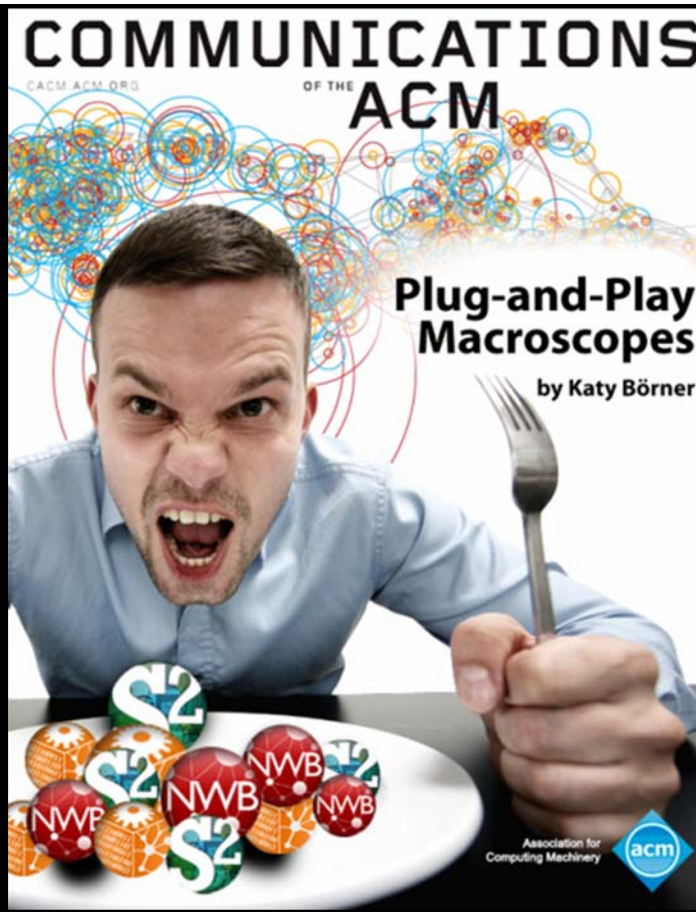
Börner, Katy. (March 2011). Plug-and-Play Macroscopes. *Communications of the ACM, 54(3), 60-69.* Video and paper are at *http://www.scivee.tv/node/27704*



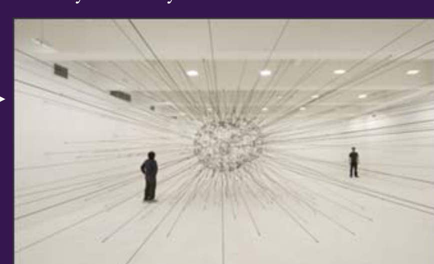
Taming Complexity
TTI VANGUARD
Conference in D.C.
2011.10.04

CNS.IU.EDU
CIShell.org
SciMaps.org

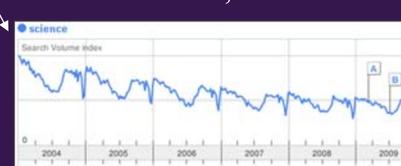
Take terra bytes of data

Plug-and-Play Macroscopes

Find your way



Find collaborators, friends



Identify trends

Macroscopes

Decision making in science, industry, and politics, as well as in daily life, requires that we make sense of data sets representing the structure and dynamics of complex systems. Macroscopes provide a "vision of the whole," helping us "synthesize" the related elements and enabling us to detect patterns, trends, and outliers while granting access to myriad details. Rather than make things larger or smaller, **macroscopes let us observe what is at once too great, slow, or complex for the human eye and mind to notice and comprehend.**

Microscopes

Telescopes

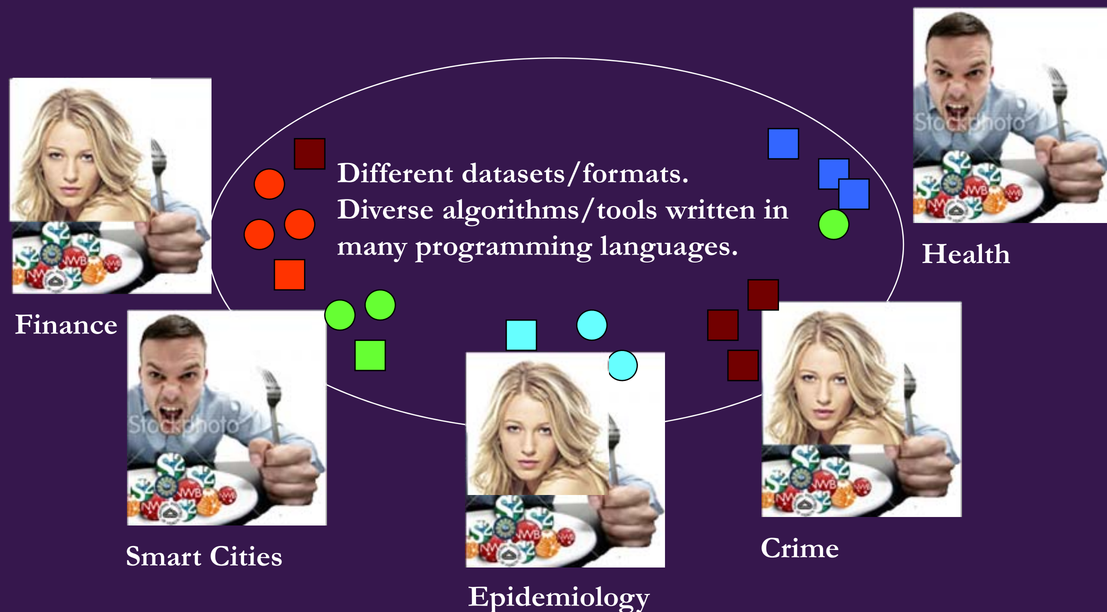
Macroscopes

Plug-and-Play Macroscopes

While microscopes and telescopes are **physical instruments**, macroscopes resemble **continuously changing bundles of software plug-ins**. Macroscopes make it easy to select and combine algorithm and tool plug-ins but also interface plug-ins, workflow support, logging, scheduling, and other plug-ins needed for scientifically rigorous yet effective work.

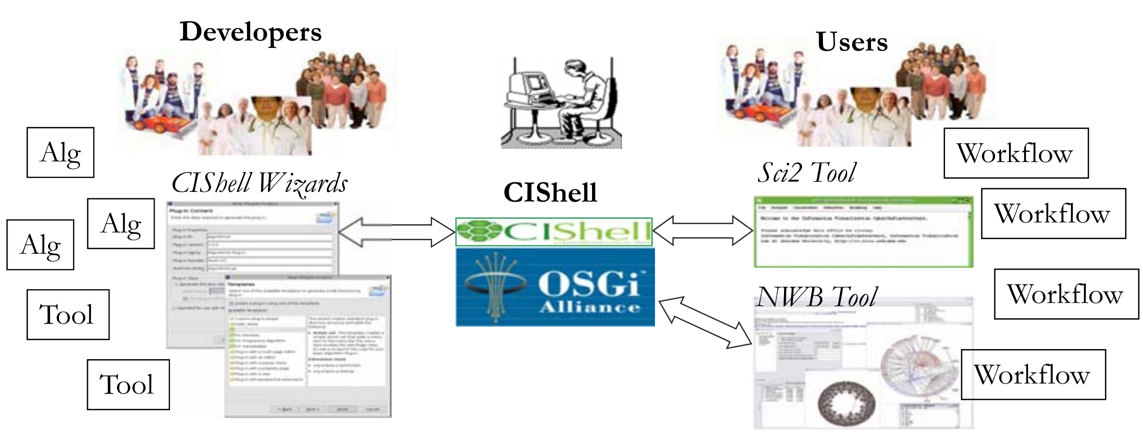
They make it easy to share plug-ins via email, flash drives, or online. To use new plugins, simply copy the files into the plug-in directory, and they appear in the tool menu ready for use. No restart of the tool is necessary. **Sharing algorithm components, tools, or novel interfaces becomes as easy as sharing images on Flickr or videos on YouTube. Assembling custom tools is as quick as compiling your custom music collection.**

Plug-and-Play Macroscopes



OSGi & Cyberinfrastructure Shell (CIShell)

- CIShell (http://cishell.org) is an open source software specification for the integration and utilization of datasets, algorithms, and tools.
- It extends the Open Services Gateway Initiative (OSGi) (http://osgi.org), a standardized, component oriented, computing environment for networked services widely used in industry since more than 10 years.
- Specifically, CIShell provides "sockets" into which existing and new datasets, algorithms, and tools can be plugged using a wizard-driven process.



Network Workbench Tool

http://nwb.cns.iu.edu

The Network Workbench (NWB) tool supports researchers, educators, and practitioners interested in the study of biomedical, social and behavioral science, physics, and other networks.

