

Multi-Scale Maps of Scholarly Activity

Katy Börner

Royal Netherlands Academy of Arts and Sciences (KNAW),
Amsterdam, The Netherlands &
Cyberinfrastructure for Network Science Center, Director
School of Library and Information Science
Indiana University, Bloomington, IN
katy@indiana.edu



Research Talk at the Organisation for Economic
Co-Operation and Development (OECD)
Paris, France
<http://www.oecd.org>



April 12, 2012

Type of Analysis vs. Level of Analysis

	Micro/Individual (1-100 records)	Meso/Local (101-10,000 records)	Macro/Global (10,000 < records)
Statistical Analysis/Profiling	Individual person and their expertise profiles	Larger labs, centers, universities, research domains, or states	All of NSF, SA, all of sci
Temporal Analysis (When)	Funding portfolio of one individual	Topic bursts of PNAS	113 Years of PNAS Research
Geospatial Analysis (Where)	Career trajectory of one individual	Wrapping a state intellectual landscape	PNAS
Topical Analysis (What)	Research topics	Research	VxOrd/Topic research, NIH funding
Network Analysis (With Whom?)	NSF network of one	NIH network	NIH's research network

Modeling Science Dynamics using

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

Katy Börner, Kevin W. Boyack, Staša Milojević, Steven Morris. (2012) An introduction to modeling science: Basic model types, key definitions, and a general framework for the comparison of process models. In Scharnhorst, Andrea, van den Besselaar, Börner (Eds) Models of Science Dynamics. Springer Verlag.

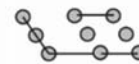
Temporal Levels

Highly dynamic processes
(download activity)

Slow processes
(citation activity)

Static structure

Data Types



Co-author network



Topic similarity network



Geospatial substrate for a set of authors

Reference Systems

Trends



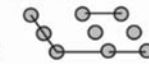
Geography



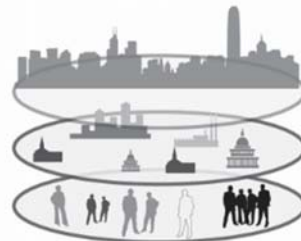
Topics



Co-authors



Levels of Aggregation



Population level

Group level

Individual level

Descriptive Models of Science

- Detect advances of scientific knowledge via "longitudinal mapping" (Garfield, 1994).
- Synthesis of specialty narratives from co-citation clusters (Small, 1986).
- Identify cross-disciplinary fertilization via "passages through science" (Small, 1999, 2000).
- Understand scholarly information foraging (Sandstrom, 2001).
- Knowledge discovery in un-connected terms (Swanson & Smalheiser, 1997).
- 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).
- 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, 1996; Martin & Irvine, 1983).
- Input (levels of funding, expertise of scientists, facilities used) - output (publications, patents, Nobel prizes, 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).

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.

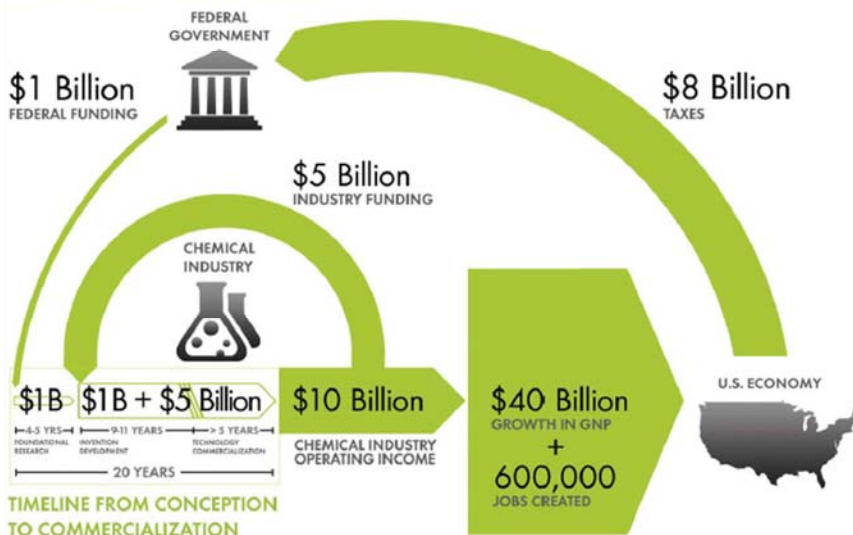
Chemical Research & Development Powers the U.S. Innovation Engine

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

The Council for Chemical Research (CCR)

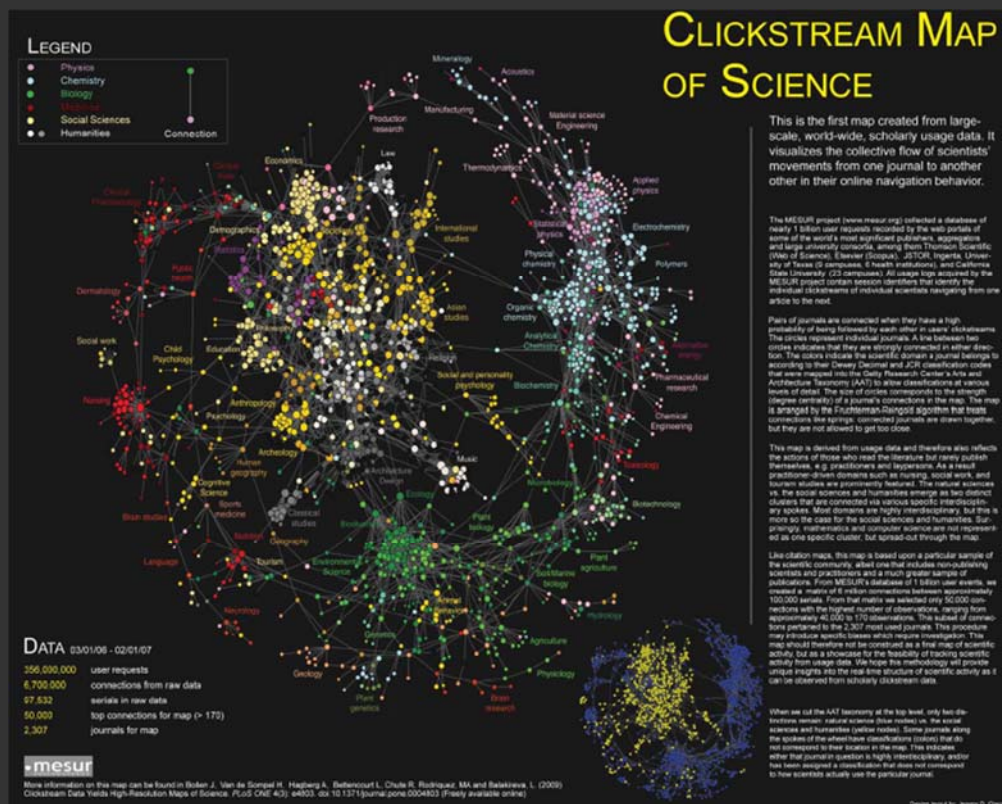
has provided the U.S. Congress and government policy makers with important results regarding the impact of Federal Research & Development (R&D) investments on U.S. innovation and global competitiveness through its commissioned 5-year two phase study. To take full advantage of typically brief access to 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.

INVESTMENT IN CHEMICAL SCIENCE R&D

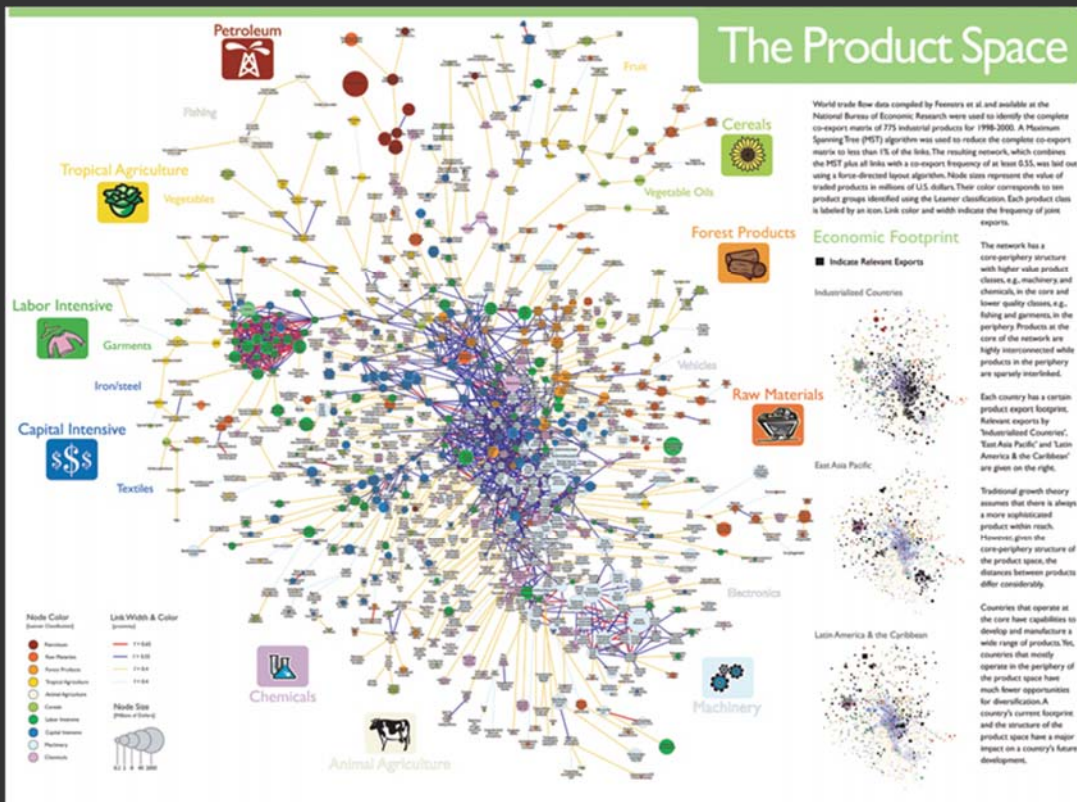


The design shows that an input of \$1B in federal investment, leveraged by \$5B industry investment, brings new technologies to market and results in \$10B of operating income for the chemical industry, \$40B growth in the Gross National Product (GNP) and further impacts the US economy by generating 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.



