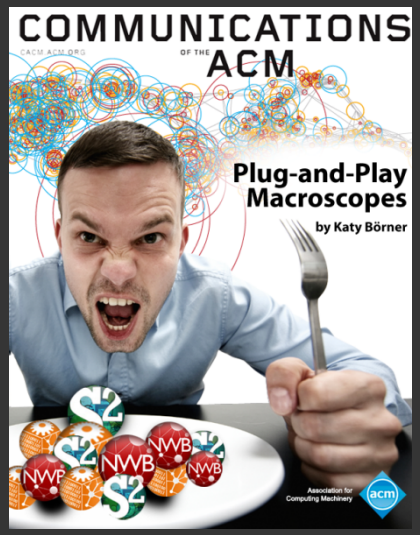
Building a Science Observatory: Research, Tools, and Maps

Katy Börner
Cyberinfrastructure for Network Science Center, Director
Information Visualization Laboratory, Director
School of Informatics and Computing
Indiana University, Bloomington, IN
katy@indiana.edu

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

Network Science Talk at IU

December 8, 2014



NSF Workshop Report on "Knowledge Management and Visualization Tools in Support of Discovery"

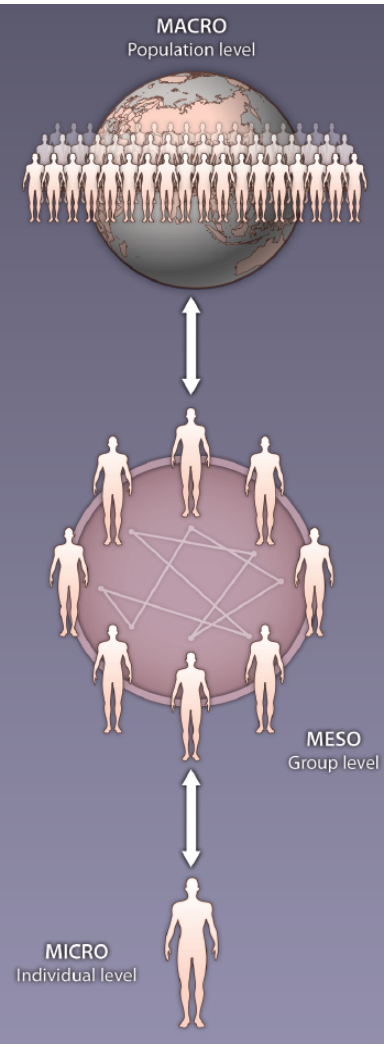
Börner, Bettencourt, Gerstein, and Uzzo (Eds.)

(http://vw.cns.iu.edu/cdi2008/whitepaper.html)

published in Dec 2009 argues for a

- A decentralized, free **"Scholarly Database"** to keep track, interlink, understand and improve the quality and coverage of Science and Technology (S&T) relevant data. (see also page 76 and 77 in Appendix D)
- A **"Science Marketplace"** that supports the sharing of expertise and resources and is fueled by the currency of science: scholarly reputation. (see page 74 in Appendix D) This marketplace might also be used by educators and the learning community to help bring science to the general public and out of the "ivory tower". (see page 89 in Appendix D)
- A **"Science Observatory"** that analyzes different datasets in real-time to assess the current state of S&T and to provide an outlook for their evolution under several (actionable) scenarios. (see page 72 in Appendix D)

- **"Validate Science [of Science Results and] Maps"** to understand and utilize their value for communicating science studies and models across scientific boundaries, but also to study and communicate the longitudinal (1980-today) impact of funding on the science system. (see page 81 in Appendix D)
- An easy to use, yet versatile, **"Science Telescope"** to communicate the structure and evolution of science to researchers, educators, industry, policy makers, and the general public at large. (see page 87 in Appendix D) The effect of this (and other science portals) on education and science perception needs to be studied in carefully controlled experiments. (see page 88 in Appendix D)
- **"Science of Science"** studies are necessary to increase our understanding and support the formation of effective research and development teams. (see page 78 and 82 in Appendix D).
- **"Success Criteria"** need to be developed that support a scientific calculation of S&T benefits for society. (see also page 88 in Appendix D)
- A **"Science Life"** (an analog to Second Life) should be created to put the scientist's face on their science. Portals to this parallel world would be installed in universities, libraries and science museums. (see page 80 in Appendix D)



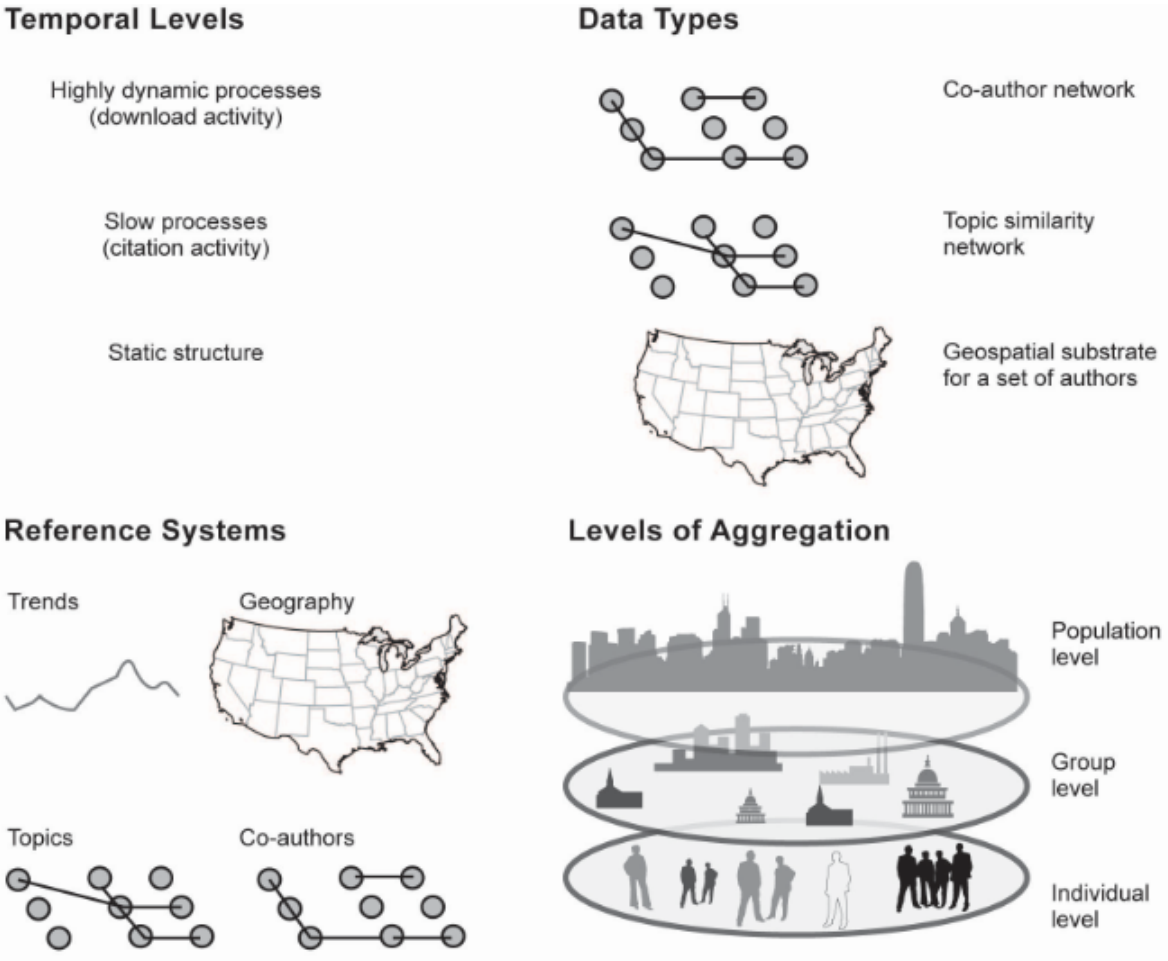
Research – Multi-level,

mixed methods approach to analyze and forecast S&T

Modeling Science Dynamics
using

- multi-level,
- mixed methods, and
- multi-perspective models

Katy Börner, Kevin W. Boyack, Staša Milojević, Steven Morris. (2011) An introduction to modeling science: Basic model types, key definitions, and a general framework for the comparison of process models. In *Scharnhorst, Andrea, Börner, van den Besselaar (Eds) Models of Science Dynamics. Springer Verlag.*



Descriptive Models of Science

- Determine areas of expertise for specific researcher, research group via "invisible colleges" (note that researchers self definition might differ from how field defines him/her) (Crane, 1972).
- Synthesis of specialty narratives from co-citation clusters (Small, 1986).
- Detect advances of scientific knowledge via "longitudinal mapping" (Garfield, 1994).
- Knowledge discovery in un-connected terms (Swanson & Smalheiser, 1997).
- Identify cross-disciplinary fertilization via "passages through science" (Small, 1999, 2000).
- Understand scholarly information foraging (Sandstrom, 2001).
- Identify profiles of authors, also called CAMEOS, to be used to for document retrieval or to map an author's subject matter and studying his/her publishing career, or to map the social and intellectual networks evident in citations to and from authors and in co-authorships (White, 2001).

Descriptive Models of Science cont.

- Identification of scientific frontiers http://www.science-frontiers.com/.
- *ISI's Essential Science Indicators* http://essentialscience.com/
- Import-export studies (Stigler, 1994).
- Evaluation of 'big science' facilities using 'converging partial indicators' (Martin & Irvine, 1983; Martin, 1996).
- Input (levels of funding, expertise of scientists, facilities used) - output (publications, patents, Nobel prices, improved health, reduced environment insults, etc. - influenced by political, economic, financial, and legal factors studies (Kostroff & DelRio, 2001).
- Determine influence of funding on research output (Boyack & Borner, 2002).
- How to write highly influential paper (van Dalen & Henkens, 2001).

The Global 'Scientific Food Web'

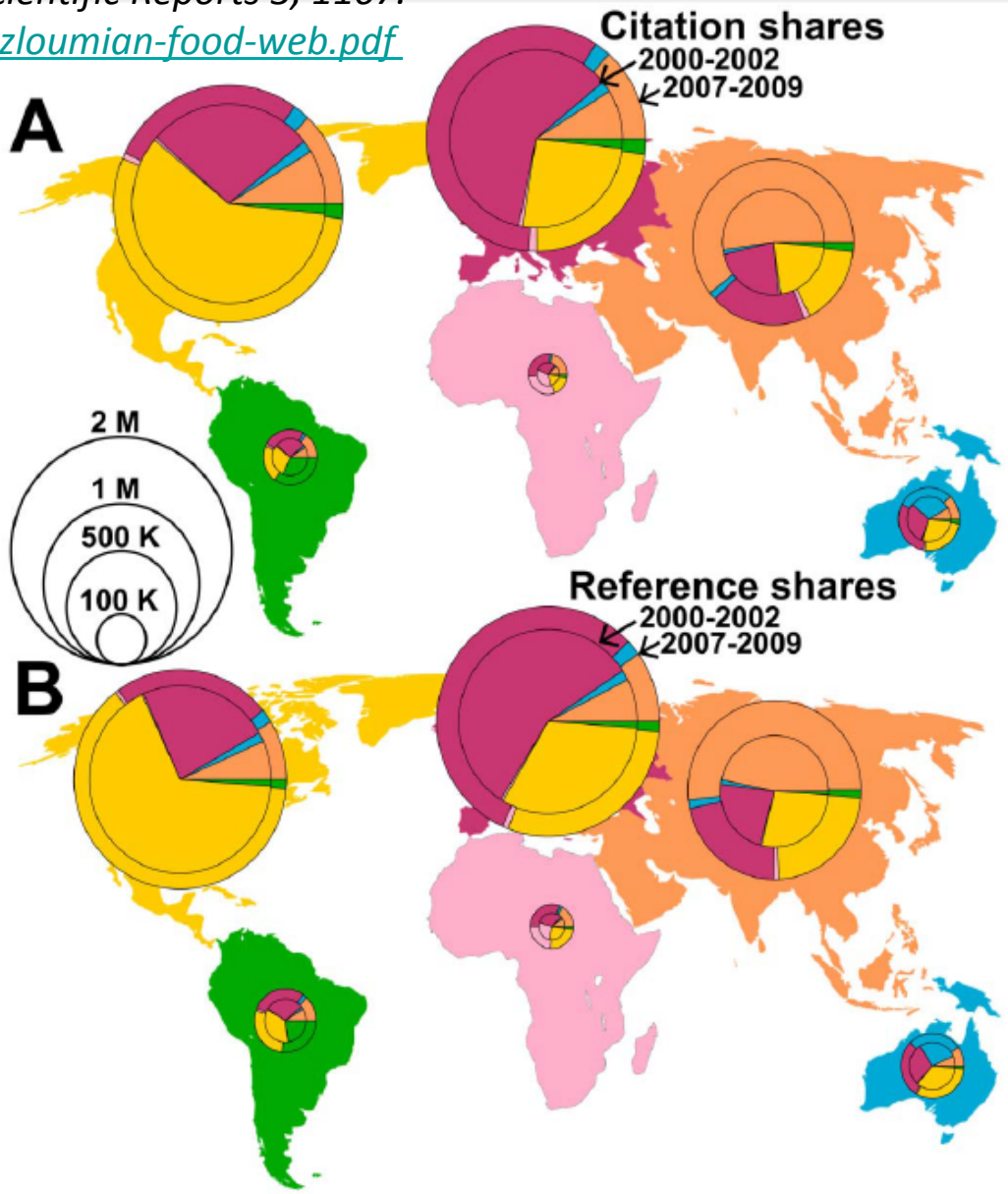
Mazloumian, Amin, Dirk Helbing, Sergi Lozano, Robert Light, and Katy Börner. 2013. "Global Multi-Level Analysis of the 'Scientific Food Web'". Scientific Reports 3, 1167.
http://cns.iu.edu/docs/publications/2013-mazloumian-food-web.pdf

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.



