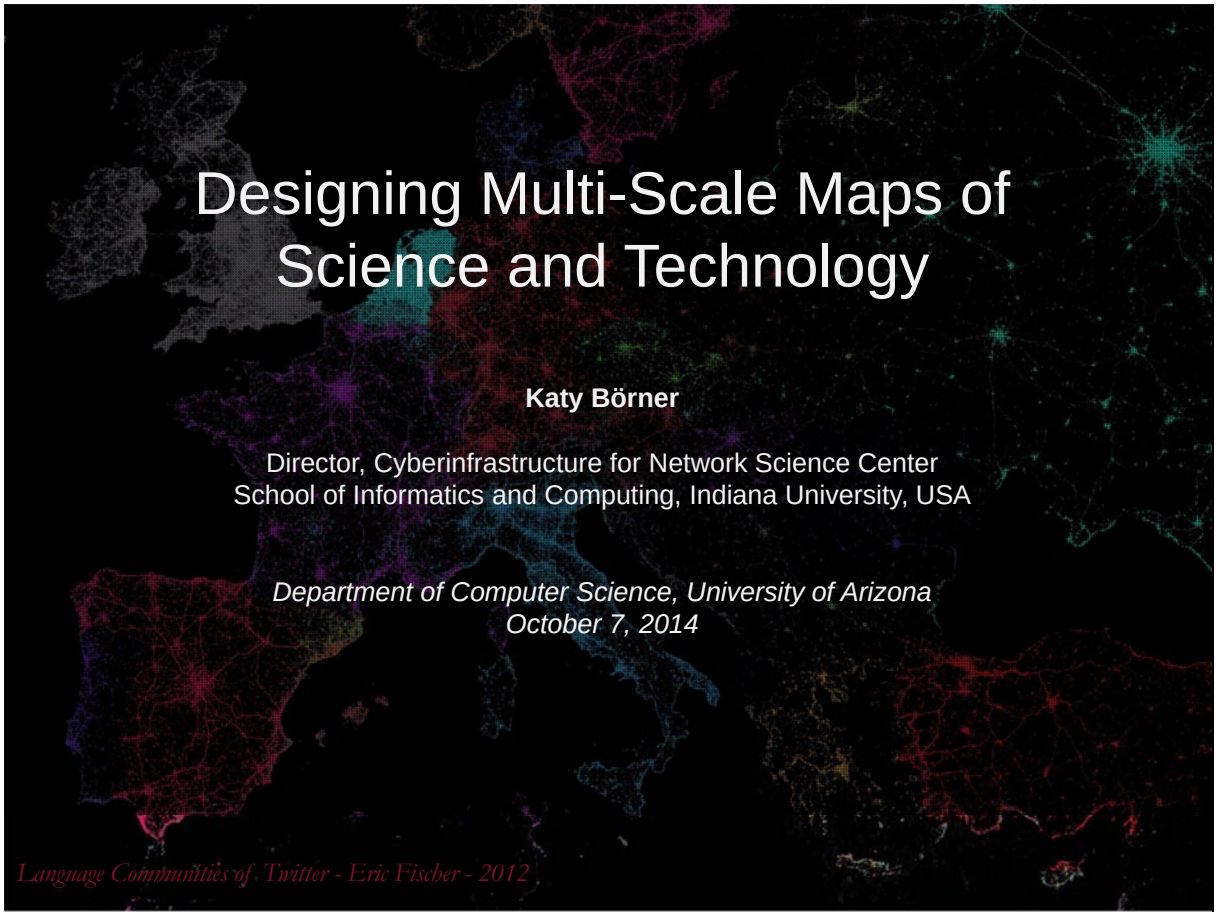


Designing Multi-Scale Maps of Science and Technology

Katy Börner

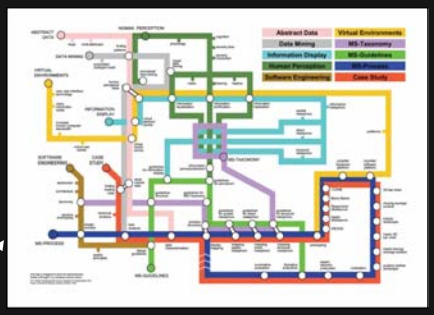
Director, Cyberinfrastructure for Network Science Center
School of Informatics and Computing, Indiana University, USA

Department of Computer Science, University of Arizona
October 7, 2014

Overview + Context

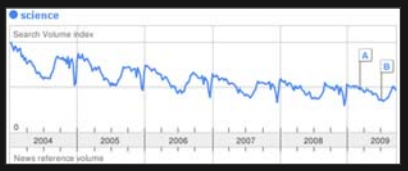
Terra bytes of data

Descriptive & Predictive Models



Find your way

Find collaborators, friends



Identify trends

Different Levels of Abstraction/Analysis

Macro/Global
Population Level

Meso/Local
Group Level

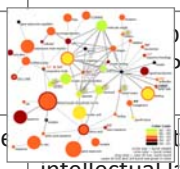
Micro
Individual Level

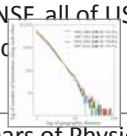
Type of Analysis vs. Level of Analysis

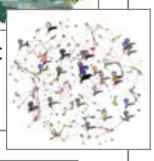
	Micro/Individual (1-100 records)	***Meso/Local (101–100,000 records)***	***Macro/Global (100,000 < records)***
Statistical Analysis/Profiling	Individual person and their expertise profiles	Larger labs, centers, universities, research domains, or states	All of NSF, all of USA, all of science.
Temporal 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

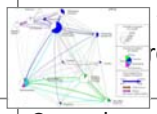
Type of Analysis vs. Level of Analysis

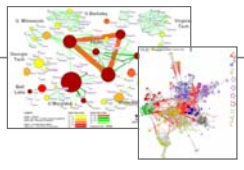
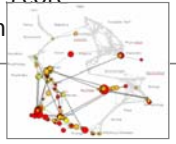
	Micro/Individual (1-100 records)	***Meso/Local (101–100,000 records)***	***Macro/Global (100,000 < records)***
Statistical Analysis/Profiling	Individual person and their expertise profiles	Larger labs, centers, universities, research domains, or states	All of NSF, all of USA, all of sc
Temporal Analysis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

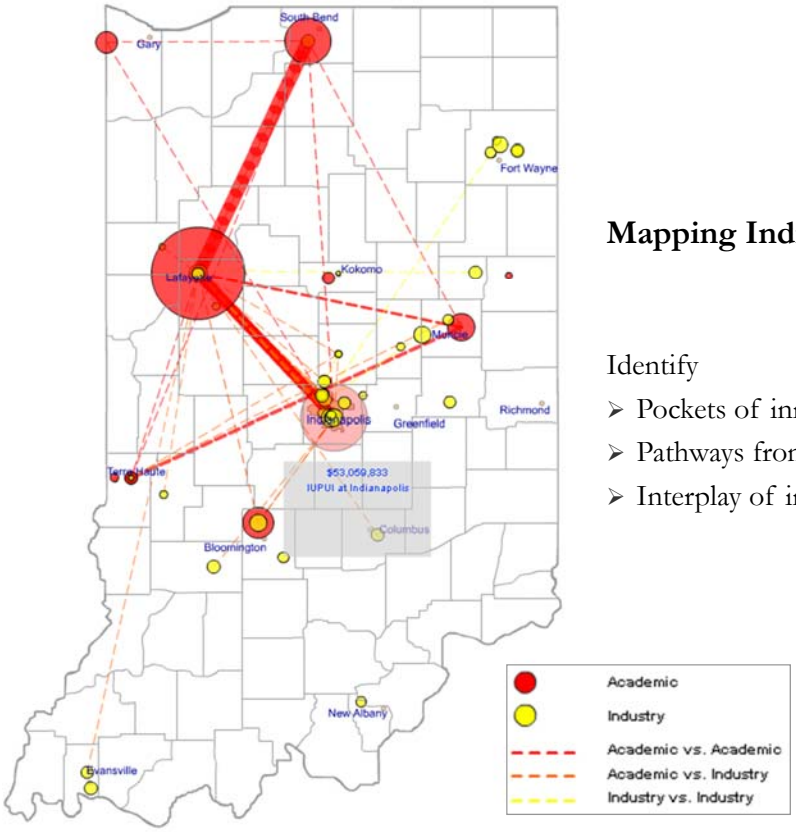










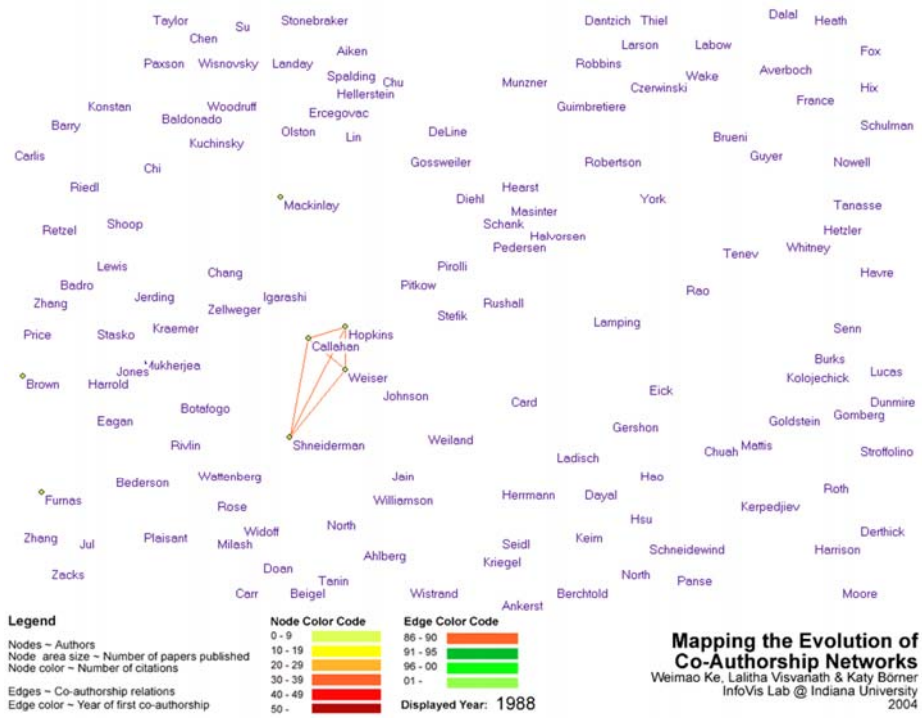
Mapping Indiana's Intellectual Space

Identify

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

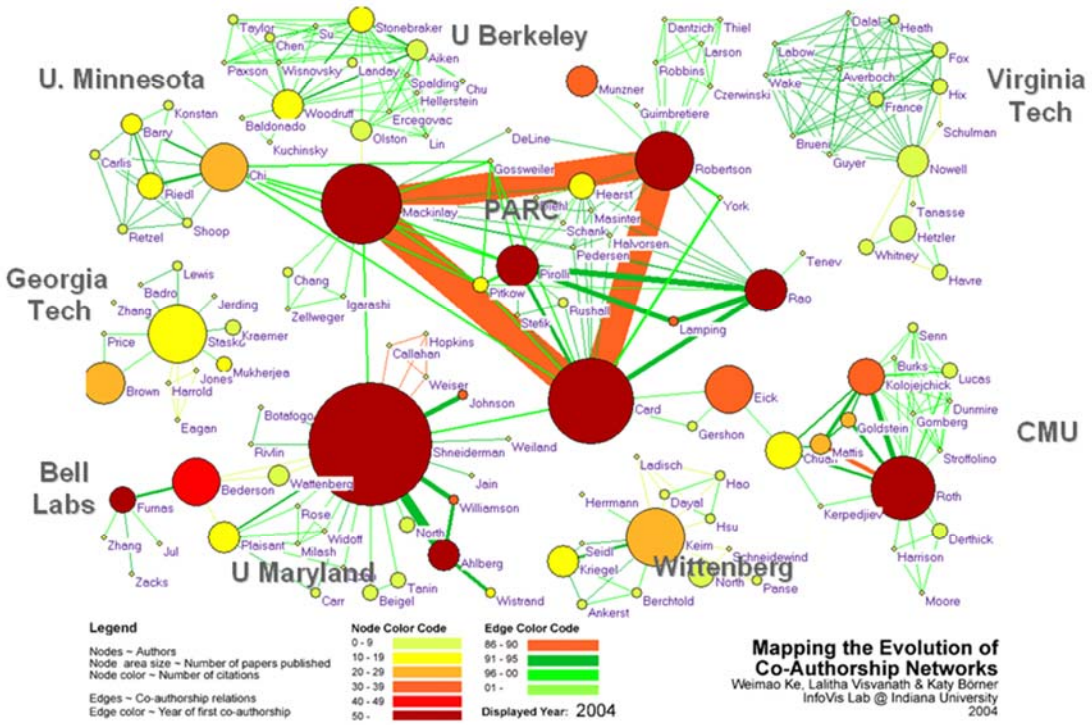
Mapping the Evolution of Co-Authorship Networks

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

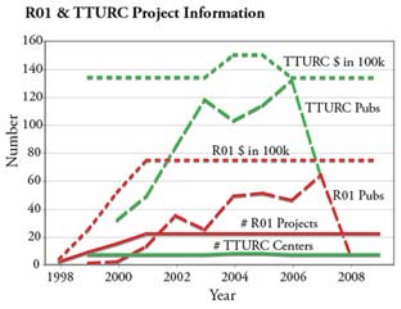


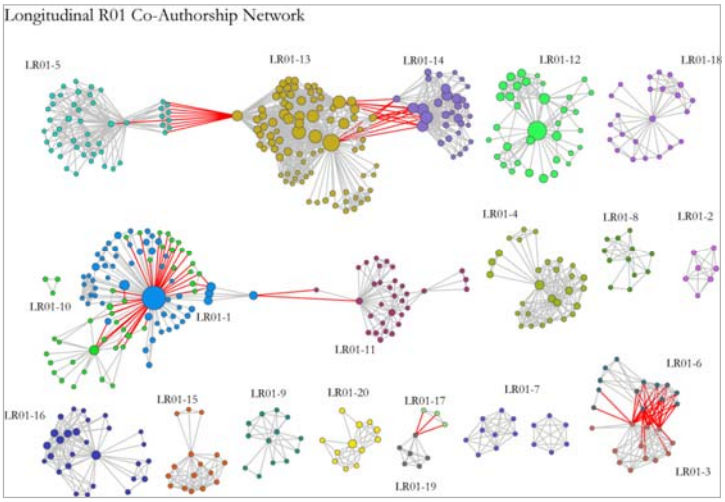
Mapping Transdisciplinary Tobacco Use Research Centers Publications

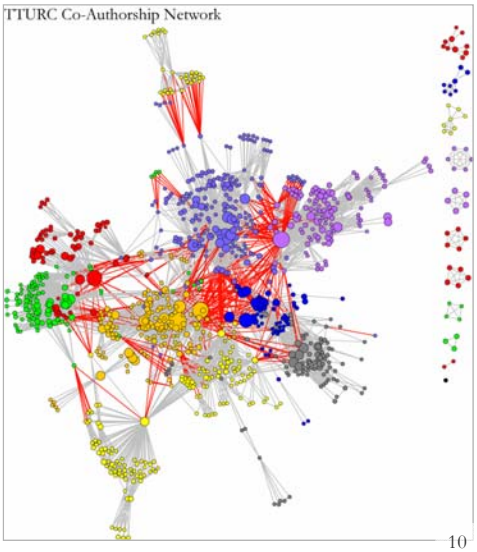
Compare R01 investigator based funding with TTURC Center awards in terms of number of publications and evolving co-author networks.

Stipelman, Hall, Zoss, Okamoto , Stokols & Börner, 2014

Supported by NIH/NCI Contract HHSN261200800812





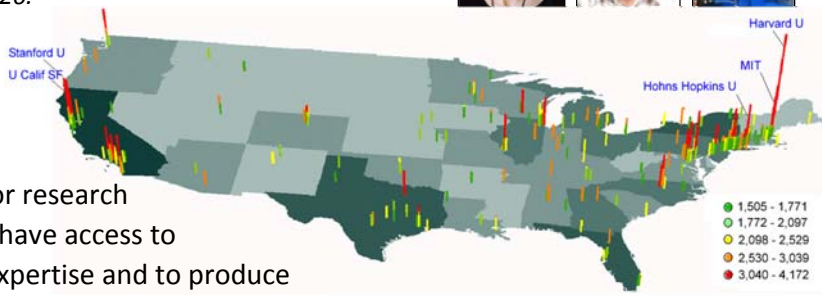


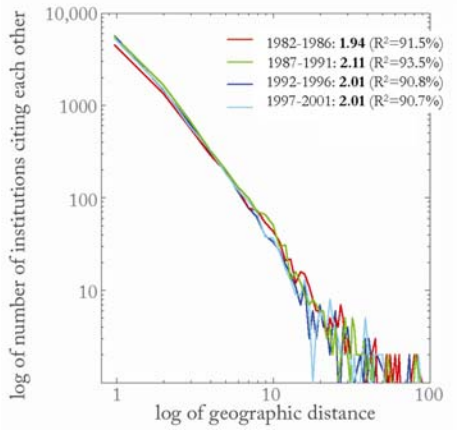
Spatio-Temporal Information Production and Consumption of Major U.S. Research Institutions

Börner, Penumarthy, Meiss & Ke (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.

The Global 'Scientific Food Web'

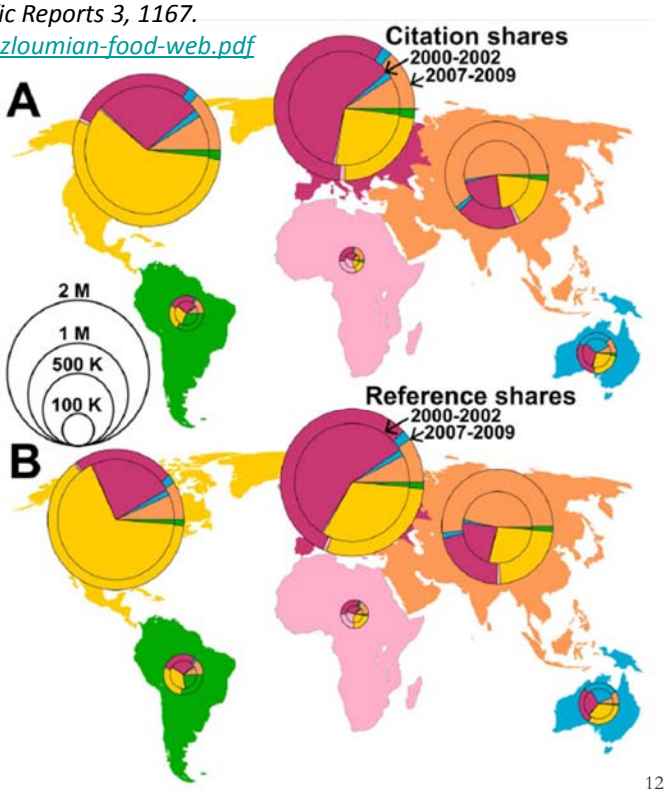
Mazloumian, Amin, Dirk Helbing, Sergi Lozano, Robert Light, and Katy Börner. 2013. "Global Multi-Level Analysis of the 'Scientific Food Web'". Scientific Reports 3, 1167.
http://cns.iu.edu/docs/publications/2013-mazloumian-food-web.pdf

Contributions:

Comprehensive global analysis of scholarly knowledge production and diffusion on the level of continents, countries, and cities.