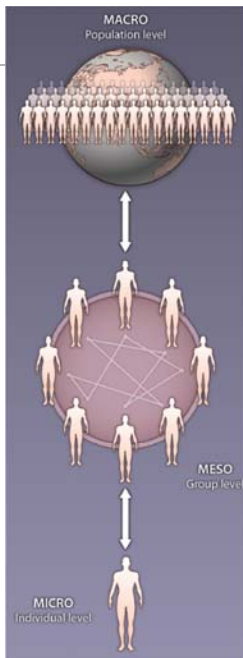
Bollen, Johan, Herbert Van de Sompel, Aric Hagberg, Luis M.A. Bettencourt, Ryan Chute, Marko A. Rodriguez, Lyudmila Balakireva. 2008. A Clickstream Map of Science.



Cesar A. Hidalgo, Bailey Klinger, Albert-László Barabási, Ricardo Hausmann. 2007. The Product Space



Adrian White and the National Geographic EarthPulse Team. 2008. A Global Projection of Subjective Well-being



TEAM SCIENCE

A Multi-Level Systems Perspective for the Science of Team Science

Katy Börner,^{1*} Noshir Contractor,² Holly J. Falk-Krzesinski,³ Stephen M. Fiore,⁴ Kara L. Hall,⁵ Joann Keyton,⁶ Bonnie Spring,⁷ Daniel Stokols,⁸ William Trochim,⁹ Brian Uzzi¹⁰

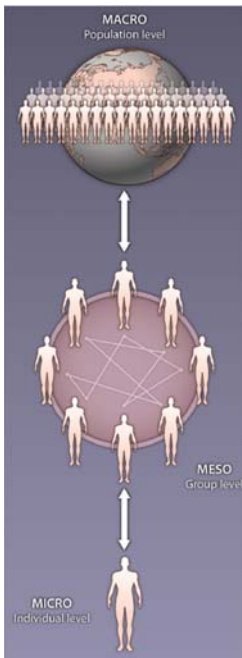
Published 15 September 2010; Volume 2 Issue 49 49cm24

This Commentary describes recent research progress and professional developments in the study of scientific teamwork, an area of inquiry termed the “science of team science” (SciTS, pronounced “sahyts”). It proposes a systems perspective that incorporates a mixed-methods approach to SciTS that is commensurate with the conceptual, methodological, and translational complexities addressed within the SciTS field. The theoretically grounded and practically useful framework is intended to integrate existing and future lines of SciTS research to facilitate the field’s evolution as it addresses key challenges spanning macro, meso, and micro levels of analysis.

Science of (team) science research and practice requires an interdisciplinary, multi-level, mixed-methods approach. **Expertise, theories, methods, data, and tools** from diverse research fields need to be applied and advanced to arrive at a holistic understanding of the science system.



Börner et al. 2010. "A Multi-Level Systems Perspective for the Science of Team Science". *Science Translational Medicine* 2 (49): 49(cm)24.



Mixed-Methods, Multi-Level Science of Science (or Team Science or SciSIP) studies need:

Expertise – identify and access it at the perfect moment using, e.g., Facebook, LinkedIn, Academia, VIVO, Harvard Profiles, Elsevier’s Collexis, Loki, Stanford’s CAP, or other systems.

Data – find, interlink, unify, merge, reformat, share them, e.g., using web sites analogous to <http://www.diggingintodata.org/Repositories/tabid/167/Default.aspx>, SDB, or LOD.

Tools – identify, learn, advance, share code, e.g., via Plug-and-Play Macroscopes, to arrive at a holistic understanding of the science system.



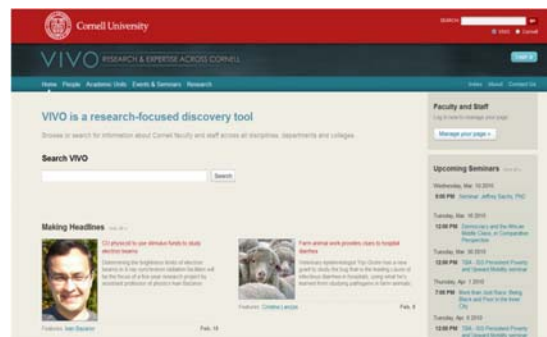
Expertise – identify and

access it at the perfect moment using, e.g., Facebook, LinkedIn, Academia, VIVO, Harvard Profiles, Elsevier's Collexis, Loki, Stanford's CAP, or other systems.

VIVO: A Semantic Approach to Creating a National Network of Researchers (<http://vivoweb.org>)



- Semantic web application and ontology editor originally developed at Cornell U.
- Integrates research and scholarship info from systems of record across institution(s).
- Facilitates research discovery and cross-disciplinary collaboration.
- Simplify reporting tasks, e.g., generate biosketch, department report.

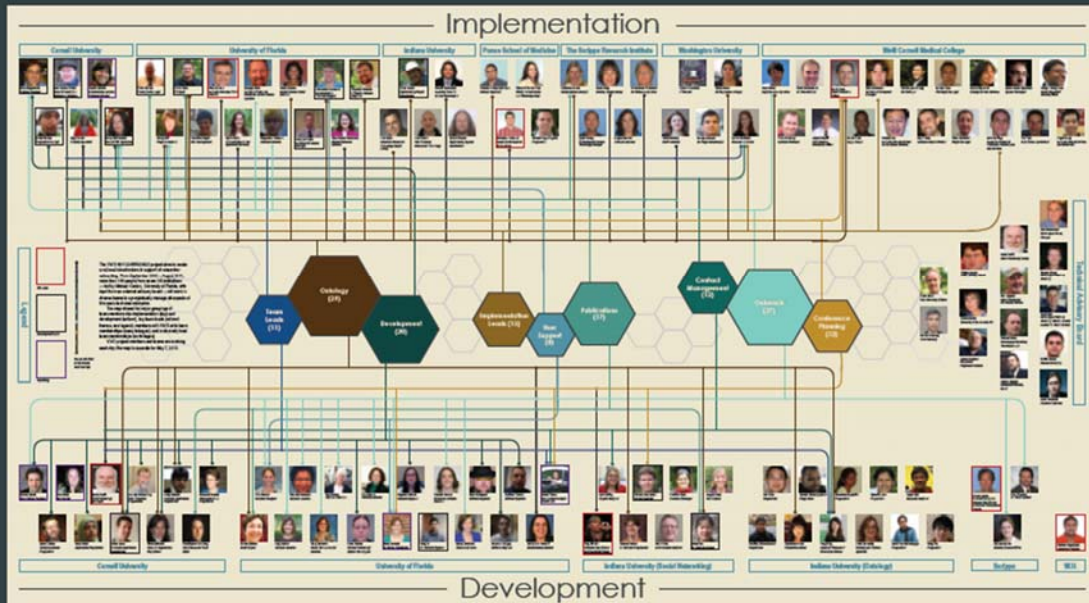


Funded by \$12 million NIH award.

Cornell University: Dean Krafft (Cornell PI), Manolo Bevia, Jim Blake, Nick Cappadona, Brian Caruso, Jon Corson-Rikert, Elly Cramer, Medha Devare, John Ferreira, Brian Lowe, Stella Mitchell, Holly Mistlebauer, Anup Sawant, Christopher Westling, Rebecca Younes. **University of Florida:** Mike Conlon (VIVO and UF PI), Cecilia Botero, Kerry Britt, Erin Brooks, Amy Buhler, Ellie Bushhousen, Chris Case, Valrie Davis, Nita Ferree, Chris Haines, Rae Jesano, Margeaux Johnson, Sara Kreinest, Yang Li, Paula Markes, Sara Russell Gonzalez, Alexander Rockwell, Nancy Schaefer, Michele R. Tennant, George Hack, Chris Barnes, Narayan Raum, Brenda Stevens, Alicia Turner, Stephen Williams. **Indiana University:** Katy Borner (IU PI), William Barnett, Shanshan Chen, Ying Ding, Russell Duhon, Jon Dunn, Micah Linnemeier, Nianli Ma, Robert McDonald, Barbara Ann O'Leary, Mark Price, Yuyin Sun, Alan Walsh, Brian Wheeler, Angela Zoss. **Ponce School of Medicine:** Richard Noel (Ponce PI), Ricardo Espada, Damaris Torres. **The Scripps Research Institute:** Gerald Joyce (Scripps PI), Greg Dunlap, Catherine Dunn, Brant Kelley, Paula King, Angela Murrell, Barbara Noble, Cary Thomas, Michaelen Trimarchi. **Washington University, St. Louis:** Rakesh Nagarajan (WUSTL PI), Kristi L. Holmes, Sunita B. Koul, Leslie D. McIntosh. **Weill Cornell Medical College:** Curtis Cole (Weill PI), Paul Albert, Victor Brodsky, Adam Cheriff, Oscar Cruz, Dan Dickinson, Chris Huang, Itay Klaz, Peter Michelini, Grace Migliorisi, John Ruffing, Jason Specland, Tru Tran, Jesse Turner, Vinay Varughese.

VIVO Enabling National Networking of Scientists

Project Members and Teams



Development

Please send comments and questions to jeff.culley@vivo-network.org (design) and john.s.boyce@ufl.edu (lead developer) and john.boyce@ufl.edu (lead developer). For more information, visit www.vivo-network.org.