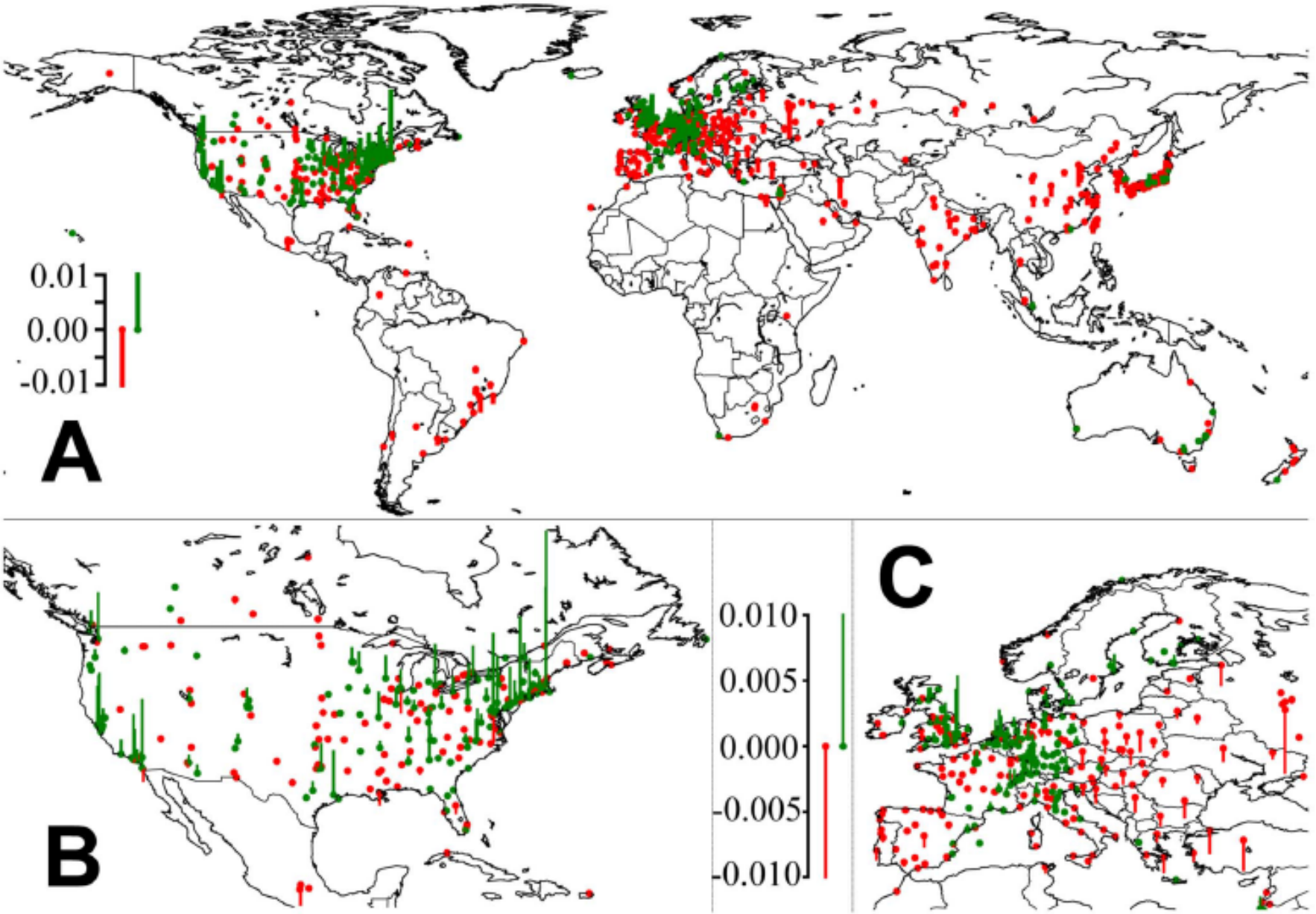


Figure 2 | World map of the greatest knowledge sources and sinks, based on our scientific fitness index. Green bars indicate that the number of citations received is over-proportional, red that the number of citations received is lower than expected (according to a homogeneous distribution of citations 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).

Process Models of Science

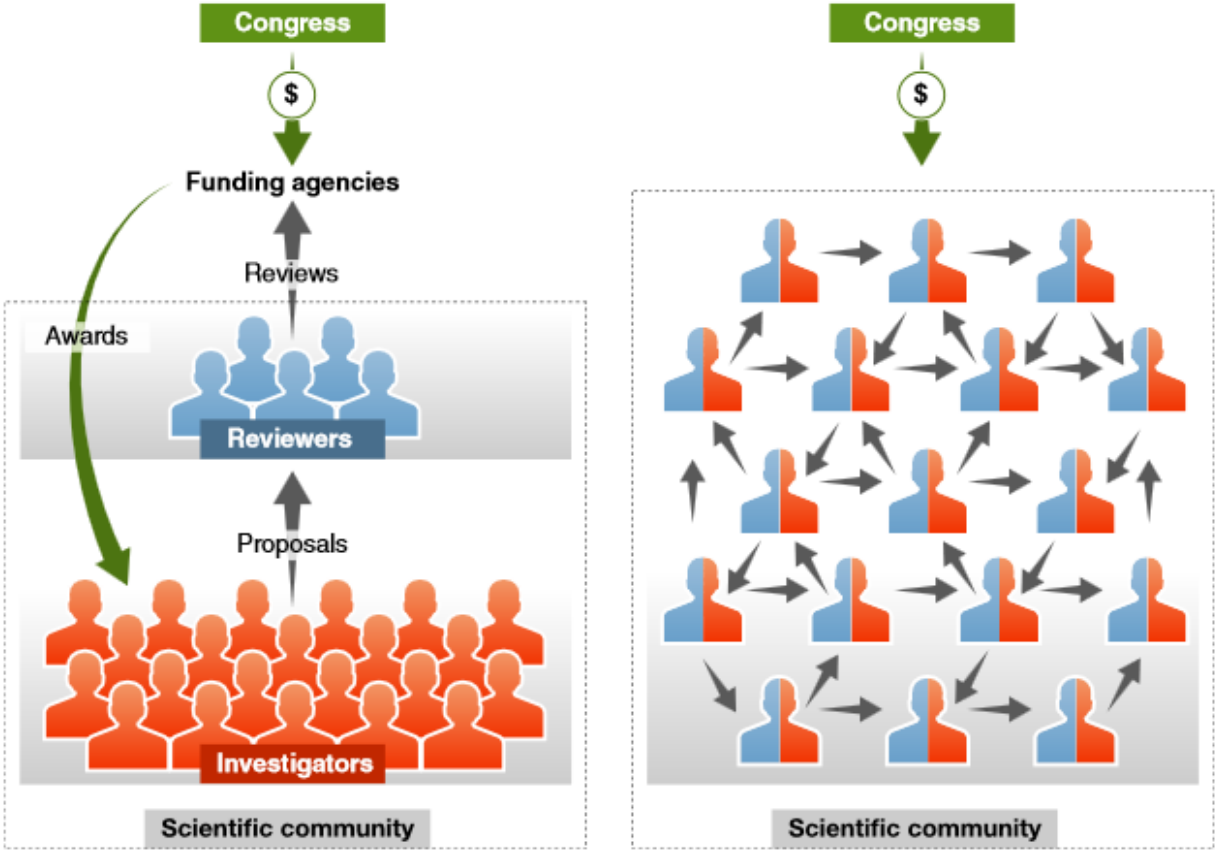
Can be used to predict the effects of

- Large collaborations vs. single author research on information diffusion.
- Different publishing mechanisms, e.g., E-journals vs. books on co-authorship, speed of publication, etc.
- Supporting disciplinary vs. interdisciplinary collaborations.
- Many small vs. one large grant on # publications, Ph.D. students, etc.
- Resource distribution on research output.
- …

In general, process model provide a means to analyze the structure and dynamics of science -- to study science using the scientific methods of science as suggested by Derek J. deSolla Price about 40 years ago.

From funding agencies to scientific agency: Collective allocation of science funding as an alternative to peer review

Bollen, Crandall, Junk, Ding & Börner. 2014. EMBO Reports 15 (1): 1-121.



Existing (left) and proposed (right) funding systems. Reviewers in blue; investigators in red.

In the proposed system, all scientists are both investigators and reviewers: every scientist receives a fixed amount of funding from the government and discretionary distributions from other scientists, but each is required in turn to redistribute some fraction of the total they received to other investigators.

From funding agencies to scientific agency: Collective allocation of science funding as an alternative to peer review

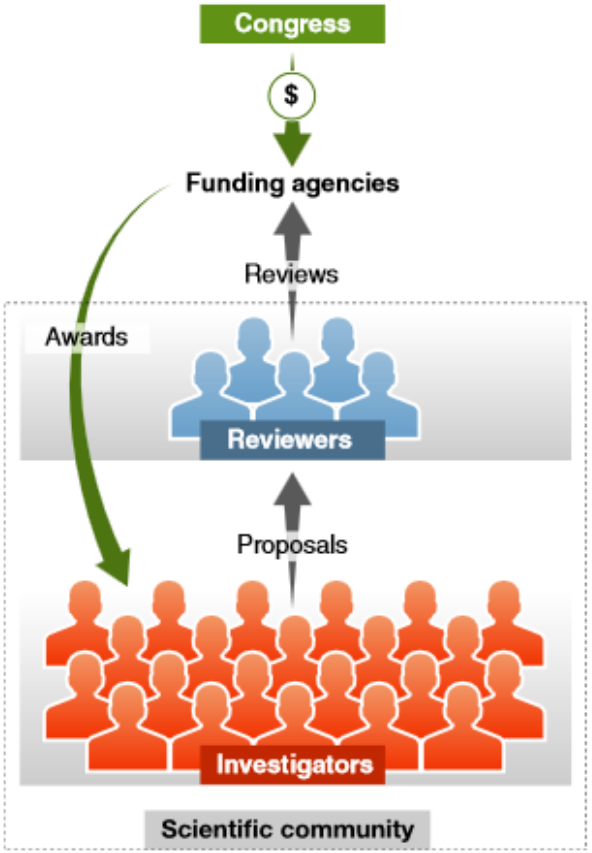
Bollen, Johan, David Crandall, Damion Junk, Ying Ding & Katy Börner. 2014. EMBO Reports 15 (1): 1-121.

Current Model is Expensive:

If four professors work four weeks full-time on a proposal submission, labor costs are about $30k [1]. With typical funding rates below 20%, about five submission-review cycles might be needed resulting in a total expected labor cost of **$150k**. The average NSF grant is **$128k** per year.

U.S. universities charge about 50% overhead (ca. $42k), leaving about **$86k**.

In other words, the four professors lose **$150k-$86k= - $64k** of paid research time by **obtaining** a grant to perform the proposed research.



To add: Time spent by researchers to review proposals. In 2012 alone, NSF convened more than 17,000 scientists to review 53,556 proposals.

[1] Taulbee Survey of Salaries Computer Science , http://cra.org/resources/taulbee

From funding agencies to scientific agency: Collective allocation of science funding as an alternative to peer review

Bollen, Crandall, Junk, Ding & Börner. 2014. EMBO Reports 15 (1): 1-121.

Assume

Total funding budget in year y is t_y

Number of qualified scientists is n

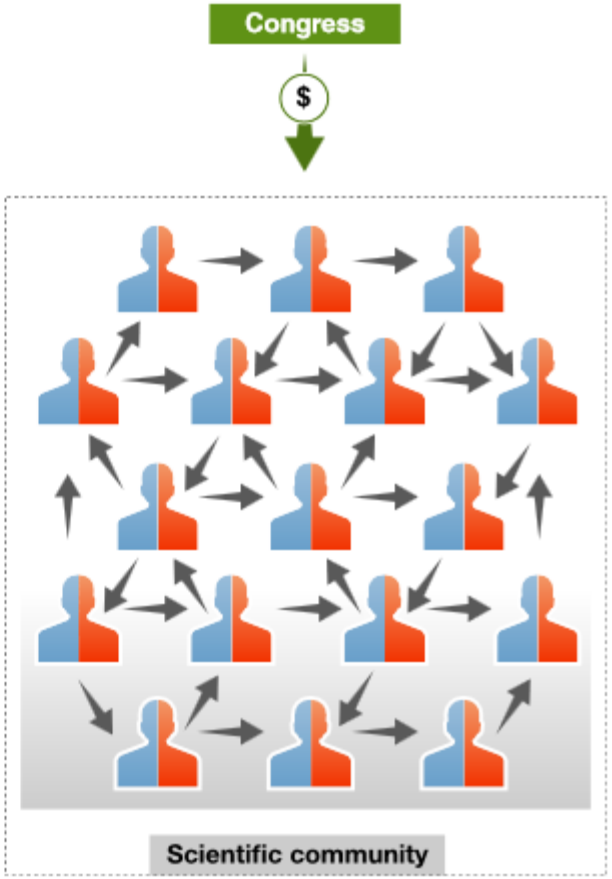
Each year,

the funding agency deposits a fixed amount into each account, equal to the total funding budget divided by the total number of scientists: t_y/n.

Each scientist must distribute a fixed fraction, e.g., 50%, of received funding to other scientists (no self-funding, COIs respected).

Result

Scientists collectively assess each others' merit based on different criteria; they "fund-rank" scientists; highly ranked scientists have to distribute more money.



From funding agencies to scientific agency: Collective allocation of science funding as an alternative to peer review

Bollen, Crandall, Junk, Ding & Börner. 2014. EMBO Reports 15 (1): 1-121.

Example:

Total funding budget per year is 2012 NSF budget

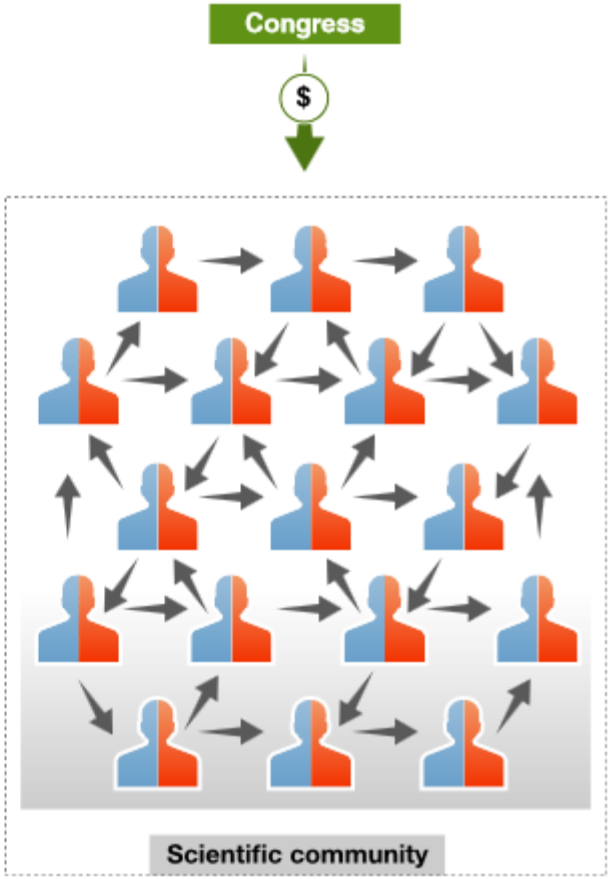
Given the number of NSF funded scientists, each receives a $100,000 basic grant.

