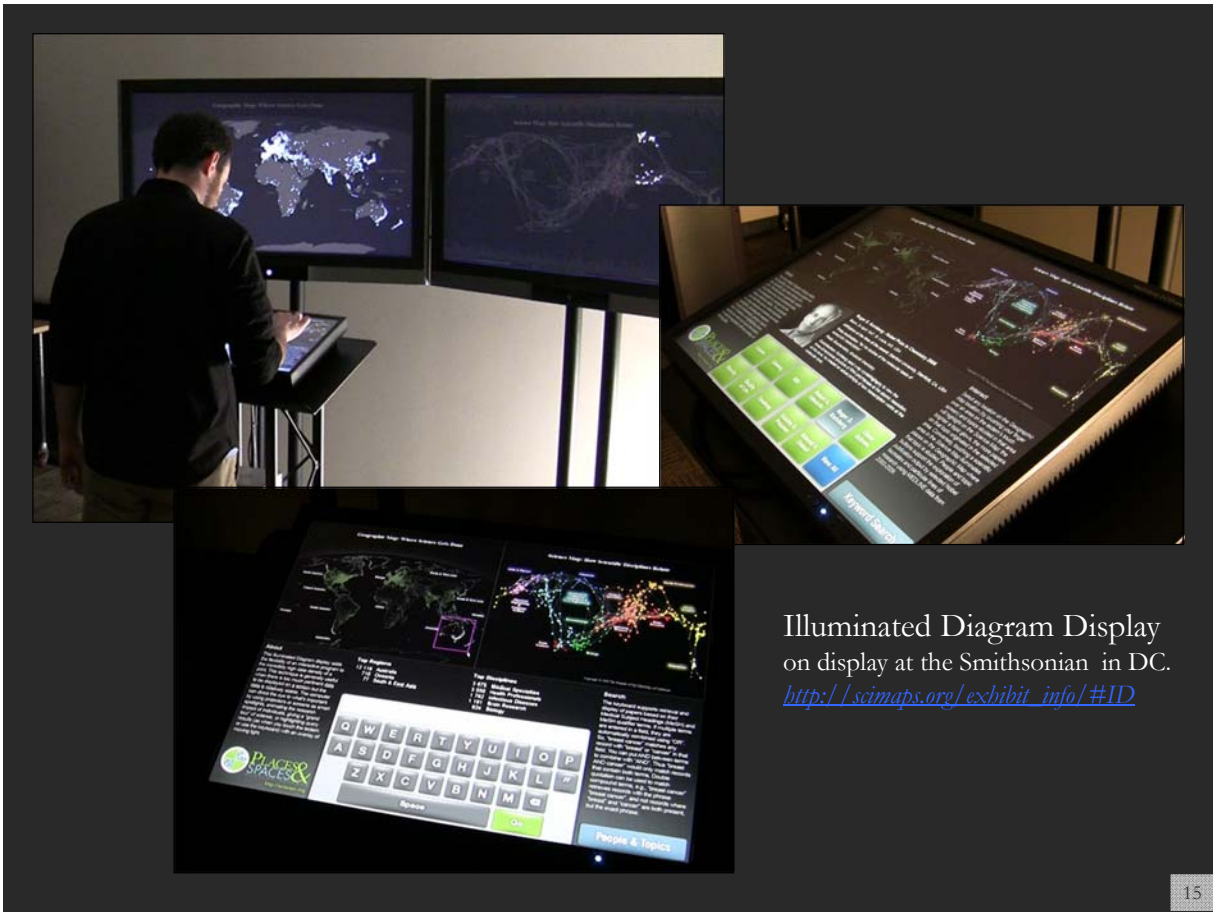
https://app.nihmaps.org

A Topic Database of NIH-Funded Grants

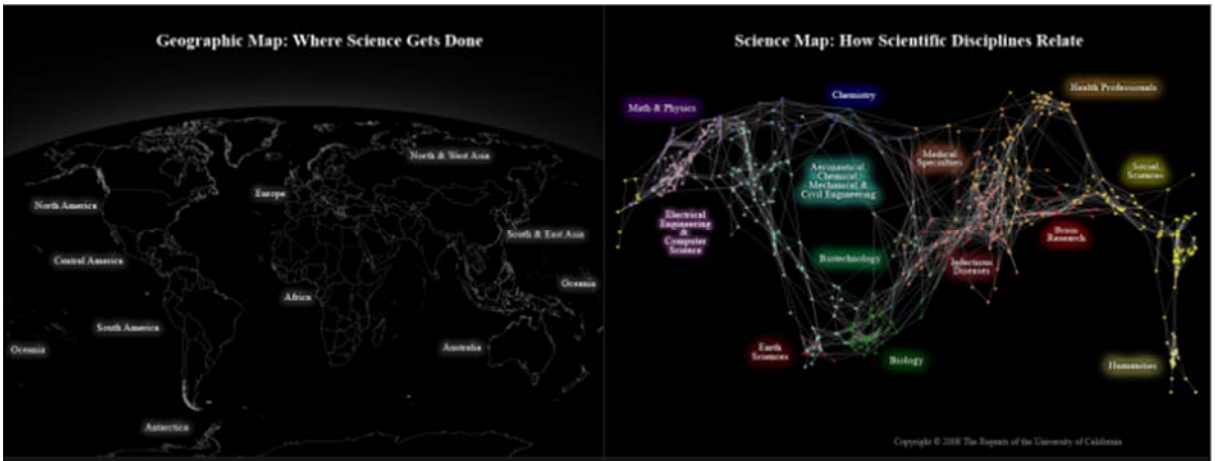


https://app.nihmaps.org

Science Maps in "Expedition Zukunft" science train visiting 62 cities in 7 months 12 coaches, 300 m long Opening was on April 23rd, 2009 by German Chancellor Merkel
http://www.expedition-zukunft.de



Illuminated Diagram Display
on display at the Smithsonian in DC.
http://scimaps.org/exhibit_info/#ID



About

This Illuminated Diagram display adds the flexibility of an interactive program to the incredibly high data density of a print. This technique is generally useful when there is too much pertinent data to be displayed on a screen but the data is relatively stable. The computer can direct the eye to what's important by using projectors or screens as smart spotlights, animating the research impact of individuals, giving a "grand tour" of science, or highlighting query results (as when you touch the lectern or use the keyboard) with an overlay of moving light.

PLACES & SPACES
http://scimaps.org

Top Five Continents

North America - 4,000 records
South & East Asia - 3,589
Australia - 2,431
Africa - 2,208
South America - 1,562

Top Five Scientific Disciplines

Math & Physics - 4,000 records
Health Professionals - 3,589
Social Sciences - 2,431
Aeronautical, Chemical, Mechanical & Civil Engineering - 2,208
Humanities - 1,562

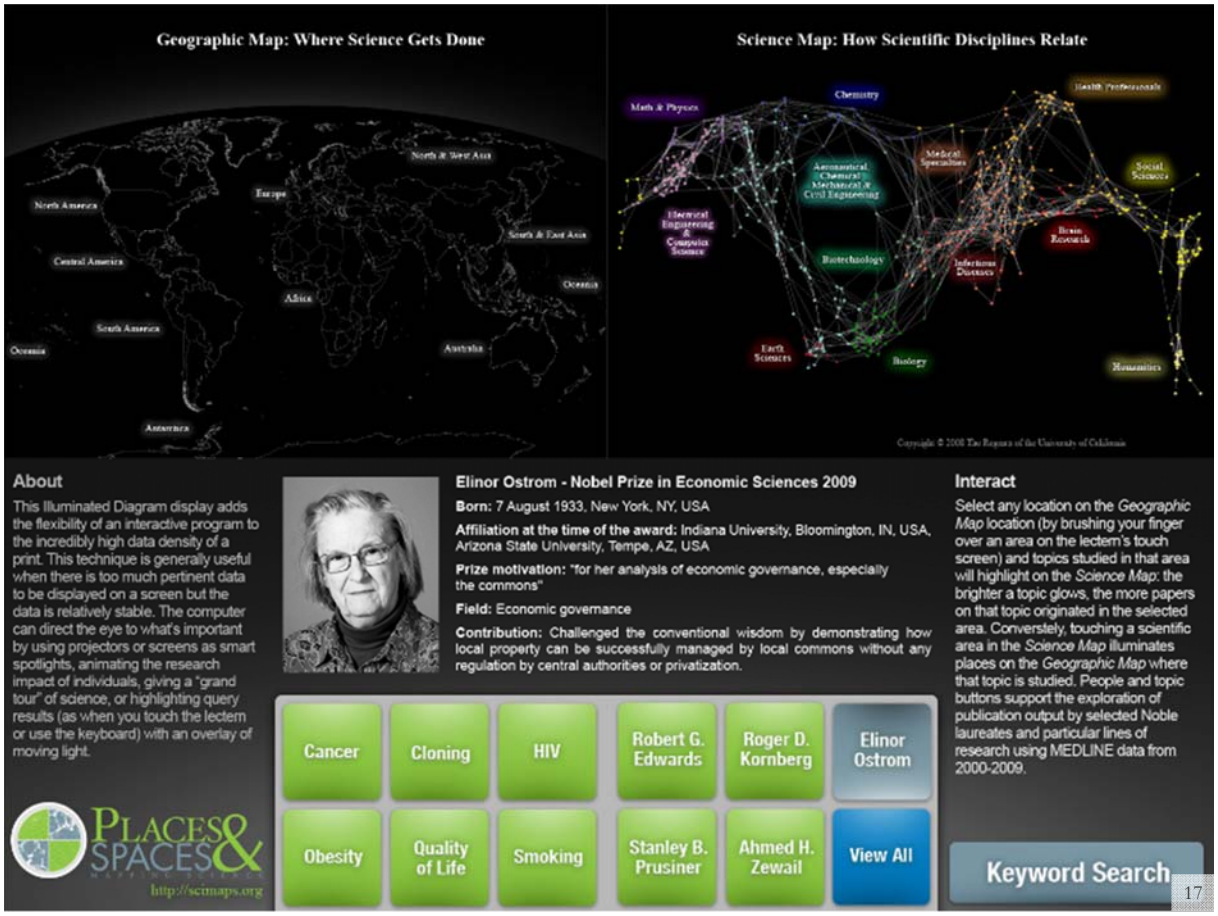
Input your search query here.

Q W E R T Y U I O P
A S D F G H J K L "
Z X C V B N M ⌫
Space Go

Search

The keyboard supports retrieval and display of papers based on their Medical Subject Headings (MeSH) and MeSH qualifier terms. If multiple terms are entered in a field, they are automatically combined using "OR". So, "breast cancer" matches any record with "breast" or "cancer" in that field. You can put AND between terms to combine with "AND". Thus "breast AND cancer" would only match records that contain both terms. Double quotation can be used to match compound terms, e.g., "breast cancer" retrieves records with the phrase "breast cancer", and not records where "breast" and "cancer" are both present, but the exact phrase.