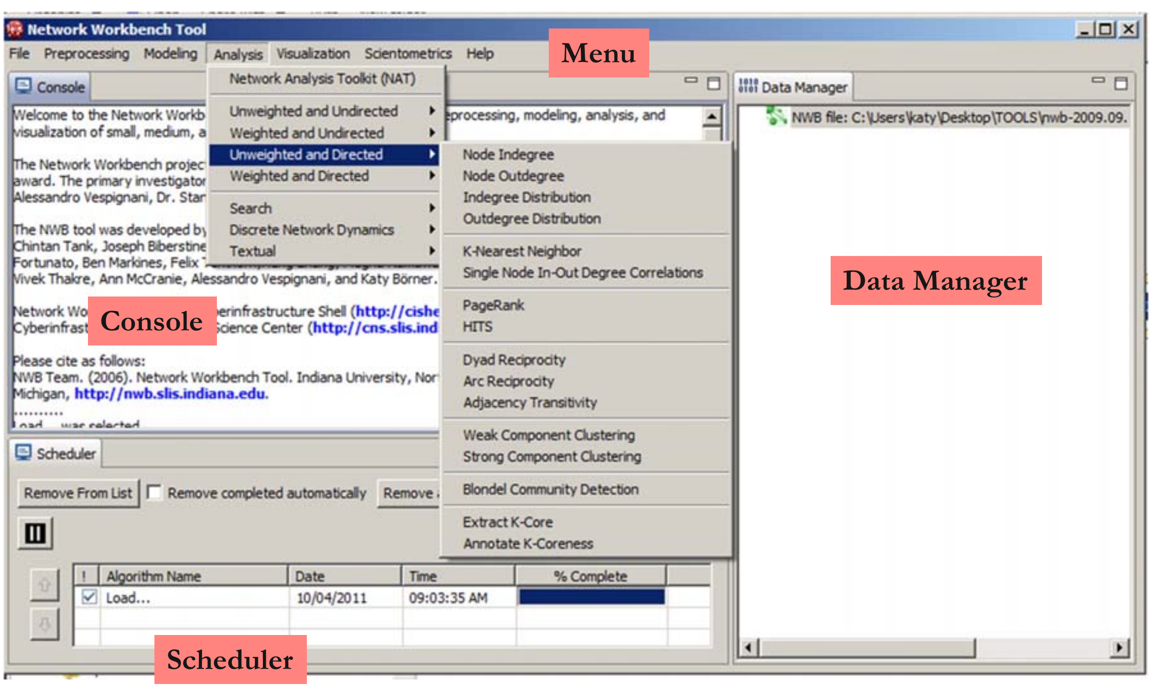
The tool provides more 160 plugins that support the preprocessing, analysis, modeling, and visualization of networks.

It has been downloaded more than 110,000 times.

Herr II, Bruce W., Huang, Weixia (Bonnie), Penumarthy, Shashikant & Börner, Katy. (2007). Designing Highly Flexible and Usable Cyberinfrastructures for Convergence. In Bainbridge, William S. & Roco, Mihail C. (Eds.), Progress in Convergence - Technologies for Human Wellbeing (Vol. 1093, pp. 161-179), Annals of the New York Academy of Sciences, Boston, MA.

Network Workbench Tool

http://nwb.cns.iu.edu



Computational Proteomics

What relationships exist between protein targets of all drugs and all disease-gene products in the human protein–protein interaction network?

Yildriim, Muhammed A., Kwan-Il Goh, Michael E. Cusick, Albert-László Barabási, and Marc Vidal. (2007). Drug-target Network. Nature Biotechnology 25 no. 10: 1119-1126.

Figure 2 Drug–target network (DT network). The DT network is generated by using the known associations between FDA-approved drugs and their target proteins. Circles and rectangles correspond to drugs and target proteins, respectively. A link is placed between a drug node and a target node if the protein is a known target of that drug. The area of the drug (protein) node is proportional to the number of targets that the drug has (the number of drugs targeting the protein). Color codes are given in the legend. Drug nodes (circles) are colored according to their Anatomical Therapeutic Chemical Classification, and the target proteins (rectangular boxes) are colored according to their cellular component obtained from the Gene Ontology database.

Computational Economics

Does the type of product that a country exports matter for subsequent economic performance?

C. A. Hidalgo, B. Klinger, A.-L. Barabási, R. Hausmann (2007) The Product Space Conditions the Development of Nations. Science 317, 482 (2007).

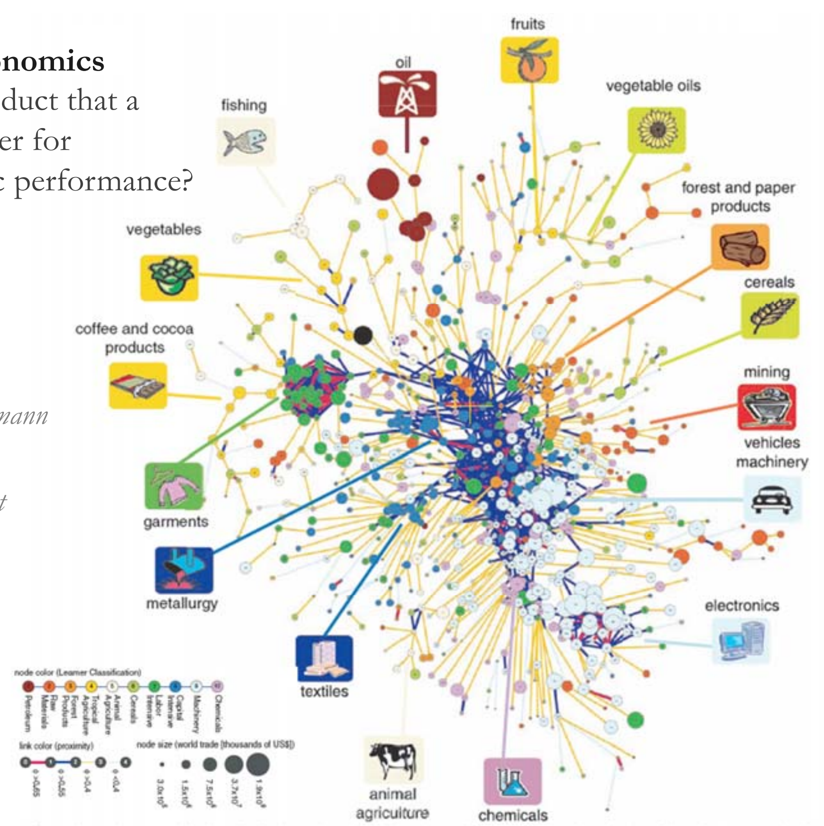
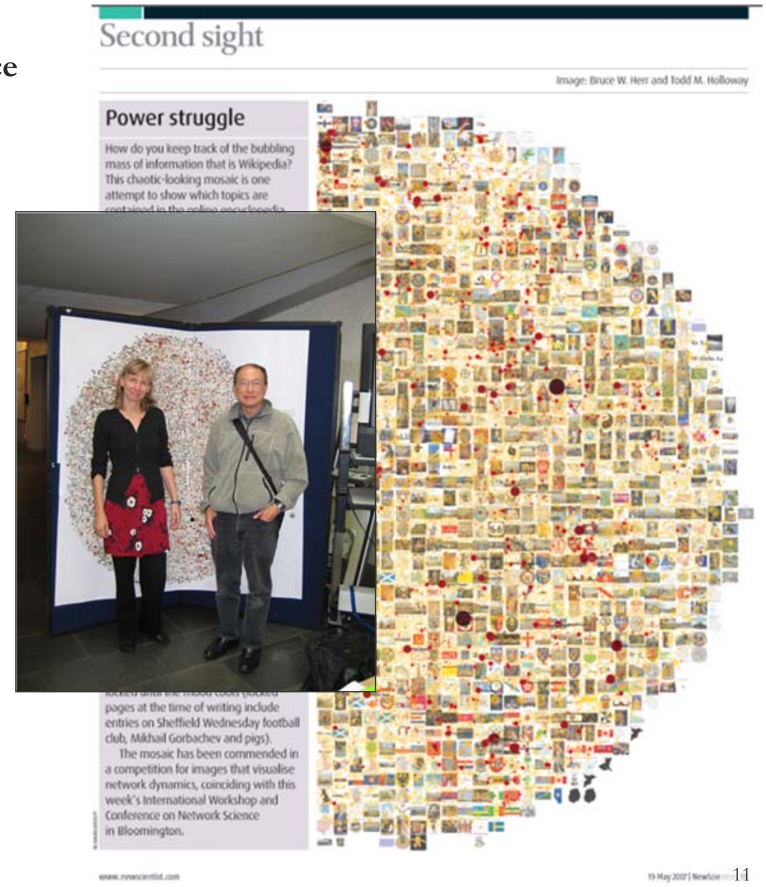


Fig. 1. The product space. **(A)** Hierarchically clustered proximity (φ) matrix representing the 775 SITC-4 product classes exported in the 1998–2000 period. **(B)** Network representation of the product space. Links are color coded with their proximity value. The sizes of the nodes are proportional to world trade, and their colors are chosen according to the classification introduced by Leamer.

Computational Social Science
Studying large scale social networks such as Wikipedia

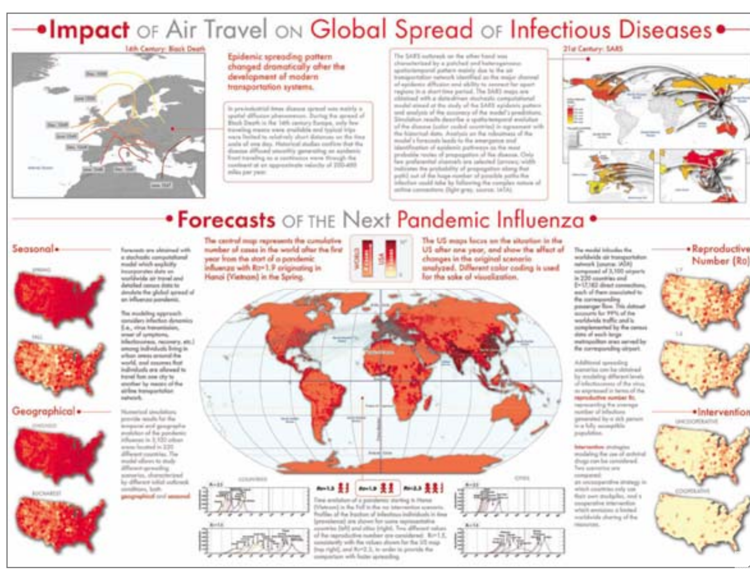


Second Sight: An Emergent Mosaic of Wikipedian Activity,
The NewScientist, May 19, 2007

Computational Epidemics
Forecasting (and preventing the effects of) the next pandemic.