Fraction is set to 50%

In 2013, scientist *S* receives a basic grant of $100,000 plus $200,000 from her peers, i.e., a total of $300,000.

In 2013, *S* can spend 50% of that total sum, $150,000, on her own research program, but must donate 50% to other scientists for their 2014 budget.

Rather than submitting and reviewing project proposals, *S* donates directly to other scientists by logging into a centralized website and entering the names of the scientists to donate to and how much each should receive.



From funding agencies to scientific agency: Collective allocation of science funding as an alternative to peer review

Bollen, Crandall, Junk, Ding & Börner. 2014. EMBO Reports 15 (1): 1-121.

Model Run and Validation:

Model is presented in http://arxiv.org/abs/1304.1067

It uses **citations as a proxy** for how each scientist might distribute funds in the proposed system.

Dataset: 37M articles from TR 1992 to 2010 Web of Science (WoS) database with **770M citations** and 4,195,734 unique author names. The **867,872 names** who had authored at least one paper per year in any five years of the period 2000–2010 were used in validation.

For each pair of authors we determined the number of times one had cited the other in each year of our citation data (1992–2010).

NIH and NSF funding records from IU's Scholarly Database provided 347,364 grant amounts for 109,919 unique scientists for that time period.

Simulation run begins in year 2000, in which every scientist was given a fixed budget of B = $100k. In subsequent years, scientists distribute their funding in proportion to their citations over the prior 5 years.

The model yields funding patterns similar to existing NIH and NSF distributions.

Making Every Scientist a Research Funder

When it comes to using peer review to distribute research dollars, Johan Bollen favors radical simplicity.

Over the years, many scientists have suggested that the current system could be improved by changing the composition of the review panels, tweaking the interactions among reviewers, or revising how the proposals are scored. But Bollen, a computer scientist at Indiana University, Bloomington, would simply award all eligible researchers a block grant—and then require them to give some of it away to colleagues they judge most deserving.

That radical step, described in a paper Bollen and four Indiana colleagues recently posted on *EMBO Reports*, retains peer review's core concept of tapping into the views of the most knowledgeable researchers. But it would eliminate the huge investment in time and money required to submit proposals and assemble panels to judge them.

Bollen's process would be almost instantaneous: In a version of expert-directed crowdsourcing, scientists would fill out a form once a year listing their favored researchers, and a predetermined portion of their annual grant money—a total of, say, 50%—would then be transferred to their choices.

"So many scientists spend so much time on peer review, and there's a high level of frustration," Bollen explains. "We already know who the best people are. And if you're doing good work, then you deserve to receive support."

Others are skeptical. "I've known Johan for a long time and have the highest regard for his ability as an out-of-the-box thinker," says Stephen Griffin, a retired National Science Foundation (NSF) program manager who's now a visiting professor of information sciences at the University of Pittsburgh in Pennsylvania. "But there are a number of issues he doesn't address."

Those sticking points include the likely mismatch between what researchers need and what their colleagues give them; the absence of any replacement for the overhead payments in today's grants, which support infrastructure at host institutions; and the dearth of public accountability for the billions of dollars that would flow from public coffers to individuals. "Scientists aren't really equipped to be a funding agency," Griffin notes.

Bollen acknowledges that the process would need safeguards to ensure that scientists don't reward their friends or punish their enemies. But his analysis suggests that the U.S. research landscape would not look all that different if his radical proposal were adopted.

Drawing upon citation data in 37 million papers over 20 years, the Indiana researchers conducted a simulation premised on the idea that scientists would reallocate their federal dollars according to how often they cited their peers. The simulation, he says, yielded a funding pattern "similar in shape to the actual distribution" at NSF and the National Institutes of Health for the past decade—at a fraction of the overhead required by the current system.

–JDM

February 7, 2014

Science 7 February 2014: Vol. 343 no. 6171 p. 598
DOI: 10.1126/science.343.6171.598
http://www.sciencemag.org/content/343/6171/598.full?sid=4f40a7f0-6ba2-4ad8-a181-7ab394fe2178

Different Stakeholder Groups and Their Needs

Funding Agencies

- Need to monitor (long-term) money flow and research developments, identify areas for future development, stimulate new research areas, evaluate funding strategies for different programs, decide on project durations, funding patterns.

Scholars

- Want easy access to research results, relevant funding programs and their success rates, potential collaborators, competitors, related projects/publications **(research push).**

Industry

- Is interested in fast and easy access to major results, experts, etc. Influences the direction of research by entering information on needed technologies **(industry-pull)**.

Advantages for Publishers

- Need easy to use interfaces to massive amounts of interlinked data. Need to communicate data provenance, quality, and context.

Society

- Needs easy access to scientific knowledge and expertise.

Scholars Have Different Roles/Needs

Researchers and Authors—need to select promising research topics, students, collaborators, and publication venues to increase their reputation. They benefit from a global view of competencies, reputation and connectivity of scholars; hot and cold research topics and bursts of activity, and funding available per research area.

Editors—have to determine editorial board members, assign papers to reviewers, and ultimately accept or reject papers. Editors need to know the position of their journals in the evolving world of science. They need to advertise their journals appropriately and attract high-quality submissions, which will in turn increase the journal's reputation and lead to higher quality submissions.

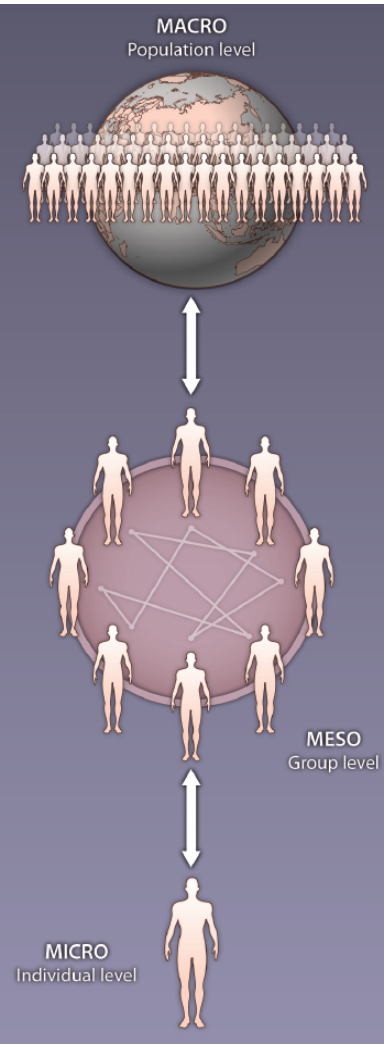
Reviewers—read, critique, and suggest changes to help improve the quality of papers and funding proposals. They need to identify related works that should be cited or complementary skills that authors might consider when selecting project collaborators.

Teachers—teach classes, train doctoral students, and supervise postdoctoral researchers. They need to identify key works, experts, and examples relevant to a topic area and teach them in the context of global science.

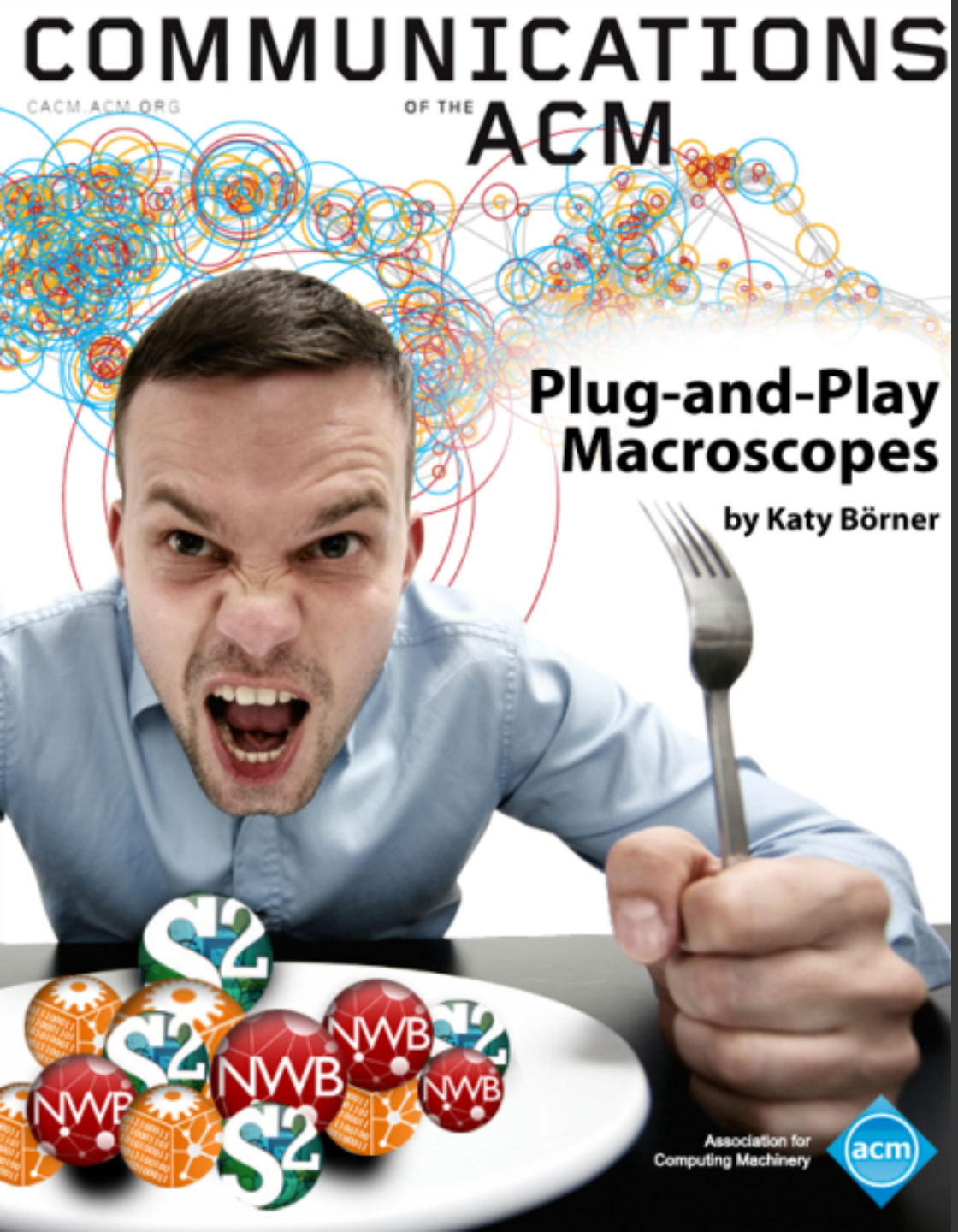
Inventors—create intellectual property and obtain patents, thus needing to navigate and make sense of research spaces as well as intellectual property spaces.

Investigators—scholars acquire funding to support students, hire staff, purchase equipment, or attend conferences. Here, research interests and proposals have to be matched with existing federal and commercial funding opportunities, possible industry collaborators and sponsors.

Team Leads and Science Administrators—many scholars direct multiple research projects simultaneously. Some have full-time staff, research scientists, and technicians in their laboratories and centers. Leaders need to evaluate performance and provide references for current or previous members; report the progress of different projects to funding agencies.



Tools – continuously identify, learn, advance, share code, e.g., via Plug-and-Play Macroscopes



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

Designing "Dream Tools"

Many of the best micro-, tele-, and macroscopes are designed by **scientists keen to observe and comprehend what no one has seen or understood before.** Galileo Galilei (1564–1642) recognized the potential of a spyglass for the study of the heavens, ground and polished his own lenses, and used the improved optical instruments to make discoveries like the moons of Jupiter, providing quantitative evidence for the Copernican theory.