People & Topics

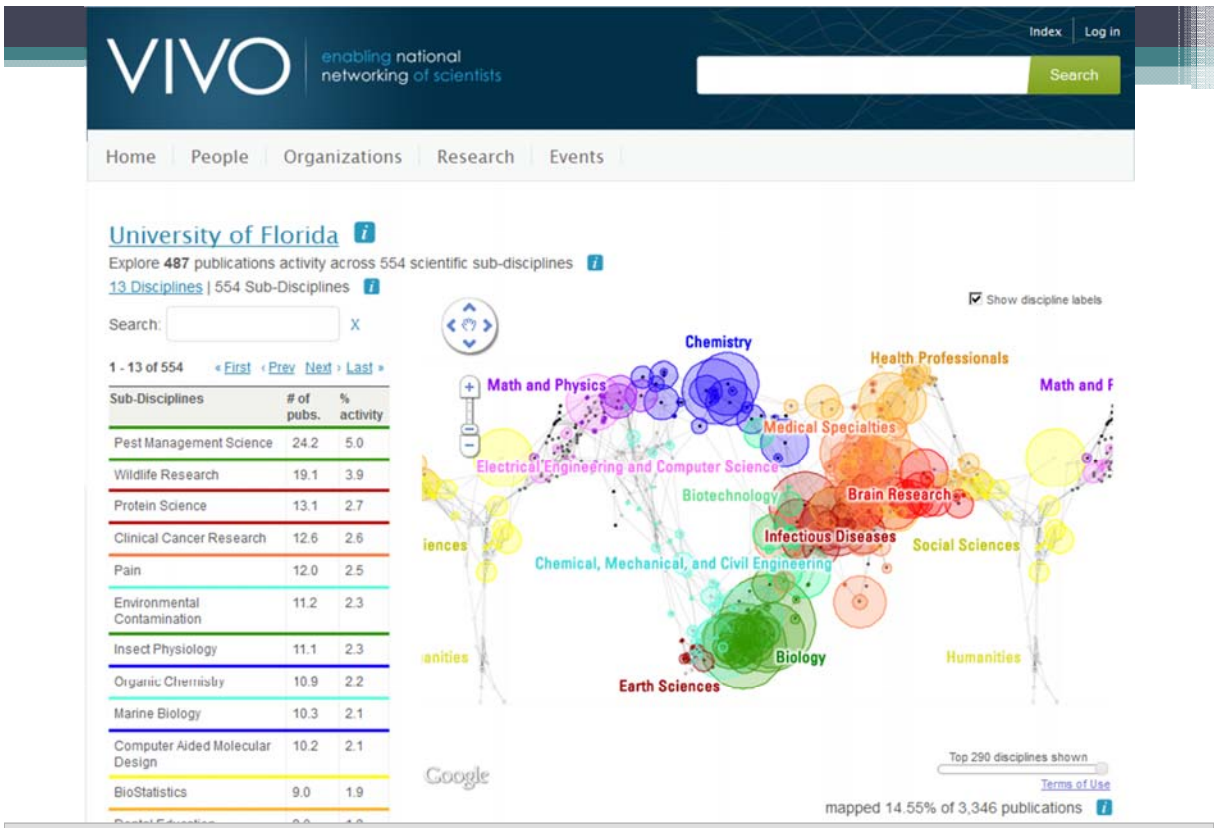


Multi-Level Science Maps

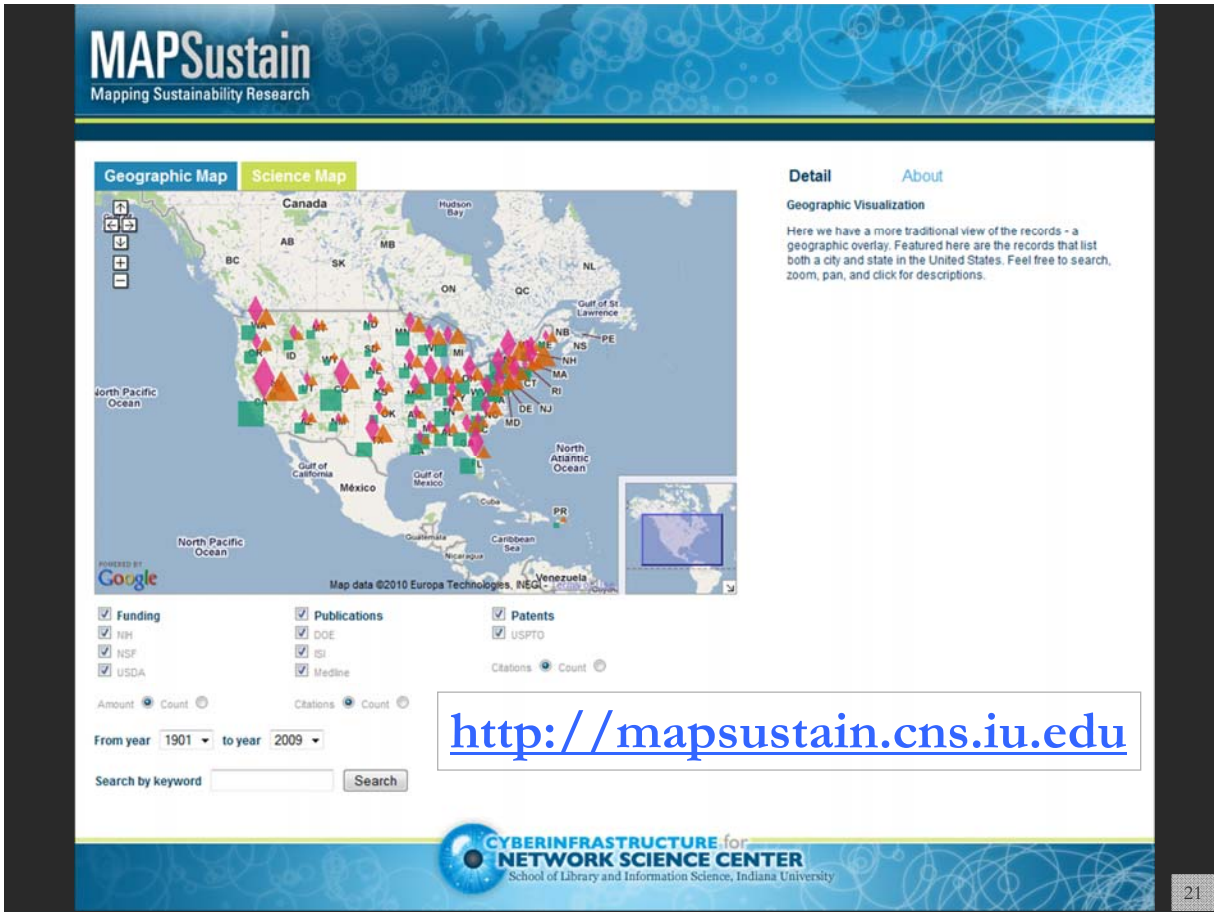
- VIVO International Researcher Network
- Sustainability Research Map
- Gene Therapy Research Map

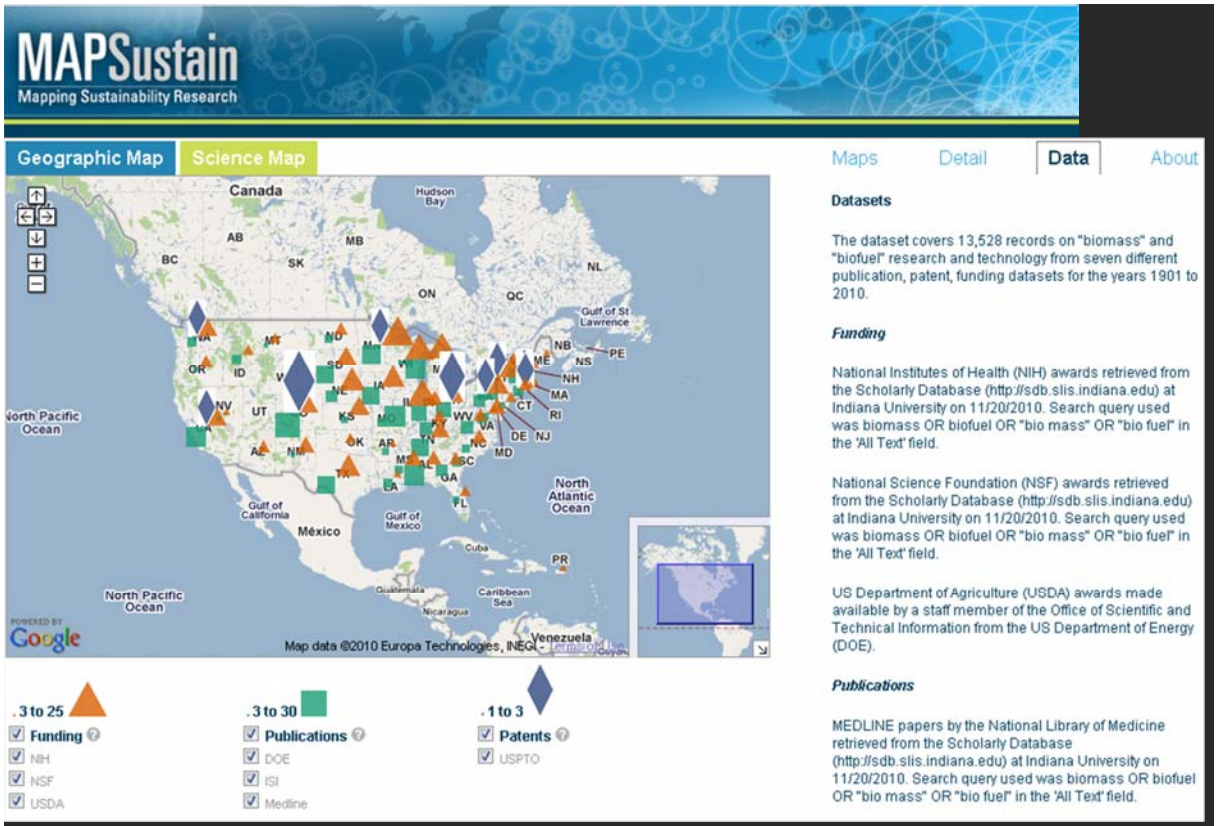
http://nrn.cns.iu.edu

Geospatial Analysis (Where) A geospatial map of the US is used to show where what science is performed by whom.



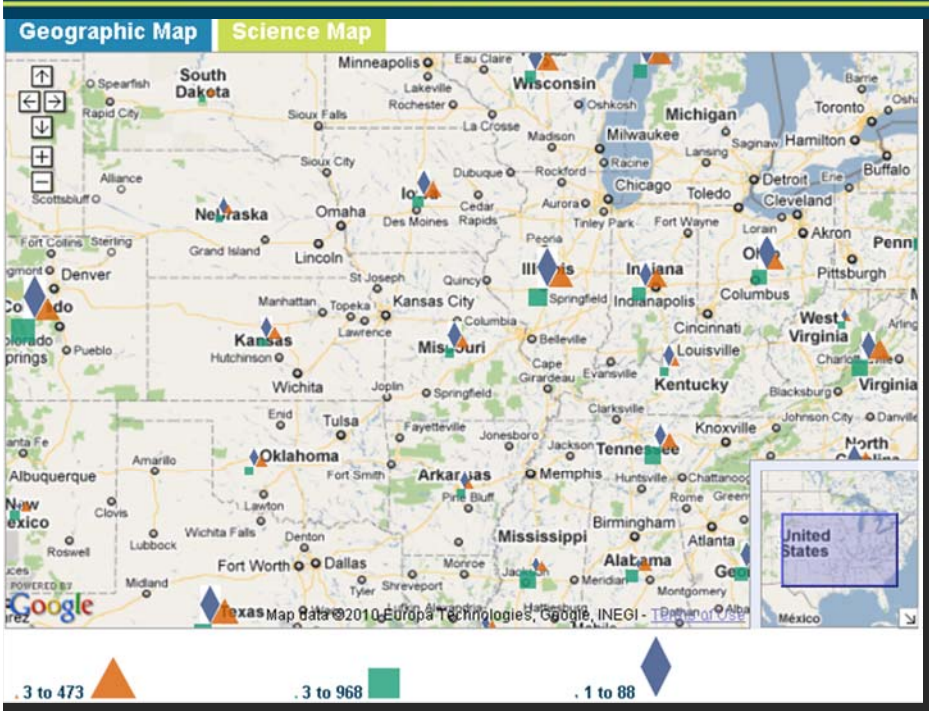
Topical Analysis (What) Science map overlays show where a person, department, or university publishes most in the world of science.





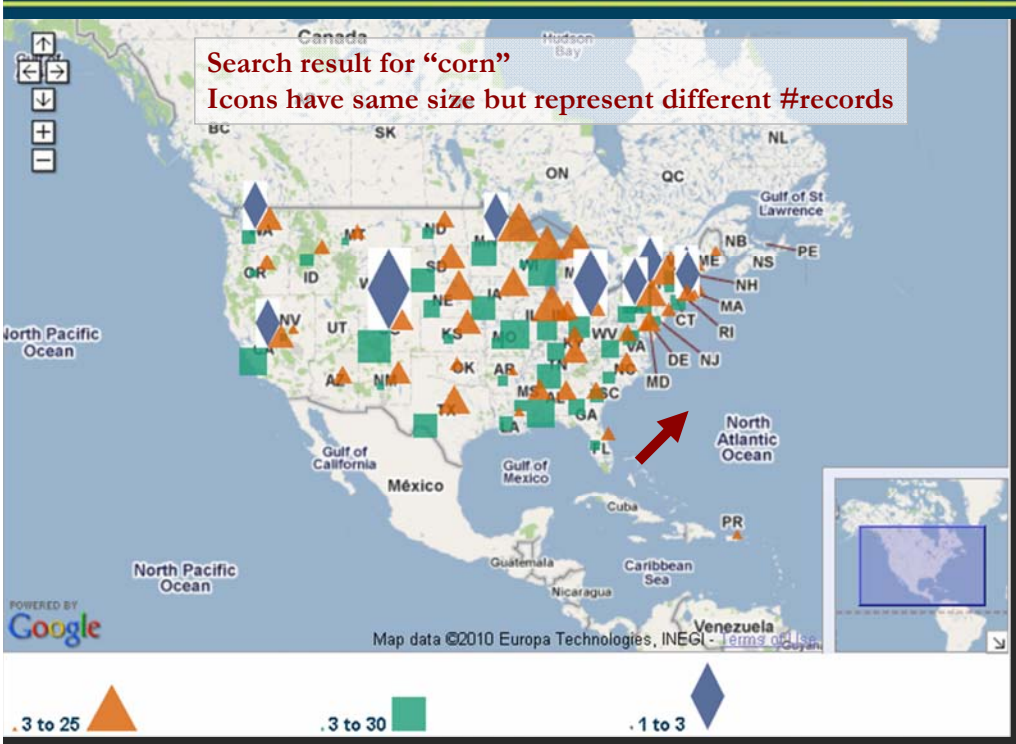
The geographic map at **state level**.