Quantifying knowledge flows between 2000 and 2009, we identify global sources and sinks of knowledge production. Our knowledge flow index reveals, where ideas are born and consumed, thereby defining a global 'scientific food web'.

While Asia is quickly catching up in terms of publications and citation rates, we find that its dependence on knowledge consumption has further increased.



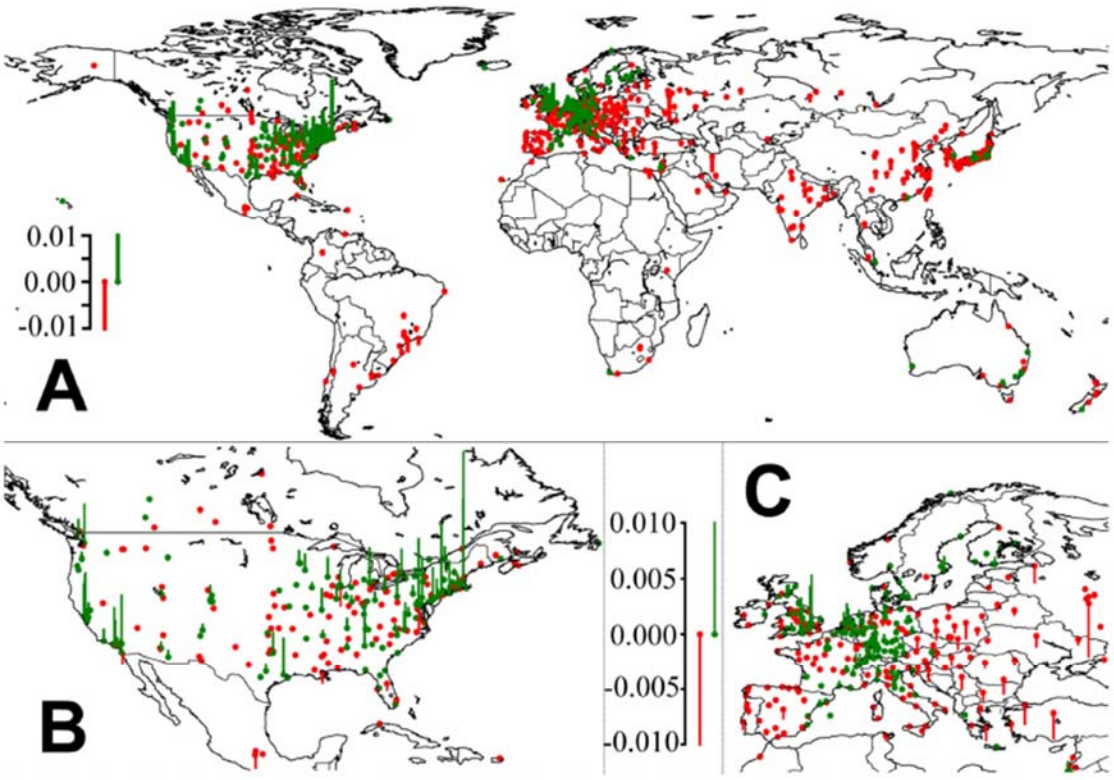


Figure 2 | World map of the greatest knowledge sources and sinks, based on our scientific fitness index. Green bars indicate that the number of citations received is over-proportional, red that the number of citations received is lower than expected (according to a homogeneous distribution of citations over all cities that have published more than 500 papers). It can be seen that most scientific activity occurs in the temperate zone. Moreover, areas of high fitness tend to be areas that are performing economically well (but the opposite does not hold).

Type of Analysis vs. Level of Analysis

	Micro/Individual (1-100 records)	***Meso/Local (101–100,000 records)***	***Macro/Global (100,000 < 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 sc
Temporal Analysis (When?)	Funding portfolio of one individual	ic bursts in PNAS	113 Years of Physics Research
Geospatial Analysis (Where?)	Career trajectory of one individual	intellectual la	PNAS publications
Topical Analysis (What?)	Base knowledge from s.	research	VxOrd/Topic NIH funding
Network Analysis (With Whom?)	NSF Co-PI network of one	Co-author network	NIH's core com

Students. Maps of science can help students gain an overview if a particular knowledge domain, identify major research areas, experts, institutions, grants, publications, patents, citations, and journals as well as their interconnections, see the influence of certain theories, and gain a global picture of the domain.
Researchers. Science maps can be used to ease access to research results, relevant funding opportunities, and potential collaborators inside and outside the fields of inquiry, and to detect social networks and invisible colleges.
Grant Agencies/R&D Managers. While maps of science cannot substitute for informed peer evaluation or expert panels, they can be used as tools to monitor (long-term) money flow and research developments, evaluate funding strategies for different programs, make informed decisions on project durations, and study funding patterns. In addition, they can also be used to identify the impact of research funding programs, scientific frontiers the dynamics (speed of growth, diversification) of scientific fields, and complementary capabilities.
Industry/National Security Agency. Maps of science can be utilized to gain access to major scientific results, knowledge carriers, etc. Information on needed technologies could be incorporated into maps, facilitating industry pulls for specific directions of research.
Data Providers. Maps provide unique visual interfaces to digital libraries.

The UCSD Map of Science and Classification System

Early Maps of the World VERSUS Early Maps of Science

3D
Physically-based
Accuracy is measurable
Trade-offs have more to do with granularity
2-D projections are very accurate at local levels
Centuries of experience
Geo-maps can be a template for other data

n-D
Abstract space
Accuracy is difficult
Trade-offs indirectly affect accuracy
2-D projections neglect a great deal of data
Decades of experience
Science maps can be a template for other data

Kevin W. Boyack, UCGIS Summer Meeting, June, 2009

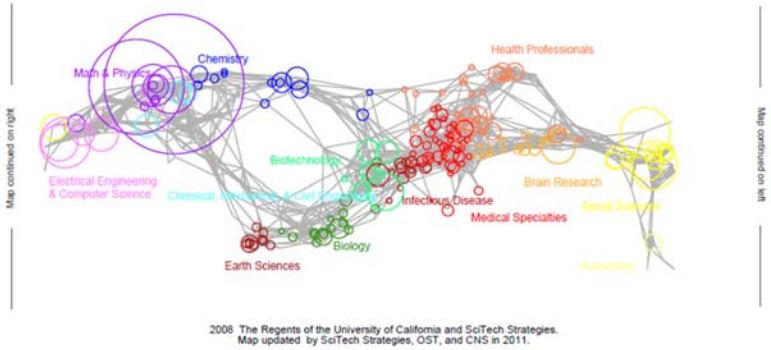
17

Design and Update of a Classification System: The UCSD Map of Science

Katy Börner[1,2*], Richard Klavans[3], Michael Patek[3], Angela M. Zoss[1], Joseph R. Biberstine[1], Robert P. Light[1], Vincent Larivière[1,4,5], Kevin W. Boyack[6]

1 Cyberinfrastructure for Network Science Center, School of Library and Information Science, Indiana University, Bloomington, Indiana, United States of America, **2** Royal Netherlands Academy of Arts and Sciences (KNAW), Amsterdam, The Netherlands, **3** SciTech Strategies, Inc., Berwyn, Pennsylvania, United States of America, **4** École de Bibliothéconomie et des Sciences de L'information, Université de Montréal, Montréal, Canada, **5** Observatoire des Sciences et des Technologies (OST), Centre Interuniversitaire de Recherche sur la Science et la Technologie (CIRST), Université du Québec à Montréal, Montréal, Canada, **6** SciTech Strategies, Inc., Albuquerque, New Mexico, United States of America

(2012) PLoS ONE 7(7): e39464.
http://sci.cns.iu.edu/ucsdmap



Design:
How was the UCSD map of science made?

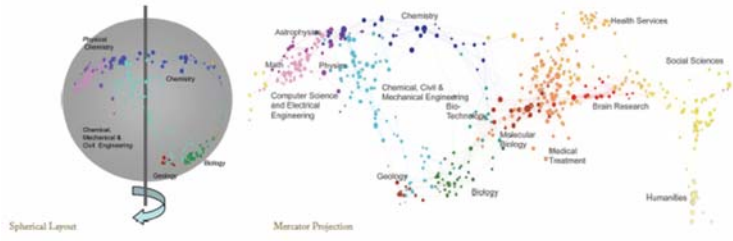
- Original Map (5 years of Scopus and WoS)
- Initial Update (5 years + Scopus)
- Full Update (10 years of Scopus and WoS)

Original Map (5 years of Scopus and WoS)

Data: The original classification and map use 7.2 million papers and their references from Elsevier's Scopus (about 15,000 source titles, 2001–2005) and Thomson Reuters' Web of Science (WoS) Science, Social Science, Arts & Humanities Citation Indexes (about 9,000 source titles, 2001–2004)–about 16,000 unique source titles

Similarity Metric: Combination of bibliographic coupling and keyword vectors.
See Supplement 1 in http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0039464 for full details.

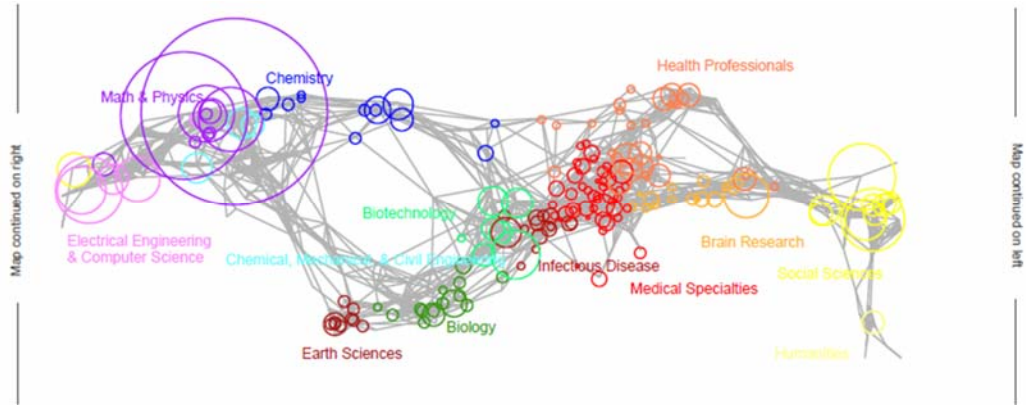
Layout: The 554 subdisciplines were laid out on the surface of a sphere; the spheric layout is then flattened using a Mercator projection to create a two-dimensional version of the map. Clusters are further aggregated into 13 main scientific disciplines that are labeled and color coded in a metaphorical way, e.g., Medicine is blood red and Earth Sciences are brown as soil.



Original Map (5 years of Scopus and WoS)

Data Overlays:
Each node is labeled and has an extensive list of key phrases as metadata, which can be used to "science locate" nonjournal data, such as patents or grants. That is, key phrases from each patent or grant (titles and abstracts) are extracted; fractional assignment to map nodes proceeds by matching the associated metadata. Thus, each grant or patent is fractionally assigned to multiple nodes. Adding the fractions allows for the number of grants, dollars by agency, or patents associated with each node to be computed.

Problem: As time passes, new journals are created, e.g., PLoS, that cannot be mapped.



Initial Update (5 years + Scopus)
by Klavans & Boyack

Data: In June 2009, 7,464 new source titles (2006–2008) from Scopus were added to to the existing category structure.
Process: Identifying all new journals that were not in the existing classification system, and assign each new journal to one of the existing categories by counting the numbers of times journals in each category were referenced by the articles in the new journals. Each journal was assigned to the category that it referenced the most, as long as it cited articles within that cluster a minimum of 10 times.
Result: Update increased the number of Scopus journals in the classification system by 47%, this only accounted for a 13% increase in the number of articles.

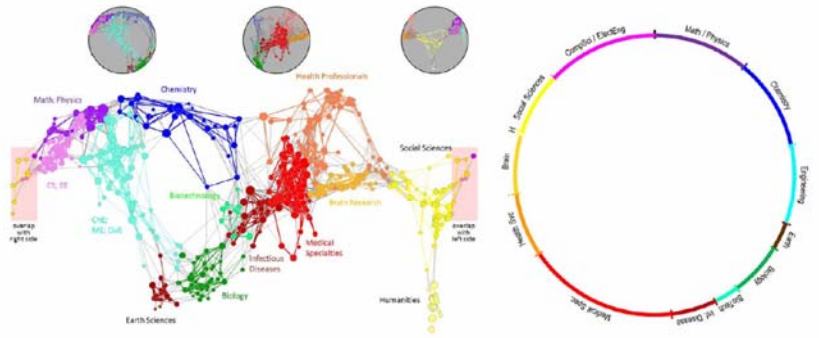


Figure 1. Visualizations of the UCSD Map: 2D Mercator projection (left) with three 3D spherical insets (top), 1D circular map (right).
Note that the left hand side of the Mercator map connects to the right hand side.
doi:10.1371/journal.pone.0039464.g001

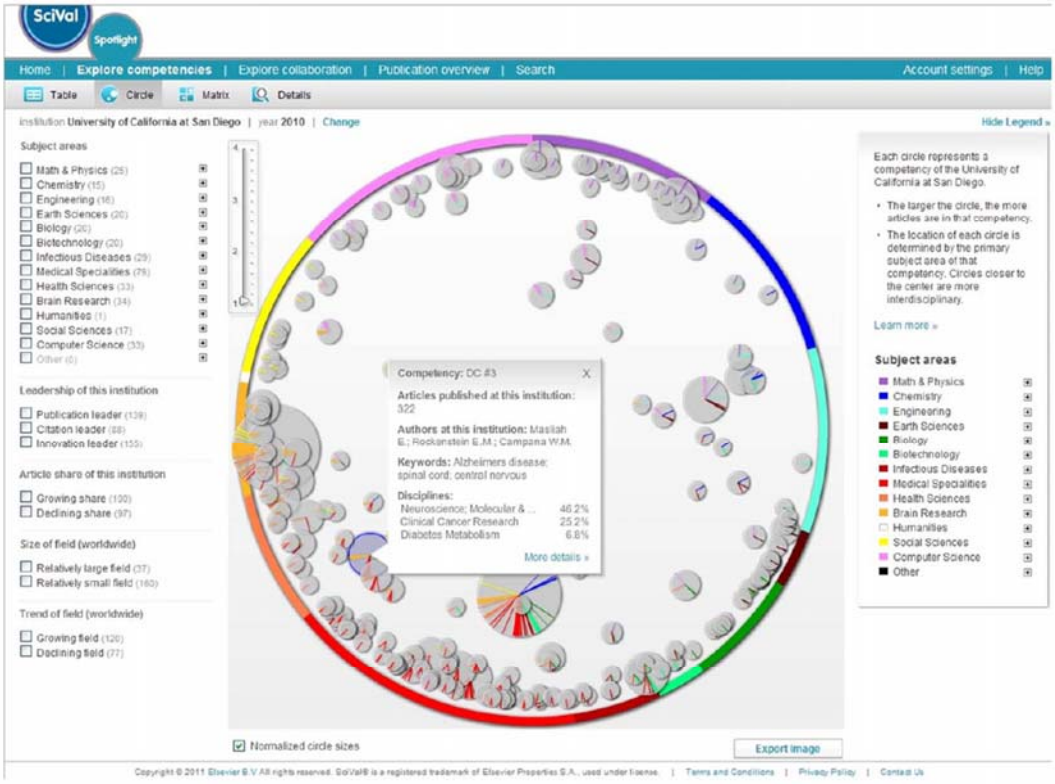


Figure 2. SciVal Spotlight map of one institution, here UCSD, showing institutional competencies. Each node within the circle map represents a competency (a group of linked topics), and is positioned at the average location of its articles. Node size reflects the number of articles. Coloured rays within each node show the disciplines that contribute to the competency.
doi:10.1371/journal.pone.0039464.g002

Full Update (10 years of Scopus and WoS)

by Börner, Klavans, Patek, Zoss, Biberstine, Light, Larivière, Boyack

Desirable features for a map of science classification system:

1. Use highest quality/coverage paper-level data to generate the science map classification system. Using journal level data or highly cited papers exclusively lead to distortions [22].
2. Employ advanced dimensionality reduction techniques to map a high dimensional sematic space to a two-dimensional map that preserves the most important data structures [23].
3. Select a clustering and layout that has easy to read, distinct clusters, e.g., subdisciplines, which have about the same number of records, are disjoint (i.e., they do not overlap or occlude one other), and have meaningful labels to ease data interpretation and communication. The map must match the typical viewer's mental model of the domain.
4. Use graphic design (color, shape, size coding) and legend that can be understood by a large audience—map must empower users to form new hypotheses and get new answers.
5. Support interactivity, e.g., zoom, filter, details on demand [24]. Multi-level maps, e.g., two-levels comprising subdisciplines aggregated into disciplines, support multi-level studies.
6. Define a mapping process to classify new data and overlay it onto the map, e.g., journals based on journal names and other records, e.g., patents, funding data based on keywords. As users have a hard time with fractional associations/counting, each record should be associated with one or few subdisciplines.

Desirable features for a map of science classification system (cont.):

7. The science map and classification system should be easy to update to capture the continuously evolving structure of science. Computational workflow should be well documented so that is easy to understand in principle and can be replicated by other experts. Updates should preserve the main structure of the map as much as possible.
8. Alignment and comparison of any new science map and classification with commonly used science classifications (e.g., classifications used by Thomson Reuters' databases, Elsevier's Scopus, the Library of Congress, Universal Decimal Classification)and the translation of major ontologies into different languages (Science-Metrix, [25]).

Full Update (10 years of Scopus and WoS)
by Börner, Klavans, Patek, Zoss, Biberstine, Light, Larivière, Boyack

Data: The updated map and classification adds six years (2005–2010) of WoS data and three years (2006–2008) from Scopus to the existing category structure–increasing the number of source titles to about 25,000.

Process:
For each of the 4,021 new journals, we counted the number of citations to/from papers published in that journal to/from each subdiscipline of the original map. This yielded for each journal an outgoing and incoming citation count for each subdiscipline of the original map. To account for the fact that some subdisciplines publish more papers than others and that, thus, the probability of citing and being cited by these subdisciplines is greater than for smaller ones, we normalized each of these citation counts by the total number of papers published among all journals assigned (even only fractionally) to that subdiscipline. The top subdiscipline citing/cited was then assigned to these new journals.

Multidisciplinary journals: PLOS ONE and SCHWEIZERISCHE MEDIZINISCHE WOCHENSCHRIFT (Swiss Medical Weekly) have the highest combined relative importance across sub-disciplines yet were assigned to one subdiscipline.
To further simplify the 2010 UCSD map, all multi-assigned journals were examined and only 34 were kept, among them *Science, Nature, the Lancet, British Medical Journal,* and *Journal of the American Medical Association.*

Results:
A comparison of the original 5-year and the new 10-year maps and classification system show (i) an increase in the total number of journals that can be mapped by 9,409 journals (social sciences had a 80% increase, humanities a 119% increase, medical (32%) and natural science (74%)), (ii) a simplification of the map by assigning all but five highly interdisciplinary journals to exactly one discipline, (iii) a more even distribution of journals over the 554 subdisciplines and 13 disciplines when calculating the coefficient of variation, and (iv) a better reflection of journal clusters when compared with paper-level citation data. When evaluating the map with a listing of desirable features for maps of science, the updated map is shown to have higher mapping accuracy, easier understandability as fewer journals are multiply classified, and higher usability for the generation of data overlays, among others.