Epidemic Modeling in Complex realities, V. Colizza, A. Barrat, M. Barthelemy, A.Vespignani, Comptes Rendus Biologie, 330, 364-374 (2007).
Reaction-diffusion processes and metapopulation models in heterogeneous networks, V.Colizza, R. Pastor-Satorras, A.Vespignani, Nature Physics 3, 276-282 (2007).

Modeling the Worldwide Spread of Pandemic Influenza: Baseline Case and Containment Interventions, V. Colizza, A. Barrat, M. Barthelemy, A.-J. Valleron, A.Vespignani, PloS-Medicine 4, e13, 95-110 (2007).



Science of Science Tool
http://sci2.cns.iu.edu

Sci2 Tool v0.5.1 Alpha (May 4th, 2011)

Can be freely downloaded for all major operating systems from
http://sci2.cns.iu.edu

Select your operating system from the pull down menu and download.
Unpack into a /sci2 directory.
Run /sci2/sci2.exe

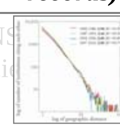
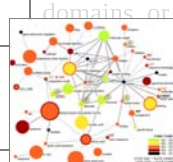
Sci2 Manual is at
http://sci2.wiki.cns.iu.edu

Cite as

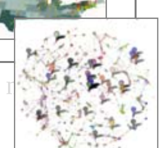
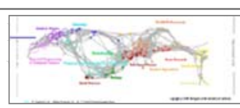
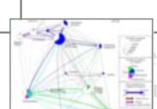
Sci² Team. (2009). Science of Science (Sci²) Tool. Indiana University and SciTech Strategies, [*http://sci2.cns.iu.edu*](http://sci2.cns.iu.edu)

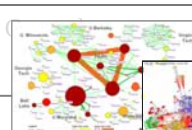
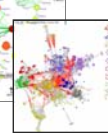
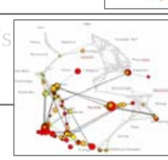
Type of Analysis vs. Level of Analysis

	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 NS SA, all of sci
Temporal Analysis (When)	Funding portfolio of one individual	ic bursts f PNAS	113 Years of P Research
Geospatial Analysis (Where)	Career trajectory of one individual	Mapping a s intellectual l	PNAS
Topical Analysis (What)	s.	research	VxOrd/Topic r NIH funding
Network Analysis (With Whom?)	NSF work of one	k	NIH's cy

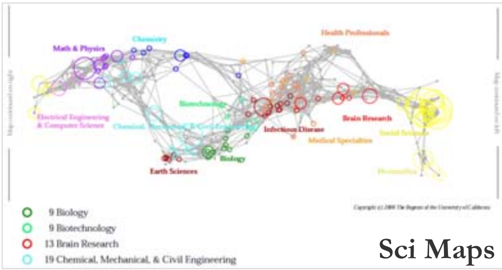


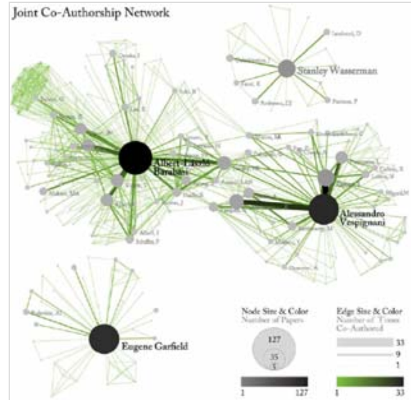




Open Code for Replicable S&T Assessment

OSGi/CIShell powered tool, see http://cishell.org
http://sci2.cns.iu.edu | http://sci2.wiki.cns.iu.edu

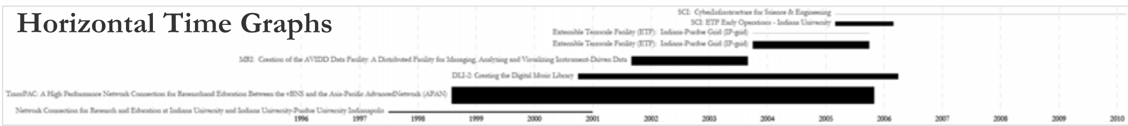


Sci Maps



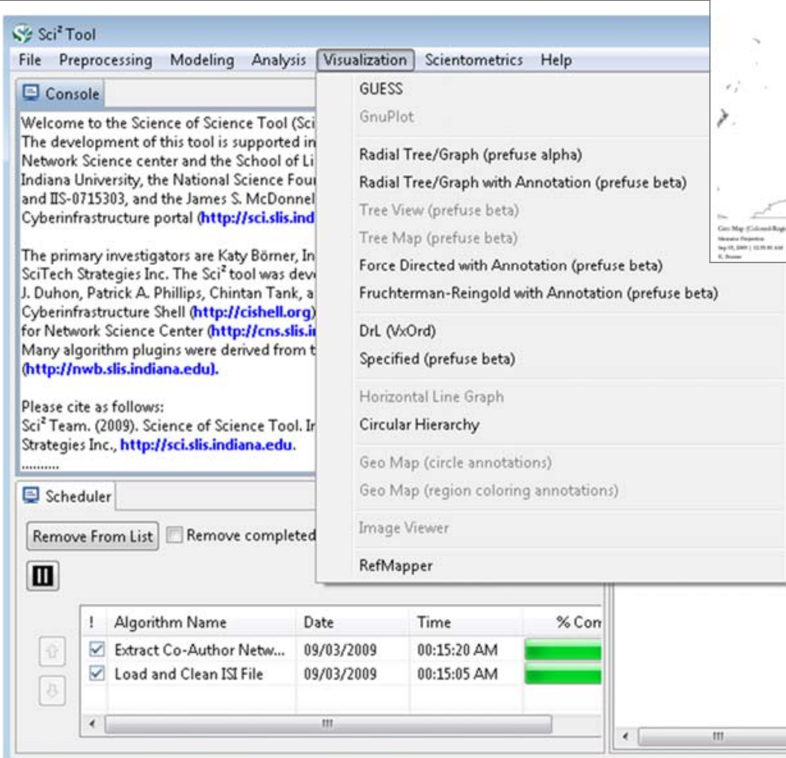
GUESS Network Vis

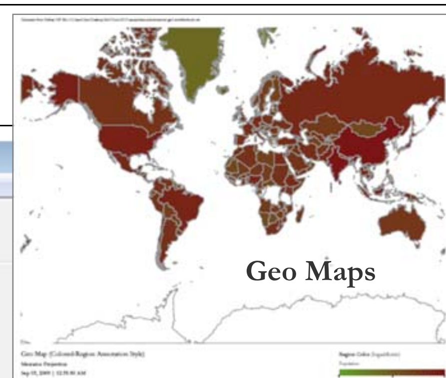
Horizontal Time Graphs

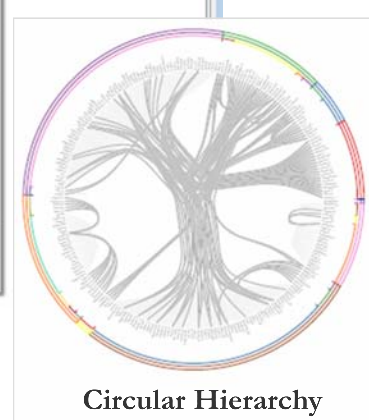


Börner, Katy, Huang, Weixia (Bonnie), Linnemeier, Micah, Duhon, Russell Jackson, Phillips, Patrick, Ma, Nianli, Zoss, Angela, Guo, Hanning & Price, Mark. (2009). Rete-Netzwerk-Red: Analyzing and Visualizing Scholarly Networks Using the Scholarly Database and the Network Workbench Tool. Proceedings of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

