

Building a Science Observatory

Katy Börner

Visiting Scientist in Dirk Helbing's Group, SOMS, ETHZ
Cyberinfrastructure for Network Science Center, Director
Information Visualization Laboratory, Director
School of Library and Information Science
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

*Fidelity Investments
Boston, MA*

July 21, 2011



Börner: Building a Science Observatory

2

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

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

(<http://www.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 (Team) 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)

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.

Temporal Levels

Highly dynamic processes
(download activity)

Slow processes
(citation activity)

Static structure

Reference Systems

Trends

Geography



Topics

Co-authors



Data Types



Co-author network

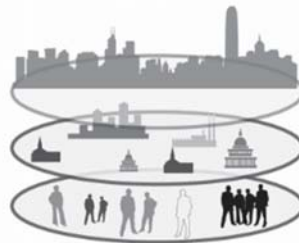


Topic similarity network



Geospatial substrate for a set of 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.

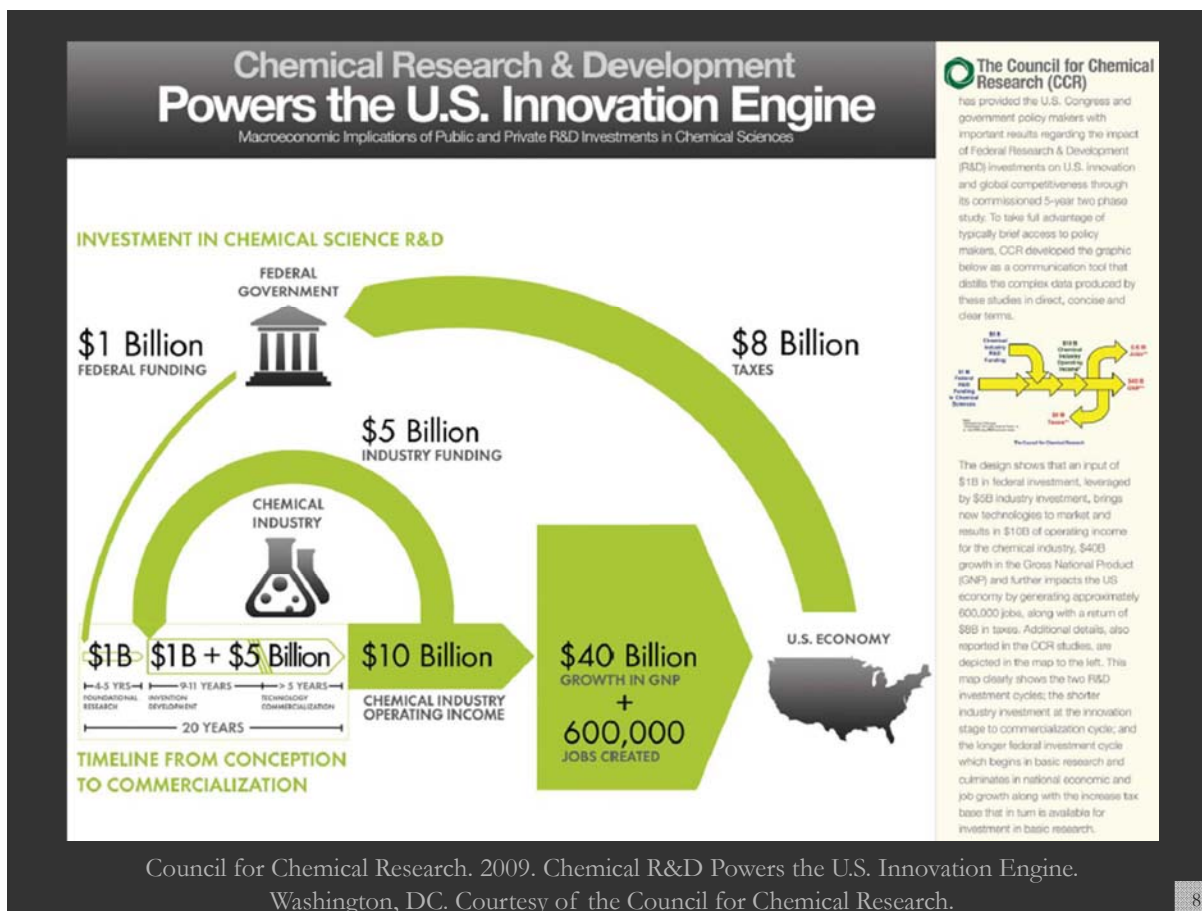
- Identification of scientific frontiers <http://www.science-frontiers.com/>.
- ISI's *Essential Science Indicators* <http://essentialscience.com/>
- Import-export studies (Stigler, 1994).
- Evaluation of 'big science' facilities using 'converging partial indicators' (Martin, 1996; Martin & Irvine, 1983).
- 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).