To our knowledge, this is the first time that a widely used map of science was updated

Full Update (10 years of Scopus and WoS)
by Börner, Klavans, Patek, Zoss, Biberstine, Light, Larivière, Boyack

#Journals	5-Year Map	10-Year Map	Difference
WoS	9,499	13,520	4,021
Scopus	14,789	22,253	7,464
WoS & Scopus	15,849	25,258	9,409

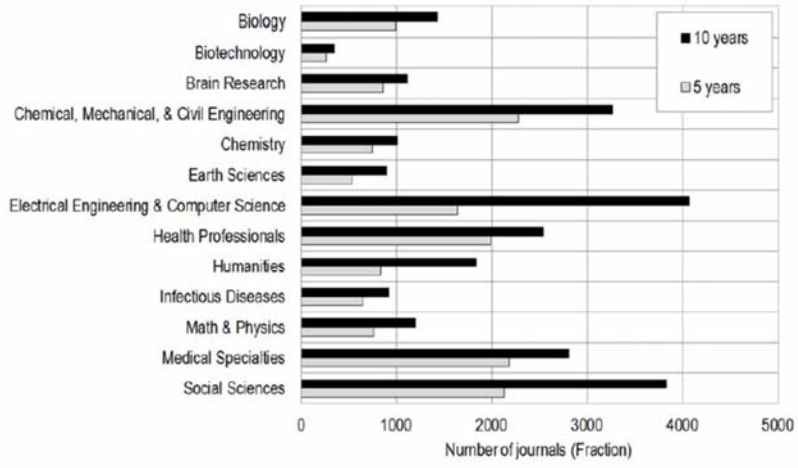


Figure 4. Number of journals per discipline for 5-year (grey) and 10-year (black) UCSD science map.
doi:10.1371/journal.pone.0039464.g004

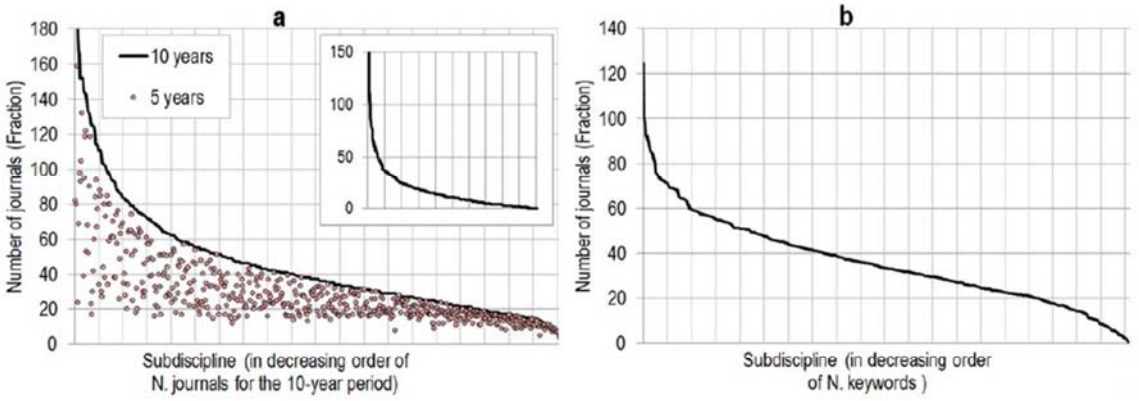


Figure 5. Number of journals per subdiscipline for 5-year (grey/red circles) and 10-year (black line) UCSD science map. Inset: distribution of the gain in number of journals for each subspecialty (a). Number of (fractionally assigned) terms per 554 subdisciplines (b). doi:10.1371/journal.pone.0039464.g005

Full Update (10 years of Scopus and WoS)
by Börner, Klavans, Patek, Zoss, Biberstine, Light, Larivière, Boyack

Deployment: The UCSD map of science data is available at http://sci.cns.iu.edu/ucsdmap/

Data
The 2010 UCSD map of science and classification system covering 10 years (2001-2010) of Web of Science data and 8 years (2001-2008) of Scopus data with subdiscipline assignments by SciTech Strategies.

1. Data as MS AccessDB and as MS Excel file (identical info as MS AccessDB) as well as data dictionary and database schema.
2. Network .net file to visually render science map. Also provided as .net file with discipline nodes and names.

Usage Conditions
This map is shared under the Creative Commons, Attribution-NonCommercial-ShareAlike 3.0 Unported (CC BY-NC-SA 3.0) license (http://creativecommons.org/licenses/by-nc-sa/3.0/). That is, you are free to share, e.g., to copy, distribute and transmit the work, and to remix, i.e., to adapt the work under the following conditions:

- Attribution — You must attribute the work in the following manner (but not in any way that suggests that they endorse you or your use of the work): Cite the above paper and use the following acknowledgment text: "*The authors wish to acknowledge The Regents of the University of California, SciTech Strategies, Observatoire des Sciences et des Technologies, and the Cyberinfrastructure for Network Science Center for making the 2010 UCSD Map of Science and Classification System available for this work.*"
- Noncommercial — You may not use this work for commercial purposes.

Please cite as: Börner, Katy, Richard Klavans, Michael Patek, Angela Zoss, Joseph R. Biberstine, Robert Light, Vincent Larivière, and Kevin W. Boyack (2012) Design and Update of a Classification System: The UCSD Map of Science. *PLoS ONE* 7(7): e39464. doi:10.1371/journal.pone.0039464

Full Update (10 years of Scopus and WoS)

by Börner, Klavans, Patek, Zoss, Biberstine, Light, Larivière, Boyack

Data: The 2010 UCSD map of science and classification system covers ten years (2001-2010) of data from Thomson Reuters' Web of Science and eight years (2001-2008) of Elsevier's Scopus, specifically the fractional assignment of about 25,000 journal names to 554 subdisciplines grouped into 13 disciplines of science.

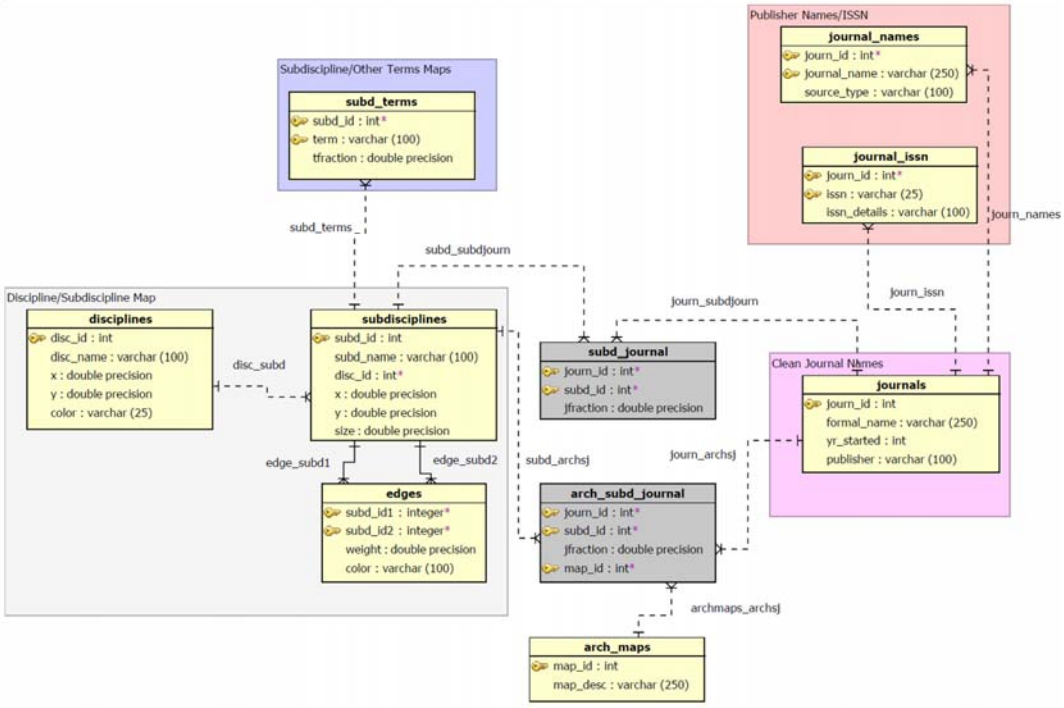
The counts for major record types are given here:

1. 13 disciplines with labels and color codes
2. 554 subdisciplines with x, y positions and size
3. 15,849 journals captured by 5-year map
4. 25,258 journals captured by 10-year map
5. 13,520 journal names used by Thomson Reuters
6. 22,253 journal names used by Scopus
7. 21,630 Scopus journal ID numbers
8. 19,988 ISSN numbers
9. 66,759 terms

See Data Dictionary in Supplement 2 in
http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0039464

Full Update (10 years of Scopus and WoS)

by Börner, Klavans, Patek, Zoss, Biberstine, Light, Larivière, Boyack



UCSD map table schema http://sci.cns.iu.edu/ucsdmap/data/UCSDmapDBSchema.pdf

Deployment: The UCSD map of science data is available at http://sci.cns.iu.edu/ucsdmap/

Data
The 2010 UCSD map of science and classification system covering 10 years (2001-2010) of Web of Science data and 8 years (2001-2008) of Scopus data with subdiscipline assignments by SciTech Strategies.

1. Data as MS AccessDB and as MS Excel file (identical info as MS AccessDB) as well as data dictionary and database schema.
2. Network .net file to visually render science map. Also provided as .net file with discipline nodes and names.

Usage Conditions
This map is shared under the Creative Commons, Attribution-NonCommercial-ShareAlike 3.0 Unported (CC BY-NC-SA 3.0) license (http://creativecommons.org/licenses/by-nc-sa/3.0/). That is, you are free to share, e.g., to copy, distribute and transmit the work, and to remix, i.e., to adapt the work under the following conditions:

- Attribution — You must attribute the work in the following manner (but not in any way that suggests that they endorse you or your use of the work): Cite the above paper and use the following acknowledgment text: "*The authors wish to acknowledge The Regents of the University of California, SciTech Strategies, Observatoire des Sciences et des Technologies, and the Cyberinfrastructure for Network Science Center for making the 2010 UCSD Map of Science and Classification System available for this work.*"
- Noncommercial — You may not use this work for commercial purposes.

Please cite as: Börner, Katy, Richard Klavans, Michael Patek, Angela Zoss, Joseph R. Biberstine, Robert Light, Vincent Larivière, and Kevin W. Boyack (2012) Design and Update of a Classification System: The UCSD Map of Science. *PLoS ONE* 7(7): e39464. doi:10.1371/journal.pone.0039464

Full Update (10 years of Scopus and WoS)
by Börner, Klavans, Patek, Zoss, Biberstine, Light, Larivière, Boyack

Note: There are no standards on how to render .net files!

Some define the zero point on the top left (e.g., GUESS), while others define the bottom left point as 0,0 (e.g., Gephi, Pajek). This only becomes important when rendering a dataset that has a predefined left and right, top and bottom such as the UCSD map of science.
Simply multiply all node's y-position with -1 to solve this issue.

Science Map Validation:

- Comparing the Accuracies of Nine Text-Based Similarity Approaches
- Consensus Map
- User Studies

Comparing the Accuracies of Nine Text-Based Similarity Approaches

We used a corpus of 2.15 million recent (2004-2008) records from MEDLINE, and generated nine different document-document similarity matrices from information extracted from their bibliographic records, including titles, abstracts and subject headings.

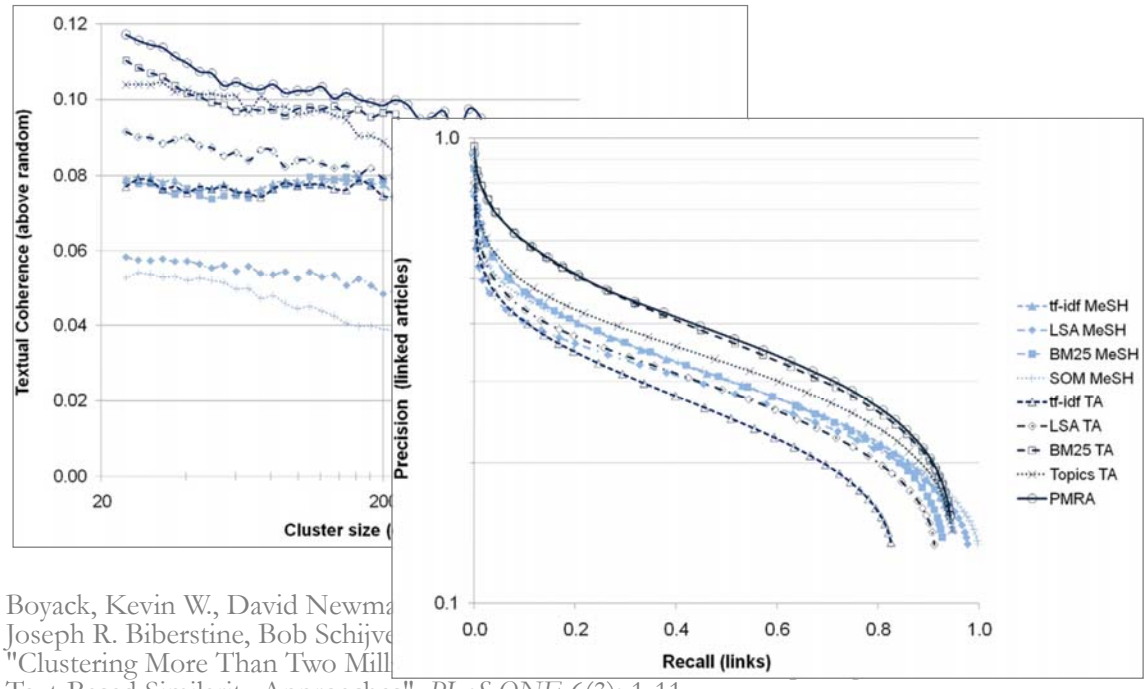
Similarity approach	Data source	
	MeSH terms	Title/abstract words
tf-idf cosine	tf-idf MeSH (Indiana U.)	tf-idf TA (Indiana U.)
Latent semantic analysis	LSA MeSH (Indiana U.)	LSA TA (Indiana U.)
Topic modeling		Topics TA (UC Irvine)
Self-organizing map	SOM MeSH (SDSU/Indiana U.)	
Poisson-based	BM25 MeSH (Collexis)	BM25 TA (Collexis) PMRA (UC Irvine/SciTech)

Cluster results from the nine similarity approaches were compared using (1) within-cluster textual coherence based on the Jensen-Shannon divergence, and (2) two concentration measures based on grant-to-article linkages indexed in MEDLINE.

Boyack, Kevin W., David Newman, Russell Jackson Duhon, Richard Klavans, Michael Patek, Joseph R. Biberstine, Bob Schijvenaars, Andre Skupin, Nianli Ma, and Katy Börner. 2011. "Clustering More Than Two Million Biomedical Publications: Comparing the Accuracies of Nine Text-Based Similarity Approaches". *PLoS ONE* 6(3): 1-11.

Data is at http://sts.cns.iu.edu

Comparing the Accuracies of Nine Text-Based Similarity Approaches

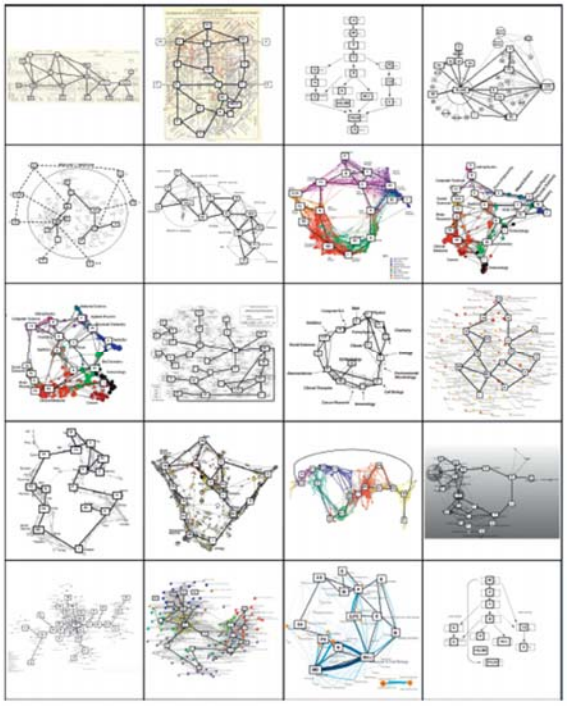


Boyack, Kevin W., David Newma
Joseph R. Biberstine, Bob Schijve
"Clustering More Than Two Mill
Text-Based Similarity Approaches". *PLoS ONE* 6(3): 1-11.

Consensus Map

20 maps of science were examined and found to have a high level of correspondence.

Source map	Year	Type
KB06-SC	2006	Paper
Backbone	2004	Jnl
UCSD	2007	Jnl
Ellingham	1948	Expert
KB-Para	2005	Paper
Bernal	1939	Expert
Scimago-I	2004	Categ
KB06-TS	2006	Paper
B03-ST	2005	Jnl
BBK02-S	2004	Jnl
Rosvall	2007	Jnl
Small99	1999	Paper
Balaban-II	2007	Pre-req
K02	2002	Jnl
L-R	2007	Categ
Balaban-I	2007	Expert
Small85	1985	Paper
Small74	1974	Paper
B-Z	1999	Jnl
Scimago-II	2007	Categ

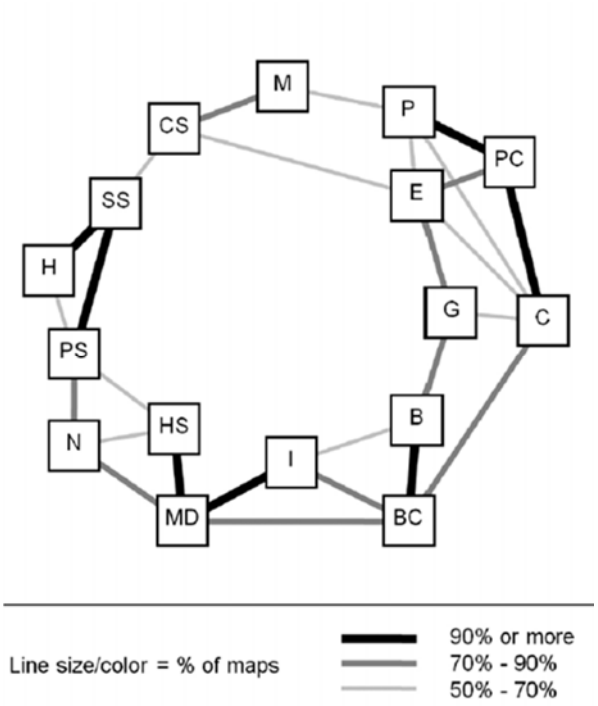


Klavans, R., & Boyack, K. W. (2009). Toward a consensus map of science. *Journal of the American Society for Information Science and Technology*, 60(3), 455-476.

Consensus Map

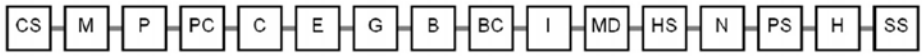
20 maps of science were examined and found to have a high level of correspondence.

M – Mathematics
CS – Computer science
P – Physics
PC – Physical chemistry
C – Chemistry
E – Engineering
G – Earth sciences (geoscience)
BC – Biochemistry
B – Biology
I – Infectious disease
MD – Medical specialties
HS – Health services
N – Brain research (neuroscience)
PS – Psychology/psychiatry
SS – Social sciences
H – Humanities



Klavans, R., & Boyack, K. W. (2009). Toward a consensus map of science. *Journal of the American Society for Information Science and Technology*, 60(3), 455-476.