MAPSustain
Mapping Sustainability Research



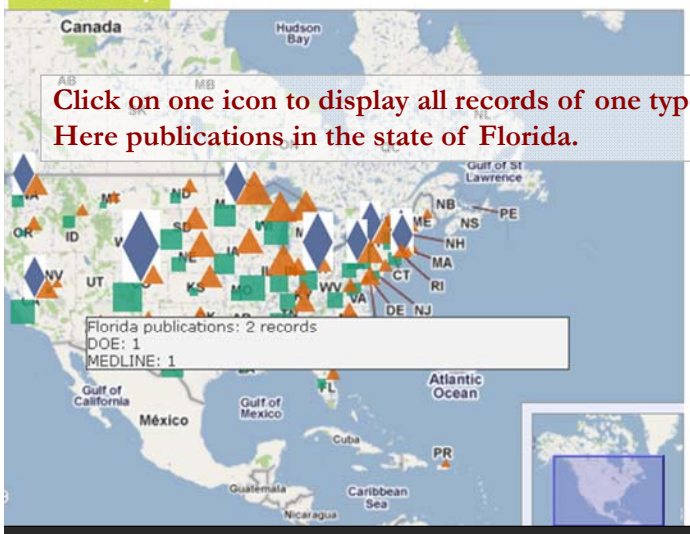
The geographic map at **city level**.

MAPSustain
Mapping Sustainability Research



MAPSustain
Mapping Sustainability Research

Science Map



Click on one icon to display all records of one type.
Here publications in the state of Florida.

Maps Detail Data About

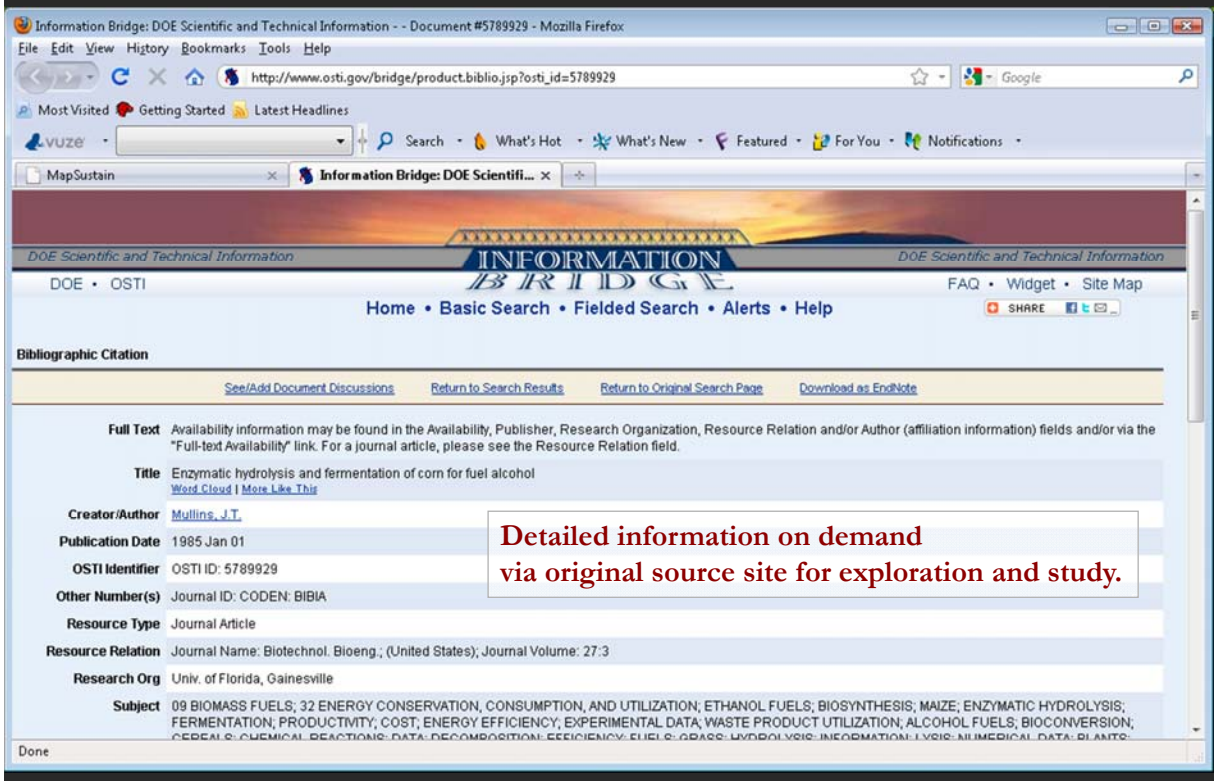
> Florida

MEDLINE
2002

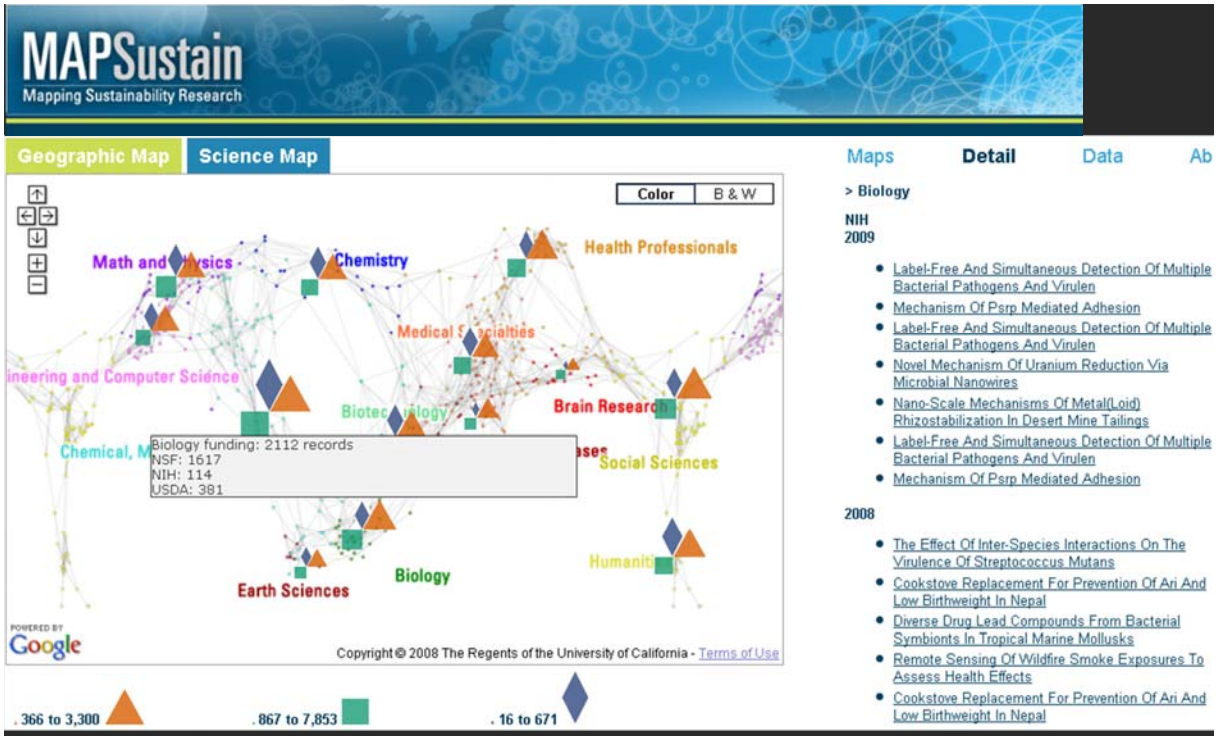
- Recovery Of Dairy Manure Nutrients By Benthic Freshwater Algae.

DOE
1985

- Enzymatic Hydrolysis And Fermentation Of Corn For Fuel Alcohol

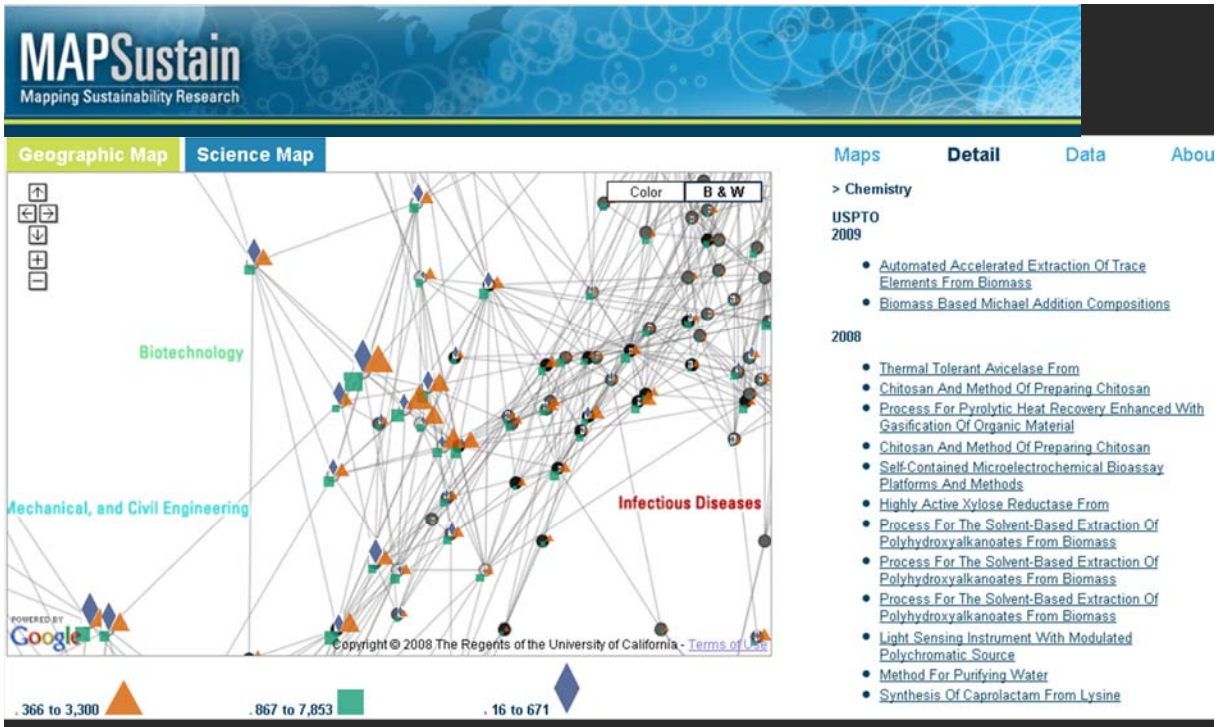


Detailed information on demand
via original source site for exploration and study.