2010.05.07

University of Florida

How do you want to compare?
by Grants

Who do you want to compare?
Search: X

Records 1 - 10 of 30

Entity Label	Grant Count	Entity Type
<input checked="" type="checkbox"/> Continuing Education	562	UF Department, Agent, Non-Academic Department, Department
<input checked="" type="checkbox"/> Florida Museum of Natural History	203	Museum, Agent
<input checked="" type="checkbox"/> College of Agricultural and Life Sciences	166	Agent, UF College, College
<input checked="" type="checkbox"/> College of Engineering	103	Agent, UF College, College
<input checked="" type="checkbox"/> Evelyn F. and William L. McKnight Brain Institute of the University of Florida	64	UF Center, Agent, Center
<input checked="" type="checkbox"/> International Center	54	UF Department, Agent, Non-Academic Department, Department
<input checked="" type="checkbox"/> Florida Sea Grant	44	UF Center, Agent, Center
<input type="checkbox"/> Whitney Laboratory for Marine Bioscience	42	UF Research Laboratory, Agent, Laboratory, Research Laboratory
<input type="checkbox"/> Water Institute	38	UF Center, Agent, Center
<input type="checkbox"/> College of Dentistry	35	Agent, UF College, College

[Save as CSV](#) [Clear](#)

Comparing Grants of Organizations in University of Florida

Total Number of Grants

You have selected 7 of a maximum 10 organizations to compare. [Clear](#)

- Florida Sea Grant 44
- International Center 54
- Evelyn F. and William L. McKnight Brain Institute of the University of Florida 64
- College of Engineering 103
- College of Agricultural and Life Sciences 166
- Florida Museum of Natural History 203
- Continuing Education 562

Temporal Analysis (When) Temporal visualizations of the number of papers/funding award at the institution, school, department, and people level

17

enabling national networking of scientists

Index Log in

Search

Home

People

Organizations

Research

Events

University of Florida

Explore 487 publications activity across 554 scientific sub-disciplines

13 Disciplines | 554 Sub-Disciplines

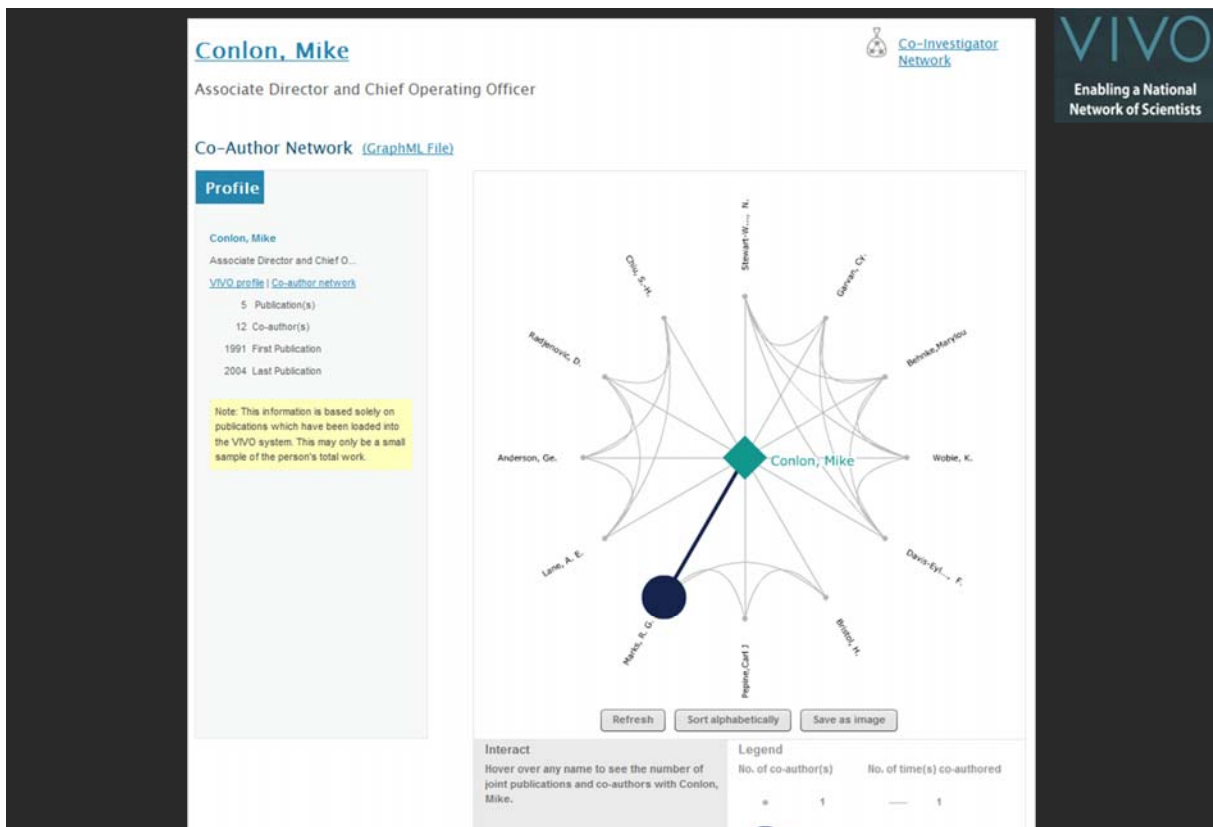
Search: X

1 - 13 of 554

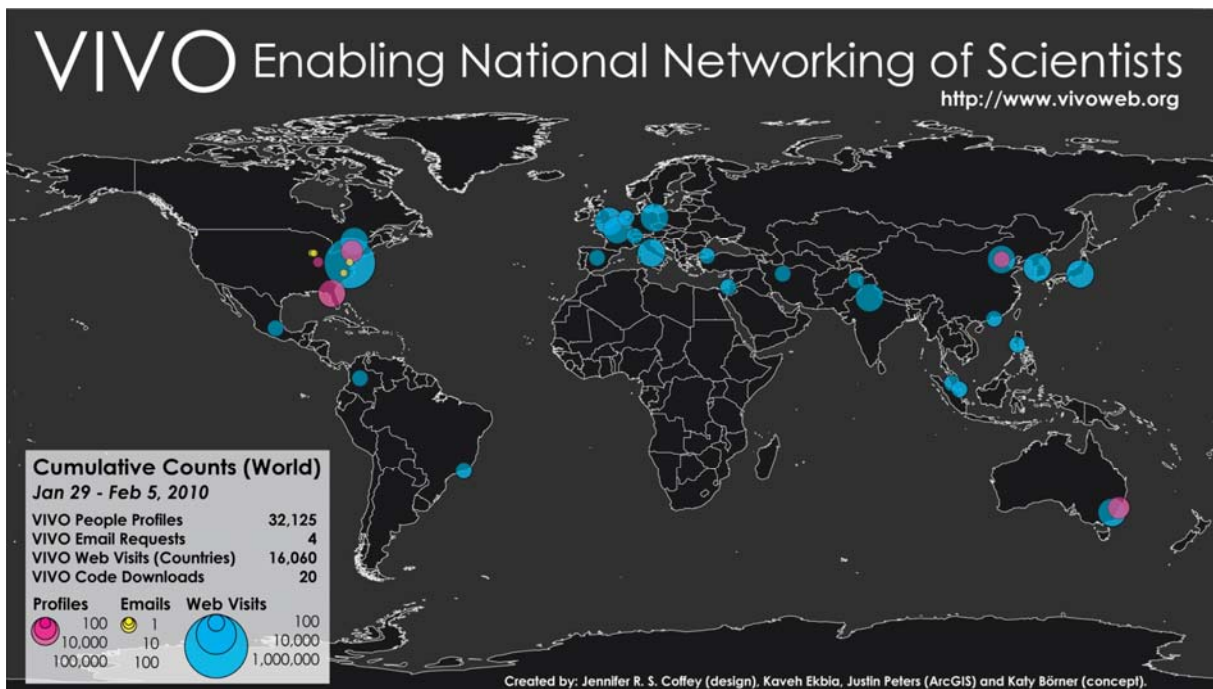
Sub-Disciplines	# of pubs.	% activity
Pest Management Science	24.2	5.0
Wildlife Research	19.1	3.9
Protein Science	13.1	2.7
Clinical Cancer Research	12.6	2.6
Pain	12.0	2.5
Environmental Contamination	11.2	2.3
Insect Physiology	11.1	2.3
Organic Chemistry	10.9	2.2
Marine Biology	10.3	2.1
Computer Aided Molecular Design	10.2	2.1
BioStatistics	9.0	1.9

mapped 14.55% of 3,346 publications

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



Network Analysis (With Whom?) Who is co-authoring, co-investigating, co-inventing with whom? What teams are most productive in what projects?

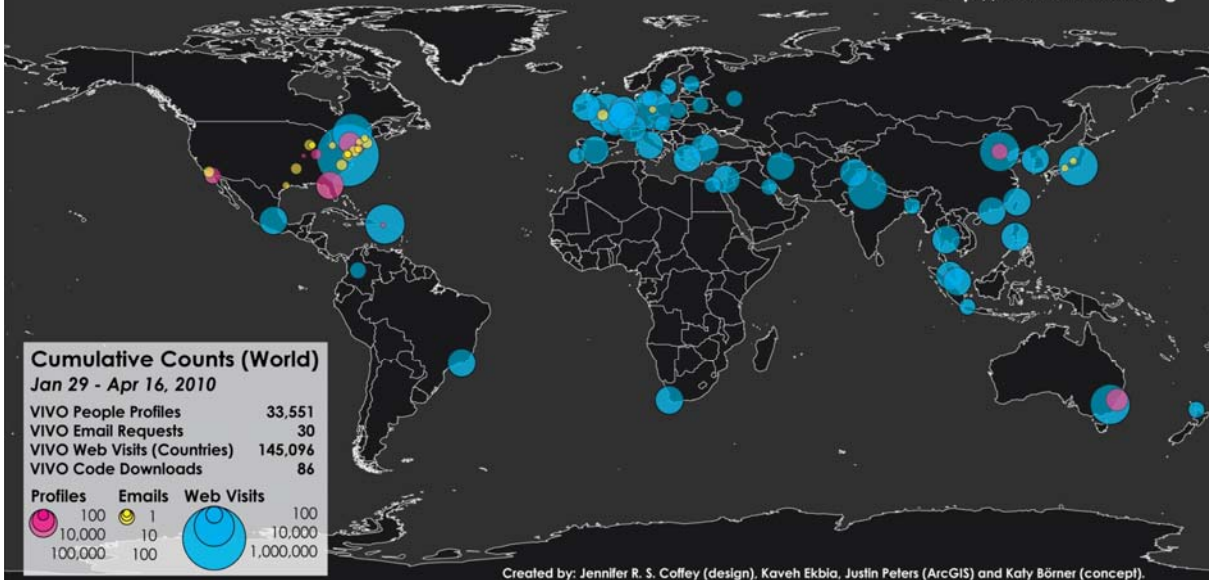


Science is global. World view of VIVO activity.
Web site visits are aggregated at the country level.