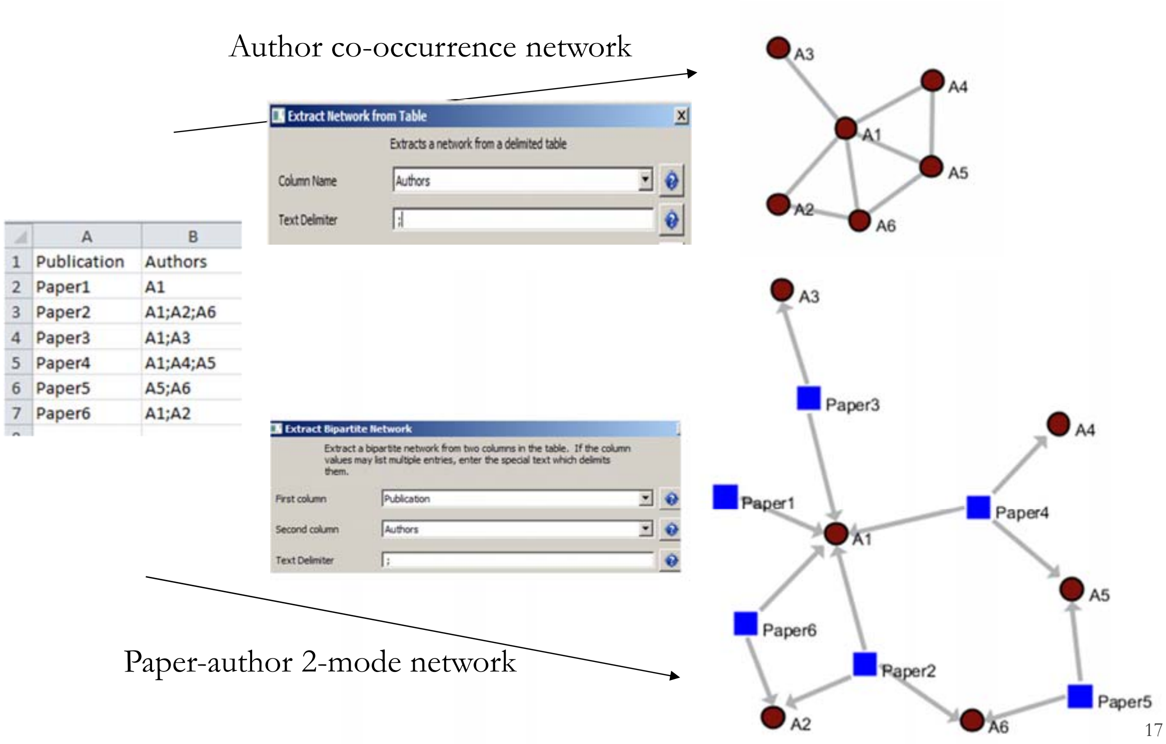
Sci² Tool




Geo Maps


Circular Hierarchy

Network Extraction: Examples



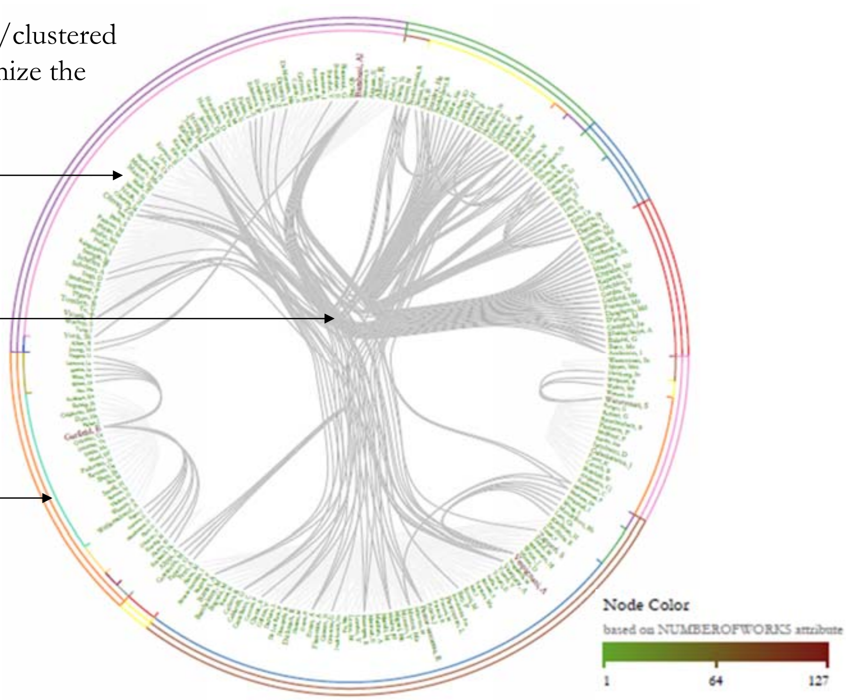
Network Visualization: Circular Hierarchy Visualization

Nodes that are interlinked/clustered are spatially close to minimize the number of edge crossings.

Node labels, e.g., author names.

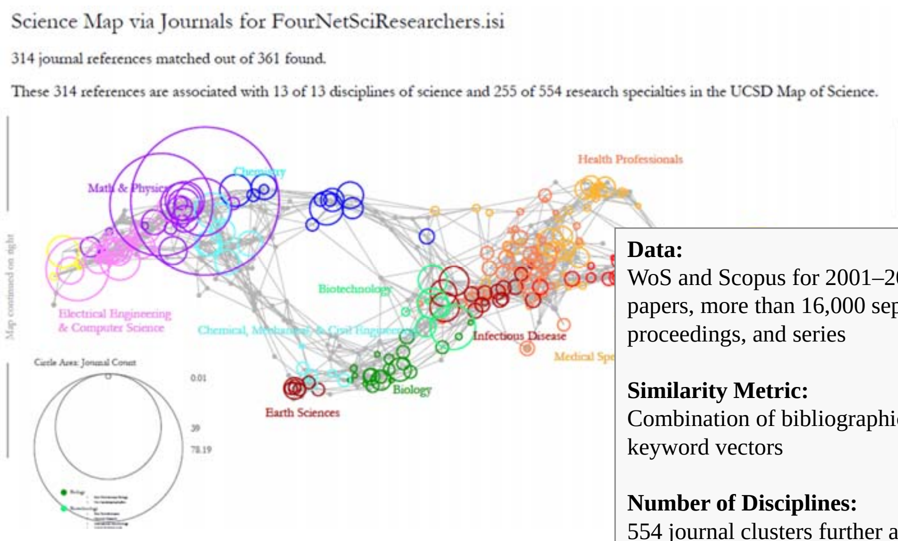
Network structure using edge bundling.

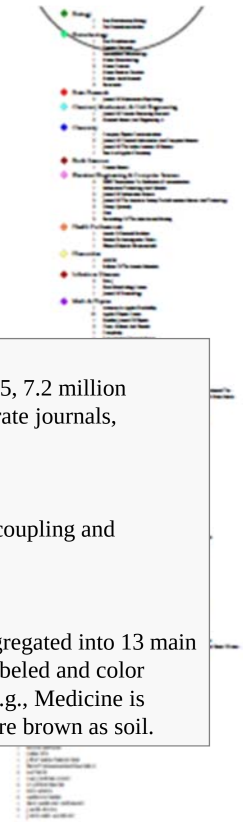
Color coded cluster hierarchy according to Blondel community detection algorithm.

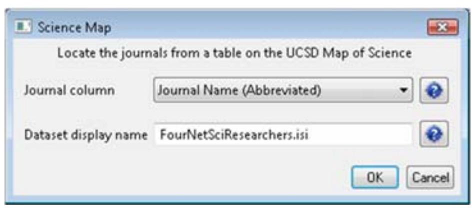


Note:
Header/footer info, legend, and more meaningful color coding are under development.

Topic Mapping: UCSD Science Map







Data:
WoS and Scopus for 2001–2005, 7.2 million papers, more than 16,000 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.

Geospatial Maps with Congressional Districts

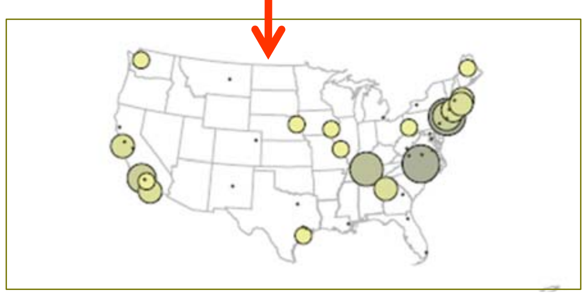
	A
1	Zip code
2	90095
3	4672
4	232980568
5	10032
6	10039242
7	46091500
8	191112434
9	27705
10	981959472
11	10065
12	10065

Identify Congressional District, Latitude, Longitude

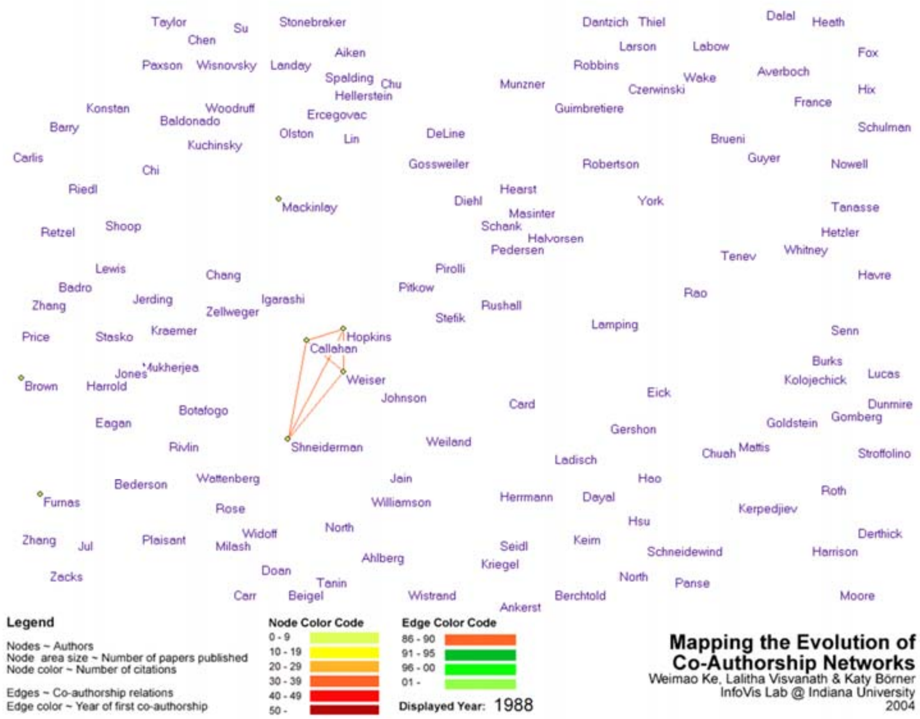
	A	B	C	D
1	Zip code	Congressional District	Latitude	Longitude
2	90095	CA-30	34.0735035	-118.6645815
3	4672	ME-02	45.818717	-69.0290345
4	232980568	VA-03	37.270472	-77.0699835