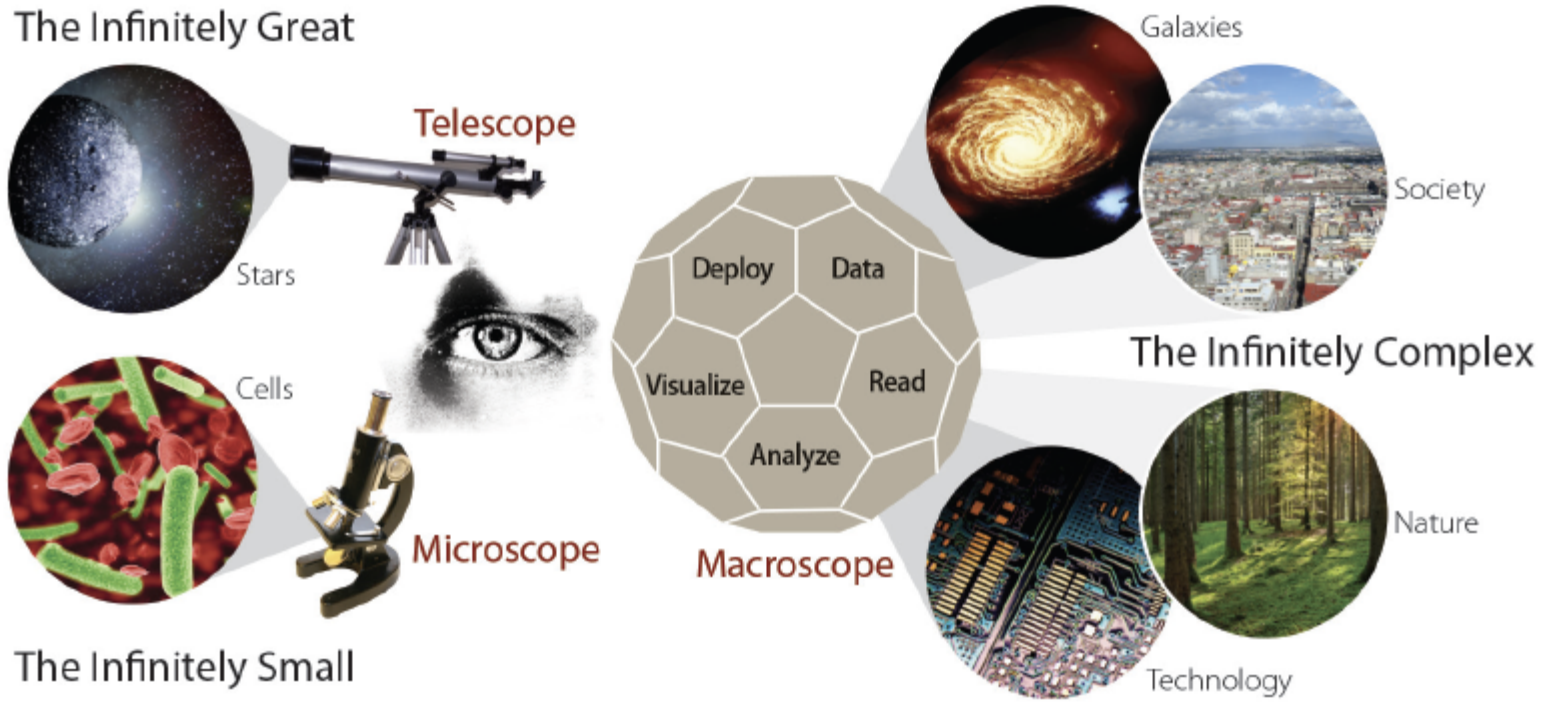
Today, scientists **repurpose, extend, and invent new hardware and software** to create **"macroscopes"** that may solve both local and global challenges.

CNS Macroscope tools **empower** me, my students, colleagues, and more than 130,000 others that downloaded them.

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. Analysis, navigation, and management of these continuously evolving data sets require a new kind of data-analysis and visualization tool we call a macroscope (from the Greek macros, or "great," and skopein, or "to observe") inspired by de Rosnay's futurist science writings.

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.**



Plug-and-Play Macroscopes

Inspire computer scientists to implement software frameworks that **empower domain scientists** to assemble their own continuously evolving macroscopes, adding and upgrading existing (and removing obsolete) plug-ins to arrive at a set that is truly relevant for their work—with little or no help from computer scientists.

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.**

Sharing Algorithms Across Disciplines

IS

CS

Bio

SNA

Physics

Different datasets/formats.
Diverse algorithms/tools written in many programming languages.

Related Work

Google Code and SourceForge.net provide special means for developing and distributing software

- In August 2009, SourceForge.net hosted more than 230,000 software projects by two million registered users (285,957 in January 2011);
- In August 2009 ProgrammableWeb.com hosted 1,366 application programming interfaces (APIs) and 4,092 mashups (2,699 APIs and 5,493 mashups in January 2011)

Cyberinfrastructures serving large biomedical communities

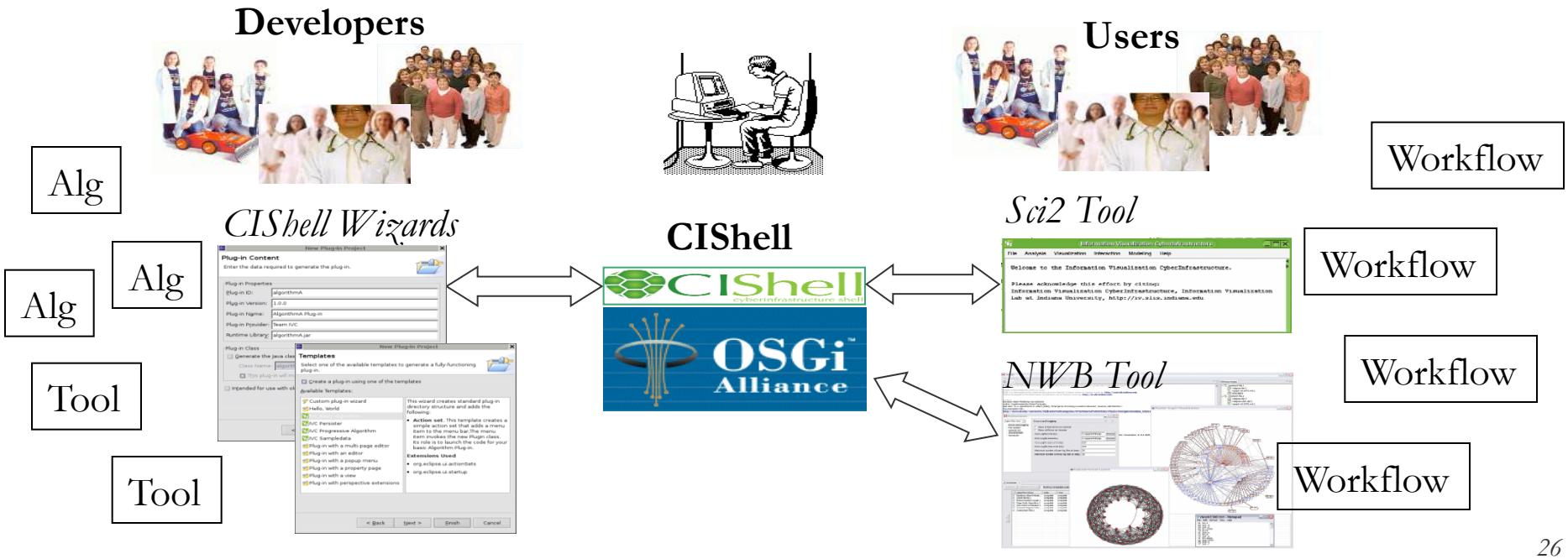
- Cancer Biomedical Informatics Grid (caBIG) (http://cabig.nci.nih.gov)
- Biomedical Informatics Research Network (BIRN) (http://nbirn.net)
- Informatics for Integrating Biology and the Bedside (i2b2) (https://www.i2b2.org)
- HUBzero (http://hubzero.org) platform for scientific collaboration uses
- myExperiment (http://myexperiment.org) supports the sharing of scientific workflows and other research objects.

Missing so far is a **common standard** for

- the design of **modular, compatible algorithm and tool plug-ins** (also called "modules" or "components")
- that can be **easily combined into scientific workflows** ("pipeline" or "composition"),
- and packaged as **custom tools.**

OSGi & 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.



CIShell Portal and Developer Guide

(http://cishell.org)

CIShell cyberinfrastructure shell

CIShell Home

Edit Add

1 Added by Micah Linnemeier, last edited by Micah Linnemeier on Mar 16, 2011 (view change)

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

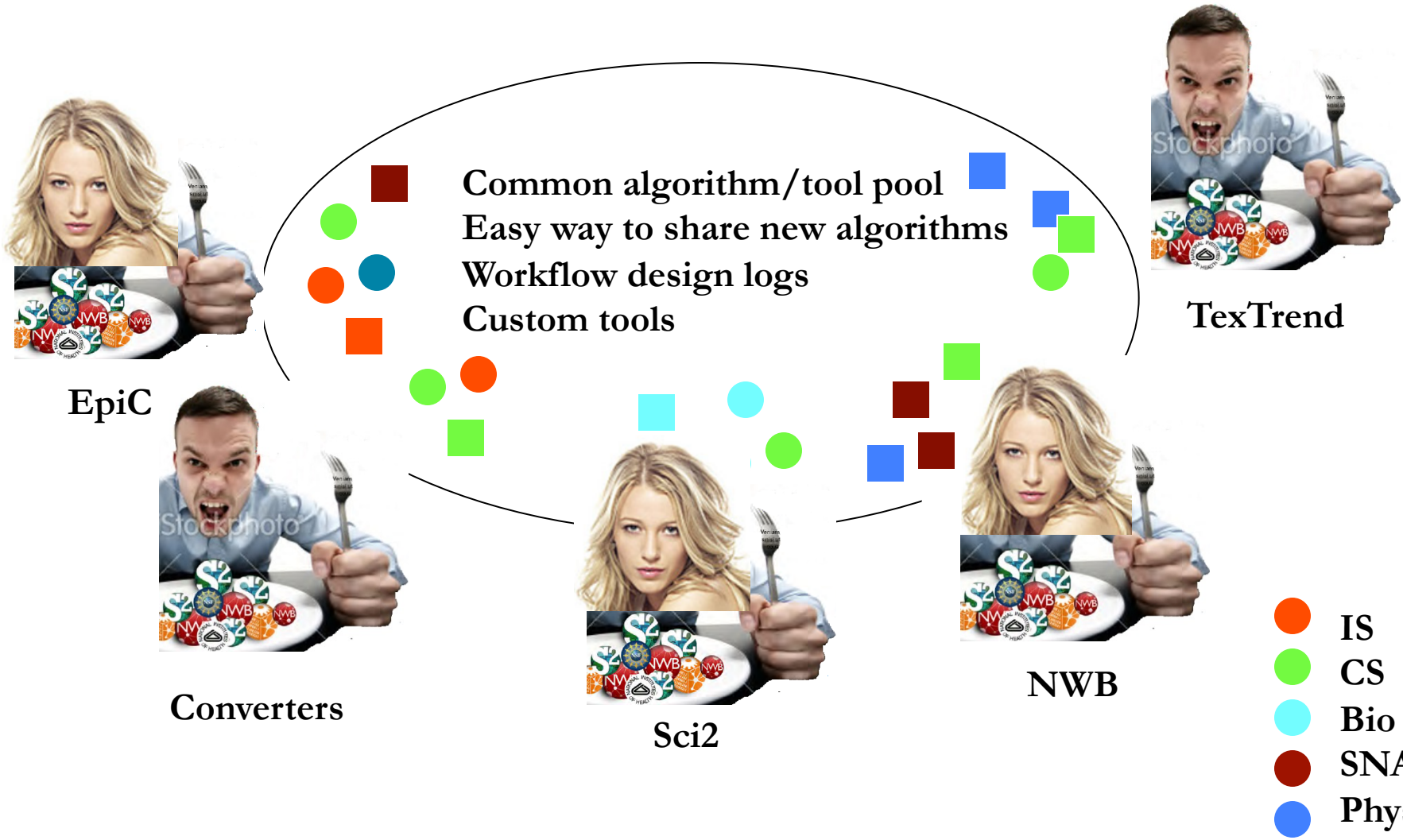
Getting Started...

- Documentation & Developer Resources
- Download

Getting Involved...

- Contact Us

Easy Creation of Custom Tools



OSGi/CIShell Adoption

CIShell/OSGi is at the core of different CIs and a total of 169 unique plugins are used in the
- *Information Visualization* (http://iv.slis.indiana.edu),
- *Network Science (NWB Tool)* (http://nwb.slis.indiana.edu),
- *Scientometrics and Science Policy (Sci² Tool)* (http://sci.slis.indiana.edu), and
- *Epidemics* (http://epic.slis.indiana.edu) research communities.

Most interestingly, a number of other projects recently adopted OSGi and one adopted CIShell:

Cytoscape (http://www.cytoscape.org) lead by Trey Ideker, UCSD 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). **Bruce visits Mike Smoot in 2009**

Taverna Workbench (http://taverna.sourceforge.net) lead by Carol Goble, University of Manchester, UK 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. **Micah, June 2010**

MAEviz (https://wiki.ncsa.uiuc.edu/display/MAE/Home) managed by Shawn Hampton, NCSA is an open-source, extensible software platform which supports seismic risk assessment based on the Mid-America Earthquake (MAE) Center research.

TEXTrend (http://www.textrend.org) lead by George Kampis, Eötvös University, Hungary develops a framework for the easy and flexible integration, configuration, and extension of plugin-based components in support of natural language processing (NLP), classification/mining, and graph algorithms for the analysis of business and governmental text corpuses with an inherently temporal component.

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.

Network Workbench Tool

http://nwb.cns.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.

In February 2009, the tool provides more 169 plugins that support the preprocessing, analysis, modeling, and visualization of networks.

More than 50 of these plugins can be applied or were specifically designed for S&T studies.

It has been downloaded more than 65,000 times since December 2006.