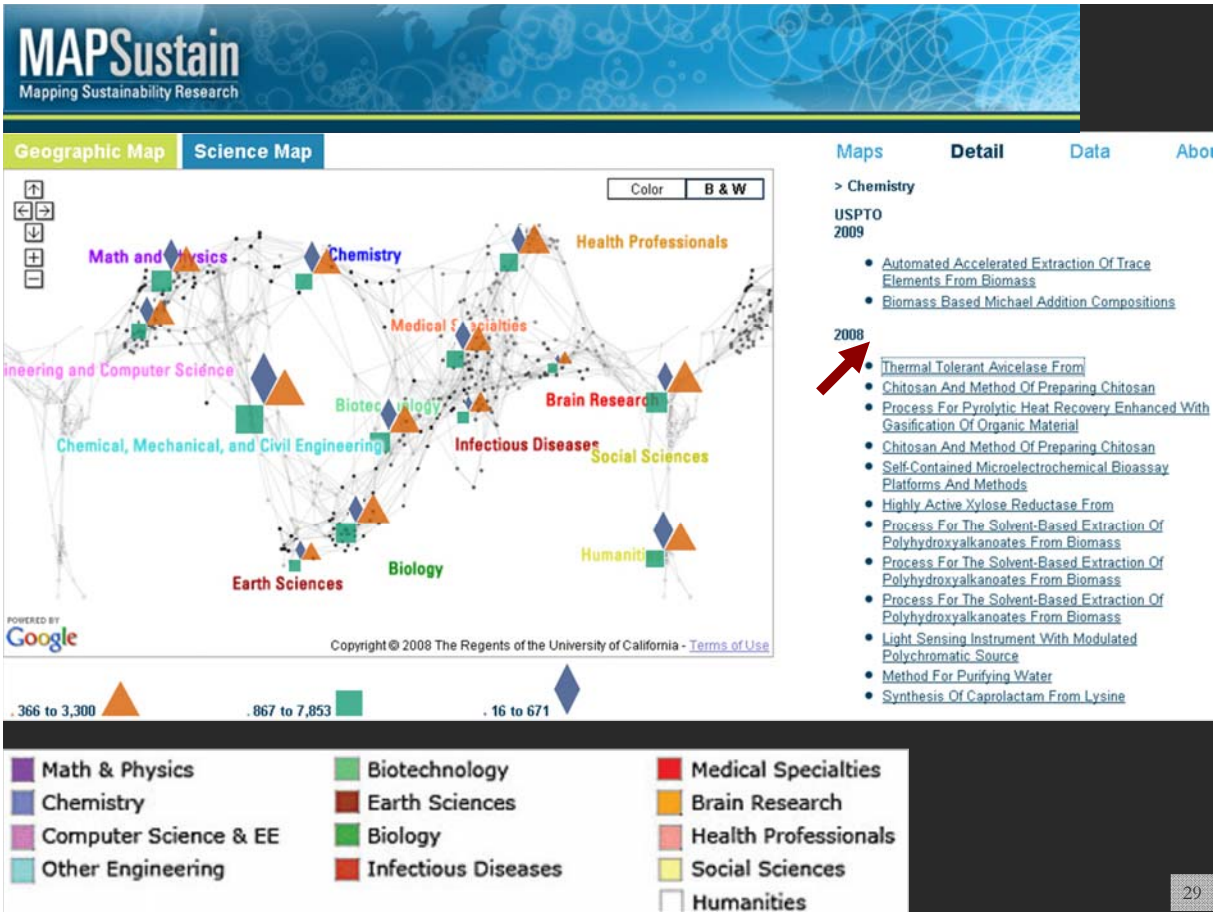


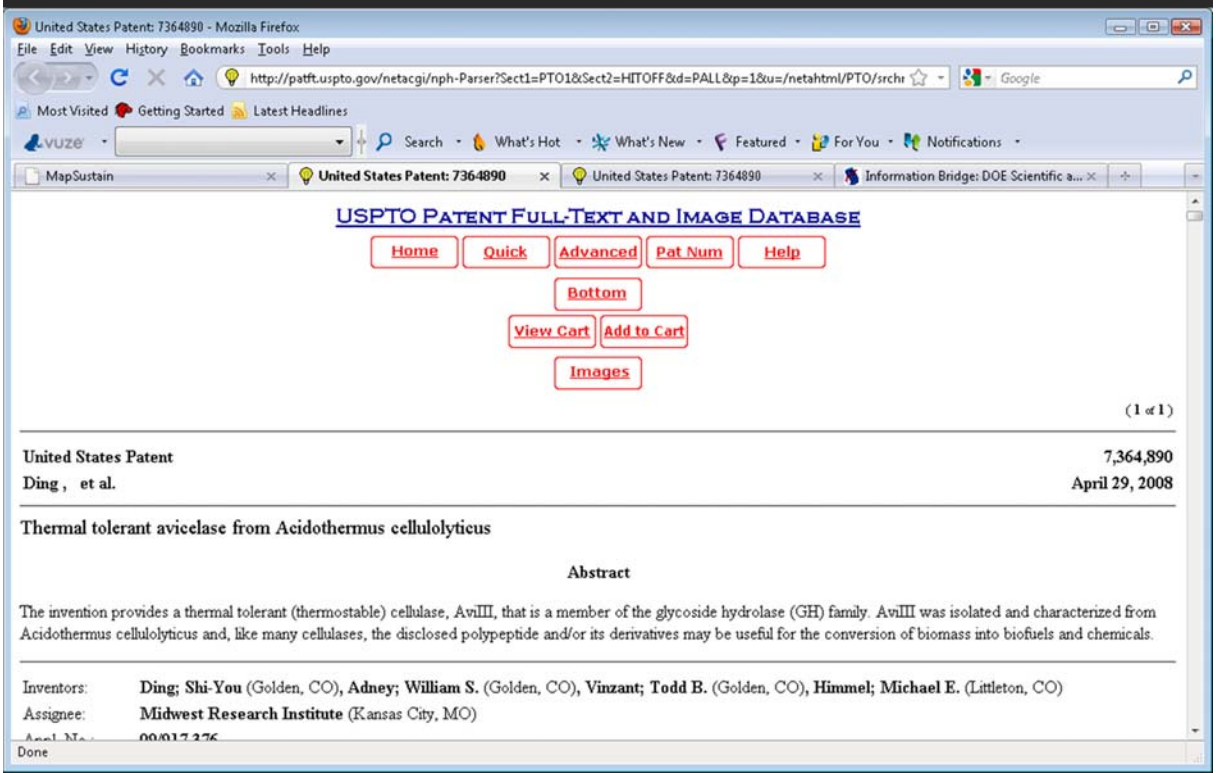
The science map at 13 top-level scientific **disciplines level**.

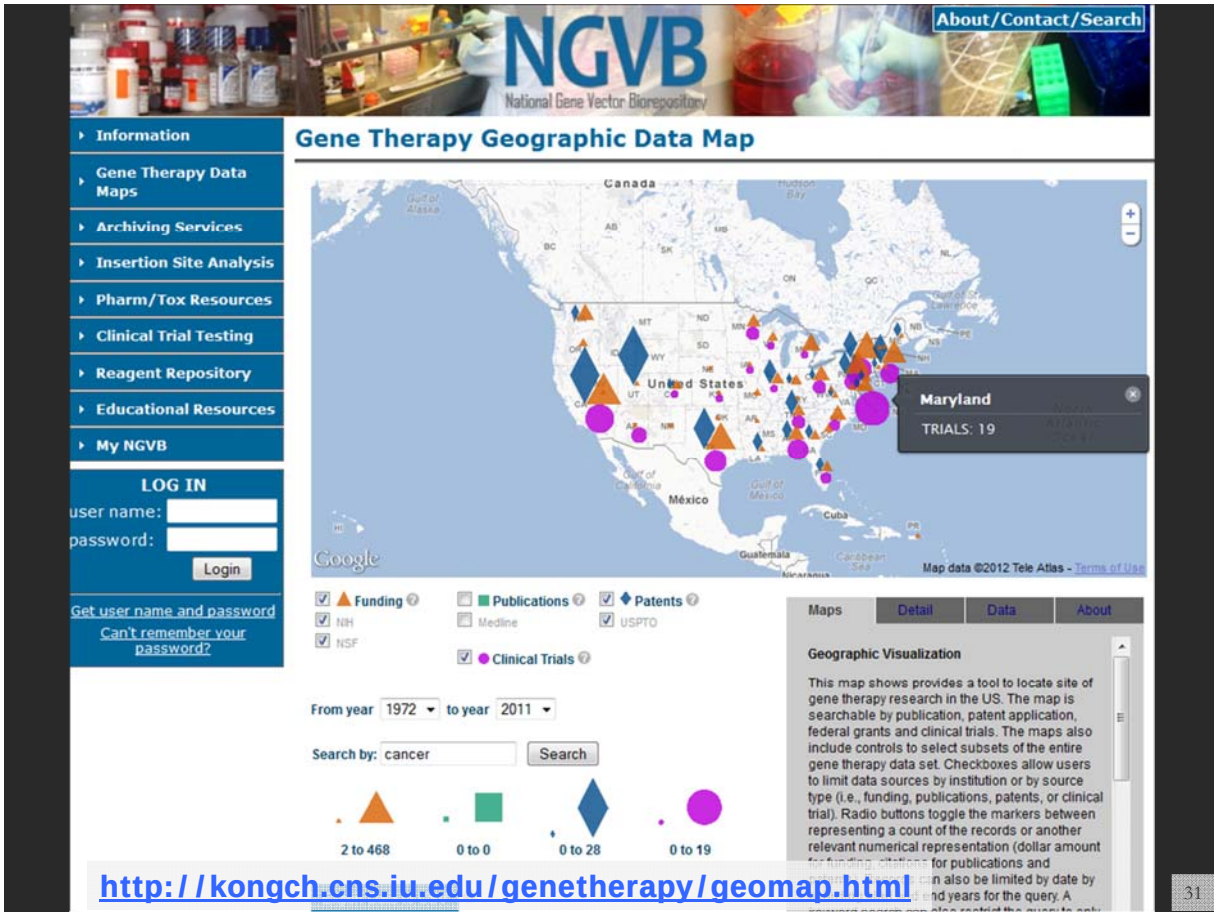
Math & Physics
Chemistry
Computer Science & EE
Other Engineering
Biotechnology
Earth Sciences
Biology
Infectious Diseases
Medical Specialties
Brain Research
Health Professionals
Social Sciences
Humanities



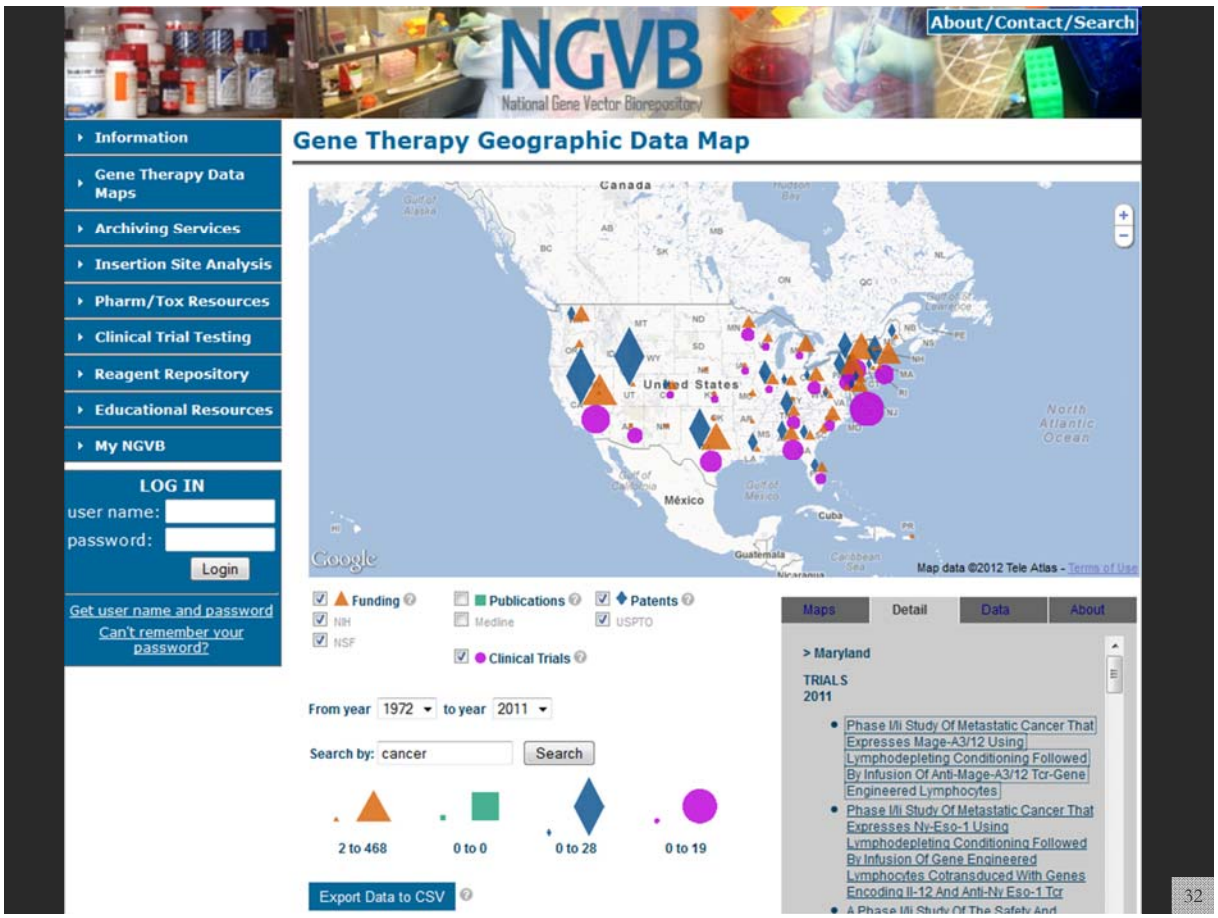
The science map at 554 **sub-disciplines level**.