Aggregate/Count identical Congressional Districts

	A	B	C	D
1	Congressional District	Latitude	Longitude	Count
2	CA-30	34.0735035	-118.6645815	4
3	ME-02	45.818717	-69.0290345	2
4	VA-03	37.270472	-77.0699835	1
5	NY-15	40.8341475	-73.9342095	4

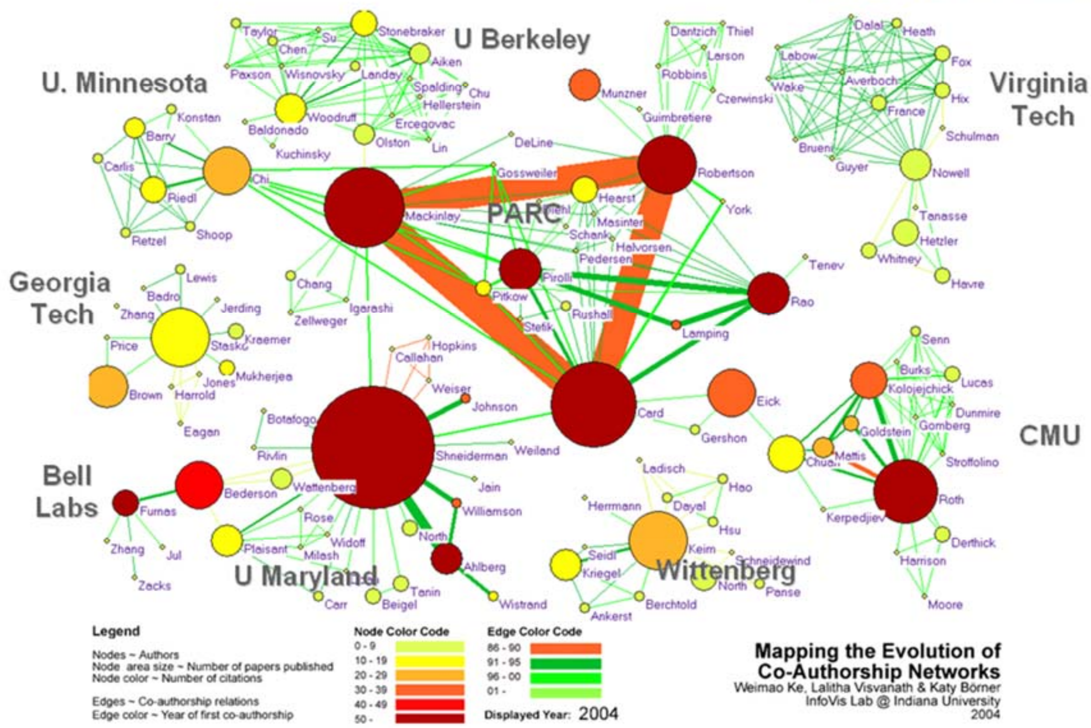


Evolving Collaboration Networks



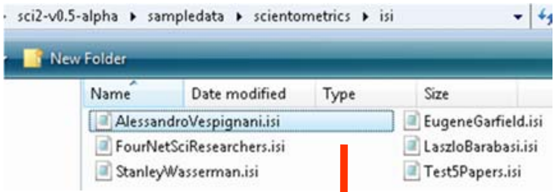
Mapping the Evolution of Co-Authorship Networks

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



Evolving Collaboration Networks

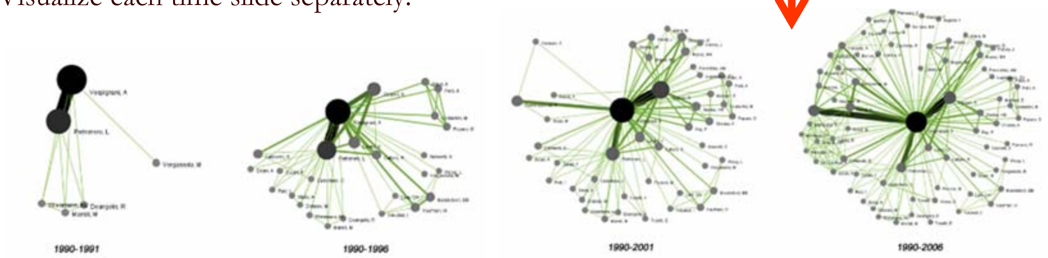
Load isi formatted file



As csv, file looks like:

	A	B	C	D	E	F	G
1	Abstract	Authors	Authors (Full Names)	Beginning	Book Serie	Book Serie	Cited Pate
2	The systematic study of	Colizza, V|Barrat, A|Barthelemy, M|Vespignani, A		2015			
3	Uncovering the hidden r	Colizza, V|Flammini, A|Serrano, MA|Vespignani, A		110			
4	Computer viruses can s	Vespignani, A		135			
5	Mapping the Internet ge	Dall'Asta, L|Alvarez-Hamelin, I|Barrat, A|Vazquez, A|Vespignani, A		140		LECTURE NOTES IN	

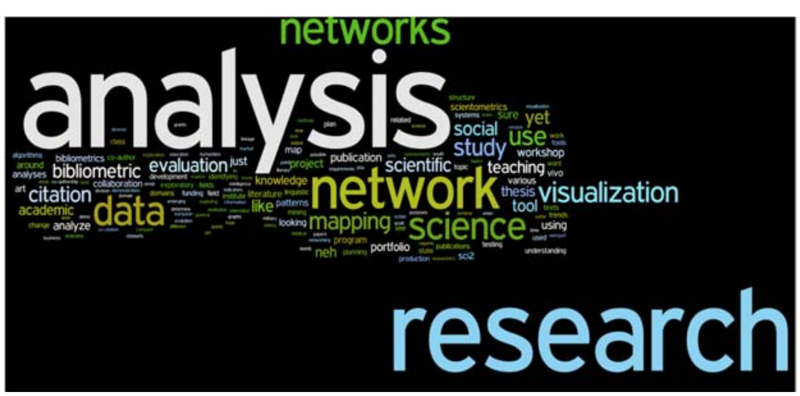
Visualize each time slide separately:

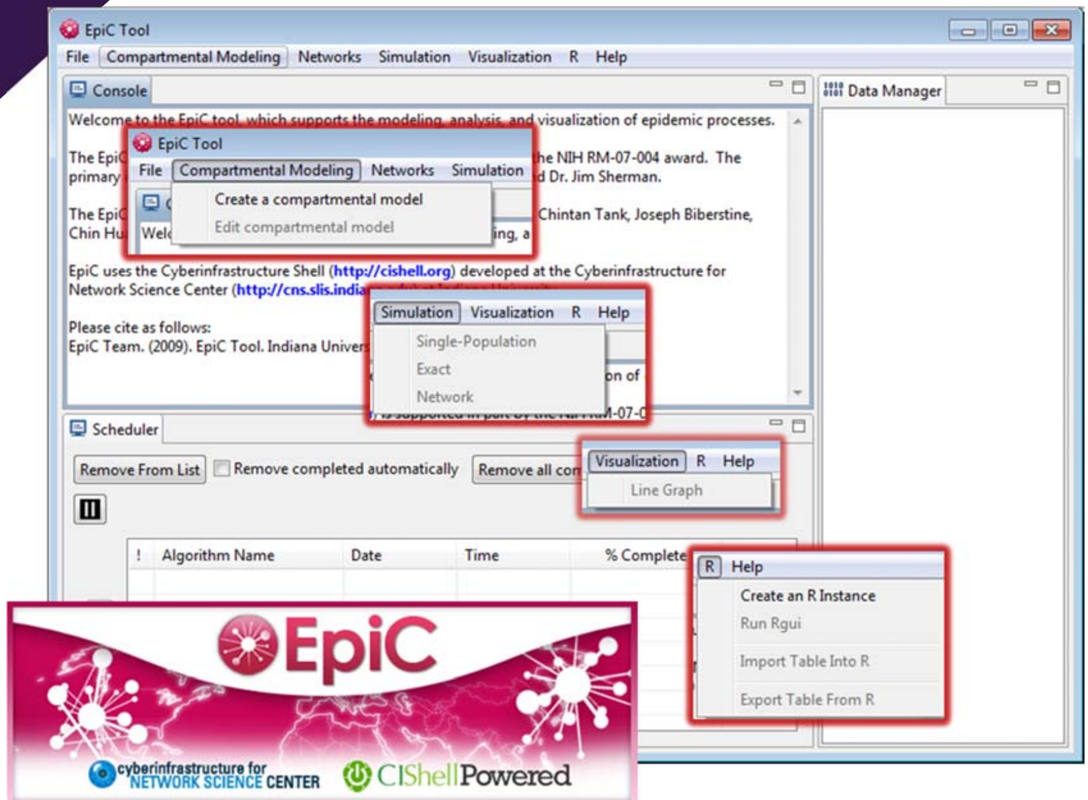


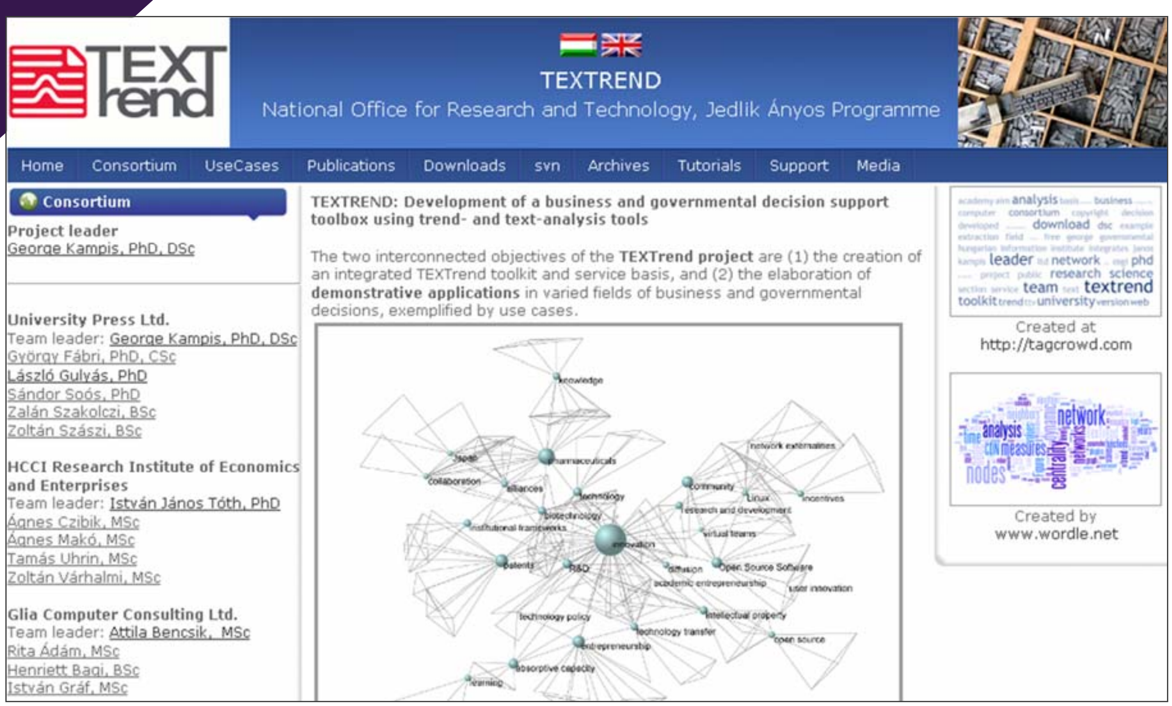
Sci2 Tool Adoption

The Sci2 Tool is used by the National Science Foundation, the National Institutes of Health, and the US Department of Agriculture.

"As a new user, I am beginning with very little knowledge of the analyses and modeling techniques that Sci2 enables. I've been able to use my own dataset and follow through some of the workflows to the point of generating the first network and time horizon visuals. That was so exciting I stayed up far past bedtime to get to the visuals."
Dr. Suzanne A. Pierce, Center for International Energy and Environmental Policy
Jackson School of Geosciences, The University of Texas at Austin







TEXTrend adds WEKA, UIMA, Wordij, CFinder, and more.

See the latest versions of TEXTrend Toolkit modules at *http://textrend.org*

CIShell – Integrate New Algorithms

About the Cyberinfrastructure Shell

The Cyberinfrastructure Shell (CIShell) is an open source, community-driven platform for the integration and utilization of datasets, algorithms, tools, and computing resources. Algorithm integration support is built in for Java and most other programming languages. Being Java based, it will run on almost all platforms. The software and specification is released under an Apache 2.0 License.

CIShell is the basis of Network Workbench, TexTrend, Sci² and the upcoming EpiC tool.

CIShell supports remote execution of algorithms. A standard web service definition is in development that will allow pools of algorithms to transparently be used in a peer-to-peer, client-server, or web front-end fashion.

CIShell Features

A framework for easy integration of new and existing algorithms written in any programming language

Using CIShell, an algorithm writer can fully concentrate on creating their own algorithm in whatever language they are comfortable with. Simple tools are provided to then take their algorithm and

Learn More...

- CIShell Papers
- CIShell Powered Tools
- Algorithms
- Plugins (coming soon)
- Misc. Tool Documentation
- CIShell Web Services (coming soon)
- Screenshots

Getting Started...

- Documentation & Developer Resources
- Download

Getting Involved...

- Contact Us

CIShell Developer Guide is at http://cishell.wiki.cns.iu.edu

Additional Sci2 Plugins are at http://sci2.wiki.cns.iu.edu/3.2+Additional+Plugins

OSGi/CIShell Adoption

A number of other projects recently adopted OSGi and/or CIShell:

USA

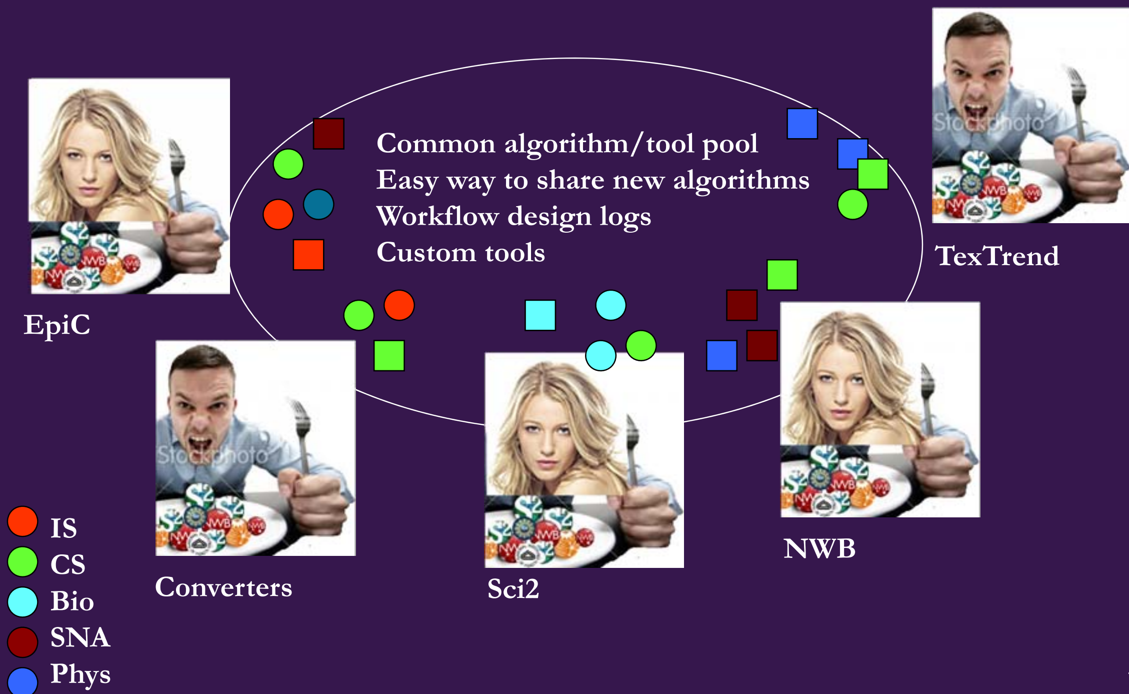
- *Cytoscape* (*http://cytoscape.org*) Led by Trey Ideker at the University of California, San Diego is an open source bioinformatics software platform for visualizing molecular interaction networks and integrating these interactions with gene expression profiles and other state data (Shannon et al., 2002).
- *MAEviz* (*https://wiki.ncsa.uiuc.edu/display/MAE/Home*) Managed by Jong Lee at NCSA is an open-source, extensible software platform which supports seismic risk assessment based on the Mid-America Earthquake (MAE) Center research.

Europe

- *Taverna Workbench* (*http://taverna.org.uk*) Developed by the myGrid team (http://mygrid.org.uk) led by Carol Goble at the University of Manchester, U.K. is a free software tool for designing and executing workflows (Hull et al., 2006). Taverna allows users to integrate many different software tools, including over 30,000 web services.
- *TEXTrend* (*http://textrend.org*) Led by George Kampis at Eötvös Loránd University, Budapest, Hungary supports natural language processing (NLP), classification/mining, and graph algorithms for the analysis of business and governmental text corpuses with an inherently temporal component.
- *DynaNets* (*http://www.dynanets.org*) Coordinated by Peter M.A. Sloot at the University of Amsterdam, The Netherlands develops algorithms to study evolving networks.
- *SISOB* (*http://sisob.lcc.uma.es*) An Observatory for Science in Society Based in Social Models.

As the functionality of OSGi-based software frameworks improves and the number and diversity of dataset and algorithm plugins increases, the capabilities of custom tools will expand.