Summary

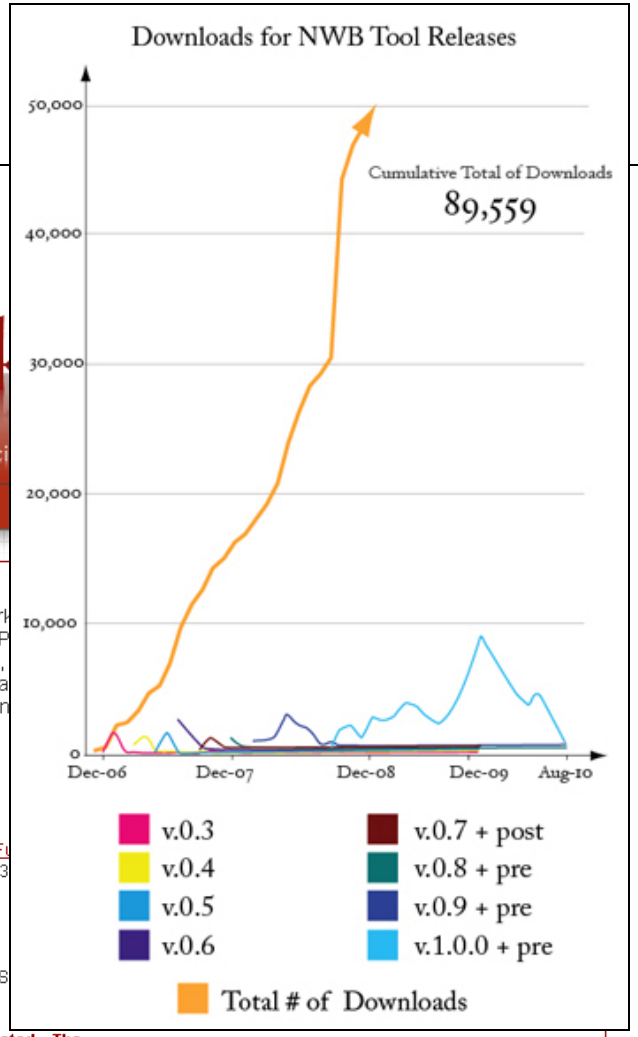
Network Workbench: A Large-Scale Network
Toolkit for Biomedical, Social Science and P
evaluate, and operate a unique distributed,
scale network analysis, modeling, and visua
(NWB). The envisioned data-code-computin
more
How to cite this project

News & Updates

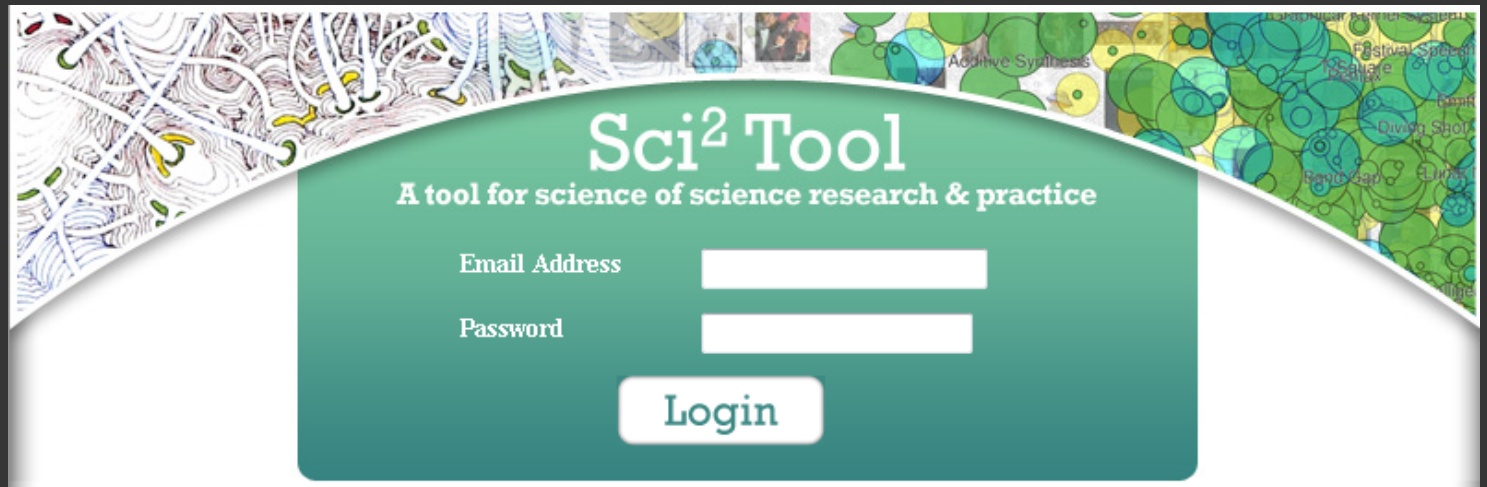
- 5.1.09 Kaelble, Steve. 2009. Mapping the Fu
Knowledge. *Research & Creative Activity*, 3
(website accessed 5/1/09)
- 3.23.09 1.0.0 beta 5 Released
- 1.23.09 Ann Mcranie's tutorial abstract for S
2009
- 11.4.08 **Two NWB PIs featured in "Connected—The Power of Six Degrees." 2008. Anna Maria Talas, Director. Australian Broadcasting Corporation, Ltd.** [YouTube] [Full Video (300MB)]

Getting Started
See more documentation

Get Involved



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.



Forgot your password?

To recover your account password, please visit our password recovery page.

Not registered yet?

Register now

Tutorials

Scott Weingart, Biberstine (2010 Science, Indiana

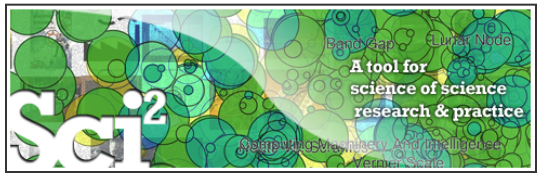
Katy Börner (2010) Science of Science Research and Tools (12 Tutorials). Reporting Branch, Office of Extramural Research/Office of the Director, National Institutes of Health, Bethesda, MD.

- Tutorial #01: Science of Science Research
- Tutorial #02: Network Science / Information Visualization
- Tutorial #03: CIShell Powered Tools: Network Workbench and Science of Science Tool
- Tutorial #04: Temporal Analysis—Burst Detection
- Tutorial #05: Geospatial Analysis and Mapping
- Tutorial #06: Topical Analysis & Mapping
- Tutorial #07: Tree Analysis and Visualization
- Tutorial #08: Network Analysis and Visualization
- Tutorial #09: Large Network Analysis and Visualization.
- Tutorial #10: Using the Scholarly Database at IU
- Tutorial #11: VIVO National Researcher Networking
- Tutorial #12: Future Developments

http://sci2.cns.iu.edu
http://sci2.wiki.cns.iu.edu

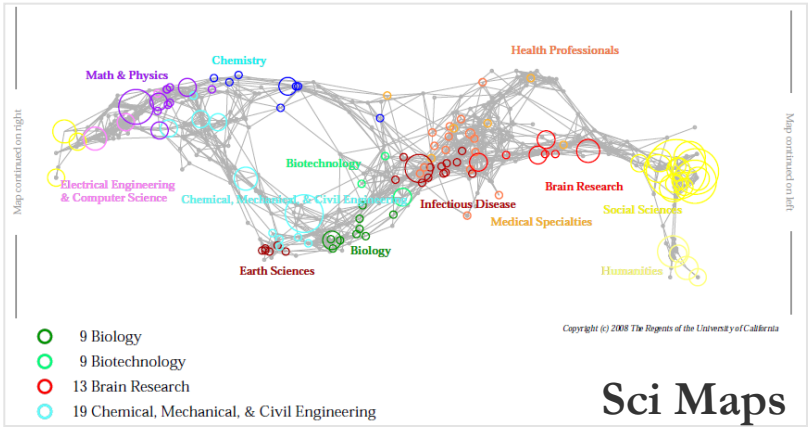
Geetha Senthil (2010). Multidisciplinary Nature of Work With Reference to PIs and ICs Within a Portfolio. PA Group at NIH.

NIH Office of Extramural Research and Katy Börner (2010) Network Visualizations Using SPIRES Data and the Sci2 Tool. Office of Extramural Research at NIH.

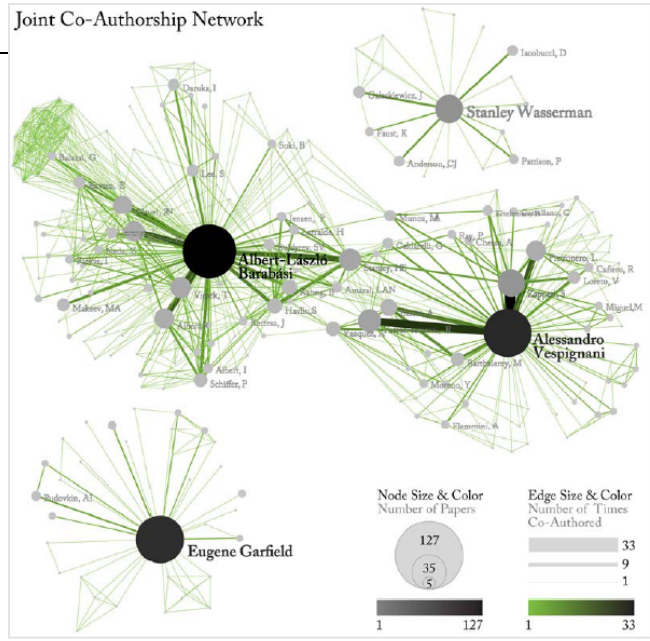


Sci² Tool – "Open Code for S&T Assessment"

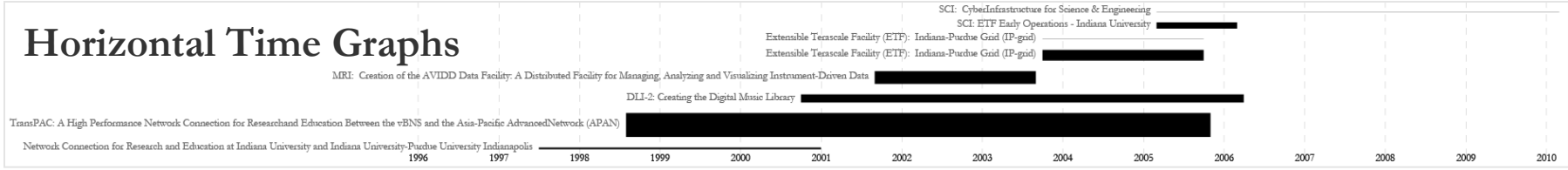
OSGi/CIShell powered tool with NWB plugins and many new scientometrics and visualizations plugins.



Sci Maps

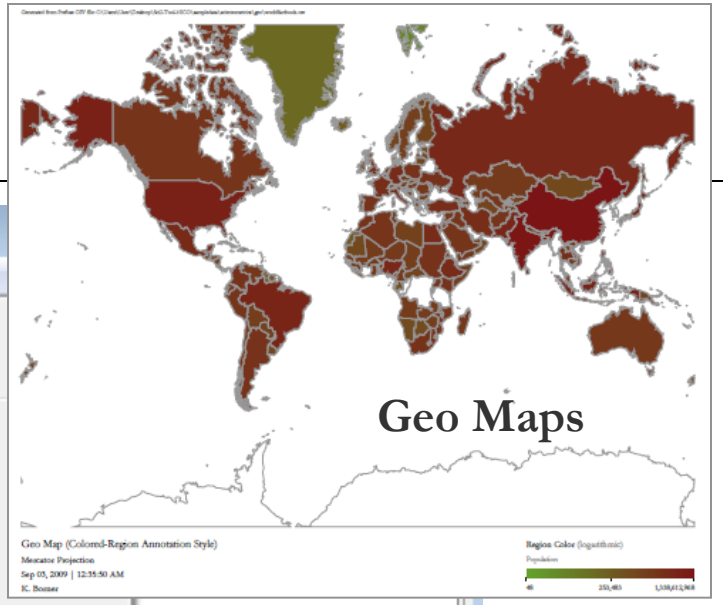


GUESS Network Vis



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



Sci² Tool

File Preprocessing Modeling Analysis Visualization Scientometrics Help

Console

Welcome to the Science of Science Tool (Sci
The development of this tool is supported in
Network Science center and the School of Li
Indiana University, the National Science Fou
and IIS-0715303, and the James S. McDonnel
Cyberinfrastructure portal (http://sci.slis.ind

The primary investigators are Katy Börner, In
SciTech Strategies Inc. The Sci² tool was dev
J. Duhon, Patrick A. Phillips, Chintan Tank, a
Cyberinfrastructure Shell (http://cishell.org)
for Network Science Center (http://cns.slis.i
Many algorithm plugins were derived from t
(http://nwb.slis.indiana.edu).

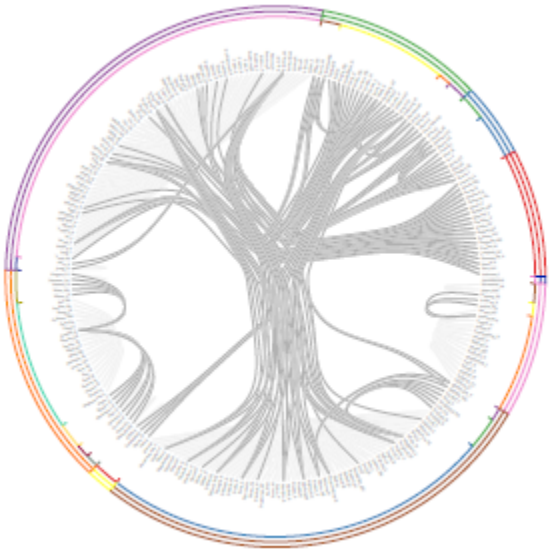
Please cite as follows:
Sci² Team. (2009). Science of Science Tool. In
Strategies Inc., http://sci.slis.indiana.edu.
.....

GUESS
GnuPlot
Radial Tree/Graph (prefuse alpha)
Radial Tree/Graph with Annotation (prefuse beta)
Tree View (prefuse beta)
Tree Map (prefuse beta)
Force Directed with Annotation (prefuse beta)
Fruchterman-Reingold with Annotation (prefuse beta)
DrL (VxOrd)
Specified (prefuse beta)
Horizontal Line Graph
Circular Hierarchy
Geo Map (circle annotations)
Geo Map (region coloring annotations)
Image Viewer
RefMapper