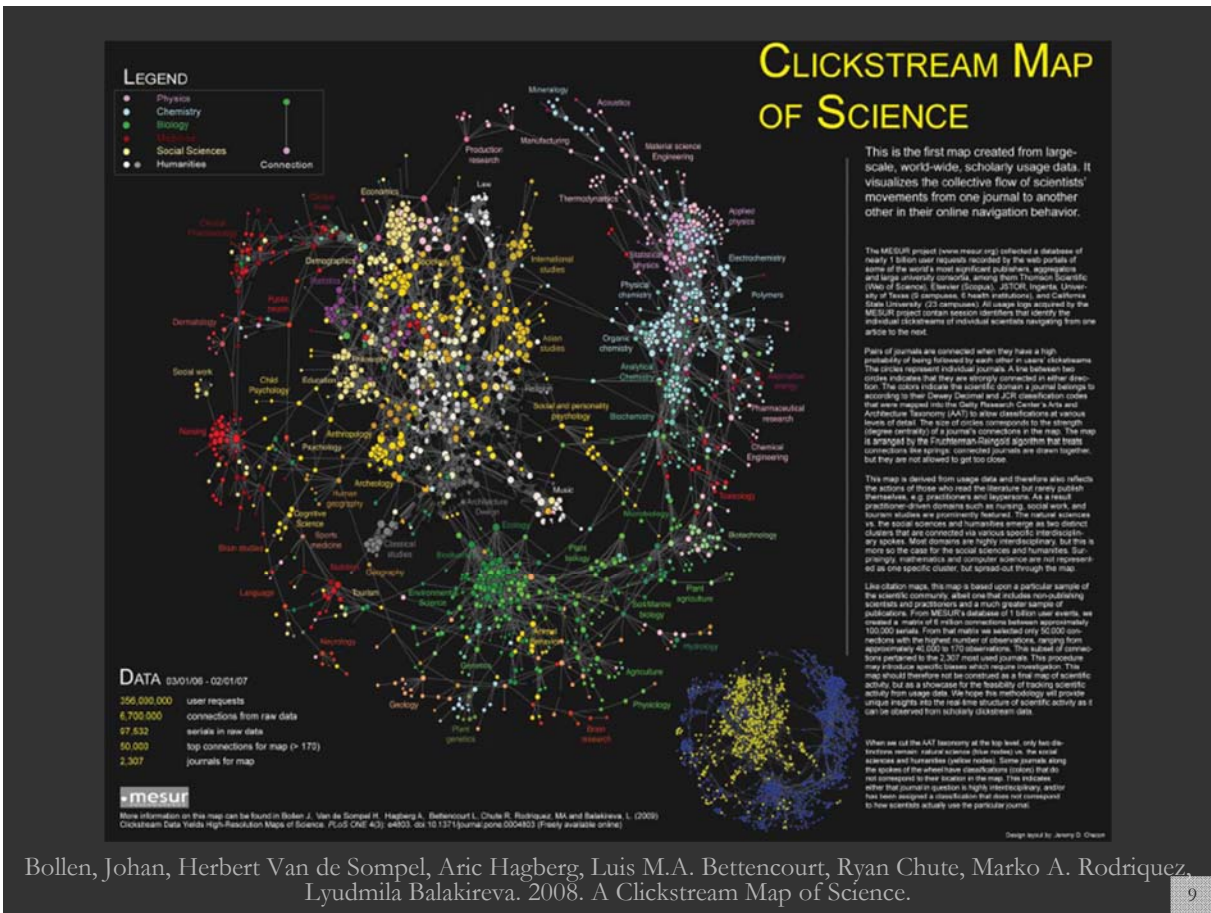
Process Models of Science

Can be used to predict the effects of

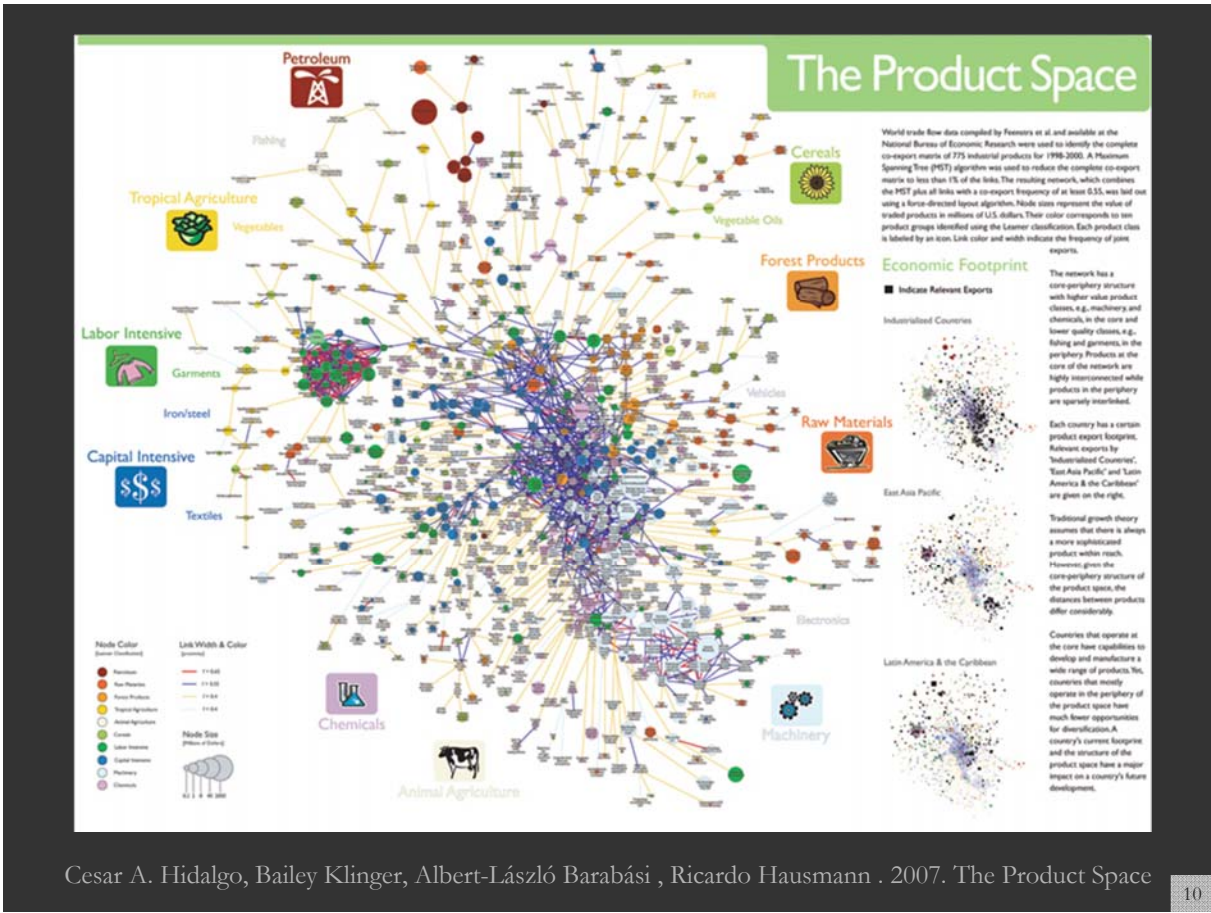
- Large collaborations vs. single author research on information diffusion.
- Different publishing mechanisms, e.g., E-journals vs. books on co-authorship, speed of publication, etc.
- Supporting disciplinary vs. interdisciplinary collaborations.
- Many small vs. one large grant on # publications, Ph.D.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





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. 9