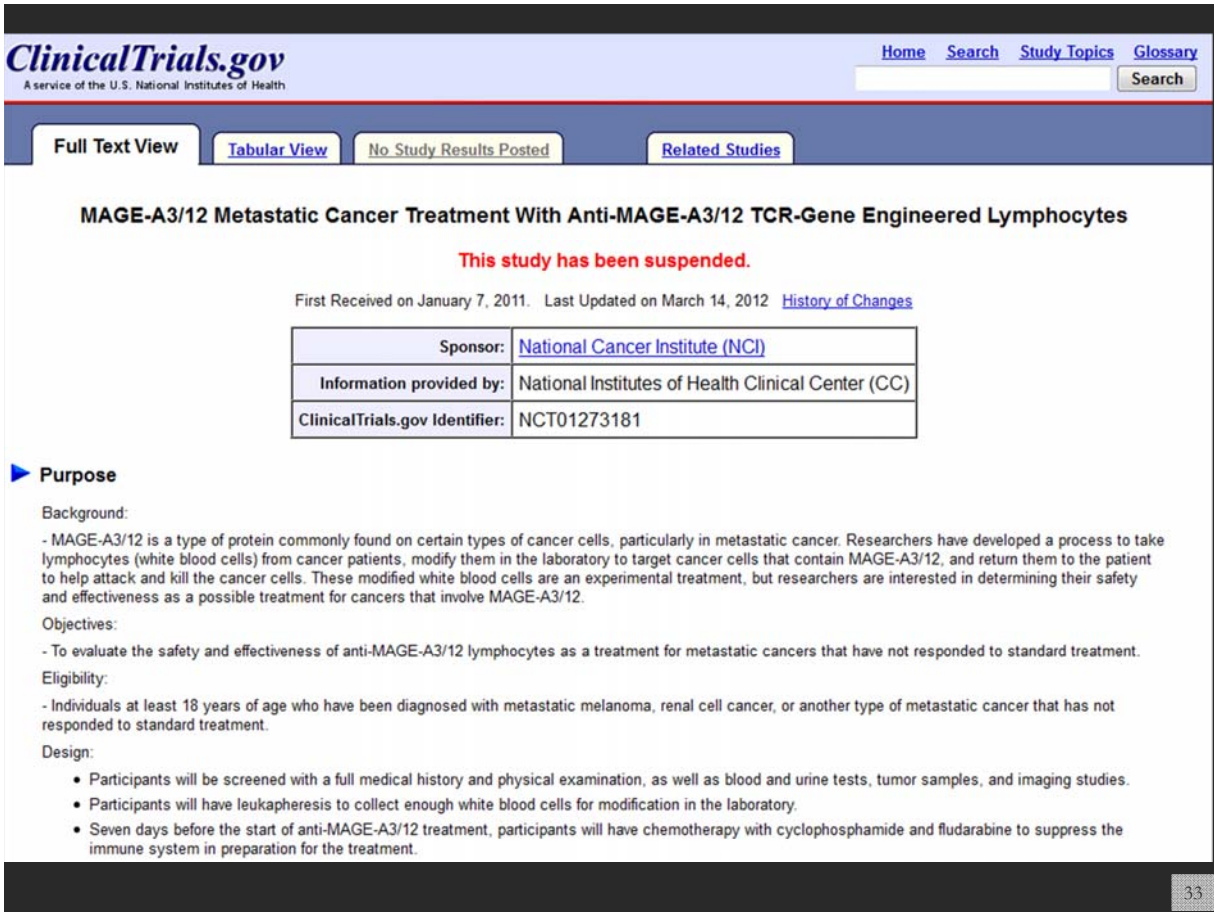

MAPSustain
Mapping Sustainability Research
Geographic Map
Science Map
Color
B & W
Math and Physics
Chemistry
Health Professionals
Medical Specialties
Engineering and Computer Science
Biotechnology
Brain Research
Chemical, Mechanical, and Civil Engineering
Infectious Diseases
Social Sciences
Humanities
Biology
Earth Sciences
Copyright © 2008 The Regents of the University of California - Terms of Use
.366 to 3,300
.867 to 7,853
.16 to 671
Maps
Detail
Data
> Chemistry
USPTO
2009
Automated Accelerated Extraction Of Trace Elements From Biomass
Biomass Based Michael Addition Compositions
2008
Thermal Tolerant Avicelase From
Chitosan And Method Of Preparing Chitosan
Process For Pyrolytic Heat Recovery Enhanced With Gasification Of Organic Material
Chitosan And Method Of Preparing Chitosan
Self-Contained Microelectrochemical Bioassay Platforms And Methods
Highly Active Xylose Reductase From
Process For The Solvent-Based Extraction Of Polyhydroxyalkanoates From Biomass
Process For The Solvent-Based Extraction Of Polyhydroxyalkanoates From Biomass
Process For The Solvent-Based Extraction Of Polyhydroxyalkanoates From Biomass
Light Sensing Instrument With Modulated Polychromatic Source
Method For Purifying Water
Synthesis Of Caprolactam From Lysine
Math & Physics
Chemistry
Computer Science & EE
Other Engineering
Biotechnology
Earth Sciences
Biology
Infectious Diseases
Medical Specialties
Brain Research
Health Professionals
Social Sciences
Humanities


United States Patent: 7364890 - Mozilla Firefox
File Edit View History Bookmarks Tools Help
http://patft.uspto.gov/netacgi/nph-Parser?Sect1=PTO1&Sect2=HITOFF&d=PALL&p=1&u=/netahtml/PTO/srchr
Most Visited Getting Started Latest Headlines
Search What's Hot What's New Featured For You Notifications
MapSustain
United States Patent: 7364890
United States Patent: 7364890
Information Bridge: DOE Scientific a...
USPTO PATENT FULL-TEXT AND IMAGE DATABASE
Home Quick Advanced Pat Num Help
Bottom
View Cart Add to Cart
Images
(1 of 1)
United States Patent
Ding , et al.
7,364,890
April 29, 2008
Thermal tolerant avicelase from Acidothermus cellulolyticus
Abstract
The invention provides a thermal tolerant (thermostable) cellulase, AviIII, that is a member of the glycoside hydrolase (GH) family. AviIII was isolated and characterized from Acidothermus cellulolyticus and, like many cellulases, the disclosed polypeptide and/or its derivatives may be useful for the conversion of biomass into biofuels and chemicals.
Inventors: Ding; Shi-You (Golden, CO), Adney; William S. (Golden, CO), Vinzant; Todd B. (Golden, CO), Himmel; Michael E. (Littleton, CO)
Assignee: Midwest Research Institute (Kansas City, MO)
Done



http://kongch.cns.iu.edu/genetherapy/geomap.html





ClinicalTrials.gov
A service of the U.S. National Institutes of Health

Home Search Study Topics Glossary
Search

Full Text View | Tabular View | No Study Results Posted | Related Studies

MAGE-A3/12 Metastatic Cancer Treatment With Anti-MAGE-A3/12 TCR-Gene Engineered Lymphocytes

This study has been suspended.

First Received on January 7, 2011. Last Updated on March 14, 2012 History of Changes

Sponsor:	National Cancer Institute (NCI)
Information provided by:	National Institutes of Health Clinical Center (CC)
ClinicalTrials.gov Identifier:	NCT01273181

Purpose

Background:

- MAGE-A3/12 is a type of protein commonly found on certain types of cancer cells, particularly in metastatic cancer. Researchers have developed a process to take lymphocytes (white blood cells) from cancer patients, modify them in the laboratory to target cancer cells that contain MAGE-A3/12, and return them to the patient to help attack and kill the cancer cells. These modified white blood cells are an experimental treatment, but researchers are interested in determining their safety and effectiveness as a possible treatment for cancers that involve MAGE-A3/12.

Objectives:

- To evaluate the safety and effectiveness of anti-MAGE-A3/12 lymphocytes as a treatment for metastatic cancers that have not responded to standard treatment.

Eligibility:

- Individuals at least 18 years of age who have been diagnosed with metastatic melanoma, renal cell cancer, or another type of metastatic cancer that has not responded to standard treatment.

Design:

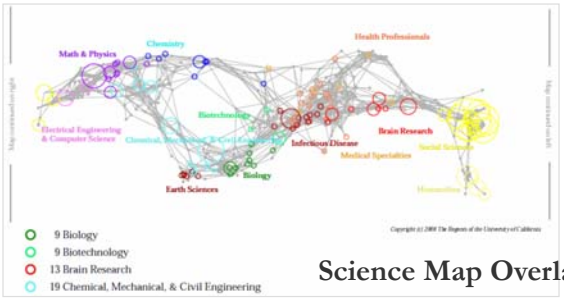
- Participants will be screened with a full medical history and physical examination, as well as blood and urine tests, tumor samples, and imaging studies.
- Participants will have leukapheresis to collect enough white blood cells for modification in the laboratory.
- Seven days before the start of anti-MAGE-A3/12 treatment, participants will have chemotherapy with cyclophosphamide and fludarabine to suppress the immune system in preparation for the treatment.

Science of Science (Sci2) Tool

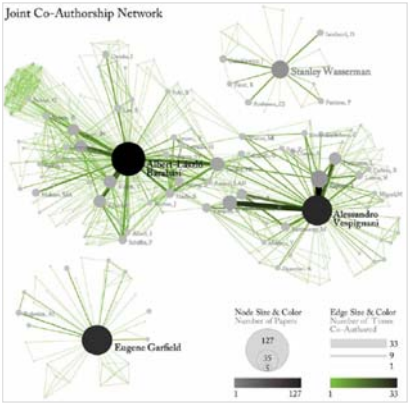
- Use your own data
- Run your own analysis
- Identify overlap, gaps and emerging areas
- Interpret results to improve decision making

Science of Science (Sci²) Tool – Open Code for S&T Assessment

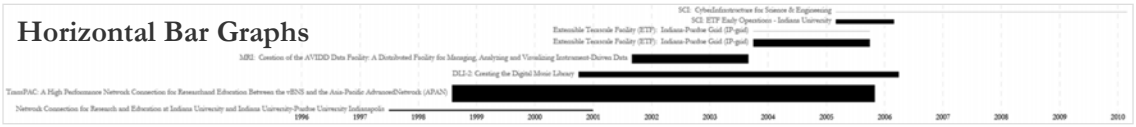
OSGi/CIShell powered tool with NWB plugins and many new scientometrics and visualizations plugins.