Geospatial Analysis (Where) Where is what science performed by whom? Science is global and needs to be studied globally. (in work)

VIVO Enabling National Networking of Scientists

<http://www.vivoweb.org>



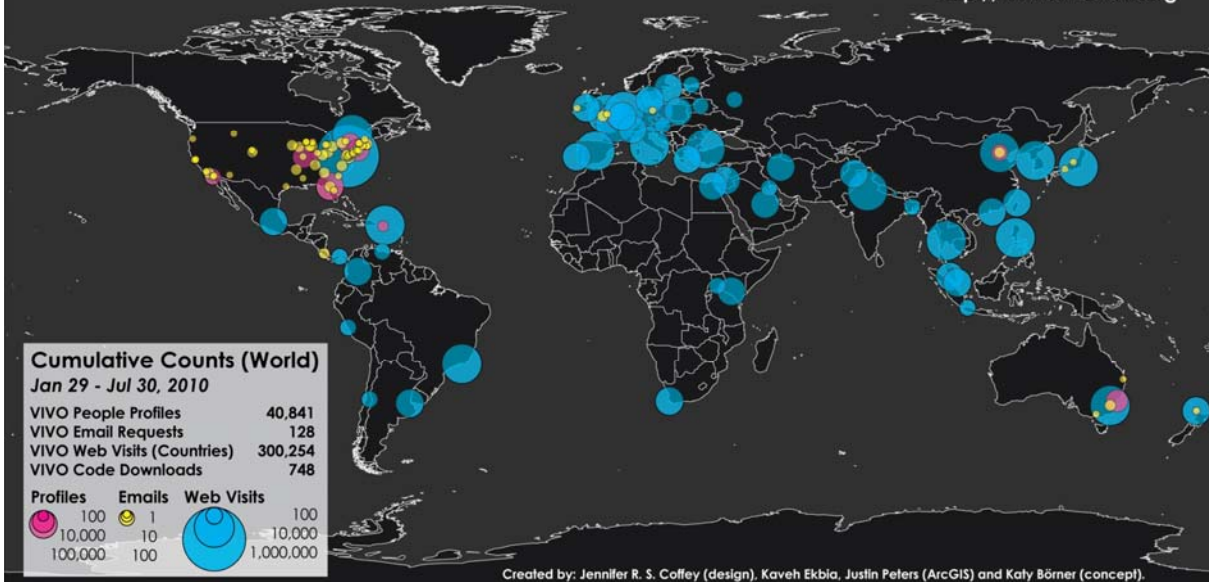
Shown are the

- Number of people profiles in the 7 different VIVO installation sites plus CAS and U Melbourne.
 - Email contacts by data and service providers as well as institutions interested to adopt VIVO.
 - The number of visitors on <http://vivoweb.org>
- Circles are area size coded using a logarithmic scale.

21

VIVO Enabling National Networking of Scientists

<http://www.vivoweb.org>



VIVO 1.0 source code was publicly released on April 14, 2010

87 downloads by June 11, 2010.

The more institutions adopt VIVO, the more high quality data will be available to understand, navigate, manage, utilize, and communicate progress in science and technology.

22

Home Index About Search

Networks and Complex Systems Research at Indiana University

This VIVO instance provides information on networks and complex systems

- [Faculty](#) and their [departments](#)
- [Publications](#)
- [Grants](#)
- [Courses](#)

at Indiana University. The site was created in support of a NSF IGERT grant application. A major intent is to cross-fertilize between research done in the social and behavioral sciences, research in natural sciences such as biology or physics, but also research on Internet technologies.

The site will be continuously updated to help

- New faculty to get in contact with relevant researchers.
- Faculty and policy makers to pool teams in response to funding solicitations.
- Faculty to coordinate research efforts – collaborations using existing funding/resources.
- Faculty to coordinate teaching.
- Students identify relevant courses, potential advisors, funding.
- Organize the Mon talk series on [Networks and Complex Systems](#).
- Arrange research meetings for visitors with relevant faculty/students

<http://vivo-netsci.cns.iu.edu>

23



Data

– find, access, interlink, unify,

merge, reformat, share them, e.g., using web sites analogous to <http://www.diggingintodata.org/Repositories/tabid/167/Default.aspx>, SDB, or LOD.

24

Supports federated search of 25 million publication, patent, grant records.

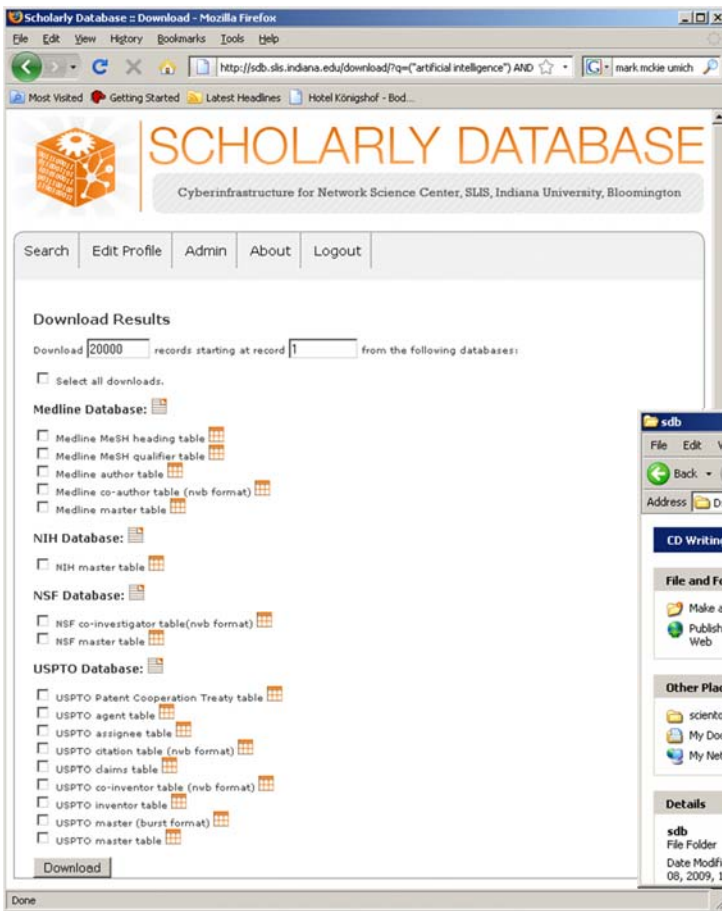
Results can be downloaded as data dump and (evolving) co-author, paper-citation networks.

The screenshot shows the Scholarly Database homepage. On the left, there are login options for 'IU User' and 'Non-IU User'. The 'IU User' section includes a 'Go to IU Login' button. The 'Non-IU User' section has fields for 'Email' and 'Password' with a 'Login' button. Below the login options, there is a 'Not Registered Yet?' section with a 'Register as an IU User' link. The main content area features a search interface with a search bar, 'Search' button, and various filters like 'Creators', 'Title', 'Abstract', 'Full Text', 'First Year', and 'Last Year'. There are also checkboxes for database sources: Medline (1998 - 2008), NIH (1961 - 2002), NSF (1985 - 2004), and USPTO (1976 - 2007). A 'Search' button is at the bottom of the search interface.

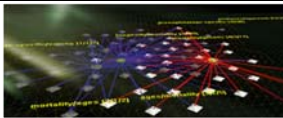
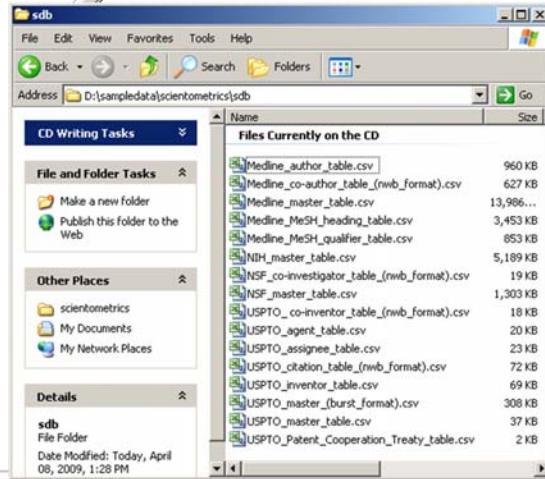
Register for free access at <http://sdb.cns.iu.edu>

The screenshot shows the search results page in a Mozilla Firefox browser. The browser's address bar shows the URL: [http://sdb.sls.indiana.edu/search/results?q=\(\"artificial intelligence\"\)](http://sdb.sls.indiana.edu/search/results?q=(\). The page header includes the Scholarly Database logo and the text 'Cyberinfrastructure for Network Science Center, SLIS, Indiana University, Bloomington'. Below the header, there are navigation links: Search, Edit Profile, Admin, About, Logout. The main content area is titled 'Browse Results' and displays the search results for 'artificial intelligence'. It shows that the search returned 13,231 results in 0.295 seconds. A 'Download' button is visible. Below this, it lists the total results per database: NIH: 2,103, Medline: 10,235, USPTO: 279, NSF: 614. The results are displayed in a table format, showing the first 20 results. The table has columns for Source, Authors/Creators, Year, Title, and Score (out of 5.71).

Source	Authors/Creators	Year	Title	Score (out of 5.71)
Medline	LaCombe	1987	Artificial intelligence.	5.71
Medline		1989	Artificial intelligence: expert systems.	5.71
Medline	Schmitt	1990	[Artificial intelligence in dentistry]	5.71
Medline	Adlassnig and Adlassnig	2002	Artificial-intelligence-augmented systems.	5.60
Medline	Touretzky	1980	Artificial intelligence.	4.86
Medline	Goldenberg	1980	Artificial intelligence.	4.86



Since March 2009:
 Users can download networks:
 - Co-author
 - Co-investigator
 - Co-inventor
 - Patent citation
 and tables for
 burst analysis in NWB.