Science Basemaps: Consensus Map



M – Mathematics
CS – Computer science
P – Physics
PC – Physical chemistry
C – Chemistry
E – Engineering
G – Earth sciences (geoscience)
BC – Biochemistry
B – Biology
I – Infectious disease
MD – Medical specialties
HS – Health services
N – Brain research (neuroscience)
PS – Psychology/psychiatry
SS – Social sciences
H – Humanities

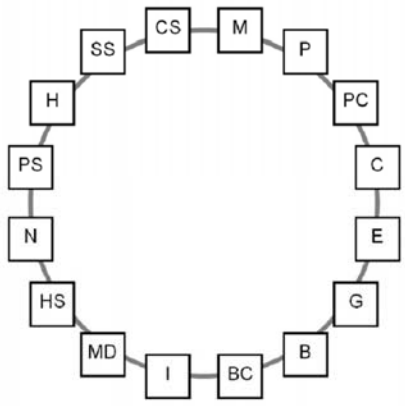


FIG. 6. One-dimensional consensus maps of science, Euclidean (top) and Riemannian (bottom).

Klavans, R., & Boyack, K. W. (2009). Toward a consensus map of science. *Journal of the American Society for Information Science and Technology*, 60(3), 455-476.

User Studies

Compare accuracies of cluster solutions of a large corpus of 2,153,769 recent articles from the biomedical literature (2004-2008) using four similarity approaches: co-citation analysis, bibliographic coupling, direct citation, and a bibliographic coupling-based citation-text hybrid approach.

Accuracies are compared using two metrics – within-cluster textual coherence as defined by the Jensen-Shannon divergence, and a new concentration factor based on the grant-to-article linkages indexed in MEDLINE.

Experts found the partitioning of our micro-structural model to be compelling at the research problem level.

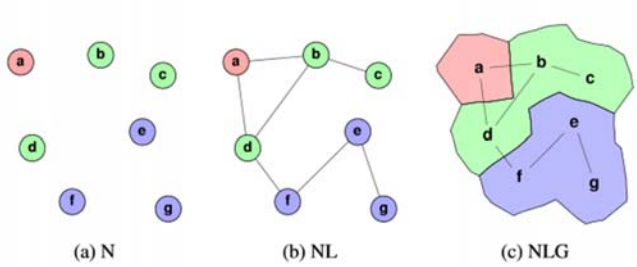
Boyack, K. W., & Klavans, R. (2010). Co-citation analysis, bibliographic coupling, and direct citation: Which citation approach represents the research front most accurately? Journal of the American Society for Information Science and Technology, 61(12), 2389-2404.

Klavans, R., Boyack, K. W. , and Small Henry (2012). Indicators and Precursors of 'Hot Science.' *17th International Conference on Science and Technology Indicators (STI), Montreal, Canada.*

User Studies on Map Layouts

Compares node diagrams, node-link diagrams, and node-link-group diagrams using nine different tasks that fall broadly in three categories:

- node-based tasks
- network-based tasks
- group-based tasks.

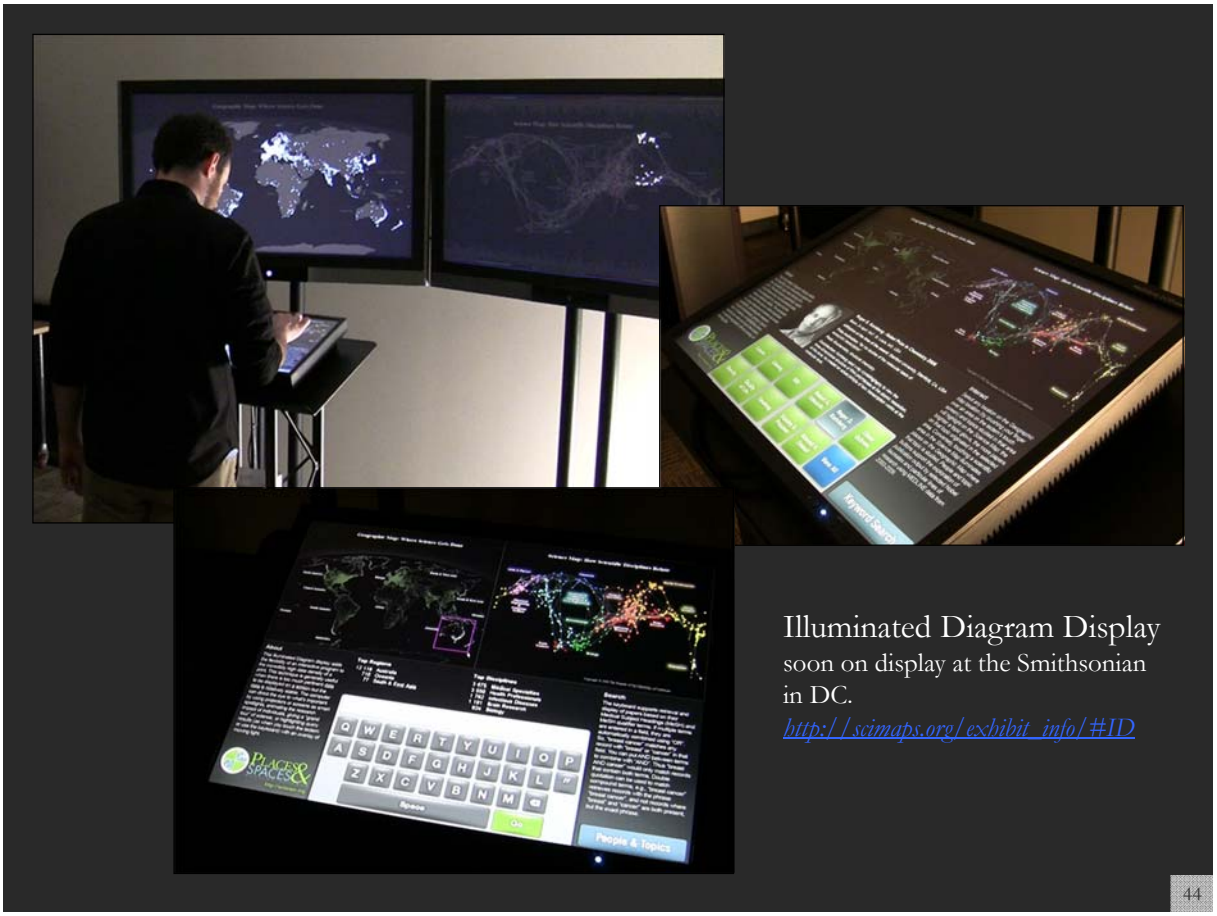


Findings indicate that adding links, or links and group representations, does not negatively impact performance (time and accuracy) of node-based tasks. Similarly, adding group representations does not negatively impact the performance of network-based tasks. Node-link-group diagrams outperform the others on group-based tasks.

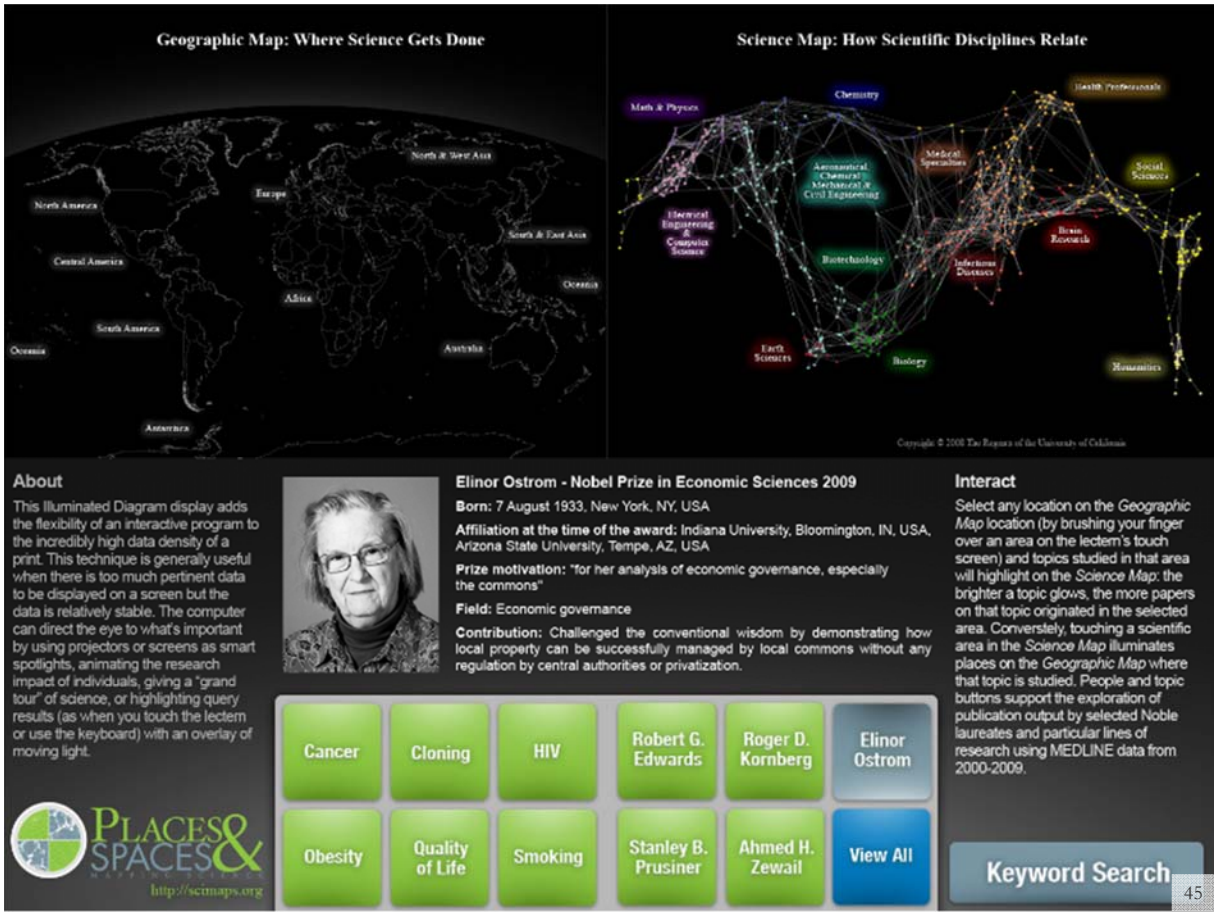
Bahador Saket, Paolo Simonetto, Stephen Kobourov and Katy Borner. (2014) Node, Node-Link, and Node-Link-Group Diagrams: An Evaluation.
http://arxiv.org/pdf/1404.1911.pdf
Accepted at IEEE InfoVis 2014.

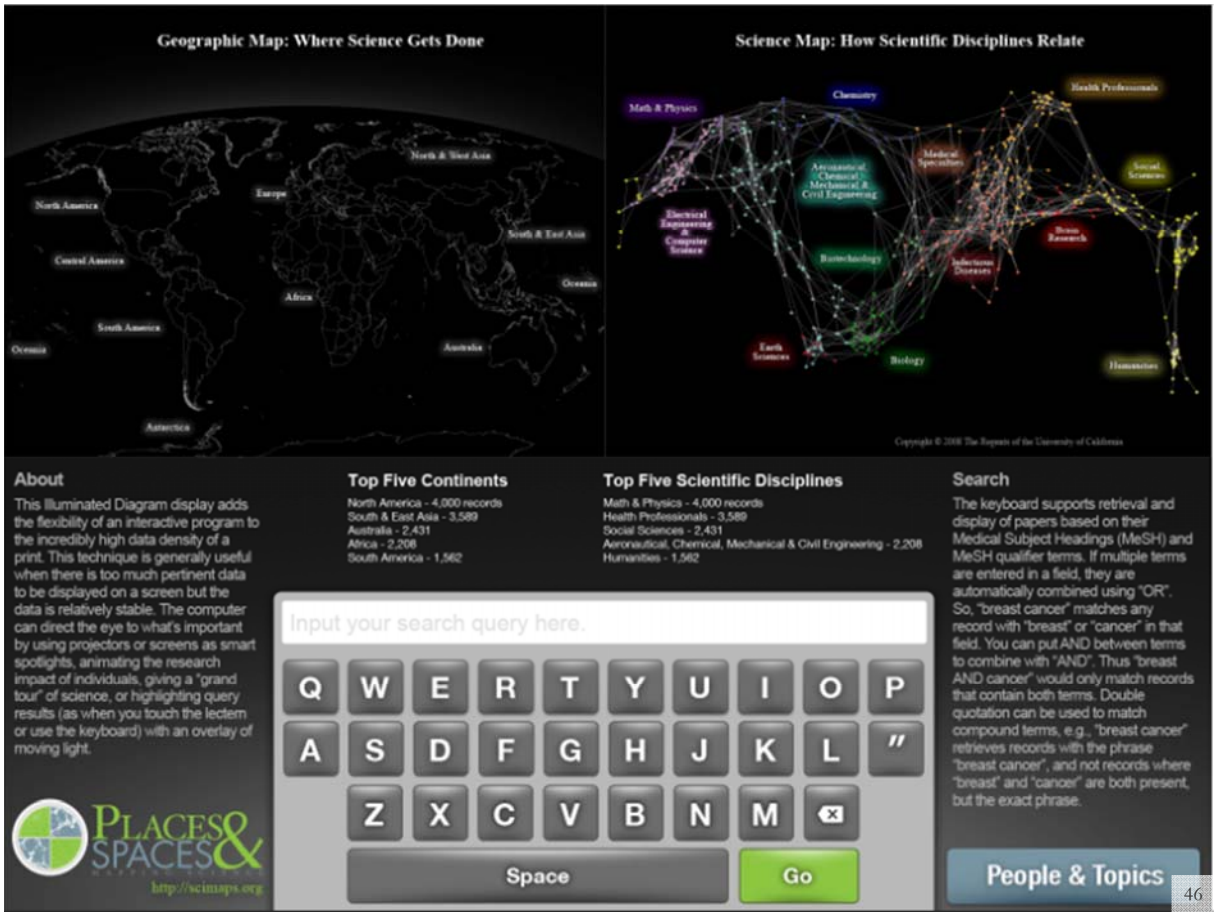
Application:
How to use the (UCSD) map of science?

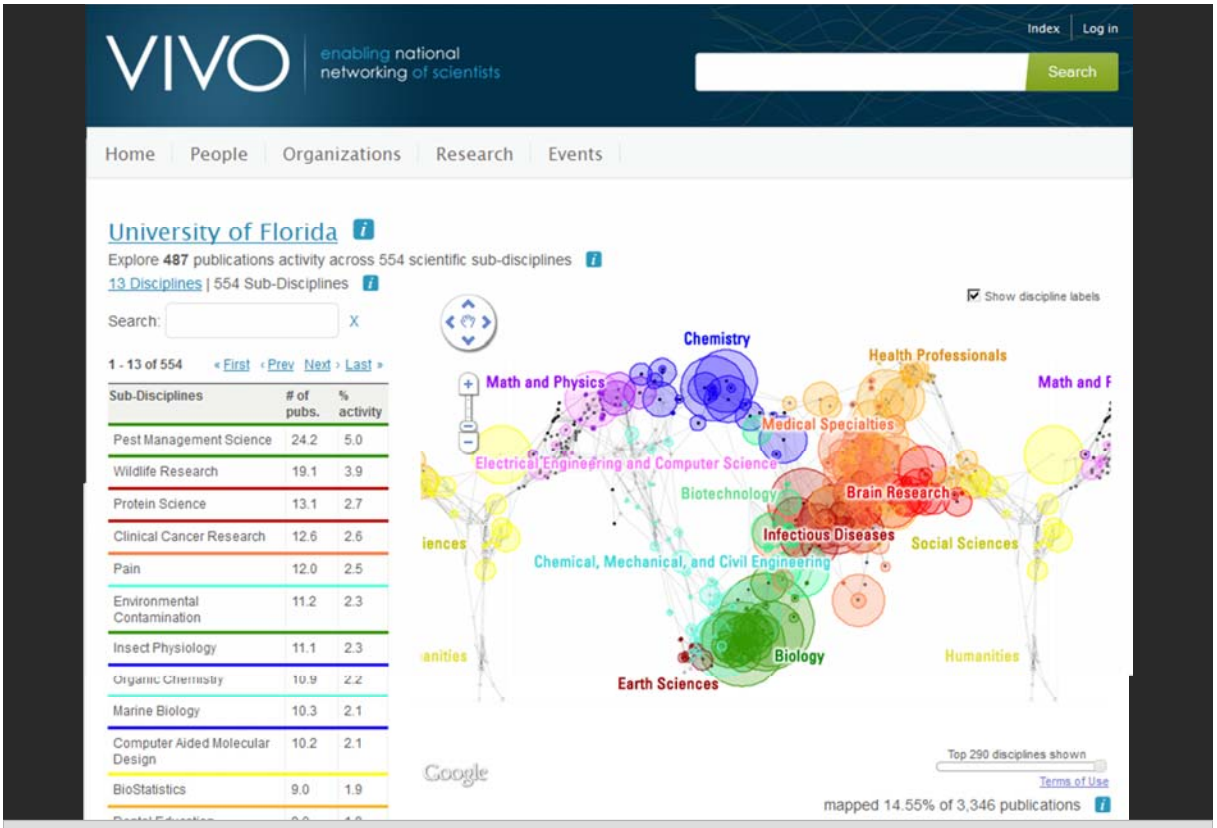
- Illuminated Diagram
- VIVO
- MapSustain
- Sci2



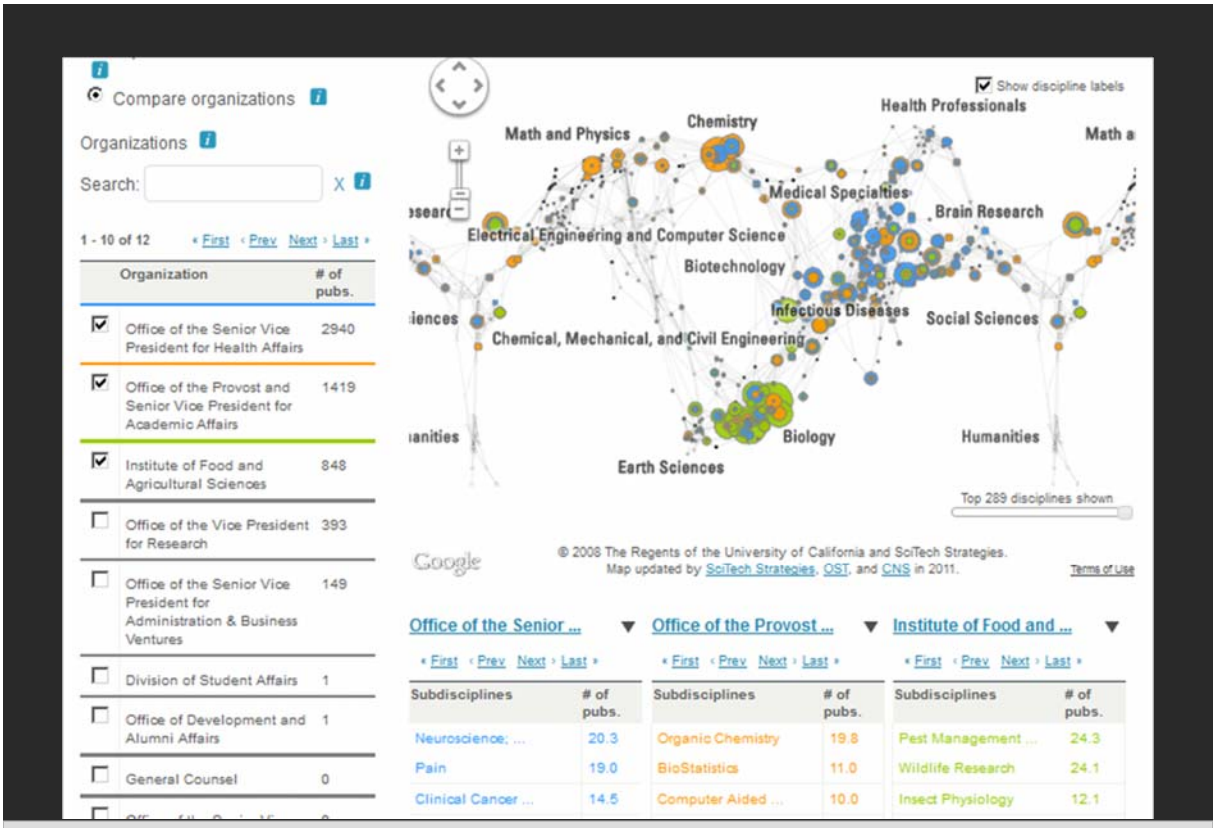
Illuminated Diagram Display
soon on display at the Smithsonian in DC.
http://scimaps.org/exhibit_info/#ID