Cesar A. Hidalgo, Bailey Klinger, Albert-László Barabási, Ricardo Hausmann. 2007. The Product Space 10

Happiness Depends on Various Factors

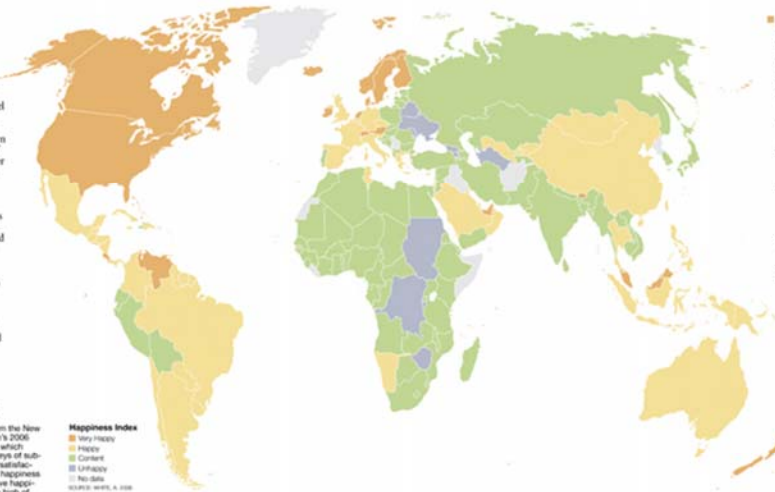
Social scientists are starting to include relative happiness with hard data on economic status, health, and other factors as they assess quality of life. They rely on surveys of "subjective well-being"—how good people feel about their lives. A world map of one "happiness index" shows many, but not all, wealthy northern countries faring well. Residents of sub-Saharan Africa and the former Soviet Union, meanwhile, report particularly low levels of contentment.