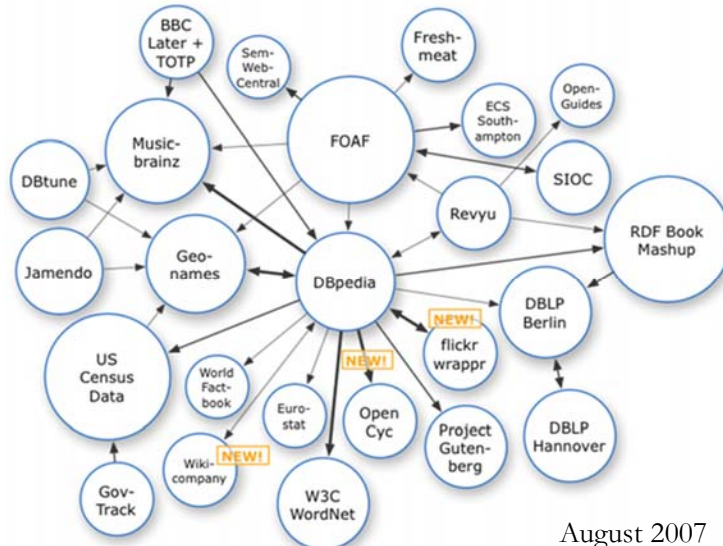
Semantic Web: Linked Open Data

- Interlinking existing data silos and
- Exposing them as structured data
- Adding new high quality data relevant for S&T studies

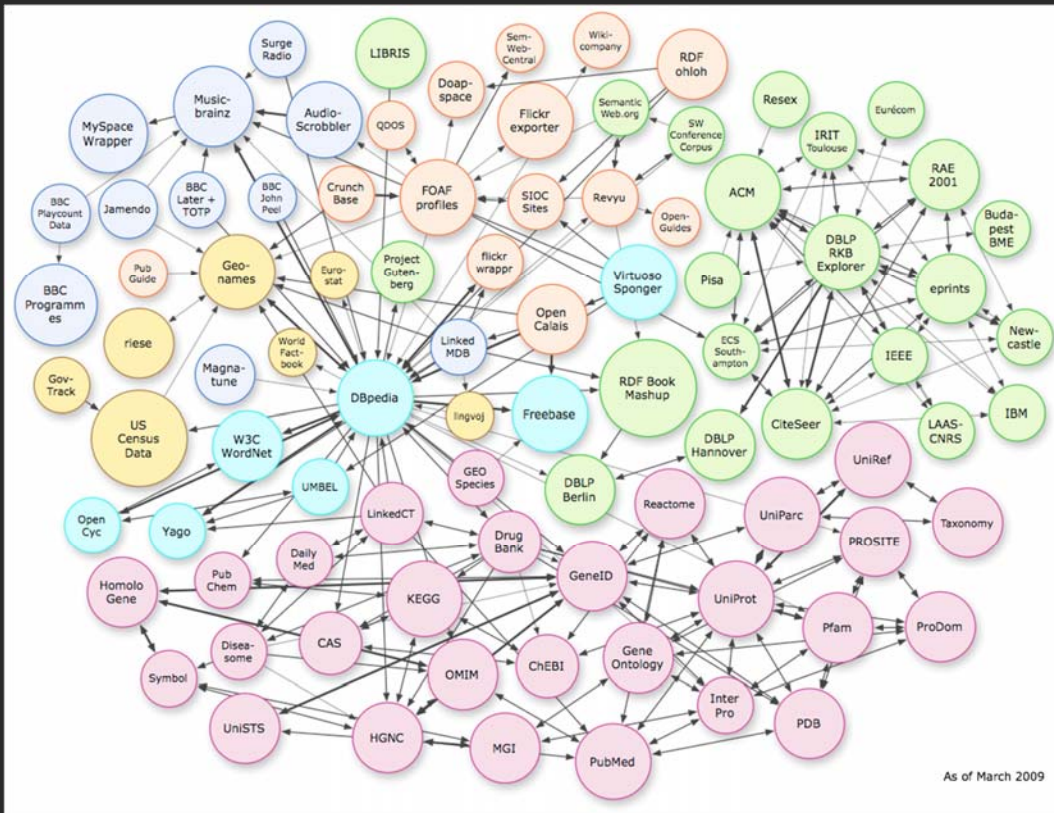
<http://linkeddata.org>

Save Data.gov, sign
 the petition at
<http://om.ly/BRPRE>

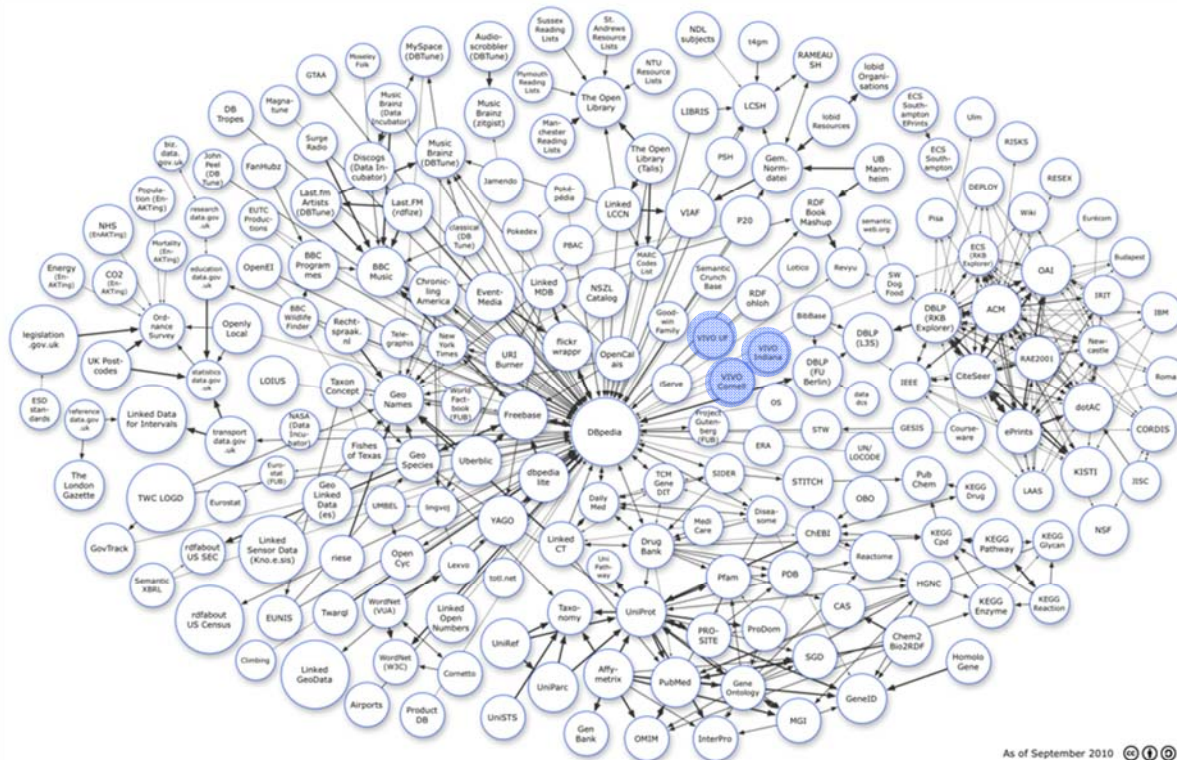
Twitter
 #savethedata



August 2007



http://www4.wirvis.fu-berlin.de/bizer/pub/lod-datasets_2009-03-05_colored.png



National Research Networking (NRN) Activity Visualization

VIVO

Enabling a National Network of Scientists



Federated Search University of Florida

Search Term: Search

Results for "cancer".

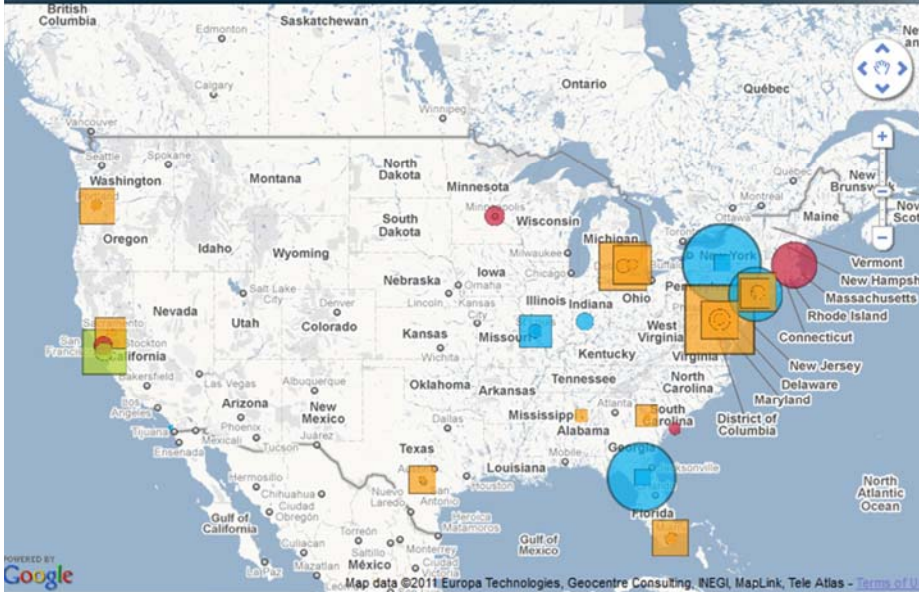
University of Florida [29 People](#)
 UF VIVO contains all 6,700 faculty and 7,400 full-time staff of the University of Florida, as well as award information for all grants, UF students, affiliates and employees of Shands HealthCare can request to be added.

Cornell University [200 People](#)
 Participants in the VIVO National Network include institutions with local installations of VIVO or those with research discovery and profiling applications that can provide subject to user-provided data.

<http://vivoexperts.csi.ufl.edu>

National Researcher Networking Visualization 1.0

cyberinfrastructure for NETWORK SCIENCE CENTER
cns.iu.edu



Data Types

Check in the boxes next to the data types you'd like to see displayed on the map.