Geographic Map: Where Science Gets Done
North America
Europe
North & West Asia
South & East Asia
Central America
Africa
Oceania
South America
Australia
Antarctica
Science Map: How Scientific Disciplines Relate
Math & Physics
Chemistry
Health Professionals
Medical Specialties
Social Sciences
Aeronautical, Chemical, Mechanical & Civil Engineering
Electrical Engineering & Computer Science
Brain Research
Biotechnology
Infectious Diseases
Earth Sciences
Biology
Humanities
Copyright © 2008 The Regents of the University of California
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
Elinor Ostrom - Nobel Prize in Economic Sciences 2009
Born: 7 August 1933, New York, NY, USA
Affiliation at the time of the award: Indiana University, Bloomington, IN, USA, Arizona State University, Tempe, AZ, USA
Prize motivation: "for her analysis of economic governance, especially the commons"
Field: Economic governance
Contribution: Challenged the conventional wisdom by demonstrating how local property can be successfully managed by local commons without any regulation by central authorities or privatization.
Cancer
Cloning
HIV
Robert G. Edwards
Roger D. Kornberg
Elinor Ostrom
Obesity
Quality of Life
Smoking
Stanley B. Prusiner
Ahmed H. Zewail
View All
Interact
Select any location on the Geographic Map location (by brushing your finger over an area on the lectern's touch screen) and topics studied in that area will highlight on the Science Map: the brighter a topic glows, the more papers on that topic originated in the selected area. Conversely, touching a scientific area in the Science Map illuminates places on the Geographic Map where that topic is studied. People and topic buttons support the exploration of publication output by selected Noble laureates and particular lines of research using MEDLINE data from 2000-2009.
Keyword Search


Geographic Map: Where Science Gets Done
North America
Europe
North & West Asia
South & East Asia
Central America
Africa
Oceania
South America
Australia
Antarctica
Science Map: How Scientific Disciplines Relate
Math & Physics
Chemistry
Health Professionals
Medical Specialties
Social Sciences
Aeronautical, Chemical, Mechanical & Civil Engineering
Electrical Engineering & Computer Science
Brain Research
Biotechnology
Infectious Diseases
Earth Sciences
Biology
Humanities
Copyright © 2008 The Regents of the University of California
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 - 1,582
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 - 1,582
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



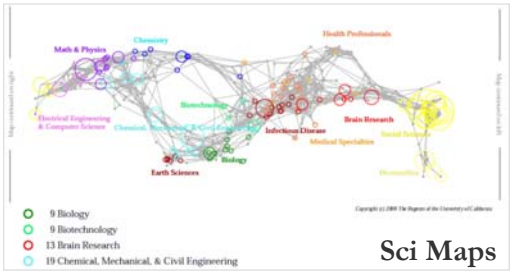
Topical Analysis (What) Science map overlays will show where a person, department, or university publishes most in the world of science. (in work)



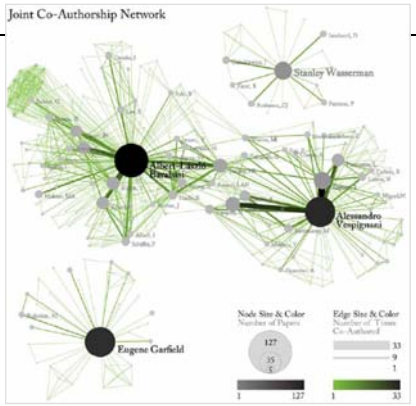
Topical Analysis (What) Science map overlays will show where a person, department, or university publishes most in the world of science. (in work)

Sci² Tool – "Open Code for S&T Assessment"

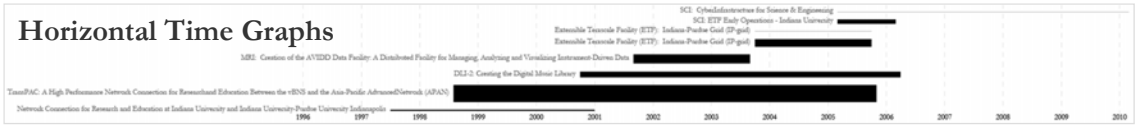
OSGi/CIShell powered tool with NWB plugins and many new scientometrics and visualizations plugins.



Sci Maps

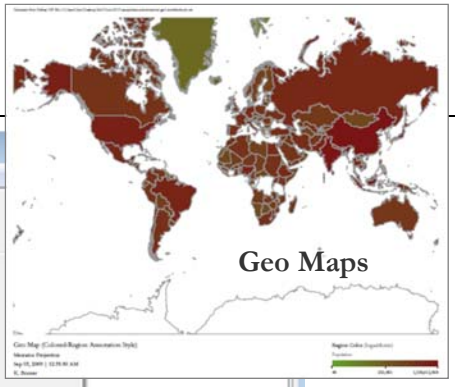


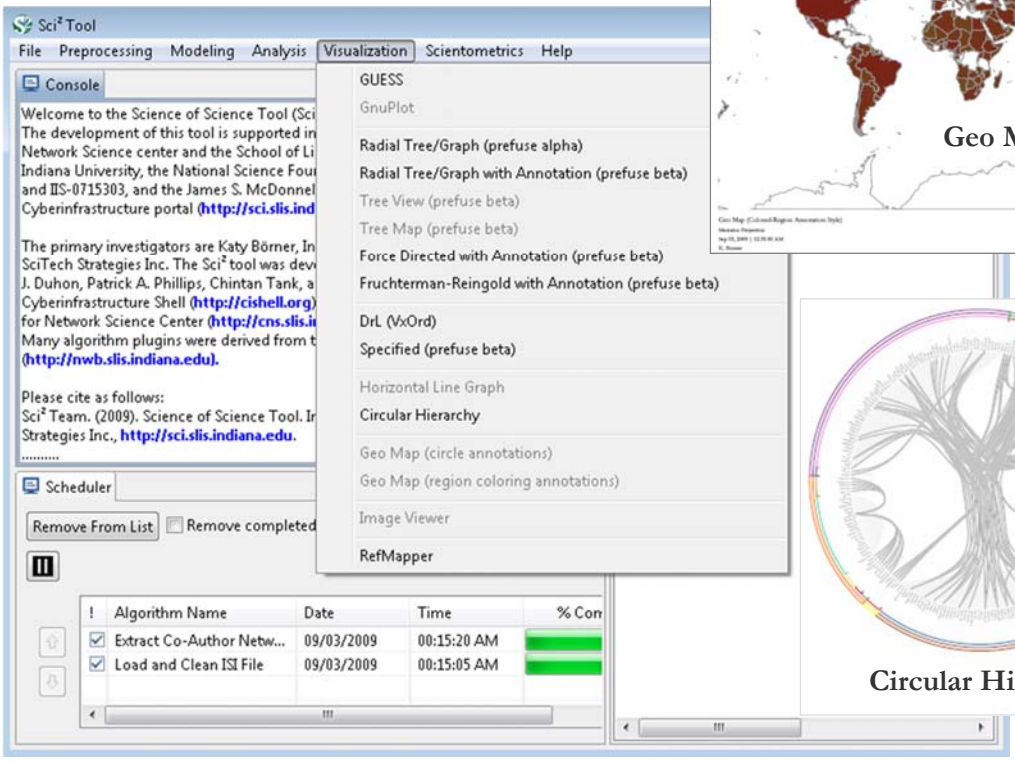
GUESS Network Vis



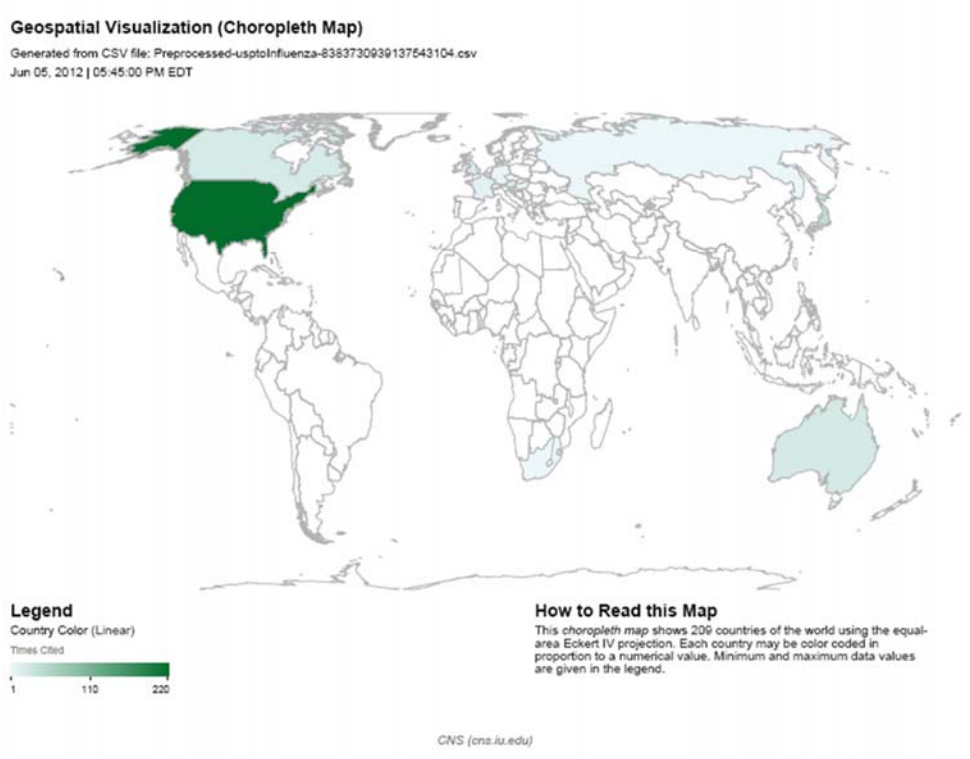
Börner, Katy, Huang, Weixia (Bonnie), Linnemeier, Micah, Duhon, Russell Jackson, Phillips, Patrick, Ma, Nianli, Zoss, Angela, Guo, Hanning & Price, Mark. (2009). Rete-Netzwerk-Red: Analyzing and Visualizing Scholarly Networks Using the Scholarly Database and the Network Workbench Tool. Proceedings of ISSI 2009: 12th International Conference on Scientometrics and Informetrics, Rio de Janeiro, Brazil, July 14-17 . Vol. 2, pp. 619-630.

Sci² Tool Vis cont.

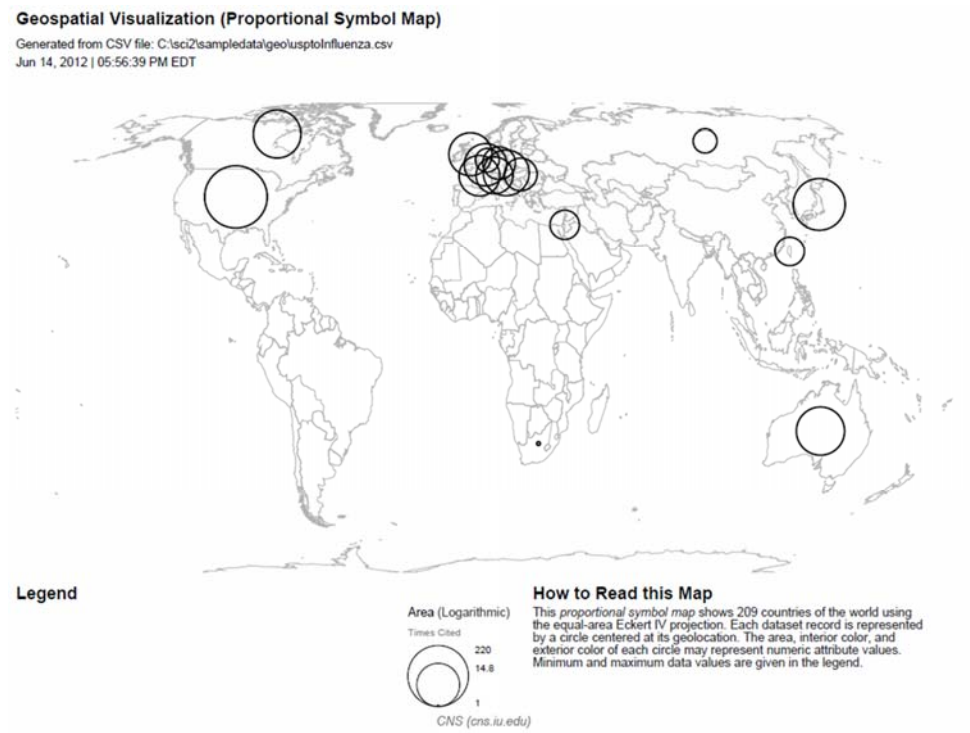




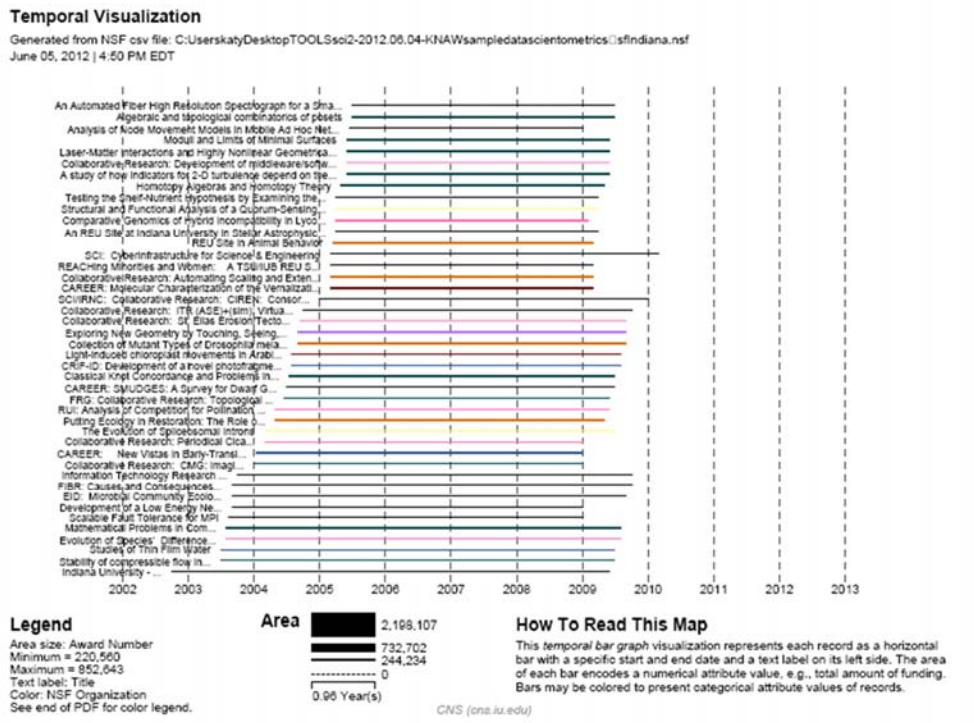
New Visualizations



New Visualizations

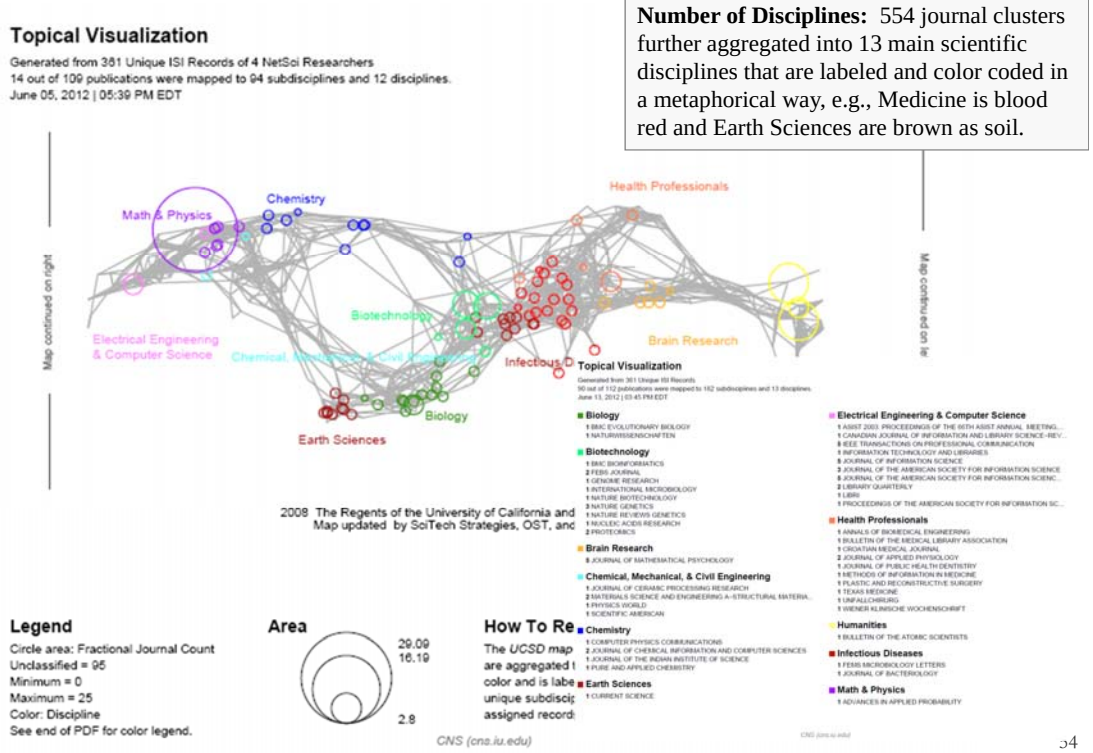


New Visualizations



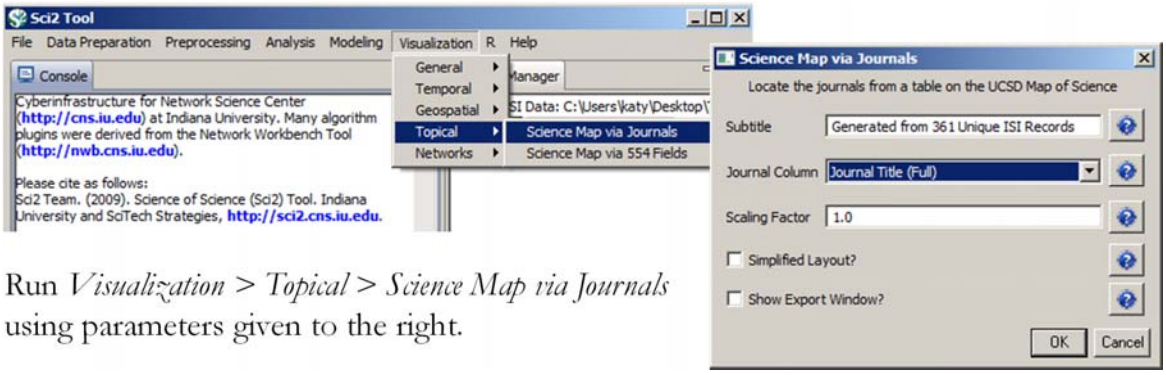
New Visualizations

Data: WoS and Scopus paper level data for 2001–2010, about 25,000 separate journals, proceedings, and series.
Similarity Metric: Combination of bibliographic coupling and keyword vectors.
Number of Disciplines: 554 journal clusters further aggregated into 13 main scientific disciplines that are labeled and color coded in a metaphorical way, e.g., Medicine is blood red and Earth Sciences are brown as soil.