Science Map Overlays

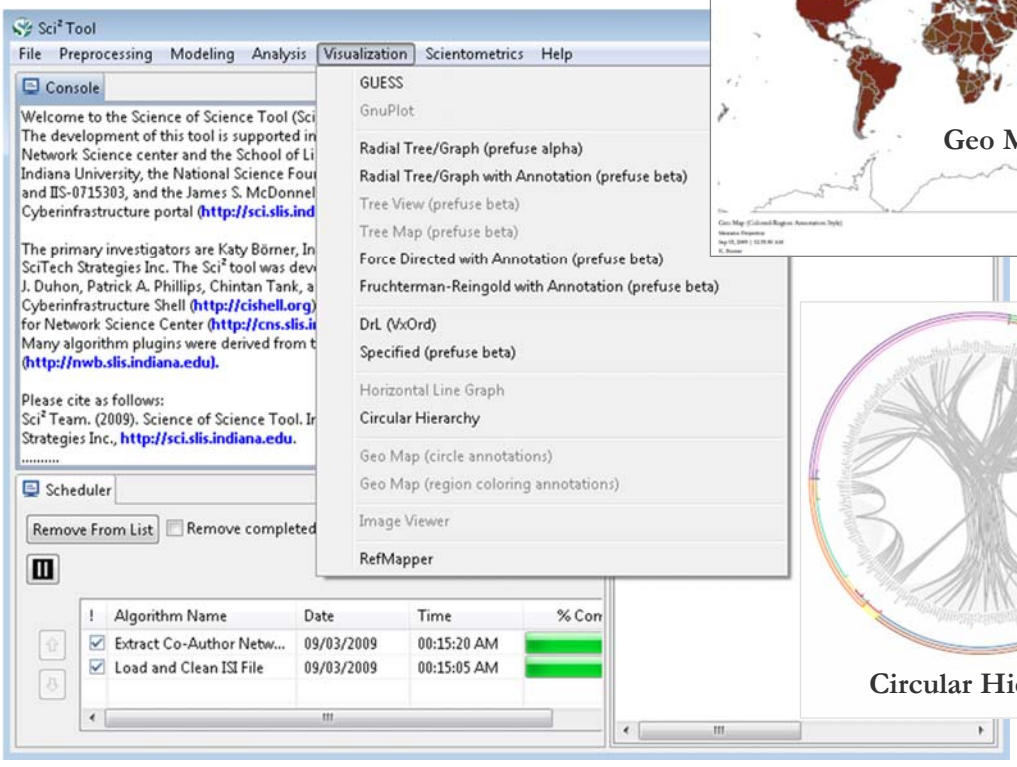


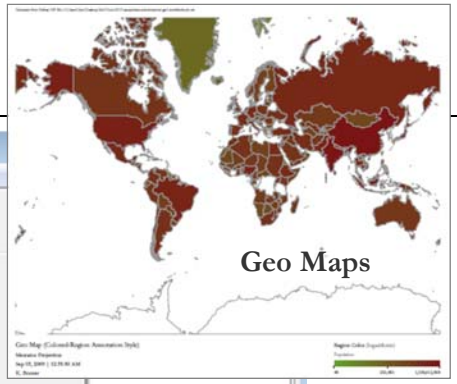
Network Visualizations



Börner, Katy. (2011). Plug-and-Play Macroscopes. Communications of the ACM, 54(3), 60-69.
Video and paper are at http://www.scivee.tv/node/27704

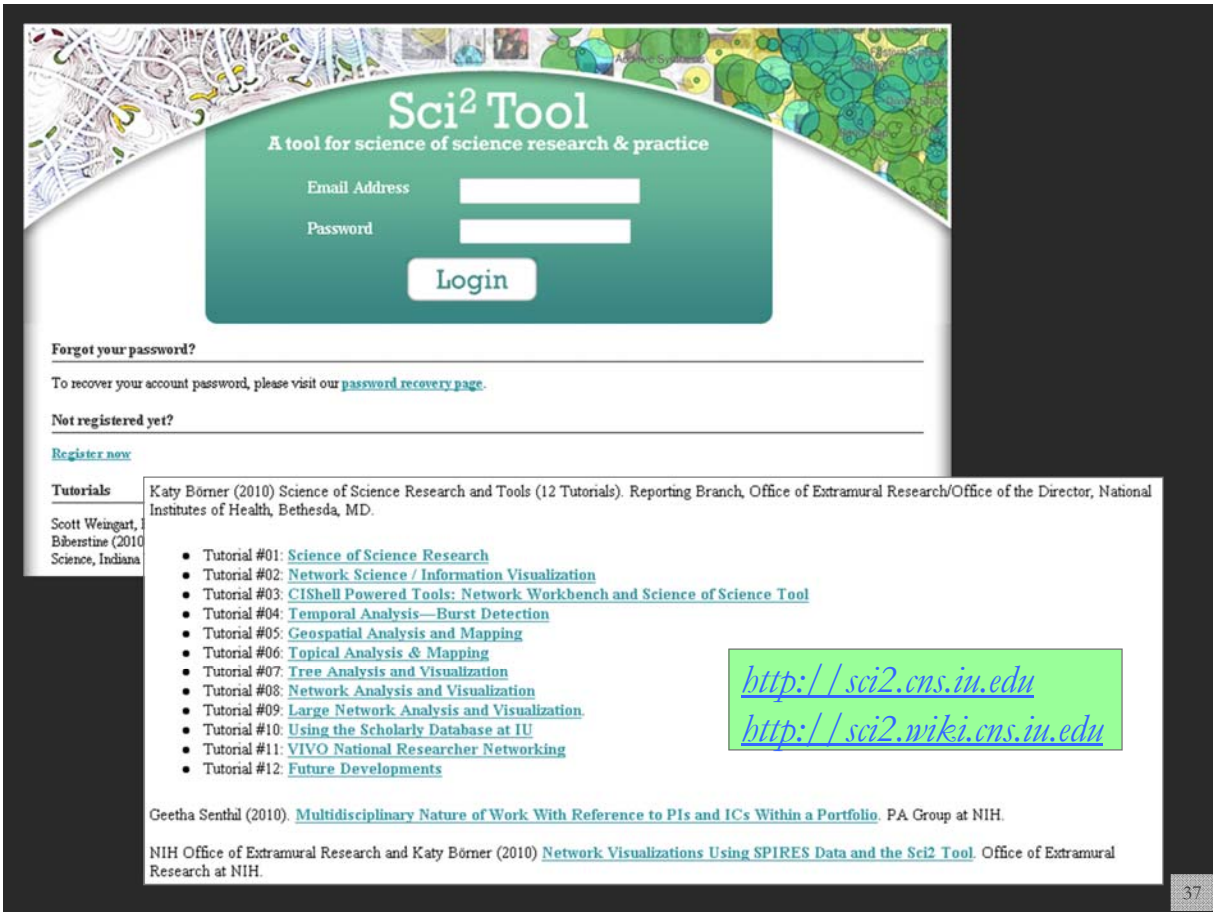
Sci² Tool cont.





Geo Maps

Circular Hierarchy



Science of Science (Sci²) Tool – Usage

The Sci² Tool is used by the

- National Science Foundation,
- National Institutes of Health,
- US Department of Agriculture, and
- National Oceanic and Atmospheric Administration

Tool registrations come from 73 countries and professions such as

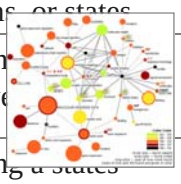
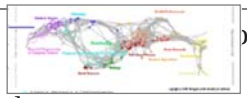
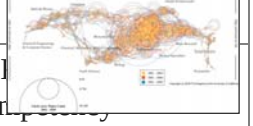
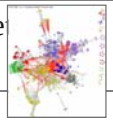
38

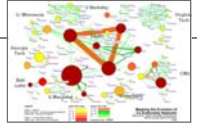
Sci² Tool – Type of Analysis vs. Level of Analysis

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

Sci² Tool – Type of Analysis vs. Level of Analysis

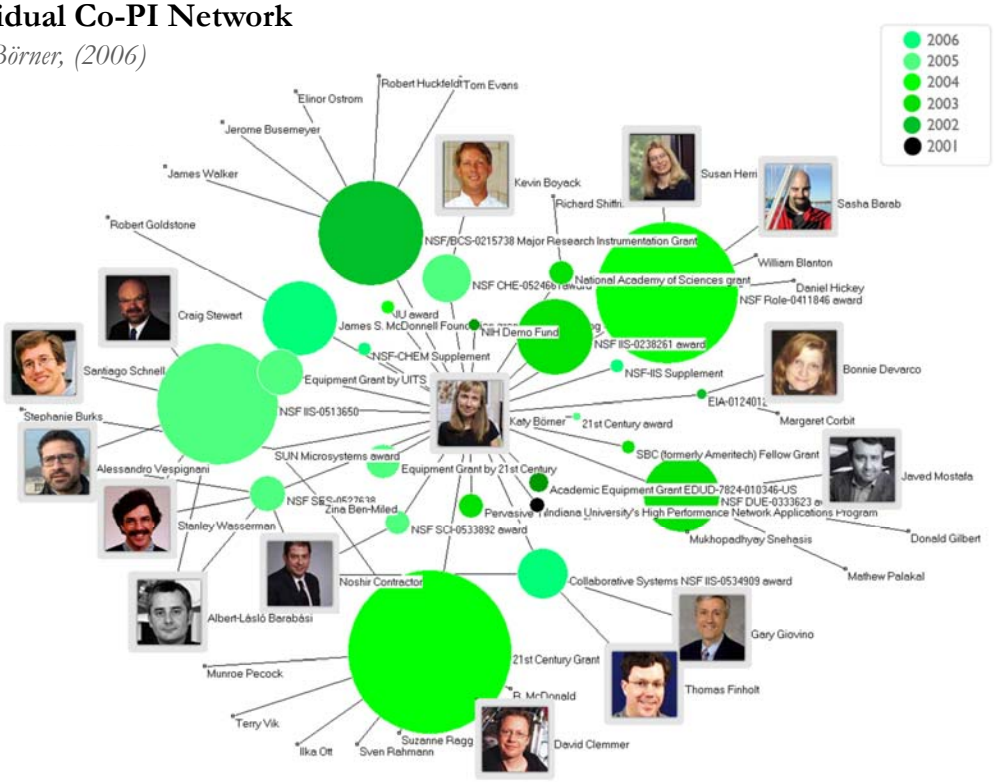
	Micro/Individual (1-100 records)	*Meso/Local (101–10,000 records)*	*Macro/Global (10,000 < records)*
Statistical Analysis/Profiling	Individual person and their expertise profiles	Larger labs, centers, universities, research domains, or states	All of NSF, all of USA, all of science.
Temporal Analysis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



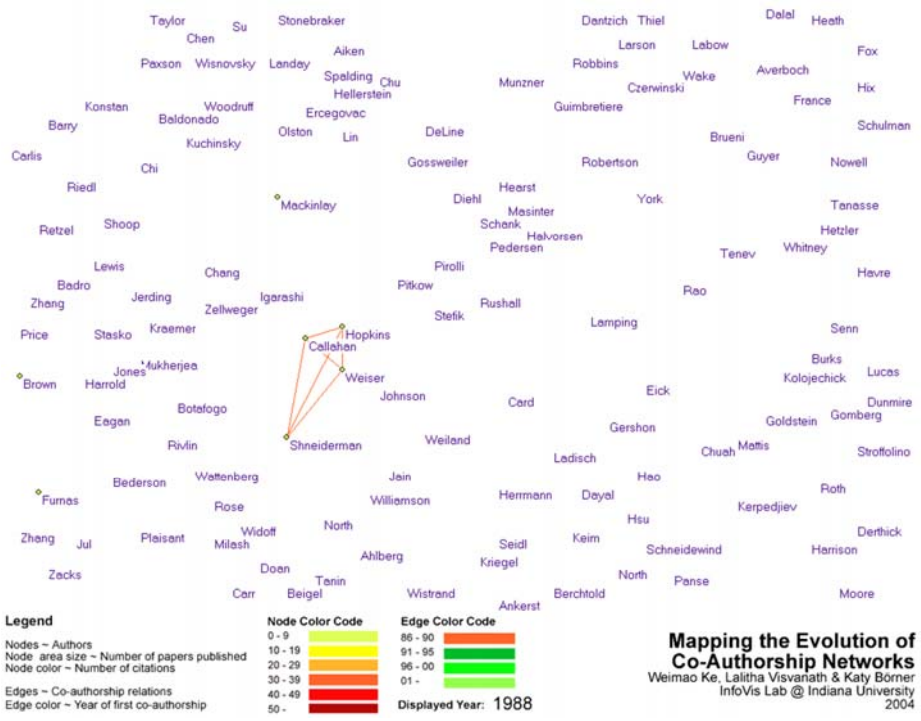
Individual Co-PI Network

Ke & Börner, (2006)