- People
- Publications
- Patents
- Funding
- Courses

System Types

- Elsevier SciVal Experts
- Harvard Catalyst Profiles
- Stanford CAP
- VIVO

Timeline

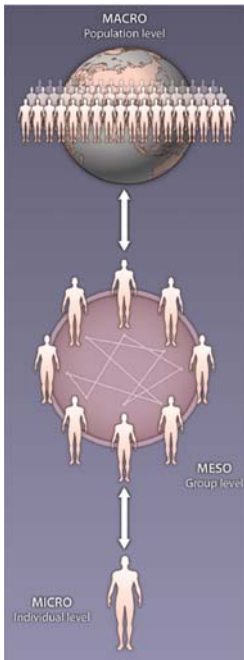
Jun, 2011

Jun, 2011 Jun, 2011

Play Pause Stop

Hide Map Data About

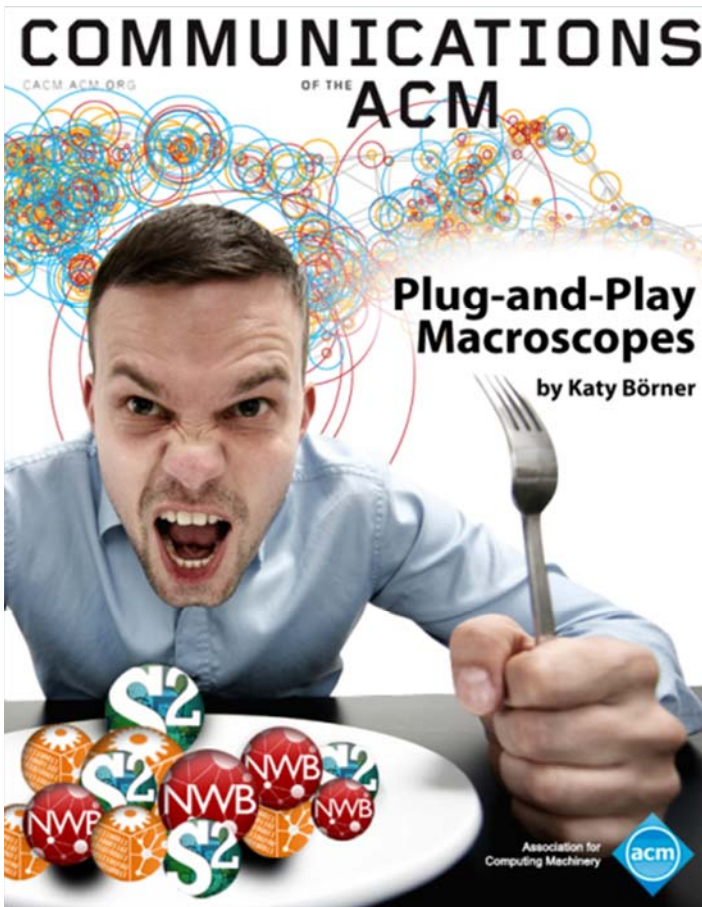
<http://nrn.cns.iu.edu>



Tools

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

33



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>

34



Designing “Dream Tools”

Many of the best micro-, tele-, and macrosopes 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 **“macrosopes”** that may solve both local and global challenges.

Plug-and-play macrosopes **empower** me, my students, colleagues, and 100,000 others that downloaded them.

35



Macrosopes

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.

Macrosopes 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, **macrosopes let us observe what is at once too great, slow, or complex for the human eye and mind to notice and comprehend.**



Microscopes



Telescopes



Macrosopes

36



Macroscopes cont.

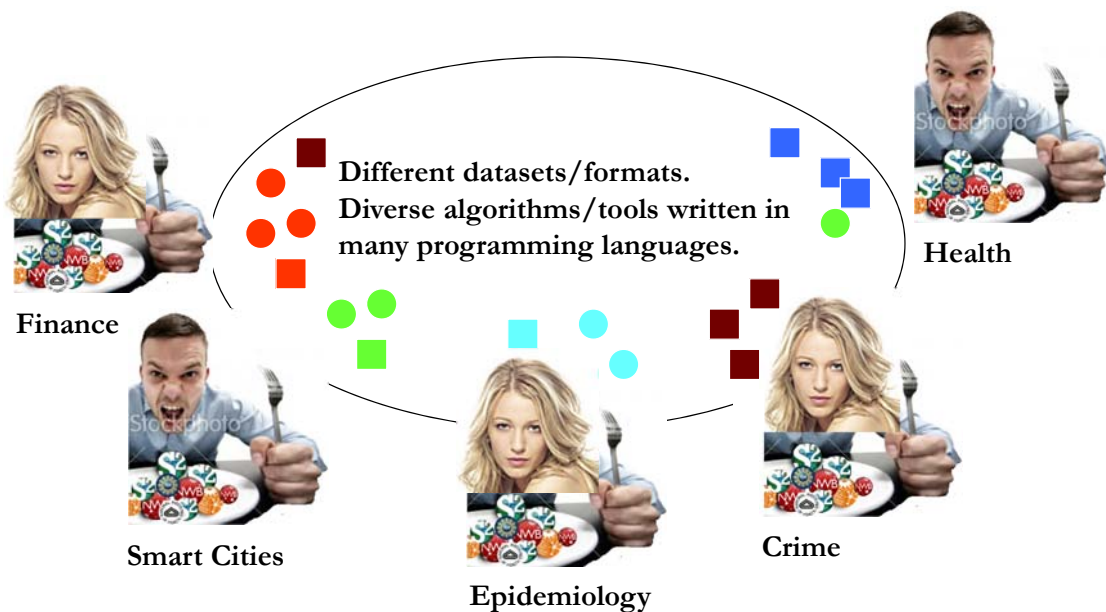
While microscopes and telescopes are physical instruments, **macroscopes resemble continuously changing bundles of software plug-ins.** Macroscopes make it easy to select and combine algorithm and tool plug-ins but also interface plug-ins, workflow support, logging, scheduling, and other plug-ins needed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

37



Macroscopes Serve the Changing Scientific Landscape



38



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

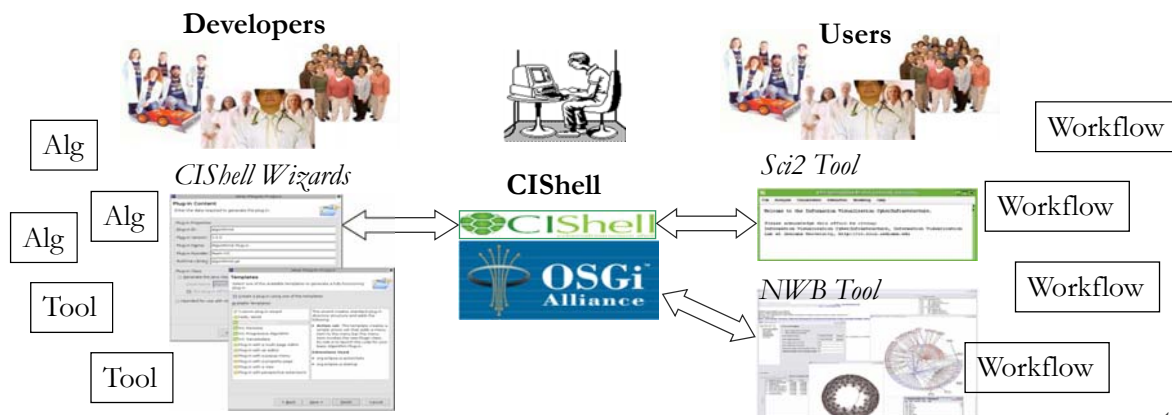
- the design of **modular, compatible algorithm and tool plug-ins** (also called “modules” or “components”)
- that can be **easily combined into scientific workflows** (“pipeline” or “composition”),
- and packaged as **custom tools**.

39



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.



40



CIShell Developer Guide

(<http://cishell.wiki.cns.iu.edu>)



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](#)

41



CIShell Portal (<http://cishell.org>)

Cyberinfrastructure Shell (CIShell)
CIShell supports the plug-and-play of datasets and algorithms and their bundling into custom tools that serve the specific needs of a user group or research community. It has been applied to develop diverse custom tools, see below. Feel free to take plugins from any of these tools to design your personal dream tool.

Provided by the [Cyberinfrastructure for Network Science Center](#) at Indiana University.

Visit the CIShell wiki
to learn more about using CIShell as a platform for your tool!

Learn more about existing CIShell-powered tools below.

Network Workbench Tool (NWB)
The NWB Tool supports researchers, educators, and practitioners interested in the study of biomedical, social and behavioral science, physics, and other networks. It comes with a 77-page [user manual](#).

Gallery

Science of Science Tool (Sci²)
The Sci² Tool was specifically developed for science policy makers and researchers that study science by scientific means. It supports the temporal, geospatial, topical, and network analysis and visualization of scholarly datasets at the micro (individual), meso (local), and macro (global) levels. There exists a [112-page user manual](#) and 24 hours of [NIM tutorials](#) in this tool.

42

Sci² Tool
A tool for science of science research & practice

Email Address

Password

Login

Forgot your password?
To recover your account password, please visit our [password recovery page](#).

Not registered yet?
[Register now](#)

Tutorials
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](#)

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

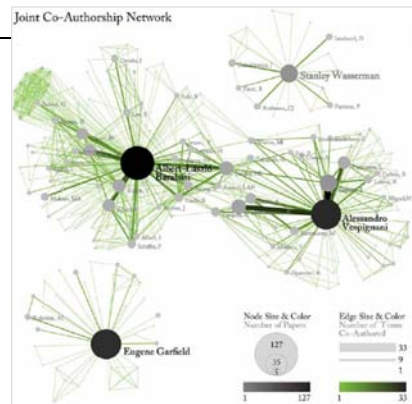
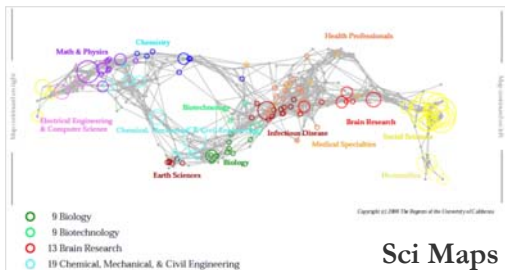
NIH Office of Extramural Research and Katy Börner (2010) [Network Visualizations Using SPIRES Data and the Sci2 Tool](#). Office of Extramural Research at NIH.

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

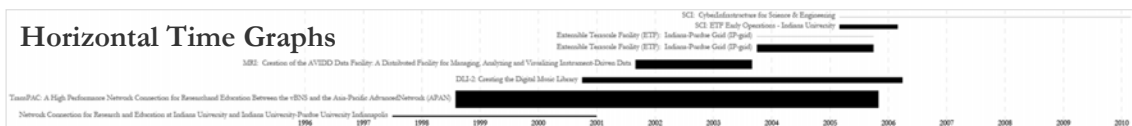


Sci² Tool – “Open Code for S&T Assessment”

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



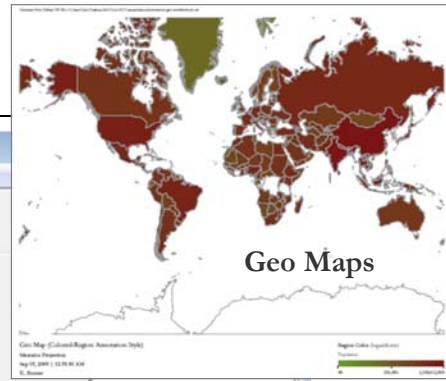
Horizontal Time Graphs



Börner, Katy, Huang, Weixia (Bonnie), Linnemeier, Micah, Dubon, Russell Jackson, Phillips, Patrick, Ma, Nianli, Zoss, Angela, Guo, Hanning & Price, Mark. (2009). *Rele-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 (Sci²). The development of this tool is supported in Network Science center and the School of Li Indiana University, the National Science Foundation and IIS-0715303, and the James S. McDonnell Cyberinfrastructure portal (<http://sci.slis.indiana.edu>).