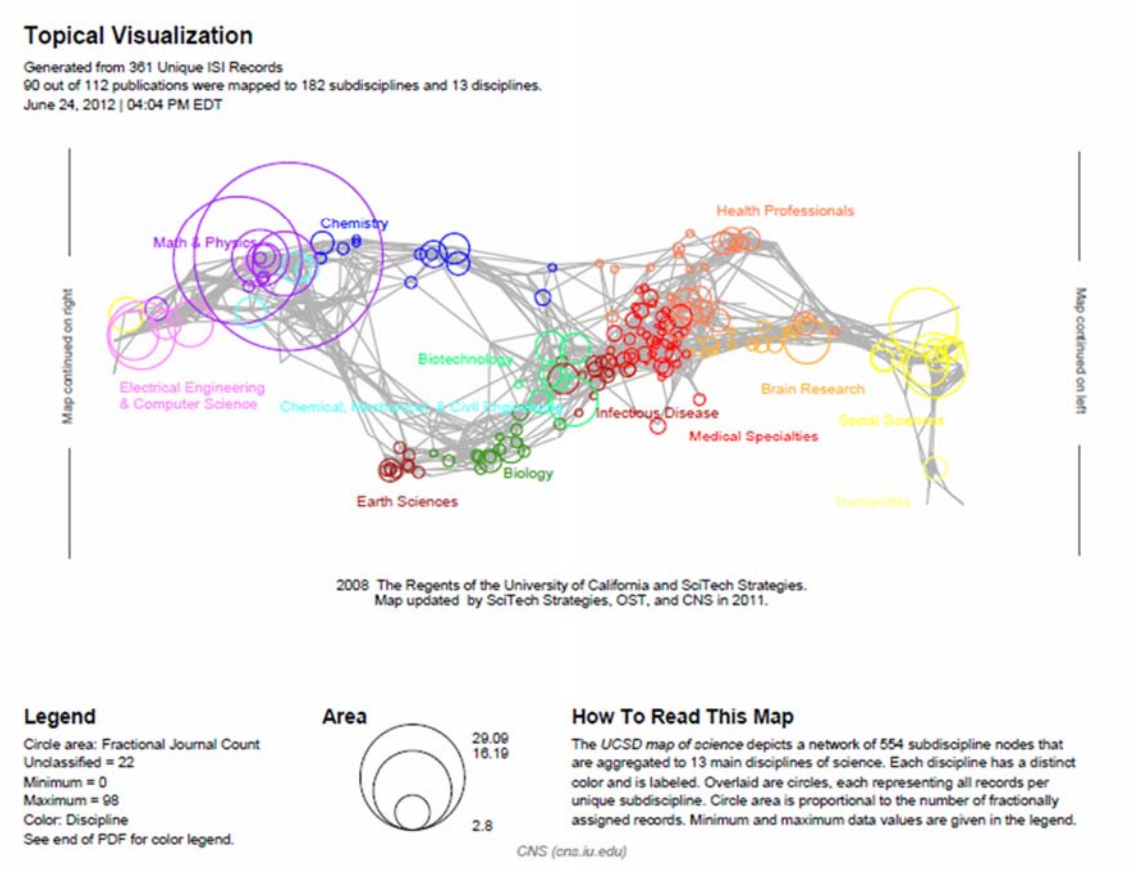
Download Sci2 Tool v1.0 Alpha (June 13, 2012) from http://sci2.cns.iu.edu
Unpack into a /sci2 directory. Run /sci2/sci2.exe
Sci2 Manual is at http://sci2.wiki.cns.iu.edu

Load an ISI (*.isi), Bibtex (*.bib), Endnote Export Format (*.enw), Scopus csv (*.scopus) file such as */sci2/sampledata/scientometrics/isi/FourNetSciResearchers.isi*



Run *Visualization > Topical > Science Map via Journals* using parameters given to the right.

Postscript file will appear in *Data Manager*. Save and open with a Postscript Viewer.



Topical Visualization

Generated from 361 Unique ISI Records
90 out of 112 publications were mapped to 182 subdisciplines and 13 disciplines.
June 24, 2012 | 04:04 PM EDT

Biology
1 BMC EVOLUTIONARY BIOLOGY
1 NATURWISSENSCHAFTEN

Biotechnology
1 BMC BIOINFORMATICS
2 FEBS JOURNAL
1 GENOME RESEARCH
1 INTERNATIONAL MICROBIOLOGY
1 NATURE BIOTECHNOLOGY
3 NATURE GENETICS
1 NATURE REVIEWS GENETICS
1 NUCLEIC ACIDS RESEARCH
2 PROTEOMICS

Brain Research
5 JOURNAL OF MATHEMATICAL PSYCHOLOGY

Chemical, Mechanical, & Civil Engineering
1 JOURNAL OF CERAMIC PROCESSING RESEARCH
2 MATERIALS SCIENCE AND ENGINEERING A-STRUCTURAL MATERIA...
1 PHYSICS WORLD
1 SCIENTIFIC AMERICAN

Chemistry
1 COMPUTER PHYSICS COMMUNICATIONS
2 JOURNAL OF CHEMICAL INFORMATION AND COMPUTER SCIENCES
1 JOURNAL OF THE INDIAN INSTITUTE OF SCIENCE
1 PURE AND APPLIED CHEMISTRY

Earth Sciences
1 CURRENT SCIENCE

Electrical Engineering & Computer Science
1 ASIST 2003: PROCEEDINGS OF THE 66TH ASIST ANNUAL MEETING,...
1 CANADIAN JOURNAL OF INFORMATION AND LIBRARY SCIENCE-REV...
5 IEEE TRANSACTIONS ON PROFESSIONAL COMMUNICATION
1 INFORMATION TECHNOLOGY AND LIBRARIES
5 JOURNAL OF INFORMATION SCIENCE
3 JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE
5 JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENC...
2 LIBRARY QUARTERLY
1 LIBRI
1 PROCEEDINGS OF THE AMERICAN SOCIETY FOR INFORMATION SC...

Health Professionals
1 ANNALS OF BIOMEDICAL ENGINEERING
1 BULLETIN OF THE MEDICAL LIBRARY ASSOCIATION
1 CROATIAN MEDICAL JOURNAL
2 JOURNAL OF APPLIED PHYSIOLOGY
1 JOURNAL OF PUBLIC HEALTH DENTISTRY
1 METHODS OF INFORMATION IN MEDICINE
1 PLASTIC AND RECONSTRUCTIVE SURGERY
1 TEXAS MEDICINE
1 UNFALLCHIRURG
1 WIENER KLINISCHE WOCHENSCHRIFT

Humanities
1 BULLETIN OF THE ATOMIC SCIENTISTS

Infectious Diseases
1 FEMS MICROBIOLOGY LETTERS
1 JOURNAL OF BACTERIOLOGY

Math & Physics
1 ADVANCES IN APPLIED PROBABILITY

Topical Visualization

Generated from 361 Unique ISI Records
90 out of 112 publications were mapped to 182 subdisciplines and 13 disciplines.
June 24, 2012 | 04:04 PM EDT

Math & Physics
10 APPLIED PHYSICS LETTERS
1 BRAZILIAN JOURNAL OF PHYSICS
3 CHAOS SOLITONS & FRACTALS
1 COMPLEXITY
1 COMPUTATIONAL MATERIALS SCIENCE
11 EUROPEAN PHYSICAL JOURNAL B
12 EUROPHYSICS LETTERS
2 INTERNATIONAL JOURNAL OF MODERN PHYSICS B
6 JOURNAL OF PHYSICS A-MATHEMATICAL AND GENERAL
1 JOURNAL OF STATISTICAL MECHANICS-THEORY AND EXPERIMENT
1 JOURNAL OF STATISTICAL PHYSICS
1 JOURNAL OF THE KOREAN PHYSICAL SOCIETY
1 MATERIALS SCIENCE AND ENGINEERING B-SOLID STATE MATERIAL...
3 NATURE PHYSICS
3 NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH SEC...
12 PHYSICA A
5 PHYSICAL REVIEW A
2 PHYSICAL REVIEW B
45 PHYSICAL REVIEW LETTERS
2 REVIEWS OF MODERN PHYSICS

Medical Specialties
1 ANNALS OF INTERNAL MEDICINE
1 REVISTA DE INVESTIGACION CLINICA

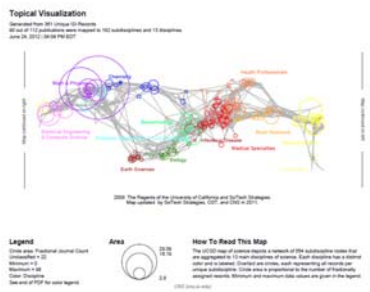
Social Sciences
1 ADMINISTRATIVE SCIENCE QUARTERLY
1 AMERICAN BEHAVIORAL SCIENTIST
1 AMERICAN SOCIOLOGICAL REVIEW
1 ANNALS OF THE AMERICAN ACADEMY OF POLITICAL AND SOCIAL S...
1 ARBOR-CIENCIA PENSAMIENTO Y CULTURA
3 BRITISH JOURNAL OF MATHEMATICAL & STATISTICAL PSYCHOLOGY
1 JOURNAL OF CLASSIFICATION

Social Sciences
2 JOURNAL OF MATHEMATICAL SOCIOLOGY
3 JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION
2 PSYCHOLOGICAL BULLETIN
5 PSYCHOMETRIKA
1 RECHERCHE
5 SCIENTOMETRICS
1 SOCIAL FORCES
6 SOCIAL NETWORKS
3 SOCIOLOGICAL METHODS & RESEARCH

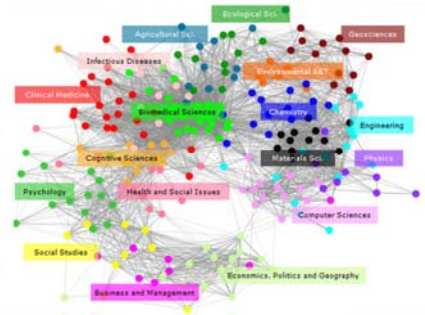
Multiple Categories
1 BRITISH MEDICAL JOURNAL
2 JAMA-JOURNAL OF THE AMERICAN MEDICAL ASSOCIATION
1 JOURNAL OF THEORETICAL BIOLOGY
18 NATURE
44 PHYSICAL REVIEW E
5 PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE ...
6 SCIENCE

Unclassified
1 ALGORITHMS AND MODELS FOR THE WEB-GRAPHS, PROCEEDINGS
2 AMERICAN DOCUMENTATION
2 ASIST 2002: PROCEEDINGS OF THE 65TH ASIST ANNUAL MEETING, ...
1 BIOLOGIYA MORYA-MARINE BIOLOGY
1 BULLETIN OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE
1 CHEMIKER-ZEITUNG
3 CHEMTECH
1 COMBINATORIAL AND ALGORITHMIC ASPECTS OF NETWORKING
7 CURRENT COMMENTS
3 CURRENT CONTENTS/LIFE SCIENCES
1 FEDERATION PROCEEDINGS
5 FRACTALS-AN INTERDISCIPLINARY JOURNAL ON THE COMPLEX GE...
1 FRONTIERS OF LIBRARIANSHIP-SYRACUSE UNIVERSITY

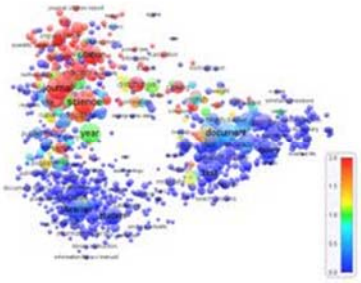
Currently Used Science Basemaps



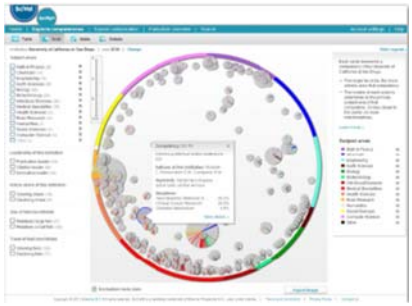
UCSD Map



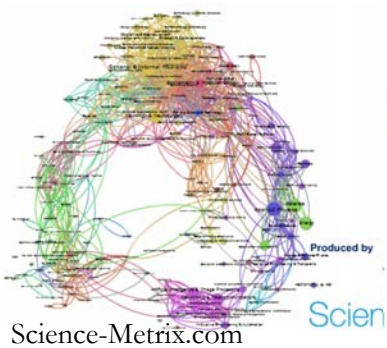
Loet et al science maps ISI categories



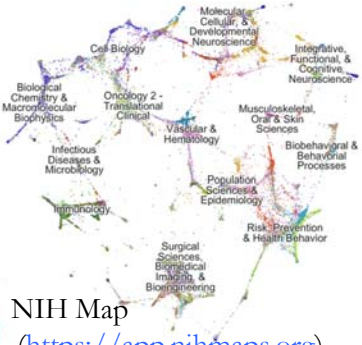
http://vosviewer.com



Elsevier's SciVal Map



Science-Metrix.com



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

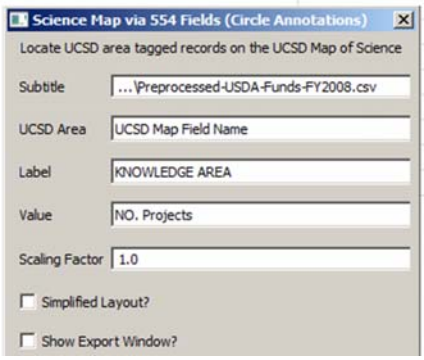
Align existing classifications/taxonomies/ hierarchies to generate science map overlays

In addition to using **journal names** to

- Map career trajectories
- Identify evolving expertise areas
- Compare expertise profiles

Existing classifications can be aligned and used to generate science map overlays.

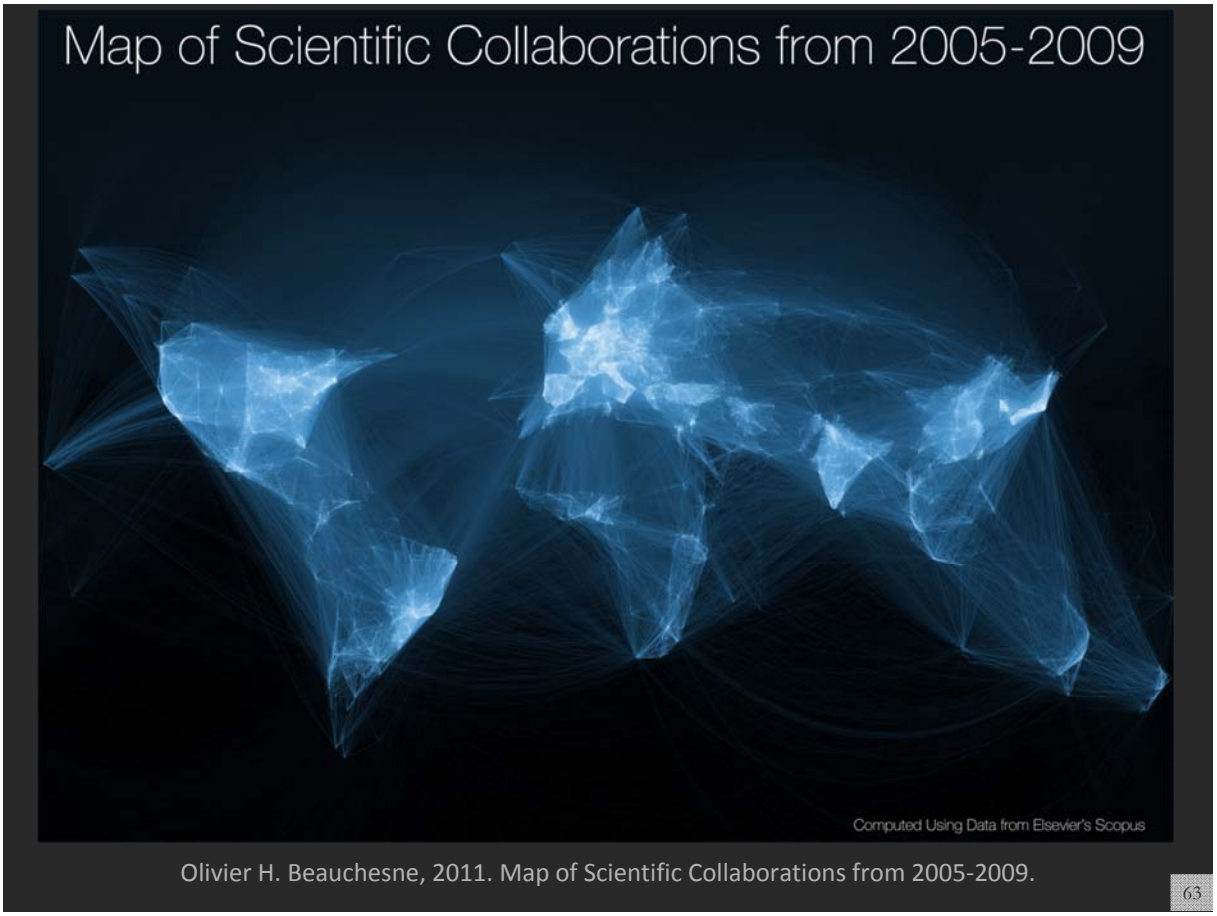
B	C	D	E	F	G
KNOWLEDGE AREA	NO. Projects	USDA Staff Years	STATE APPR	TOTAL FUNDS	UCSD Map Field Name
101 Appraisal of Soil Resources					315
102 Soil, Plant, Water, Nutrient Relationships					227
103 Management of Saline and Sodic Soils and Salinity					158
104 Protect Soil from Harmful Effects of Natural Elements					120
111 Conservation and Efficient Use of Water					245
112 Watershed Protection and Management					245
121 Management of Range Resources					520
122 Management and Control of Forest and Range Fires					520
123 Management and Sustainability of Forest Resources					231
124 Urban Forestry					231
125 Agroforestry					231