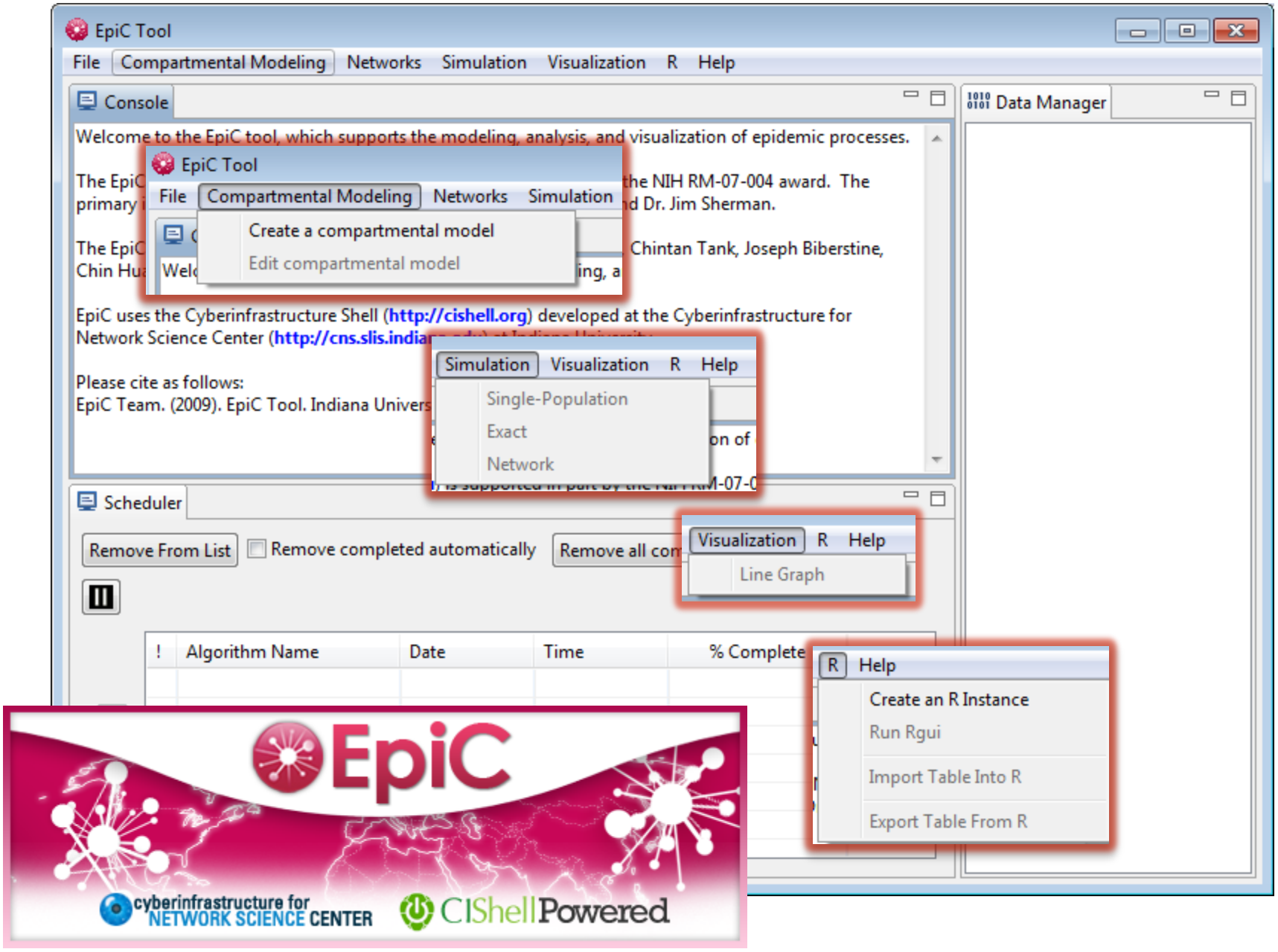
Scheduler

Remove From List Remove completed

!	Algorithm Name	Date	Time	% Com
☑	Extract Co-Author Netw...	09/03/2009	00:15:20 AM	
☑	Load and Clean ISI File	09/03/2009	00:15:05 AM	



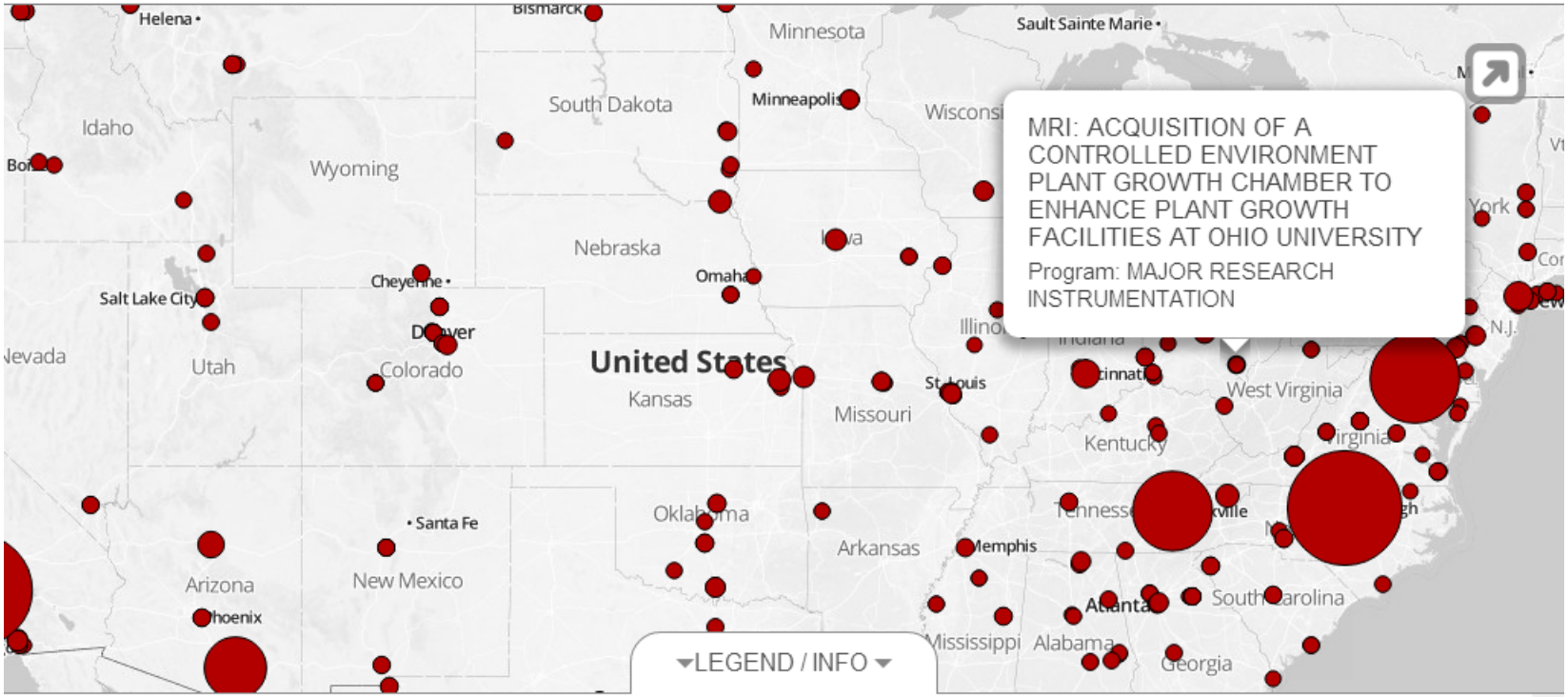
Circular Hierarchy


EpiC Tool
File
Compartmental Modeling
Networks
Simulation
Visualization
R
Help
Console
Welcome to the EpiC tool, which supports the modeling, analysis, and visualization of epidemic processes.
Create a compartmental model
Edit compartmental model
EpiC uses the Cyberinfrastructure Shell (http://cishell.org) developed at the Cyberinfrastructure for
Please cite as follows:
Single-Population
Exact
Network
Scheduler
Remove From List
Remove completed automatically
Line Graph
Algorithm Name
Date
Time
% Complete
Create an R Instance
Run Rgui
Import Table Into R
Export Table From R
Data Manager
EpiC
cyberinfrastructure for NETWORK SCIENCE CENTER
CIShell Powered

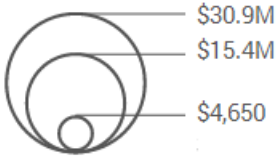
NETE® A|V
ANALYZER | VISUALIZER

Proportional Symbol Map

Relationship between Projects and External Organizations - Larry E. Humes, Bernice A. Pescosolido; Generated by NETE March 5, 2014 | 9:34 AM EST



Amount Awarded



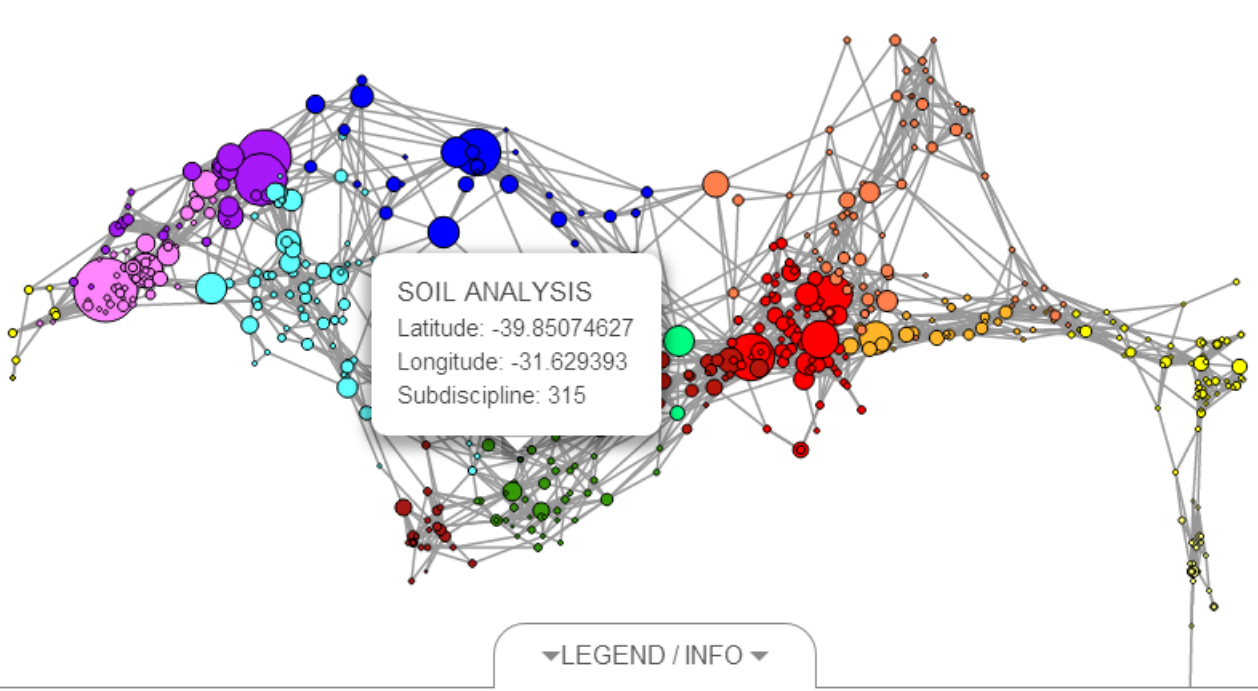
How To Read This Map

This proportional symbol map shows 52 U.S. states and other jurisdictions using the Albers equal-area conic projection with Alaska, Puerto Rico, and Hawaii inset. Each dataset record is represented by a circle centered at its geolocation. The area, interior color, and exterior color of each circle may represent numeric attribute values. Minimum and maximum data values are given in the legend.

NETE® AV
ANALYZER | VISUALIZER

Topic Analysis - Map of Science

Generated from Publications for top 20 projects - Jeffrey R. Alberts, Larry E. Humes, Bernice A. Pescosolido and 9 others; Generated by NETE.



LEGEND / INFO

Total Awards

How To Read This Map

This map is a visual representation of 554 sub-disciplines within 13 disciplines of science and their relationships to one another, shown as points and lines connecting those points respectively. Over top this visualization is drawn the result of mapping a dataset's journals to the underlying sub-discipline(s) those journals contain. Mapped sub-disciplines are shown with size relative to the number of matching journals and color from the discipline.

Information Visualization MOOC

INDIANA UNIVERSITY CNS

Overview

This course provides an overview about the state of the art in information visualization. It teaches the process of producing effective visualizations that take the needs of users into account.

This year, the course can be taken for three Indiana University credits as part of the Online Data Science Program just announced by the School of Informatics and Computing. Students interested in applying to the program can find more information here.

Among other topics, the course covers:

- Data analysis algorithms that enable extraction of patterns and trends in data
- Major temporal, geospatial, topical, and network visualization techniques
- Discussions of systems that drive research and development.

Just like last year, students will have the opportunity to collaborate on real-world projects for a variety of clients. Click here to see this year's list of clients and projects.

Everyone who registers gains free access to the Scholarly Database (26 million paper, patent, and grant records) and the Sci2 Tool (100+ algorithms and tools).

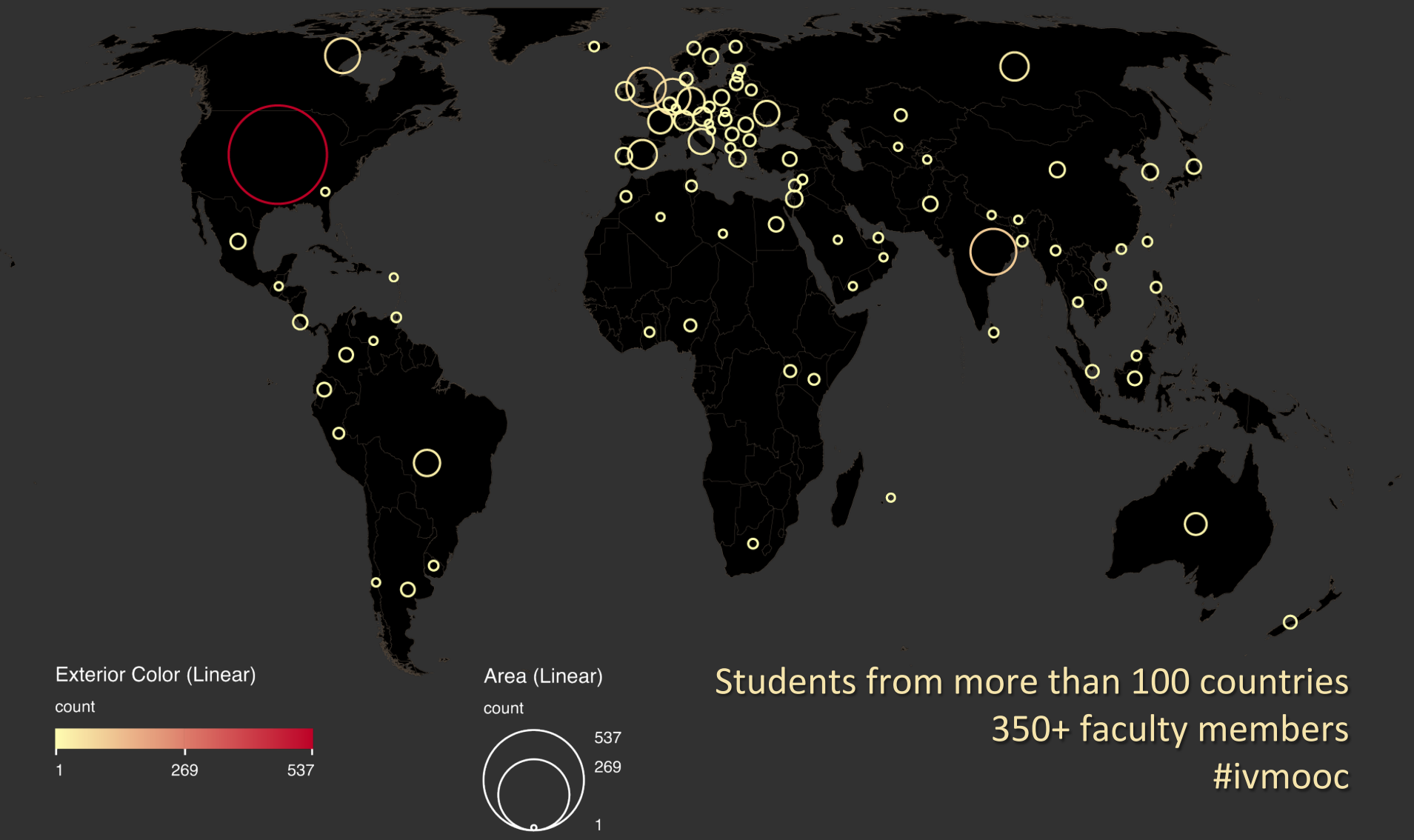
Please watch the introduction video to learn more.