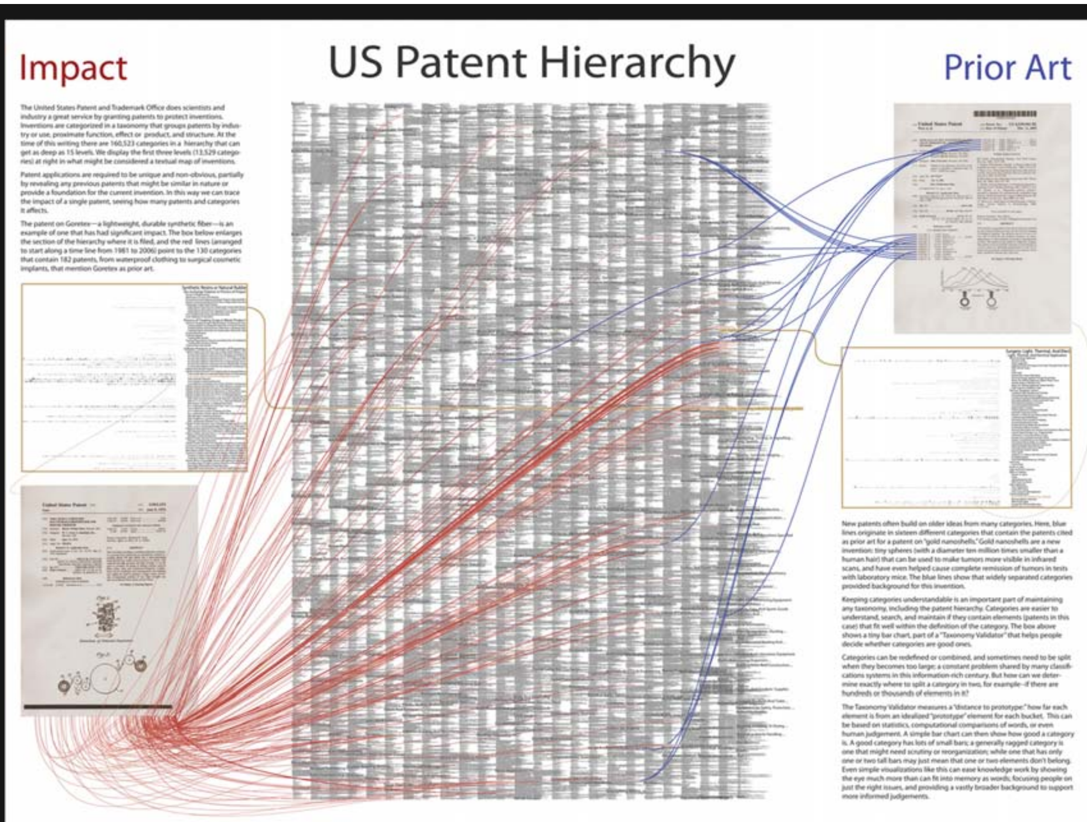
Plug-and-Play Macroscopes



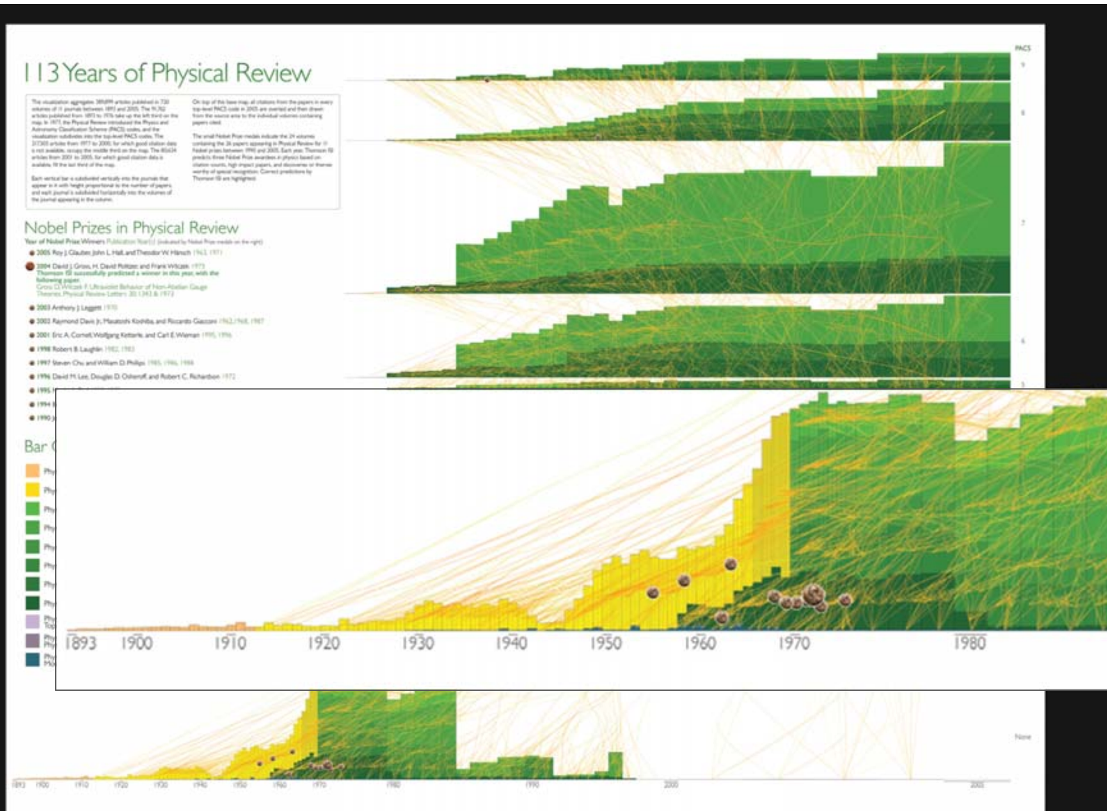
CNS.IU.EDU, CIShell.org, SciMaps.org

Maps created using Sci2 are travelling in the “Expedition Zukunft” science train visiting 62 cities in 7 months, 12 coaches, 300 m long. http://www.expedition-zukunft.de

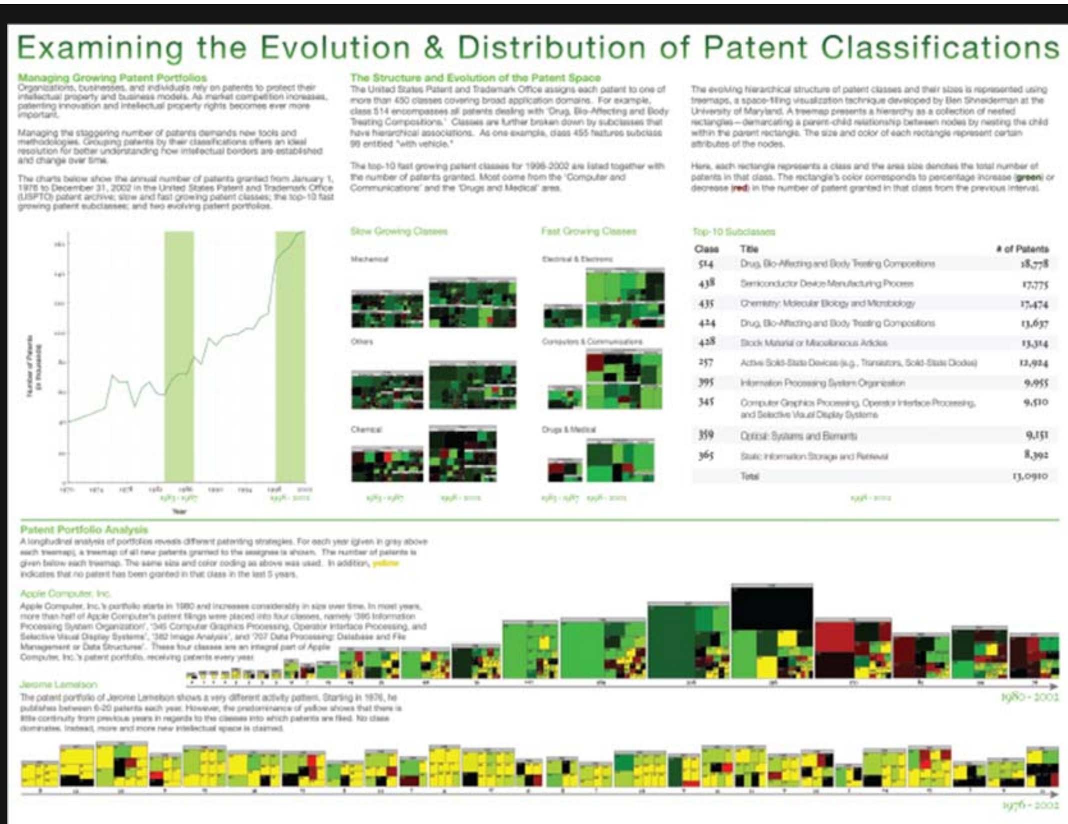
And they are part of the international **Mapping Science exhibit**: http://scimaps.org



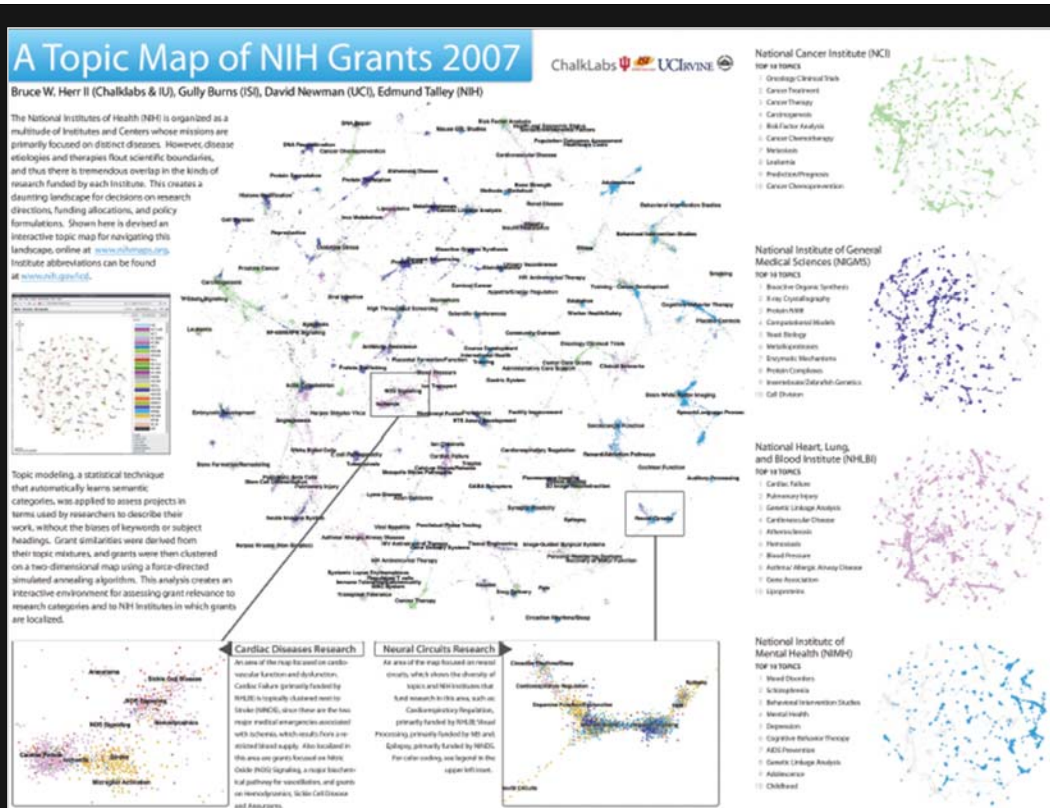
Katy Börner, Elisha Hardy, Bruce W. Herr II, Todd Holloway, and W. Bradford Paley (2006)
Taxonomy Visualization of Patent Data.