Any attempt to measure happiness will fall short—each life is a series of joys, struggles, and sorrows, and satisfaction can depend as much on outlook as on circumstances. Averages obscure the happy moments in struggling nations, as well as people who suffer from poor health, poverty, or discrimination in countries that rank high. Still, happiness indices can help researchers move beyond simple economics as they track progress—or backsliding—over time.

MEASURING THE INTANGIBLE

The map is based on the New Economics Foundation's 2006 "Happy Planet Index," which draws on over 100 surveys of subjective well-being. Its "satisfaction with life scale"—a happiness index—rates the relative happiness of nations, from a high of 273 (Denmark and Switzerland) to a low of 100 (Burundi).

Happiness Index
 Very Happy
 Happy
 Content
 Unhappy
 No data source with a link



RANKING THE WORLD'S HAPPIEST PLACES

Northern Europe, North America, and several wealthy countries make the list, but so do many less prosperous island nations.

- 1 DENMARK
SWITZERLAND
- 2 AUSTRIA
ICELAND
- 3 BAHAMAS
FINLAND
SWEDEN
- 4 BHUTAN
BRUNEI
CANADA
IRELAND
LUXEMBOURG
- 5 COSTA RICA
MALTA
NETHERLANDS
- 6 ANTIQUA AND BARBUDA
MALAYSIA
NEW ZEALAND
MOROCCO
SEYCHELLES
ST. KITTS AND NEVIS
UNITED ARAB EMIRATES
UNITED STATES
VANUATU
VENEZUELA

DEFINING WELL-BEING

By comparing the happiness index to data from the UN, the CIA, and other sources, a U.K. psychologist determined that good health and health care, enough money for fundamental needs, and access to basic education are the most important factors for subjective well-being. European countries top all three measures.



HEALTH

Japan boasts the world's longest life expectancy—a measure of overall health. Swaziland, at the other end of the scale, is plagued by poverty, disease, and violence. Disparities in access to health care divide many countries into haves and have-nots.



WEALTH

Money still can't buy love, or happiness, and wealthier people aren't always more content. Still, tiny Luxembourg, which takes top rank in per capita Gross Domestic Product (GDP), also rates a 213 on the happiness index. Real poverty means real misery, a fate shared by billions.



EDUCATION

Residents of Australia can expect to spend more time in school—an average of almost 21 years—than citizens of any other country. But only a basic education is needed to see a significant jump in overall happiness. Around the world, hundreds of millions lack even that.

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

11

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.

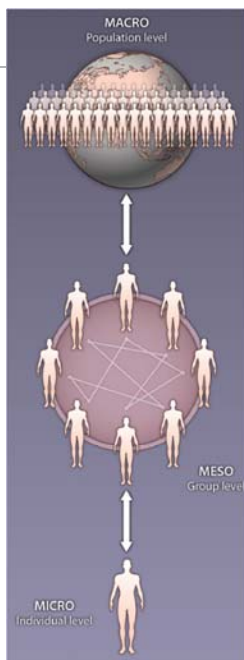
Reviewers—read, critique, and suggest changes to help improve the quality of papers and funding proposals. They need to identify related works that should be cited or complementary skills that authors might consider when selecting project collaborators.

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.



TEAM SCIENCE

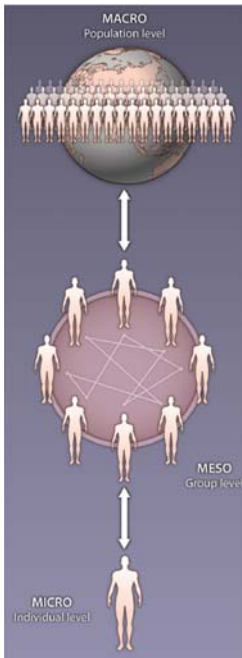
A Multi-Level Systems Perspective for the Science of Team Science

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.