Register for Course

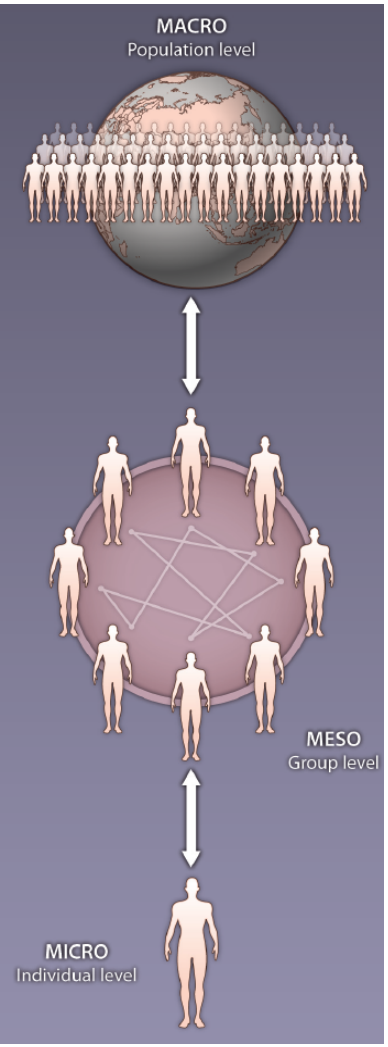
IVMOOC 2014 course materials will be available until end of November 2014. The IVMOOC 2015 will open in January 2015 with new materials and a cloud computing setup.

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

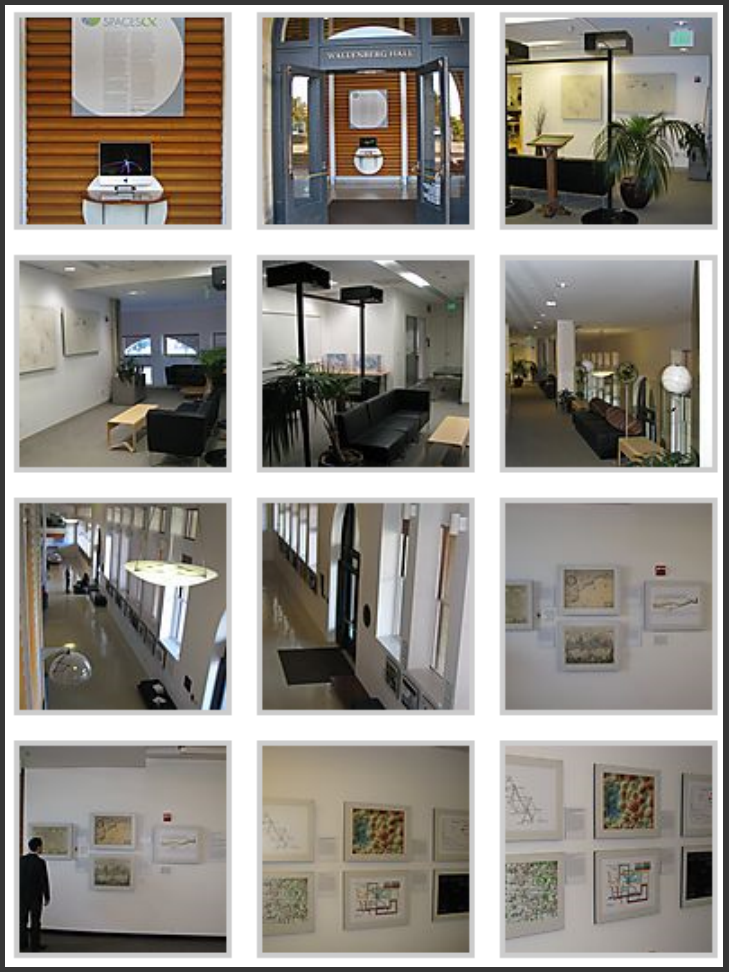
The Information Visualization MOOC
ivmooc.cns.iu.edu



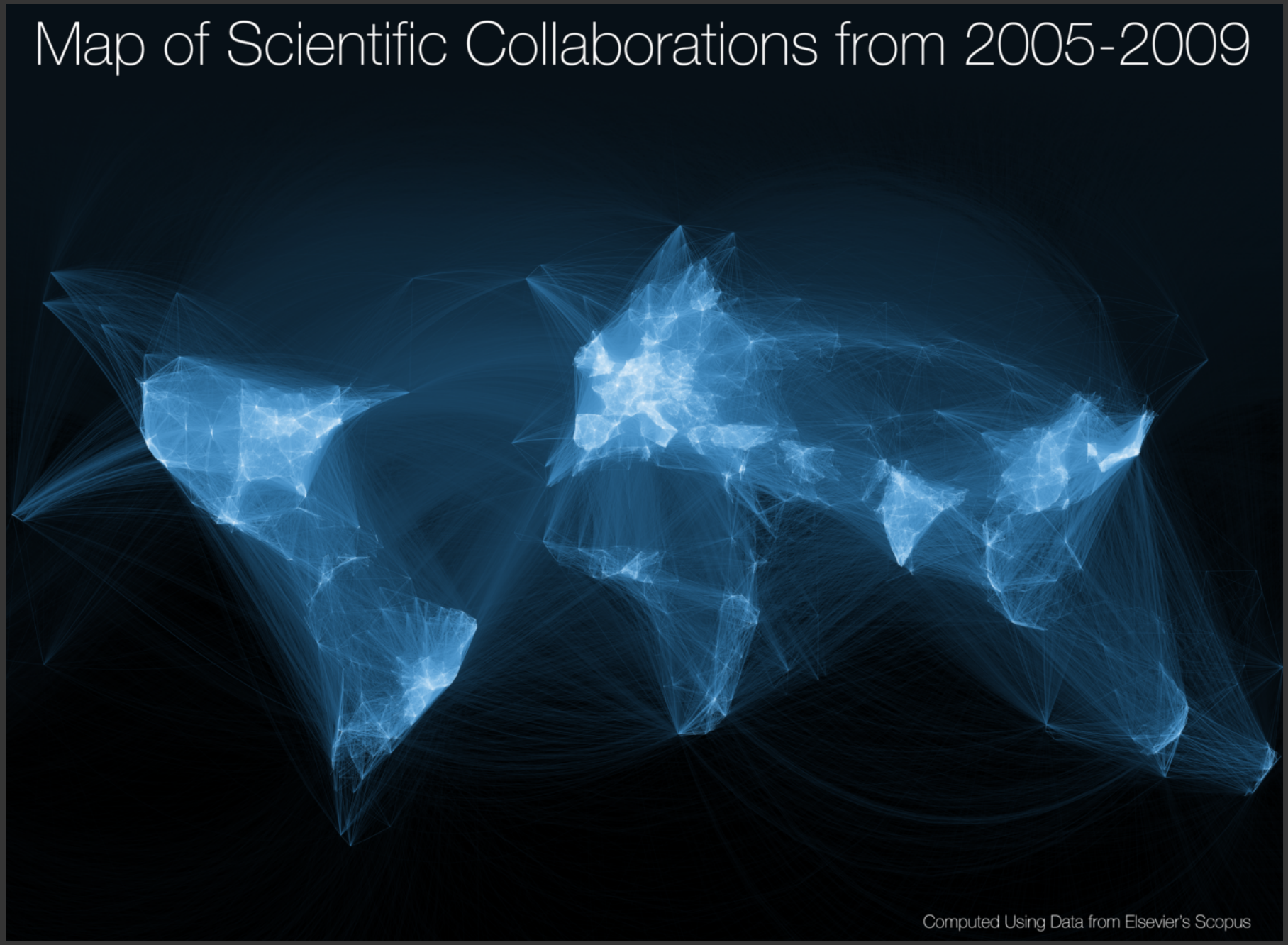
Students from more than 100 countries
350+ faculty members
#ivmooc



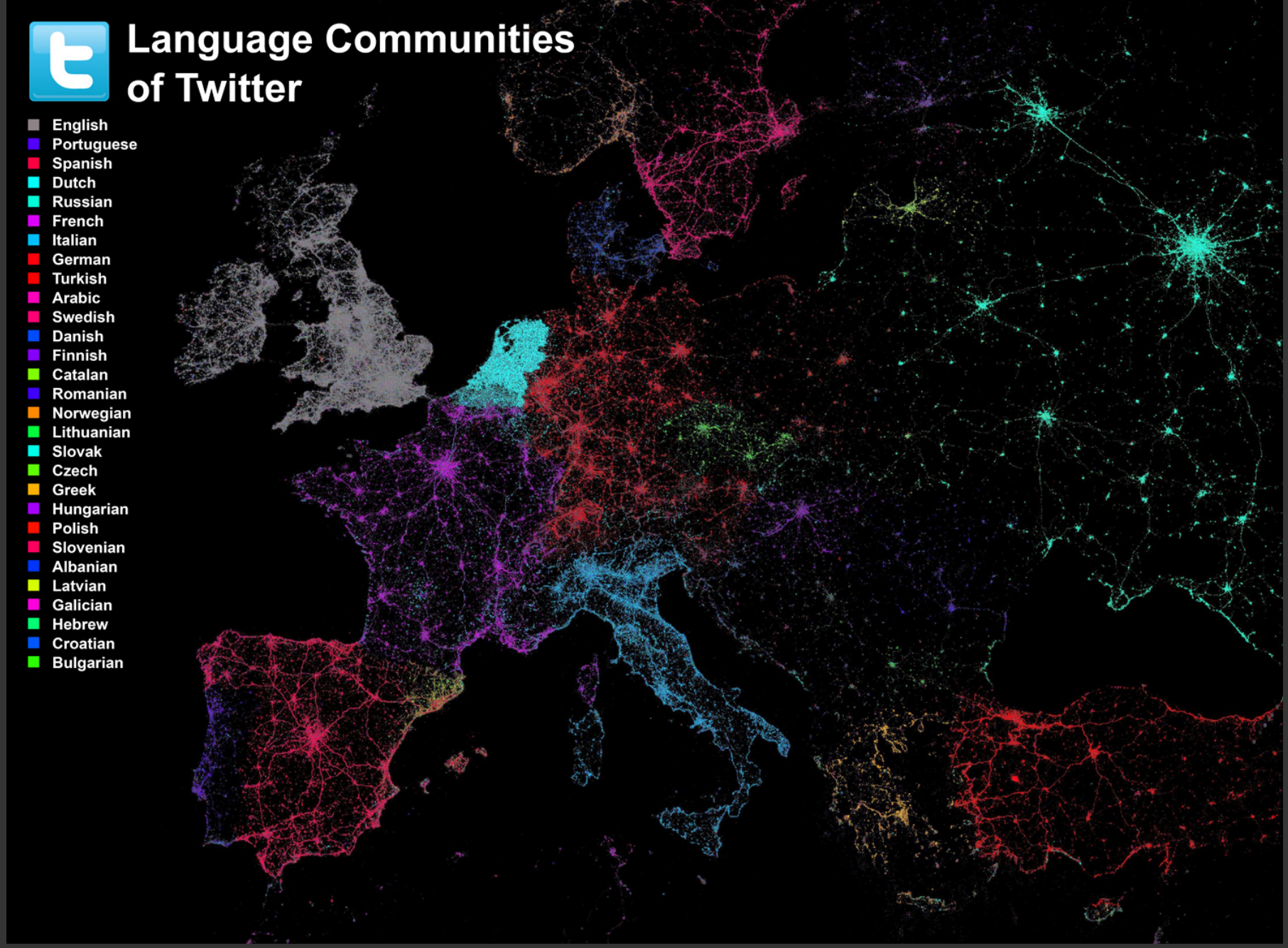
Map – effectively communicate the structure and dynamics of science to different stakeholders using (interactive) visualizations.