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

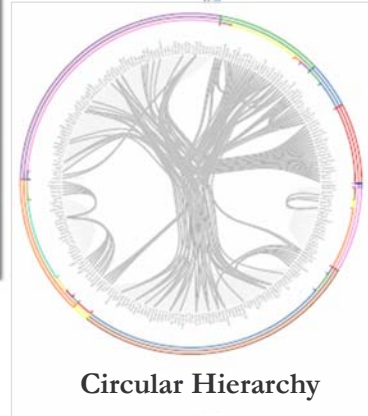
Please cite as follows:
Sci² Team. (2009). Science of Science Tool. Indiana University. Network Science Center, <http://sci.slis.indiana.edu>.

Scheduler

Remove From List Remove completed

!	Algorithm Name	Date	Time	% Complete
<input checked="" type="checkbox"/>	Extract Co-Author Network	09/03/2009	00:15:20 AM	100%
<input checked="" type="checkbox"/>	Load and Clean ISI File	09/03/2009	00:15:05 AM	100%

- 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



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.

The EpiC tool is supported in part by the NIH RM-07-004 award. The primary investigators are Chintan Tank, Joseph Biberstine, and Dr. Jim Sherman.

The EpiC tool is supported in part by the NIH RM-07-004 award. The primary investigators are Chintan Tank, Joseph Biberstine, and Dr. Jim Sherman.

EpiC uses the Cyberinfrastructure Shell (<http://cishell.org>) developed at the Cyberinfrastructure for Network Science Center (<http://cns.slis.indiana.edu>).

Please cite as follows:
EpiC Team. (2009). EpiC Tool. Indiana University. Network Science Center, <http://cns.slis.indiana.edu>.

Scheduler

Remove From List Remove completed automatically Remove all completed

!	Algorithm Name	Date	Time	% Complete
---	----------------	------	------	------------

Data Manager

Simulation Visualization R Help

- Single-Population
- Exact
- Network

Visualization R Help

- Line Graph

R Help

- Create an R Instance
- Run Rgui
- Import Table Into R
- Export Table From R



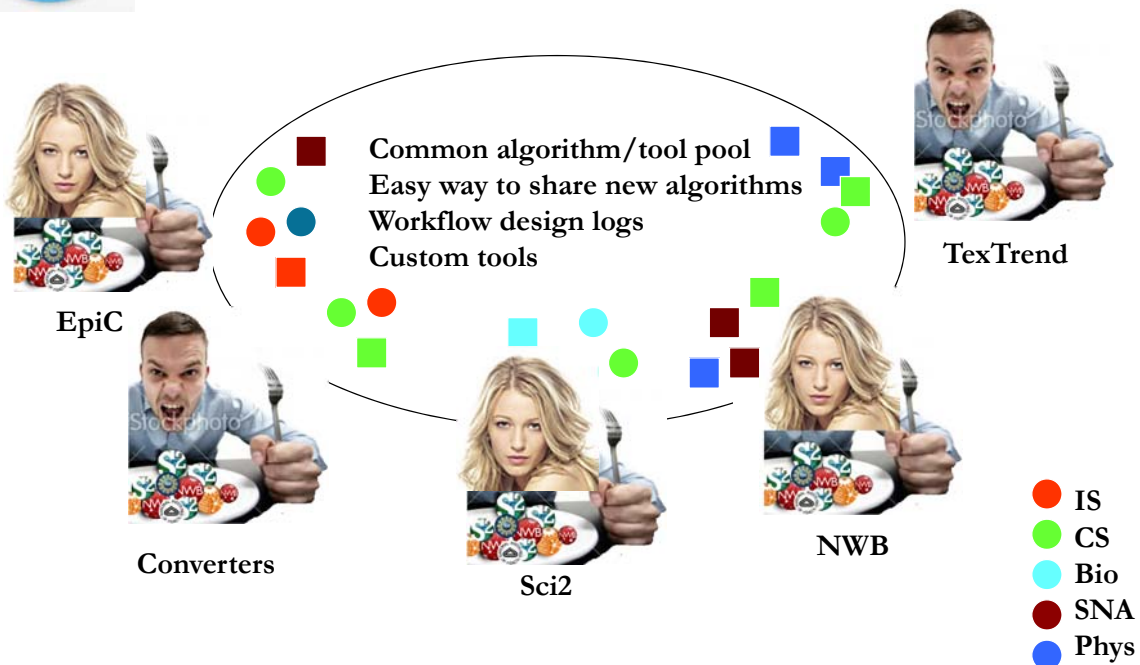
OSGi/CIShell Adoption

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.



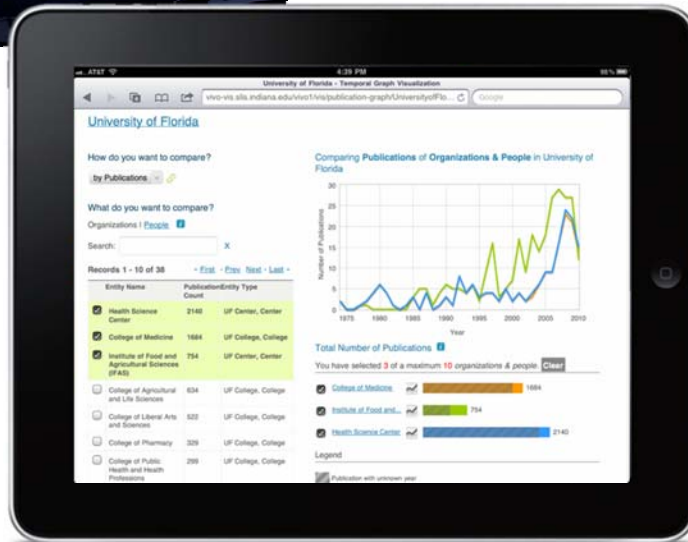
Embrace the Changing Scientific Landscape





Few have access to or time to visit
“Visualization Domes”

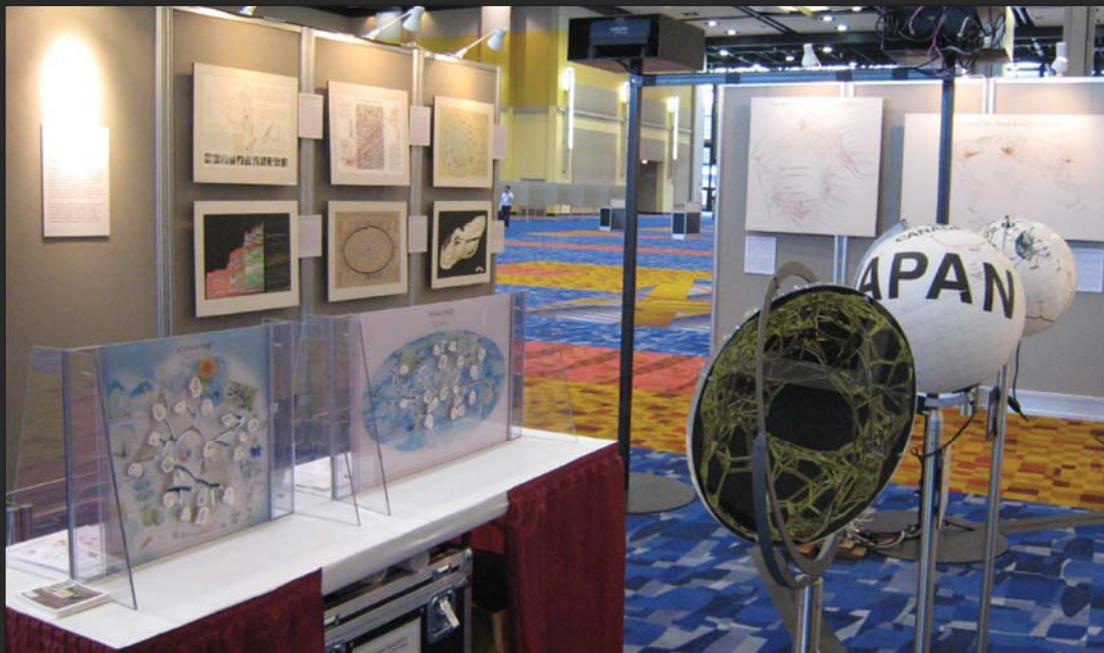
Overview, Interactivity,
Details on Demand
must come to
commonly
used devices
and environments