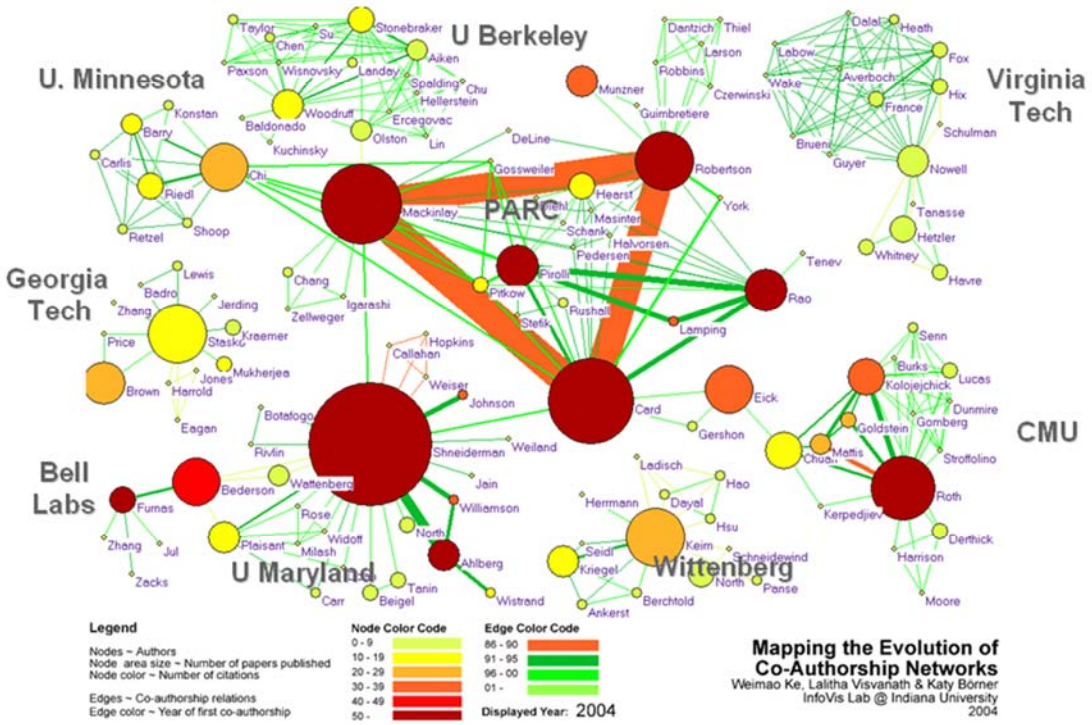
Mapping the Evolution of Co-Authorship Networks

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



Mapping the Evolution of Co-Authorship Networks

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

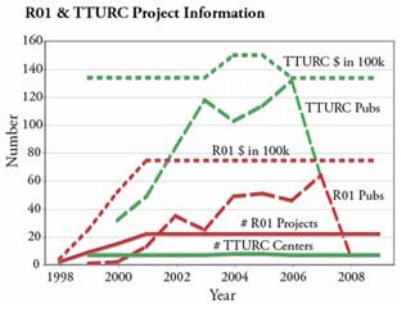


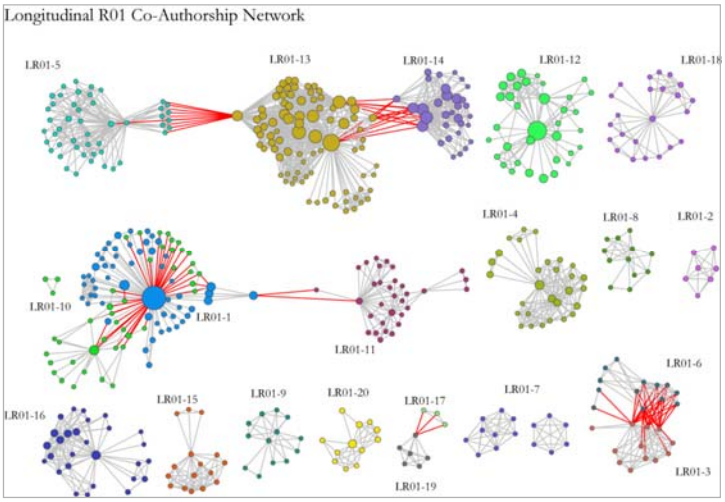
Mapping Transdisciplinary Tobacco Use Research Centers Publications

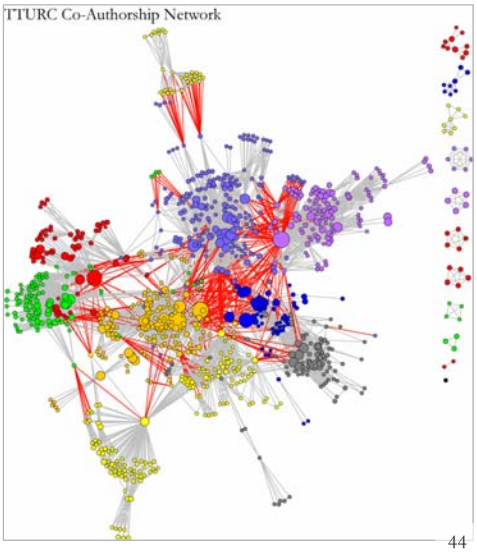
Compare R01 investigator based funding with TTURC Center awards in terms of number of publications and evolving co-author networks.

Zoss & Börner, forthcoming.

Supported by NIH/NCI Contract HHSN261200800812

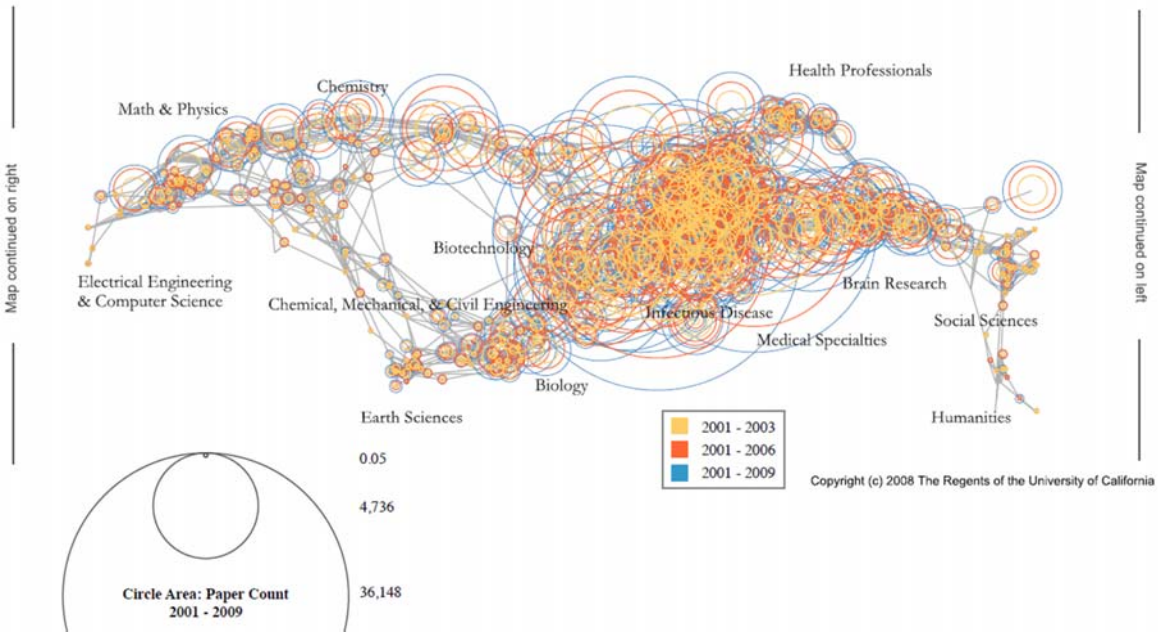




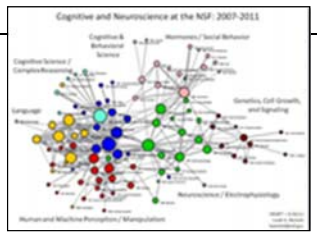


MEDLINE Publication Output by The National Institutes of Health (NIH) Using Nine Years of ExPORTER Data

Katy Börner, Nianli Ma, Joseph R. Biberstine, Cyberinfrastructure for Network Science Center, SLIS, Indiana University, Robin M. Wagner, Rediet Berhane, Hong Jiang, Susan E. Ivey, Katrina Pearson and Carl McCabe, Reporting Branch, Division of Information Services, Office of Research Information Systems, Office of Extramural Research, Office of the Director, National Institutes of Health (NIH), Bethesda, MD.



Sci² Tool Usage at National Science Foundation

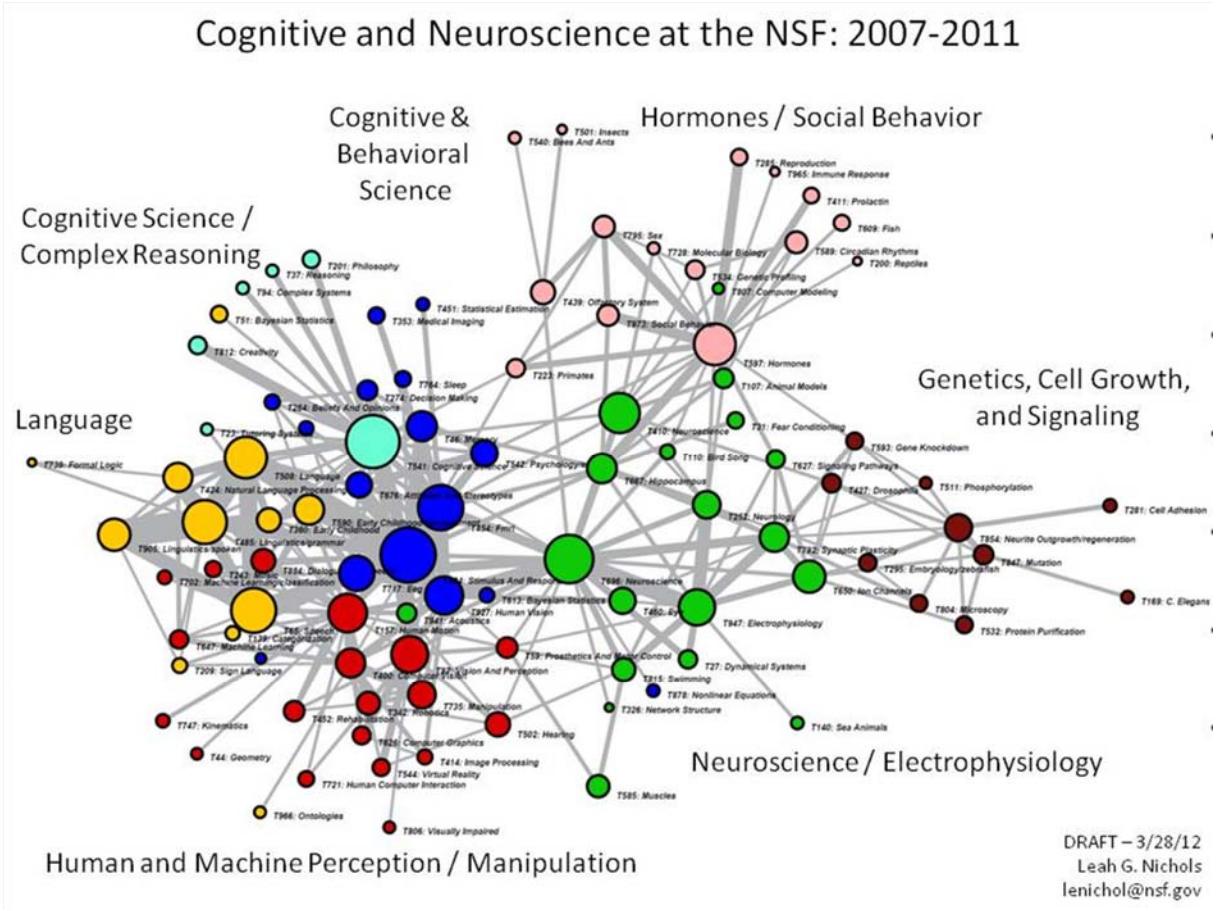


Topic co-occurrence network of the 2885 cognitive and neuroscience NSF projects funded between 2007 and 2011. Statistical text mining (Topic model) was used to identify topics from NSF awards and proposals from 2000-2011.

Each award is tagged with up to 4 topics. Lines represent the co-occurrence of the connected topics within an award(s). The nodes are scaled by number of awards (max = 355) and the lines are scaled on number of co-occurrences (max =91). The node colors differentiate the nodes via the level-0 Blondel communities.

This is … an **entirely new way of characterizing and understanding the NSF portfolio**. This is in part because this enables **analysis of the content of the awards/proposals independent of the institutional structure.** One can quickly identify ALL of the Cog/Neuro awards throughout the entire NSF portfolio – so it captures research in all of the unexpected institutional places. This method also allows one to **easily identify areas of parallel or potentially collaborative research being funded by different institutional structures** and … to identify potential areas for advancing science by facilitating collaborations.