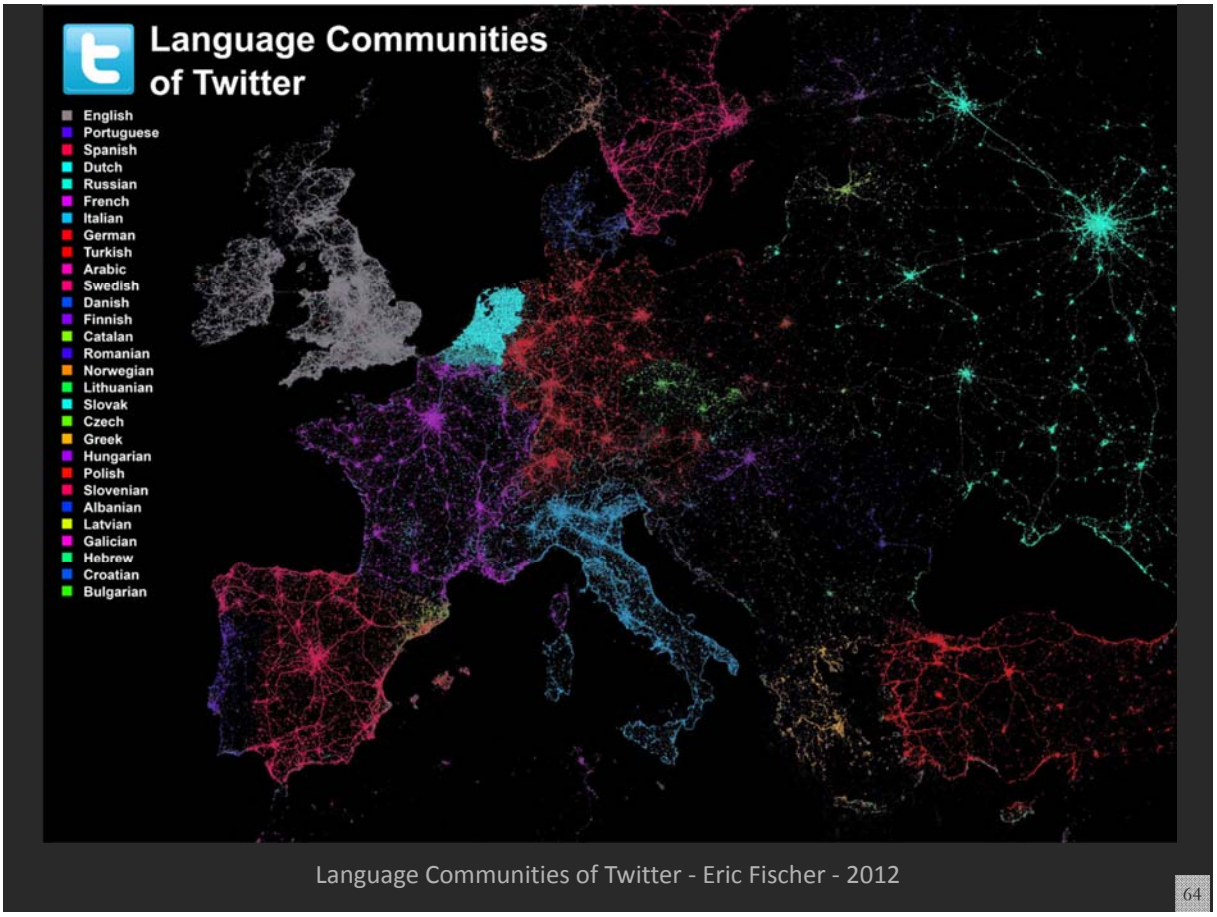
Run *Visualization > Topical > Science Map via 554 Fields* using parameters given to the right.
Postscript file will appear in *Data Manager*.
Save and open with a Postscript Viewer.

Places & Spaces: Mapping Science Exhibit

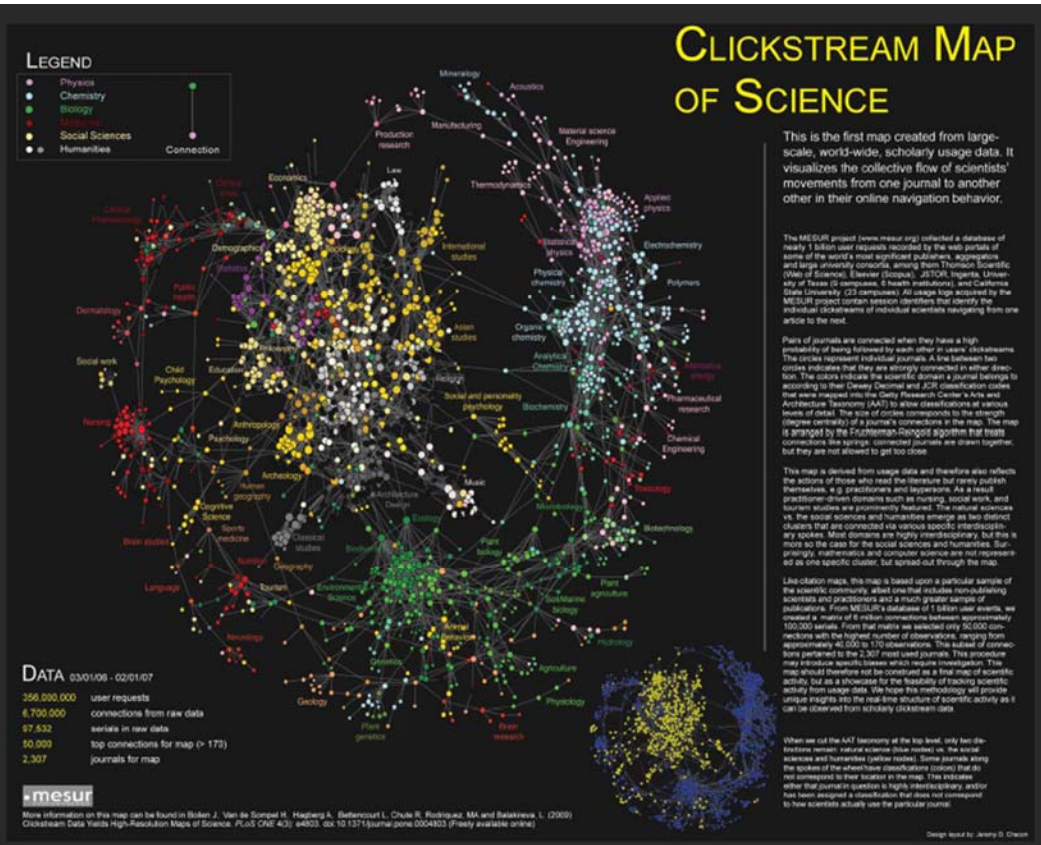
Mapping Science Exhibit on display at MEDIA X, Stanford University
http://mediax.stanford.edu, *http://scaleindependentthought.typepad.com/photos/scimaps*



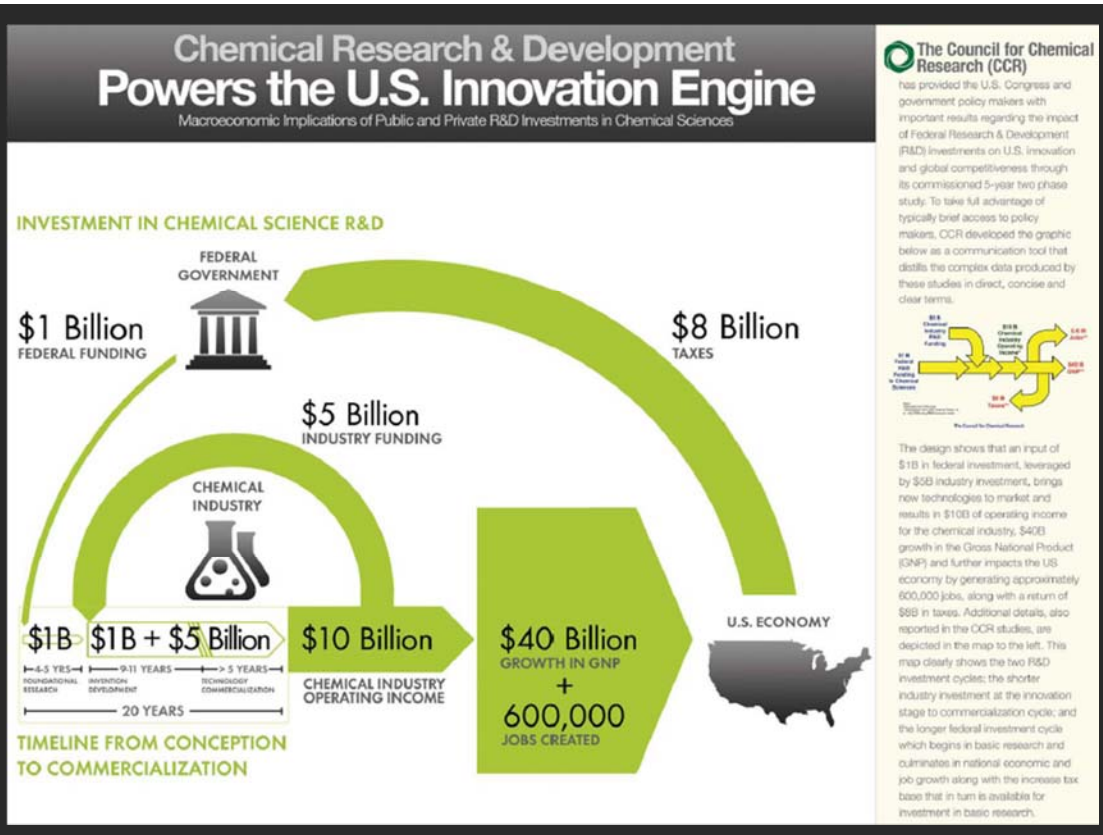
Olivier H. Beauchesne, 2011. Map of Scientific Collaborations from 2005-2009.



Language Communities of Twitter - Eric Fischer - 2012

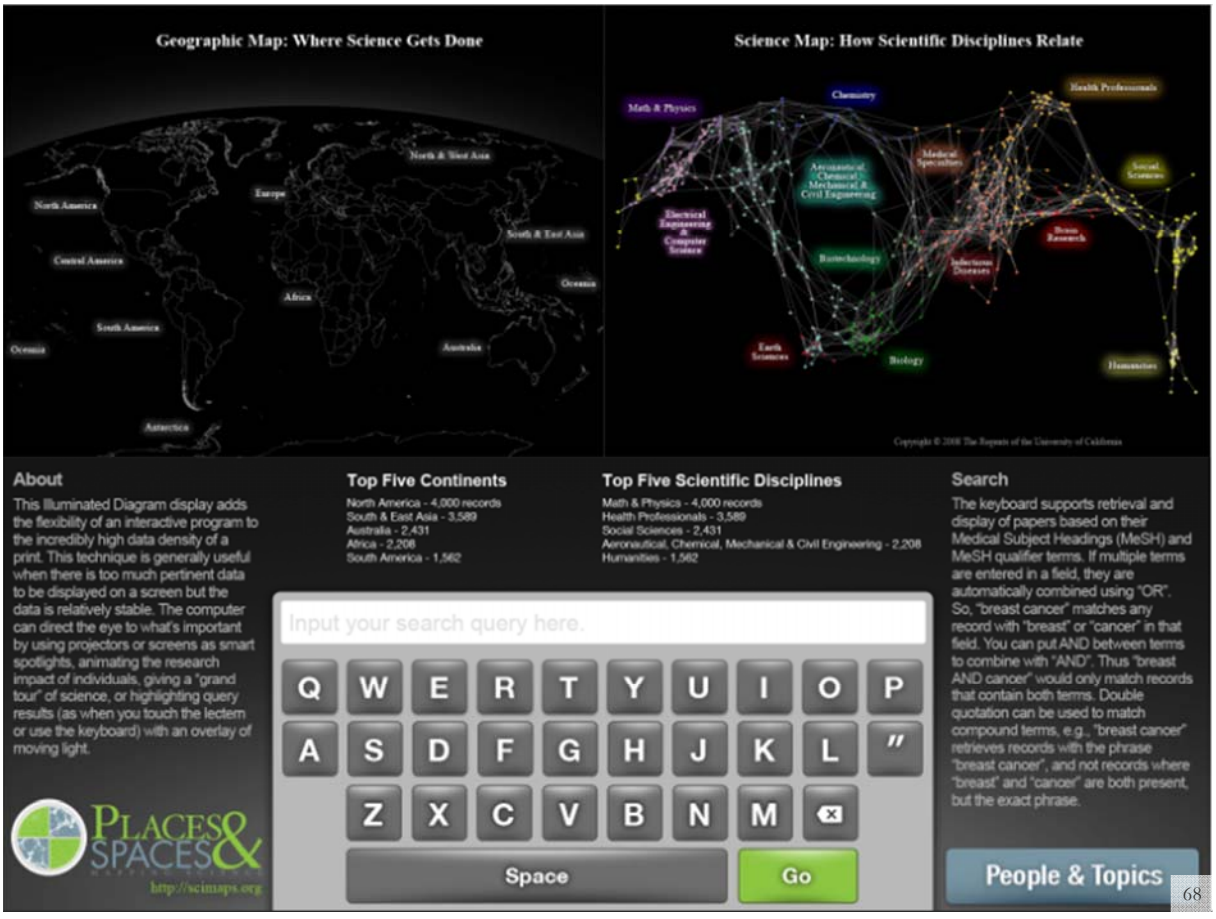


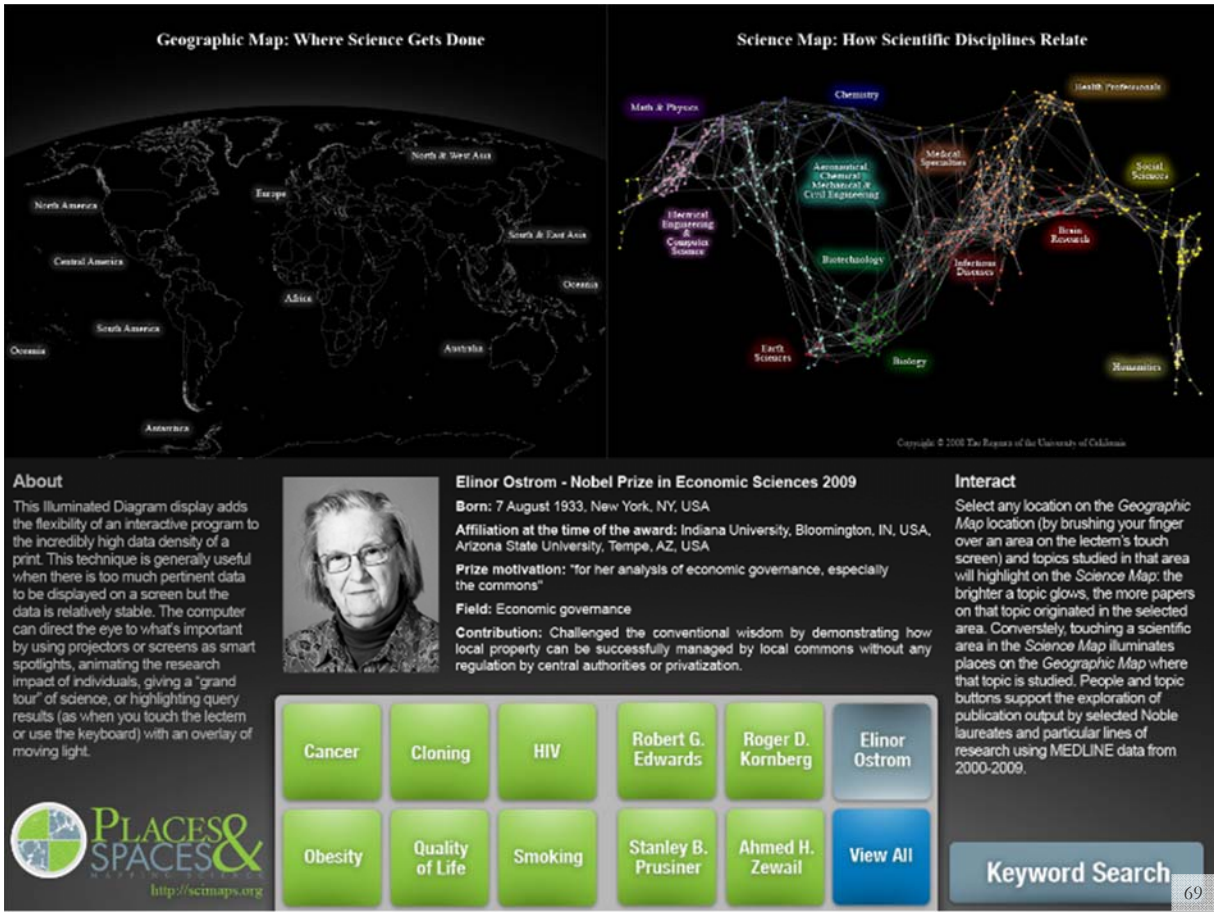
Bollen, Johan, Herbert Van de Sompel, Aric Hagberg, Luis M.A. Bettencourt, Ryan Chute, Marko A. Rodriquez, Lyudmila Balakireva. 2008. A Clickstream Map of Science.