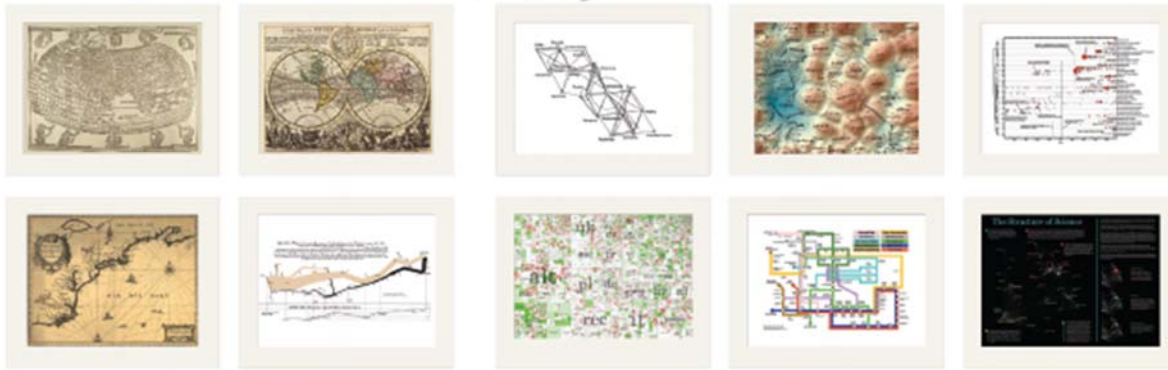
 INDIANA UNIVERSITY

Mapping Science Exhibit – 10 Iterations in 10 years

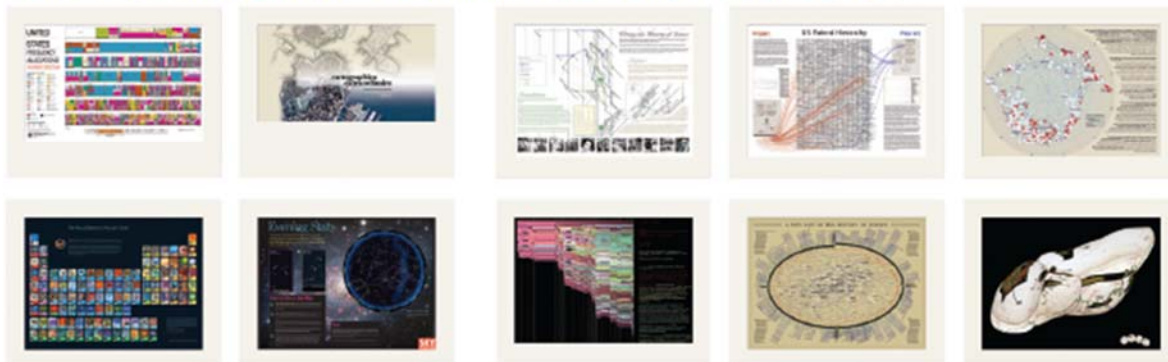
<http://scimaps.org/>



THE POWER OF MAPS 2005



THE POWER OF REFERENCE SYSTEMS 2006

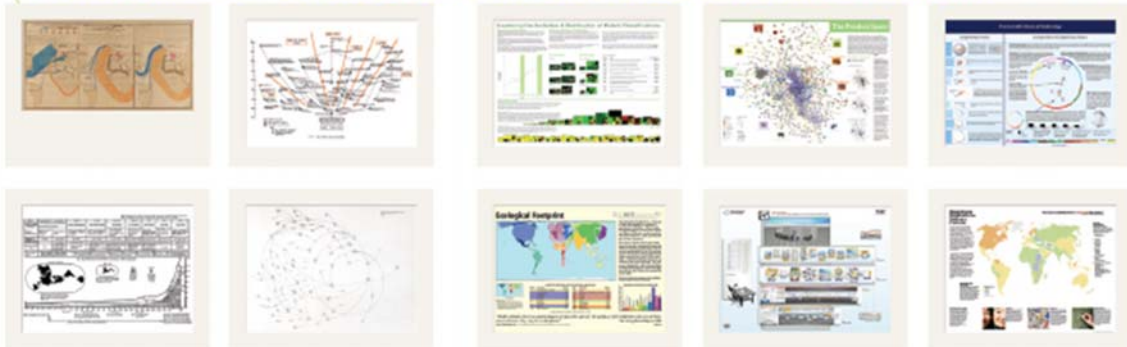


51

THE POWER OF FORECASTS 2007



SCIENCE MAPS FOR ECONOMIC DECISION MAKERS 2008



52

SCIENCE MAPS FOR SCIENCE POLICY MAKERS 2009



SCIENCE MAPS FOR SCHOLARS 2010



53

SCIENCE MAPS AS VISUAL INTERFACES TO DIGITAL LIBRARIES 2011



Read about and zoom into maps at http://scimaps.org/exhibit_info

Call for Maps for the 8th Iteration of the Places & Spaces: Mapping Science Exhibit on "Science Maps for Kids" (2012)

<http://scimaps.org/call>

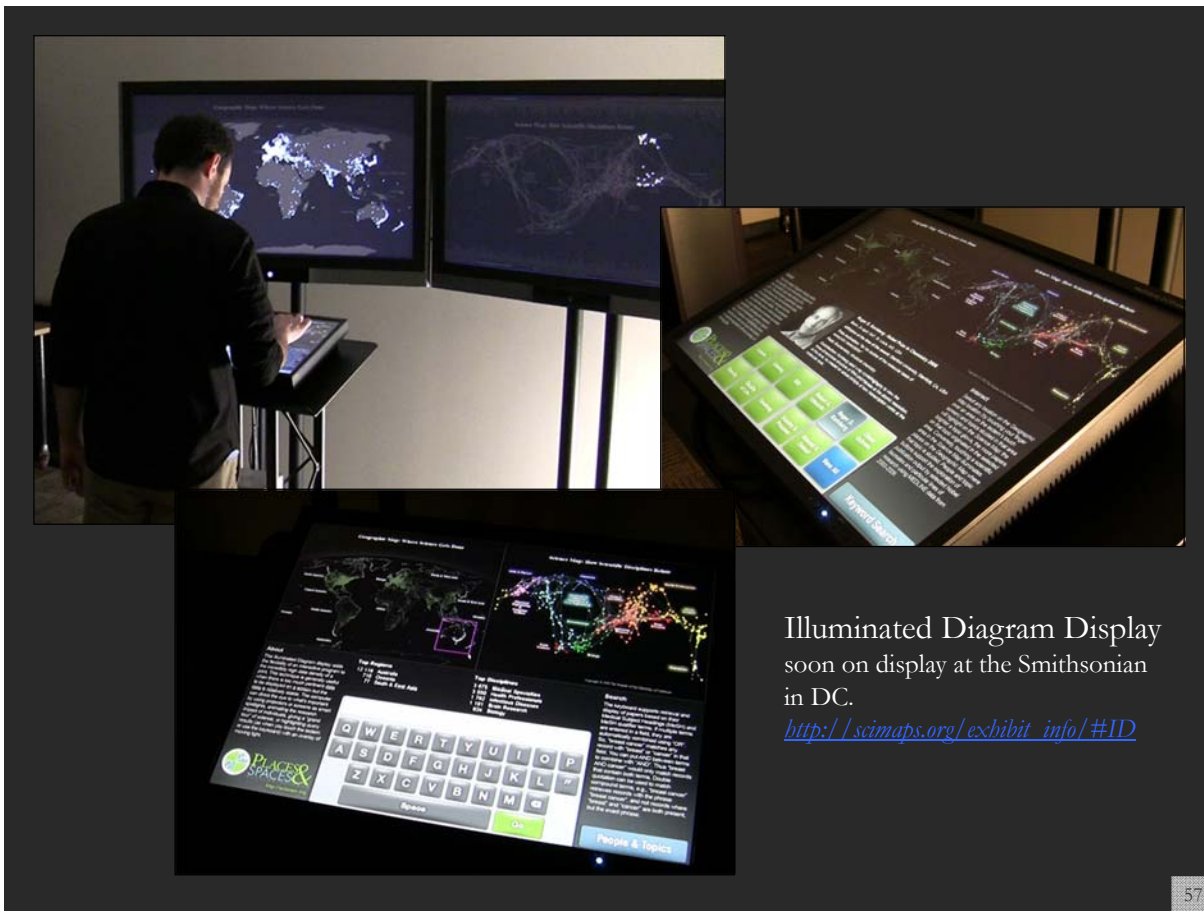
54



Mapping Science Exhibit at MEDIA X was on May 18, 2009 at Wallenberg Hall, Stanford University,
<http://mediax.stanford.edu>, <http://scaleindependentthought.typepad.com/photos/scimaps>



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



Illuminated Diagram Display soon 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.

Cloning

Cloning is the process of making a genetically identical copy. Cloning can refer to the technique of producing a genetic copy of an organism by replacing the nucleus of an unfertilized egg with the nucleus of a body cell from the organism. The reconstructed egg containing the DNA from a donor cell must be treated with chemicals or electric current in order to stimulate cell division. Once the cloned embryo reaches a suitable stage, it is transferred to the uterus of a female host where it continues to develop until birth.

The first adult mammal cloned was Dolly the Sheep in 1997.

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

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.

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.

Keyword Search

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

59

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.

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 Professions - 3,589
- Social Sciences - 2,431
- Aeronautical, Chemical, Mechanical & Civil Engineering - 2,208
- Humanities - 1,562

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

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

60



Ingo Gunther's Worldprocessor globe design now on display at the Giant Geo Cosmos OLED Display at the Museum of Emerging Science and Innovation in Tokyo, Japan



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/

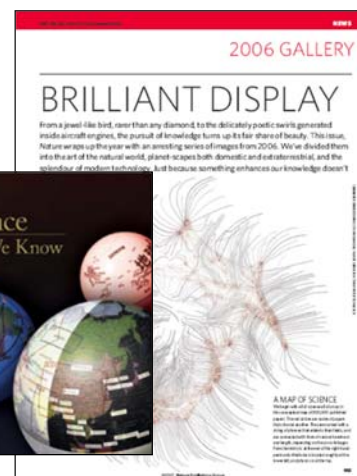
Börner, Katy, Sanyal, Soma and Vespignani, Alessandro (2007). **Network Science**. In Blaise Cronin (Ed.), *ARIST*, Information Today, Inc., Volume 41, Chapter 12, pp. 537-607.

<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.



63

Scientometrics (2011) 89:421–435
DOI 10.1007/s11192-011-0433-7

Mixed-indicators model for identifying emerging research areas

Hanning Guo · Scott Weingart · Katy Börner

Abstract This study presents a mixed model that combines different indicators to describe and predict key structural and dynamic features of emerging research areas. Three indicators are combined: sudden increases in the frequency of specific words; the number and speed by which new authors are attracted to an emerging research area, and changes in the interdisciplinarity of cited references. The mixed model is applied to four emerging research areas: RNAi, Nano, h-Index, and Impact Factor research using papers published in the *Proceedings of the National Academy of Sciences of the United States of America* (1982–2009) and in *Scientometrics* (1978–2009). Results are compared in terms of strengths and temporal dynamics. Results show that the indicators are indicative of emerging areas and they exhibit interesting temporal correlations: new authors enter the area first, then the interdisciplinarity of paper references increases, then word bursts occur. All workflows are reported in a manner that supports replication and extension by others.

Keywords Burst detection · Prediction · Emerging trend · Temporal dynamics · Science of science (Sci²) tool

<http://ivl.cns.in.edu/km/pub/2011-guo-mixed-indicators.pdf>



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>