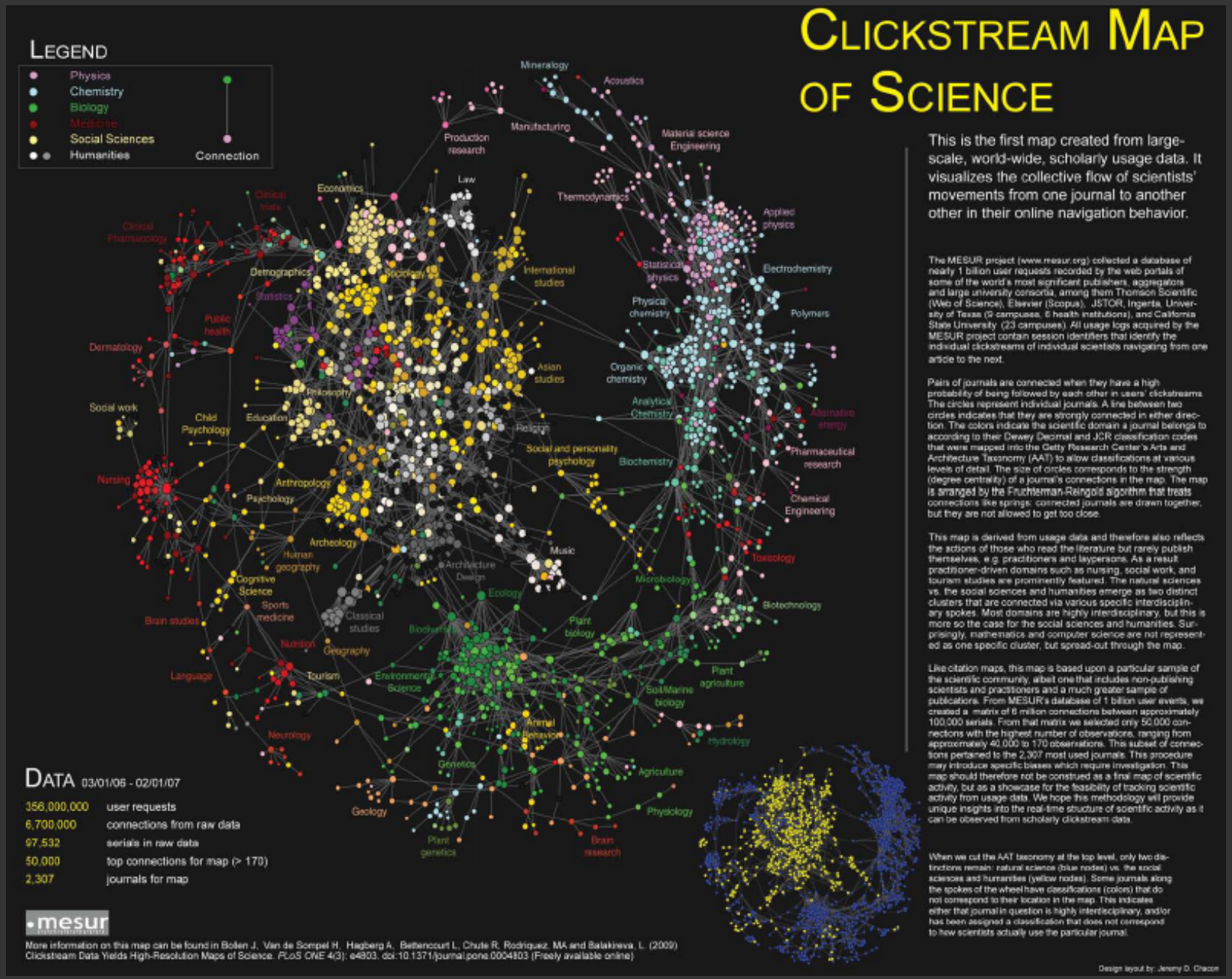
Mapping Science Exhibit on display at MEDIA X, Stanford University
http://mediax.stanford.edu, *http://scaleindependentthought.typepad.com/photos/scimaps*



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



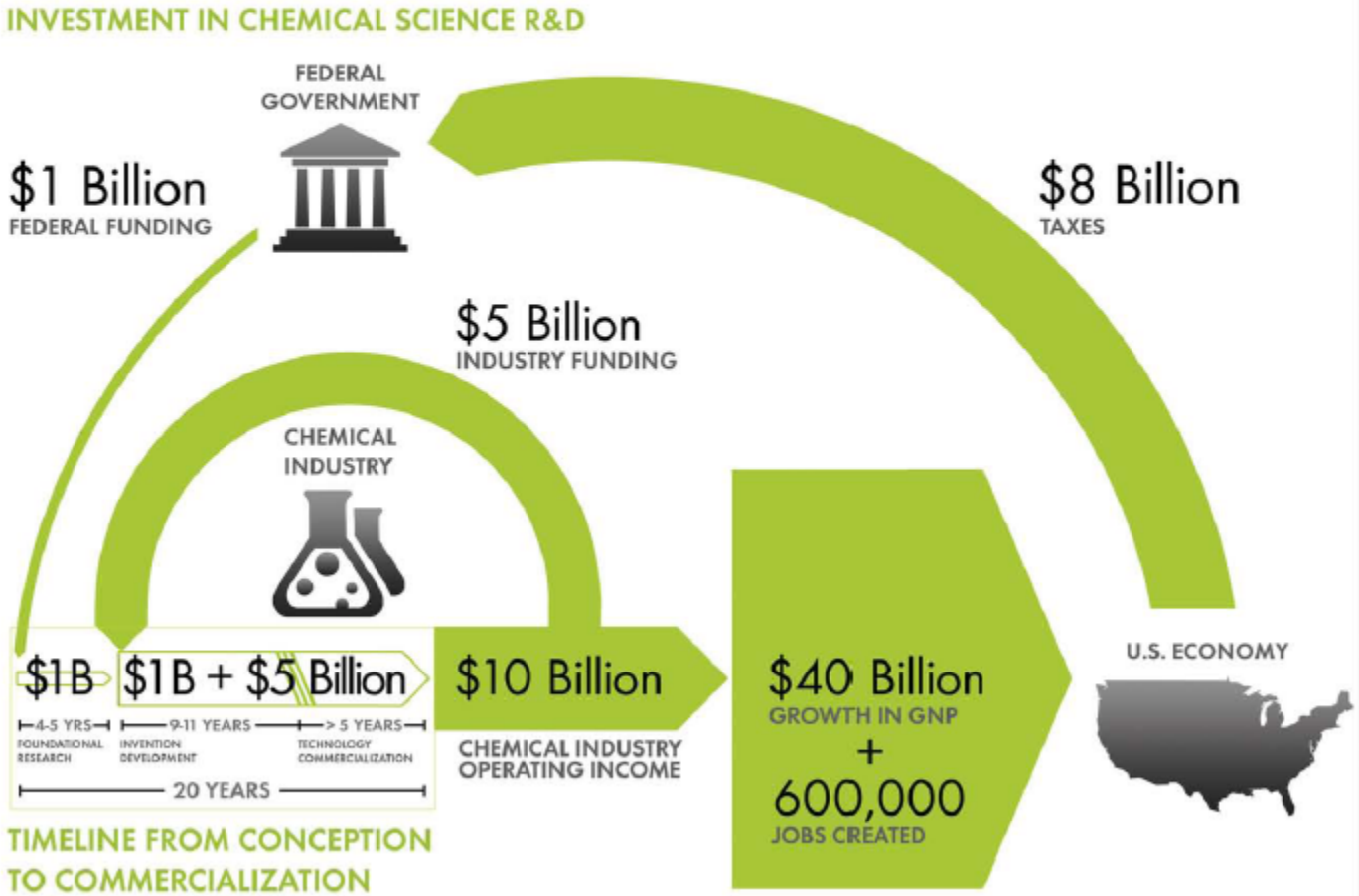
Language Communities of Twitter - Eric Fischer - 2012



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.

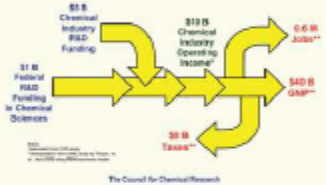
Chemical Research & Development Powers the U.S. Innovation Engine

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



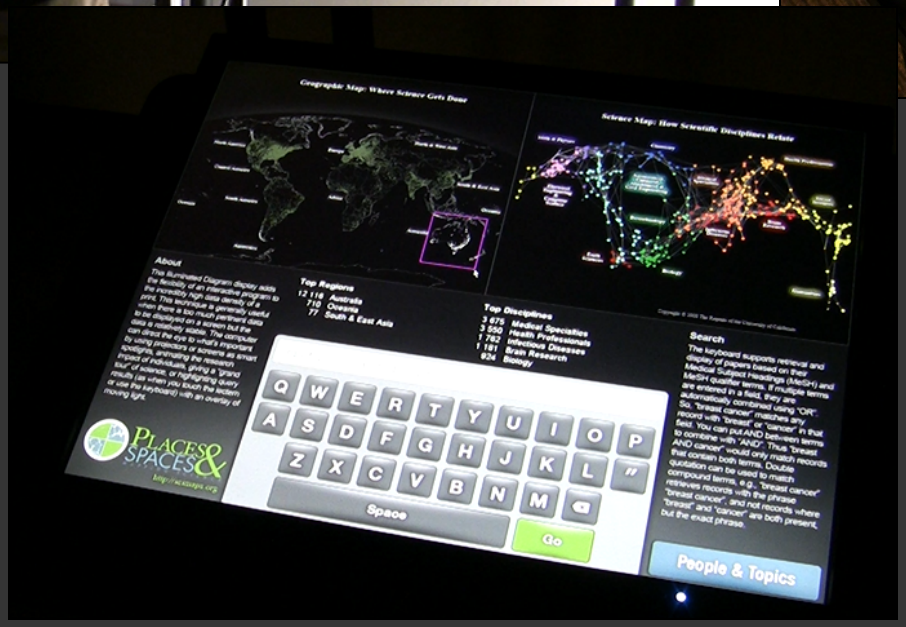
The Council for Chemical Research (CCR)

has provided the U.S. Congress and government policy makers with important results regarding the impact of Federal Research & Development (R&D) investments on U.S. innovation and global competitiveness through its commissioned 5-year two phase study. To take full advantage of typically brief access to policy makers, CCR developed the graphic below as a communication tool that distills the complex data produced by these studies in direct, concise and clear terms.



The design shows that an input of $1B in federal investment, leveraged by $5B industry investment, brings new technologies to market and results in $10B of operating income for the chemical industry, $40B growth in the Gross National Product (GNP) and further impacts the US economy by generating approximately 600,000 jobs, along with a return of $8B in taxes. Additional details, also reported in the CCR studies, are depicted in the map to the left. This map clearly shows the two R&D investment cycles; the shorter industry investment at the innovation stage to commercialization cycle; and the longer federal investment cycle which begins in basic research and culminates in national economic and job growth along with the increase tax base that in turn is available for investment in basic research.

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

Geographic Map: Where Science Gets Done

Science Map: How Scientific Disciplines Relate

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
MAPPING SCIENCE
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

Geographic Map: Where Science Gets Done

Science Map: How Scientific Disciplines Relate

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.

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.

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.

Cancer	Cloning	HIV	Robert G. Edwards	Roger D. Kornberg	Elinor Ostrom
Obesity	Quality of Life	Smoking	Stanley B. Prusiner	Ahmed H. Zewail	View All

Keyword Search

PLACES & SPACES MAPPING SCIENCE

http://scimaps.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

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


INDIANA
UNIVERSITY
PLACES AND SPACES
EXHIBIT

Science & Technology Forecasts @ Times Square in 2020

This is the only mockup in this slide show.
Everything else is available today.

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/

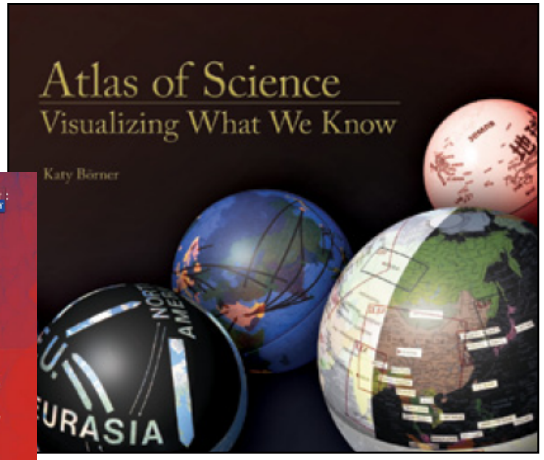
Börner, Katy (2010) **Atlas of Science: Visualizing What We Know**. The MIT Press. http://scimaps.org/atlas

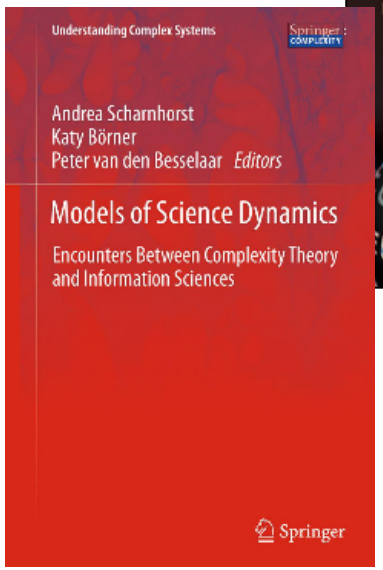
Scharnhorst, Andrea, Börner, Katy, van den Besselaar, Peter (2012) **Models of Science Dynamics**. Springer Verlag.

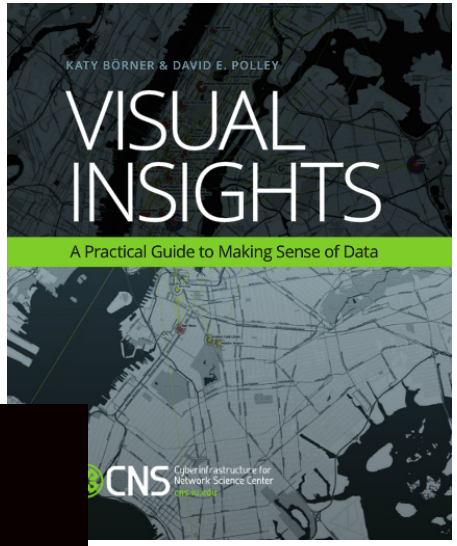
Katy Börner, Michael Conlon, Jon Corson-Rikert, Cornell, Ying Ding (2012) **VIVO: A Semantic Approach to Scholarly Networking and Discovery**. Morgan & Claypool.

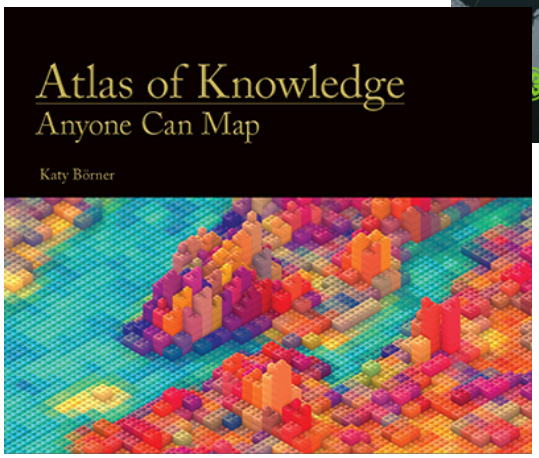
Katy Börner and David E Polley (2014) **Visual Insights: A Practical Guide to Making Sense of Data**. The MIT Press.

Börner, Katy (2015) **Atlas of Knowledge: Anyone Can Map**. The MIT Press. http://scimaps.org/atlas2









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