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

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





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

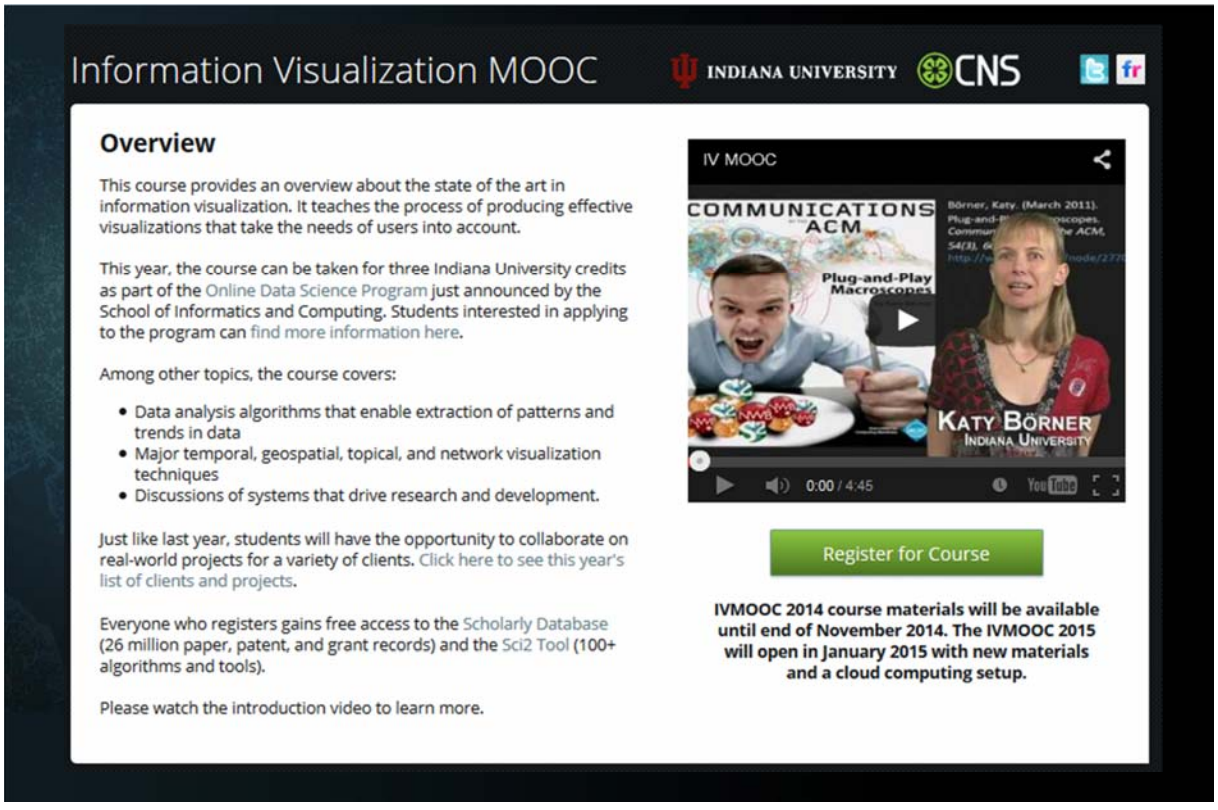
Places & Spaces *Digital Display in North Carolina State's brand new Immersion Theater*

Places & Spaces: Mapping Science Exhibit

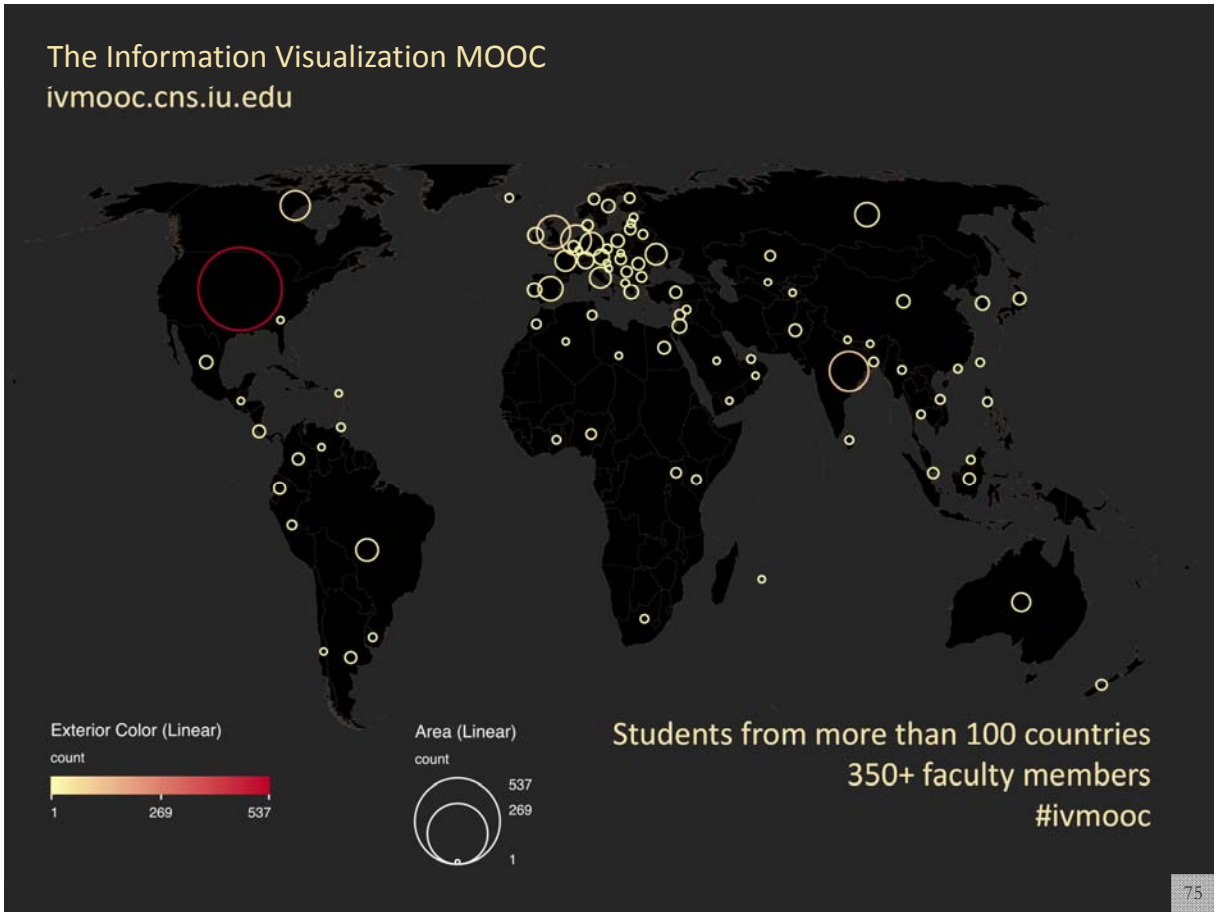
http://scimaps.org

Maps are available for sale and the exhibit can be hosted by anyone.

The Information Visualization MOOC



Register for free at http://ivmooc.cns.iu.edu. Class will restart in January 2015.



Course Schedule

- **Session 1** – Workflow design and visualization framework
- **Session 2** – "When:" Temporal Data
- **Session 3** – "Where:" Geospatial Data
- **Session 4** – "What:" Topical Data

Mid-Term

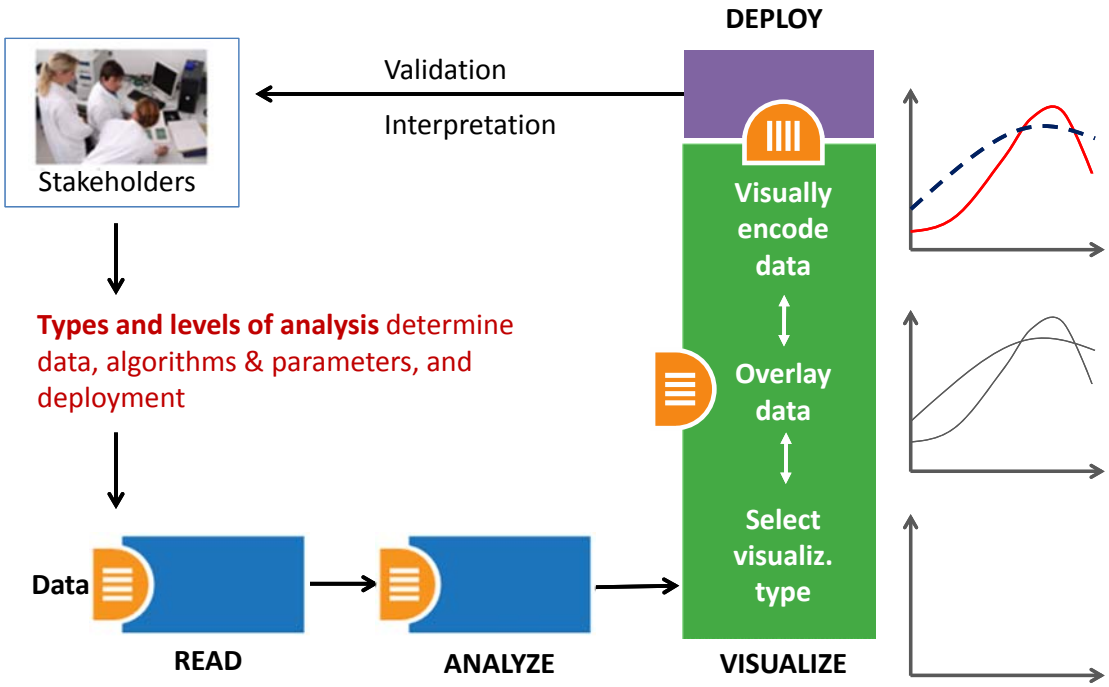
Students work in teams with clients.

- **Session 5** – "With Whom:" Trees
- **Session 6** – "With Whom:" Networks
- **Session 7** – Dynamic Visualizations and Deployment

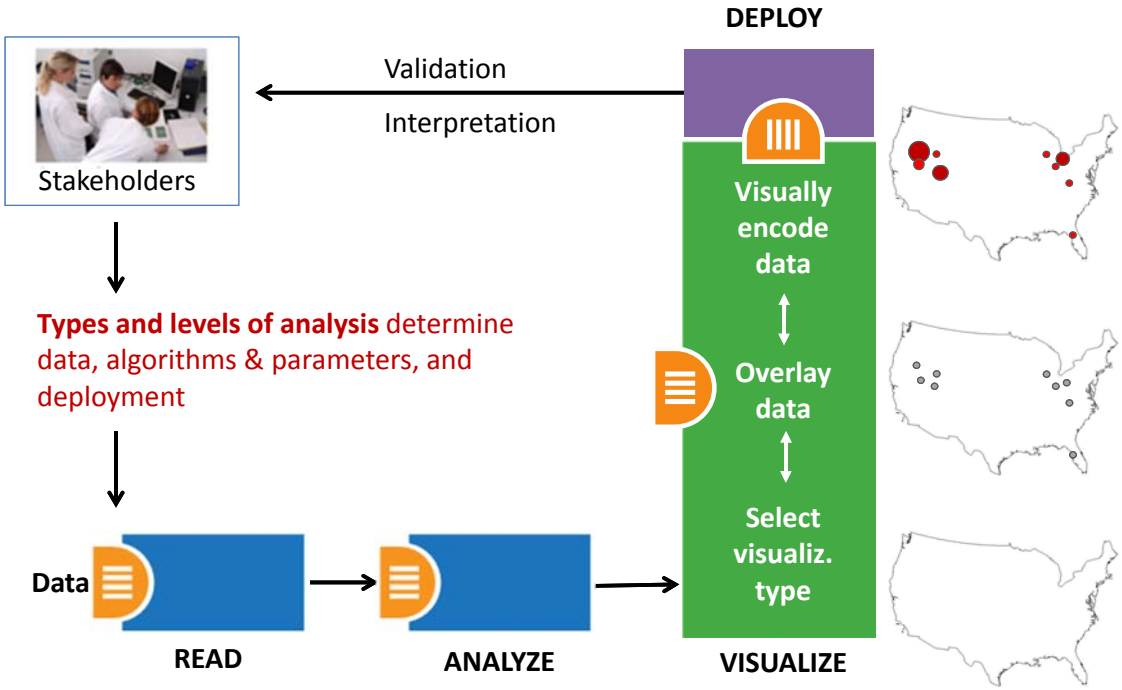
Final Exam

Final grade is based on Midterm (**30%**), Final (**40%**), Client Project (**30%**).

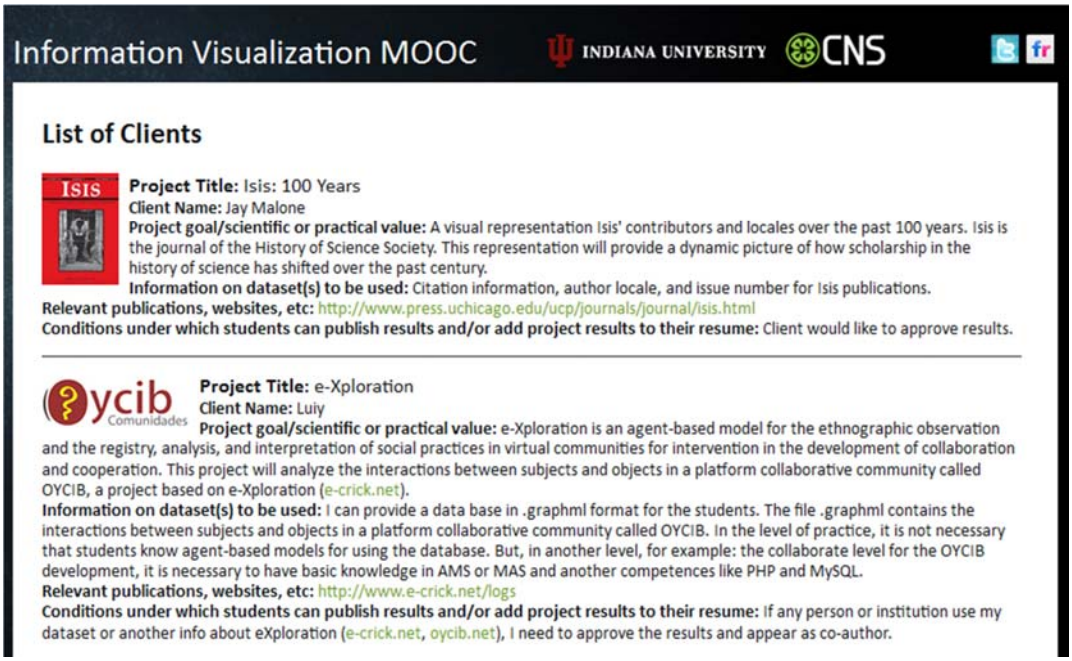
Needs-Driven Workflow Design



Needs-Driven Workflow Design

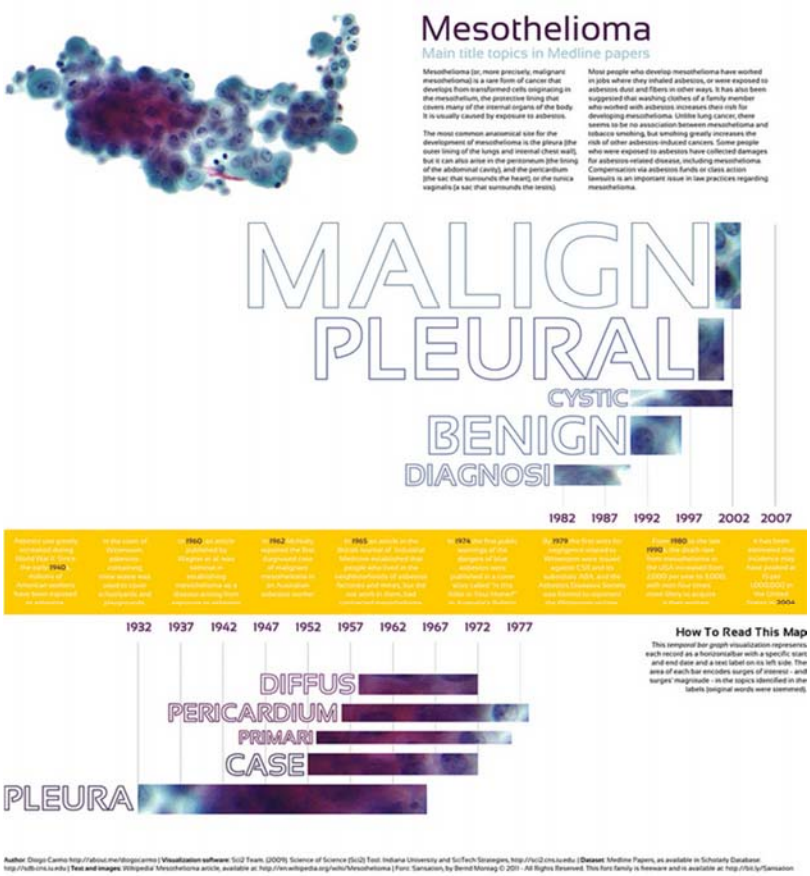


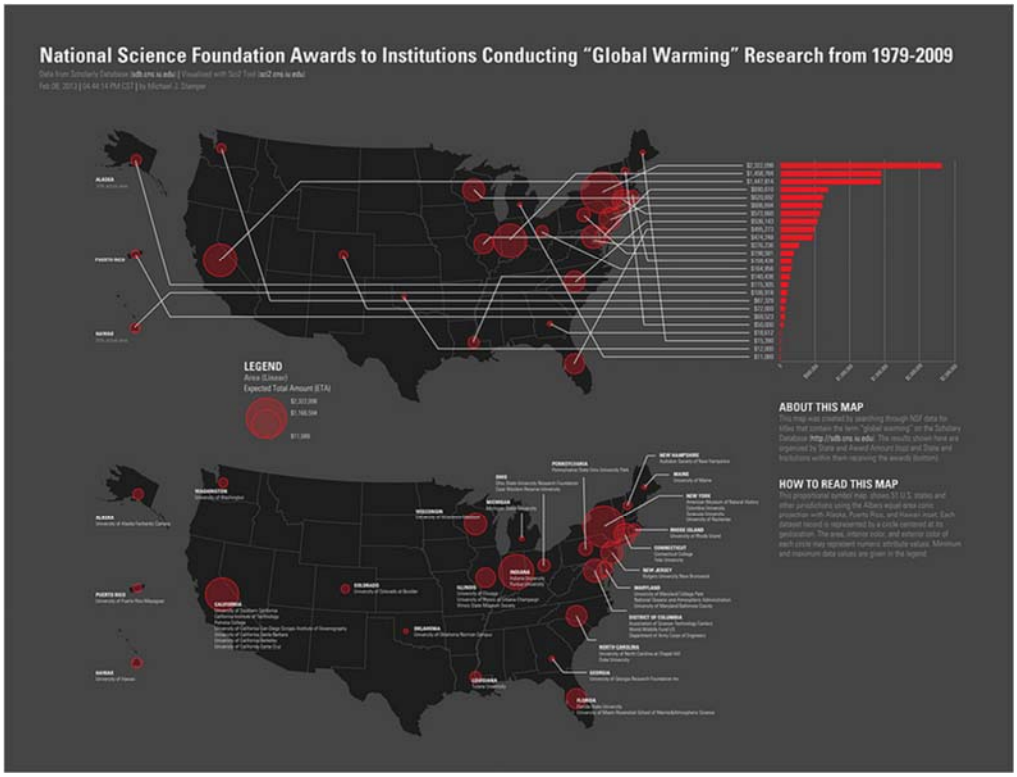
Clients



http://ivmooc.cns.iu.edu/clients.html

Diogo Carmo





mjstamper_ivmooc

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/

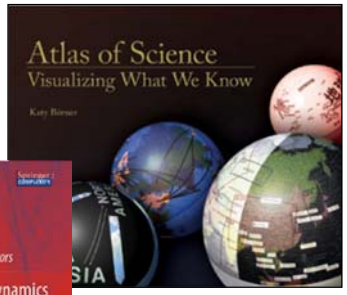
Börner, Katy, Sanyal, Soma and Vespignani, Alessandro (2007). **Network Science.** In Blaise Cronin (Ed.), *ARIST*, Information Today, Inc., Volume 41, Chapter 12, pp. 537-607.
http://ivl.slis.indiana.edu/km/pub/2007-borner-arist.pdf

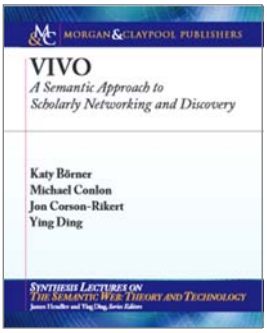
Börner, Katy (2010) **Atlas of Science**. MIT Press. http://scimaps.org/atlas

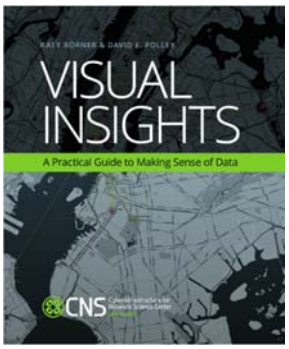
Scharnhorst, Andrea, Börner, Katy, van den Besselaar, Peter (2012) **Models of Science Dynamics**. Springer Verlag.

Katy Börner, Michael Conlon, Jon Corson-Rikert, Cornell, Ying Ding (2012) **VIVO: A Semantic Approach to Scholarly Networking and Discovery**. Morgan & Claypool.

Katy Börner and David E Polley (2014) **Visual Insights: A Practical Guide to Making Sense of Data**. MIT Press.







All papers, maps, tools, talks, press are linked from http://cns.iu.edu
These slides will soon be at http://cns.iu.edu/docs/presentations

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