Leah G. Nichols, NSF



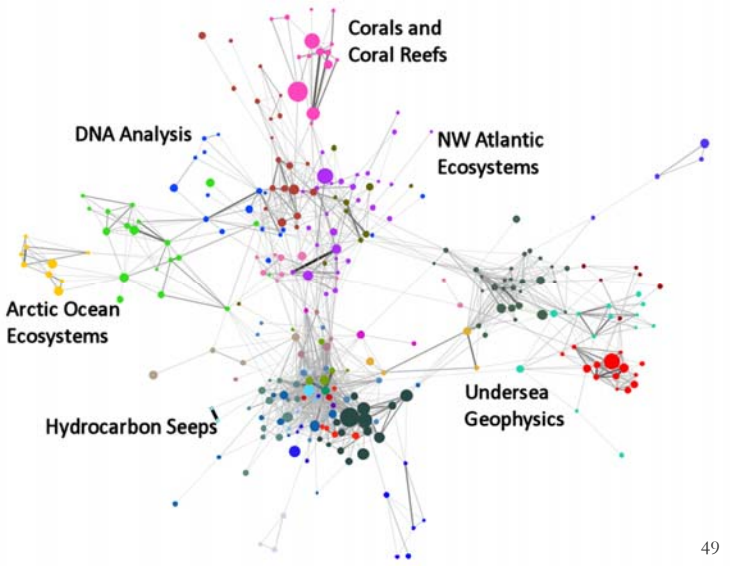
Sci² Tool Usage at National Institutes of Health

Sci2 Tool now supports Web services and serves as a visual interface to publically available NIH RePORT Expenditure and Results RePORTER)/ RePORTER data provided by NIH.

Sci² Tool Usage at the National Oceanic and Atmospheric Administration (NOAA)

Co-author network generated from publications supported by NOAA's Office of Ocean Exploration and Research (OER). Nodes are sized based on the number of publications produced and colored to highlight clustering. Edges are sized and colored based on the number of collaborations between authors.

For details, see "Visualizing Networks of Scientific Research"
by Chris Belter
http://www.infotoday.com/online/may12/Belter-Visualizing-Networks-of-Scientific-Research.shtml



Sci² Tool Usage at James S. McDonnell Foundation

How did cognitive neuroscience of attention emerge from neurobiology and psychology, 1980–2005? Author co-citation analysis and Pfnet is used to **trace prospectively the development of the field from its precursor disciplines**: cognitive psychology, single cell neurophysiology, neuropsychology, and evoked potential research.

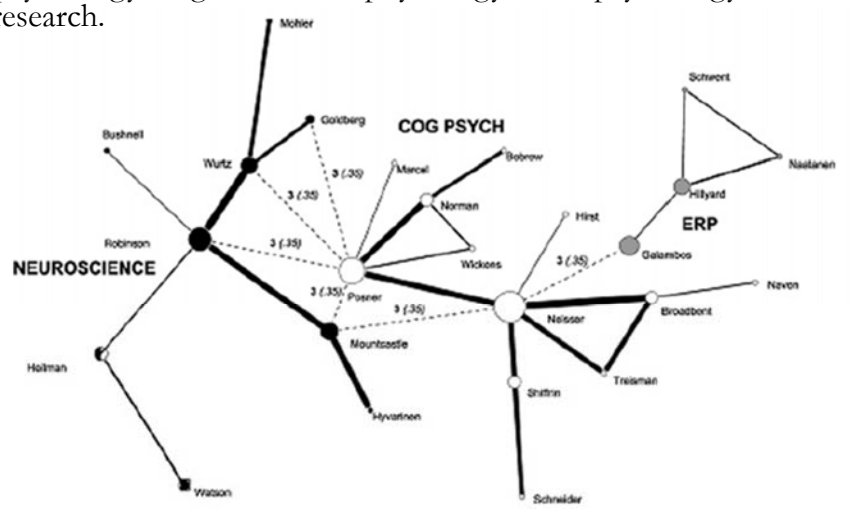


Fig. 1 In the 1980 net, neuroscience (*black* nodes and *black–white* nodes) and cognitive psychology (*white* nodes) develop as clusters with high internal co-citation rates. ERP (*grey* nodes) develops later in net construction. These clusters are connected by secondary edges at very low levels of co-citation

By 1990 a distinct cognitive neuroscience specialty cluster emerges, dominated by authors engaged in brain imaging research.

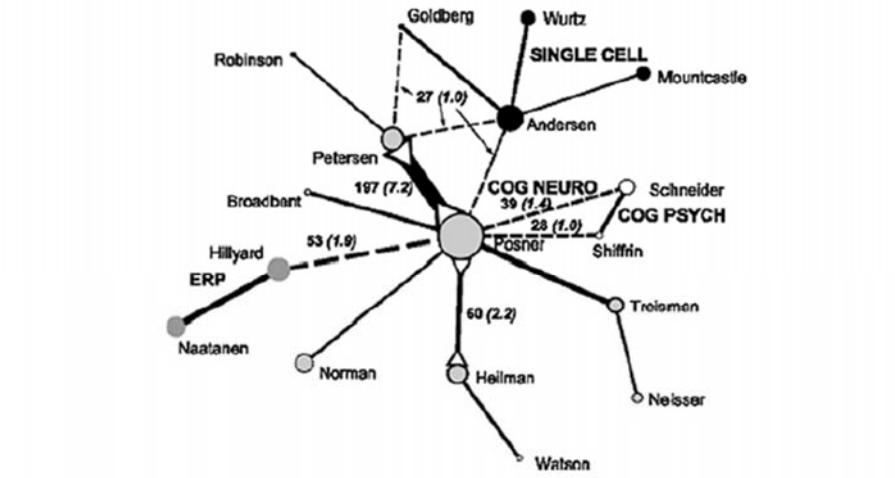


Fig. 5 The strongest link in the 1995 net is a primary edge linking Posner and Petersen. ERP and single cell neurophysiology are linked to cognitive neuroscience cluster by secondary edges

Bruer, John T. (2010). Can we talk? How the cognitive neuroscience of attention emerged from neurobiology and psychology, 1980.2005. Scientometrics, 83(3), 751-764. http://ivl.cns.iu.edu/km/tools/2010-bruer-scientometrics.pdf

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.

Mixed-Indicators Model for Identifying Emerging Research Areas

Guo, Hanning, Scott B. Weingart, and Katy Börner. 2011. Scientometrics 89 (1): 421-435.

Three indicators are combined:

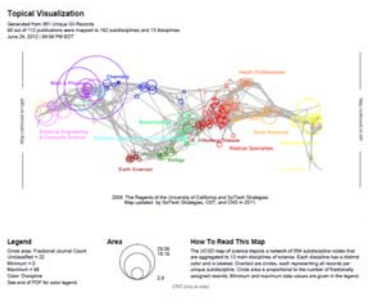
- sudden increases in the frequency of specific words,
- number and speed by which new authors are attracted to an emerging area,
- changes in the interdisciplinarity of cited references.

The model is validated using four emerging research areas and two datasets:
"RNAi", "Nano", "h-Index", and "Impact Factor" research using papers published in the *PNAS* (1982–2009) and *Scientometrics* (1978–2009).

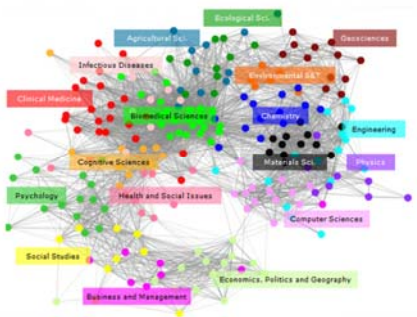
Results:

- Keyword bursts occurred 8 years later for "Nano*," 7 years later for "RNAi" and only 1 year later for "h-Index." *(different from full text words)*
- Appearance of new authors always signifies the beginning of an emerging area.
- In "Nano*," "RNAi" and "h-Index" datasets, a sudden increase in the diversity of cited references occurred with the appearance of new authors simultaneously. The correlation between increasing new authors and diversity of cited references suggests that new authors are coming from diverse established areas rather than some already nascent cohort with a pre-existing body of research.

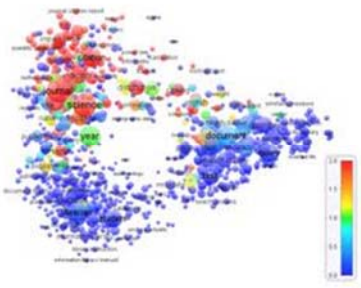
Aligning Science Basemaps using the Sci2 Tool



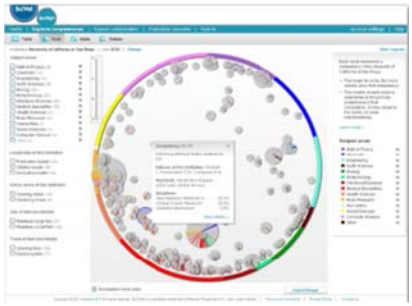
UCSD Map



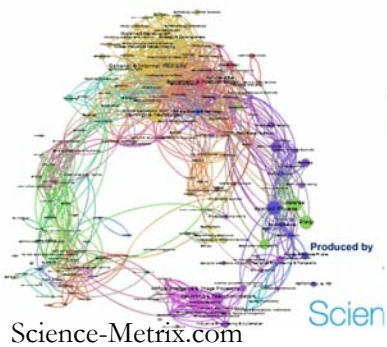
Loet et al science maps ISI categories



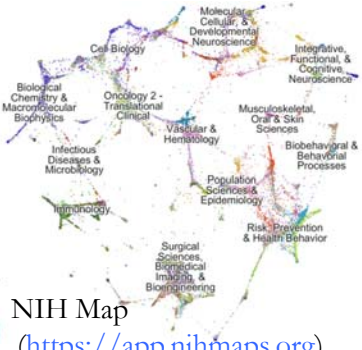
http://vosviewer.com



Elsevier's SciVal Map



Science-Metrix.com



NIH Map
(https://app.nihmaps.org)

References

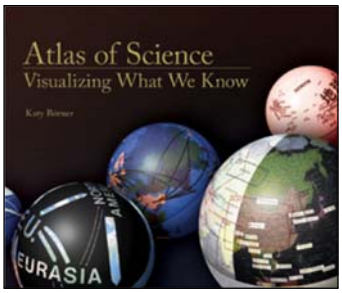
Börner, Katy, Chen, Chaomei, and Boyack, Kevin. (2003). **Visualizing Knowledge Domains.** In Blaise Cronin (Ed.), *ARIST*, Medford, NJ: Information Today, Volume 37, Chapter 5, pp. 179-255.
http://ivl.slis.indiana.edu/km/pub/2003-borner-arist.pdf

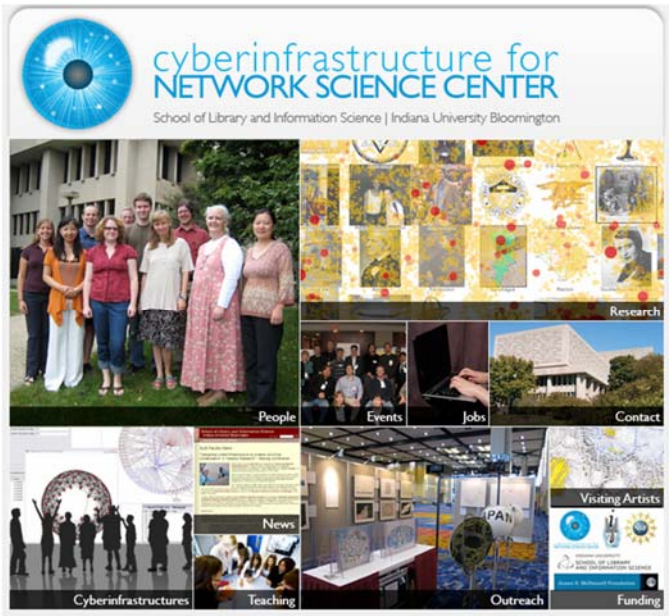
Shiffrin, Richard M. and Börner, Katy (Eds.) (2004). **Mapping Knowledge Domains**. *Proceedings of the National Academy of Sciences of the United States of America*, 101(Suppl_1).
http://www.pnas.org/content/vol101/suppl_1/

Börner, Katy, Sanyal, Soma and Vespignani, Alessandro (2007). **Network Science.** In Blaise Cronin (Ed.), *ARIST*, Information Today, Inc., Volume 41, Chapter 12, pp. 537-607.
http://ivl.slis.indiana.edu/km/pub/2007-borner-arist.pdf

Börner, Katy (2010) **Atlas of Science**. MIT Press.
http://scimaps.org/atlas

Scharnhorst, Andrea, Börner, Katy, van den Besselaar, Peter (2012) **Models of Science Dynamics**. Springer Verlag.





All papers, maps, tools, talks, press are linked from http://cns.iu.edu

CNS Facebook: http://www.facebook.com/cnscenter
Mapping Science Exhibit Facebook: http://www.facebook.com/mappingscience