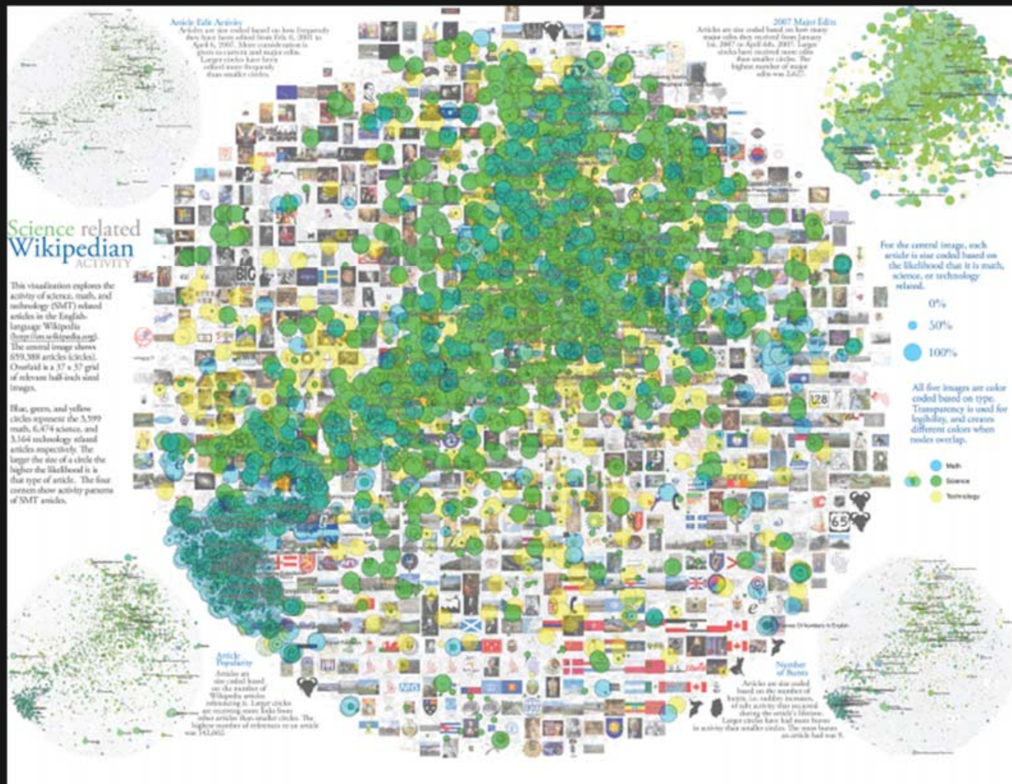
Bruce W. Herr, Russell Duhon, Elisha F. Hardy, Shashikant Penumarthy, and Katy Börner (2007)
113 Years of Physical Review.



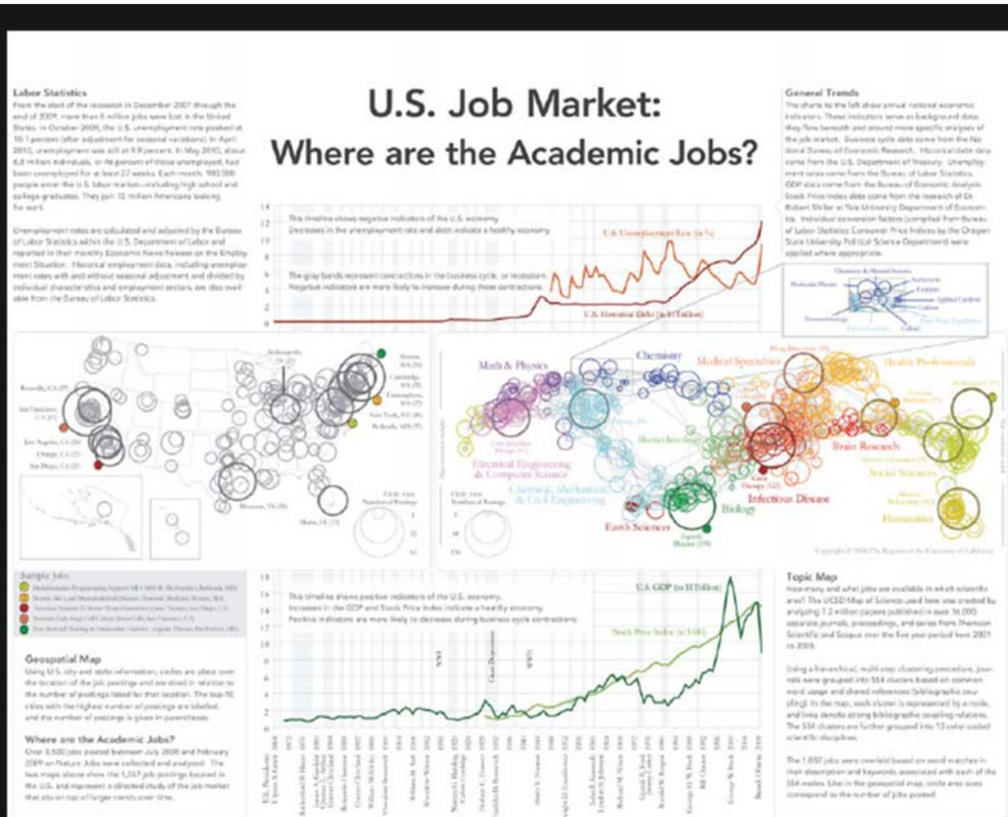
Daniel O. Kutz, Katy Börner & Elisha Hary (2004) Examining the Evolution and Distribution of Patent Classifications.



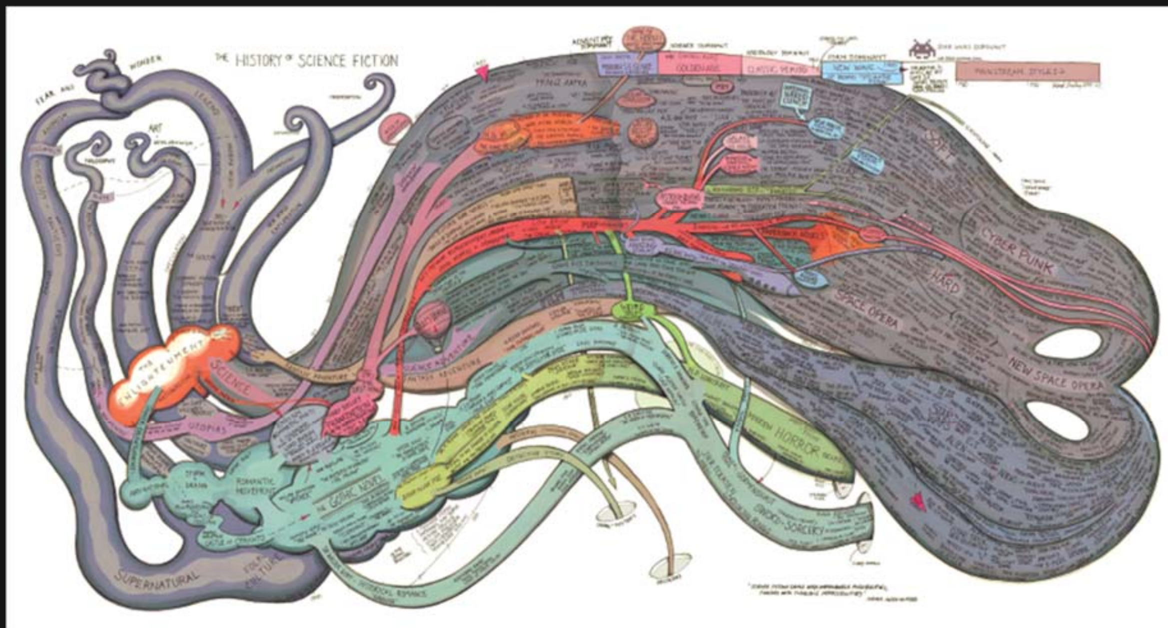
Bruce W. Herr II, Gully Burns, David Newman, Edmund Talley (2007) A Topic Map of NIH Grants.



Bruce W. Herr II, Todd M. Holloway, Elisha F. Hardy, Kevin W. Boyack, and Katy Börner (2007) Science Related Wikipedian Activity.



Angela M. Zoss and Katy Börner (2010) U.S. Job Market: Where are the Academic Jobs?



Ward Shelley. 2011. History of Science Fiction.

Debut of 5th Iteration of the Mapping Science Exhibit at MEDIA X took place at Wallenberg Hall, Stanford University, http://mediax.stanford.edu, http://scaleindependentthought.typepad.com/photos/scimaps

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 (2011) **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