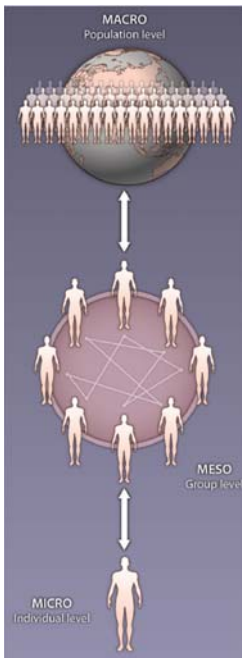


Mixed-Methods, Multi-Level Science of Science (or Team Science or SciSIP) studies need:

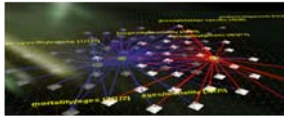
Expertise – identify and access it at the perfect moment using, e.g., Facebook, LinkedIn, Academia, VIVO, Harvard Profiles, Elsevier’s Collexis, Loki, Stanford’s CAP, or other systems.

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.

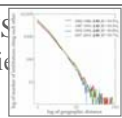



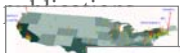
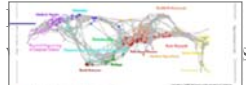


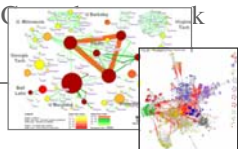
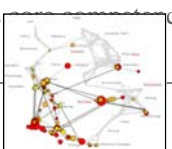
Tools – identify, learn, advance, share code, e.g., via Plug-and-Play Microscopes, to arrive at a holistic understanding of the science system.



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.



Type of Analysis vs. Level of Analysis

	Micro/Individual <i>(1-100 records)</i>	Meso/Local <i>(101-10,000 records)</i>	Macro/Global <i>(10,000 < records)</i>
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 P Research 
Geospatial Analysis (Where)	Career trajectory of one individual	Mapping a st intellectual k 	PNAS 
Topical Analysis (What)	 s.	flows in research	VxOrd/Topic r NIH funding 
Network Analysis (With Whom?)	NSF one work of 	work 	NIH's 

How do you want to compare?

by Grants

Who do you want to compare?

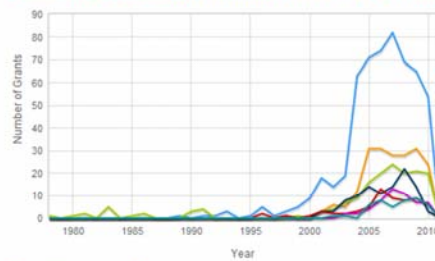
Search: X

Records 1 - 10 of 30 < First < Prev Next > Last >

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

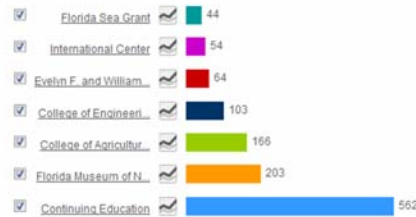
Save as CSV Clear

Comparing Grants of Organizations in University of Florida



Total Number of Grants

You have selected 7 of a maximum 10 organizations to compare. Clear



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

Search

Home People Organizations Research Events

University of Florida

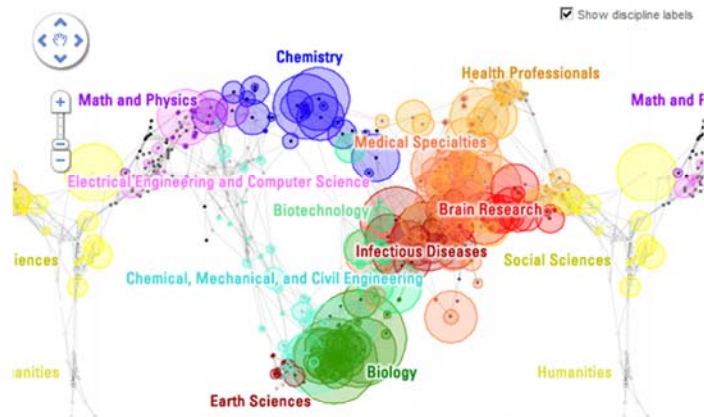
Explore 487 publications activity across 554 scientific sub-disciplines

13 Disciplines | 554 Sub-Disciplines

Search: X

1 - 13 of 554 < First < Prev Next > Last >

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



Show discipline labels

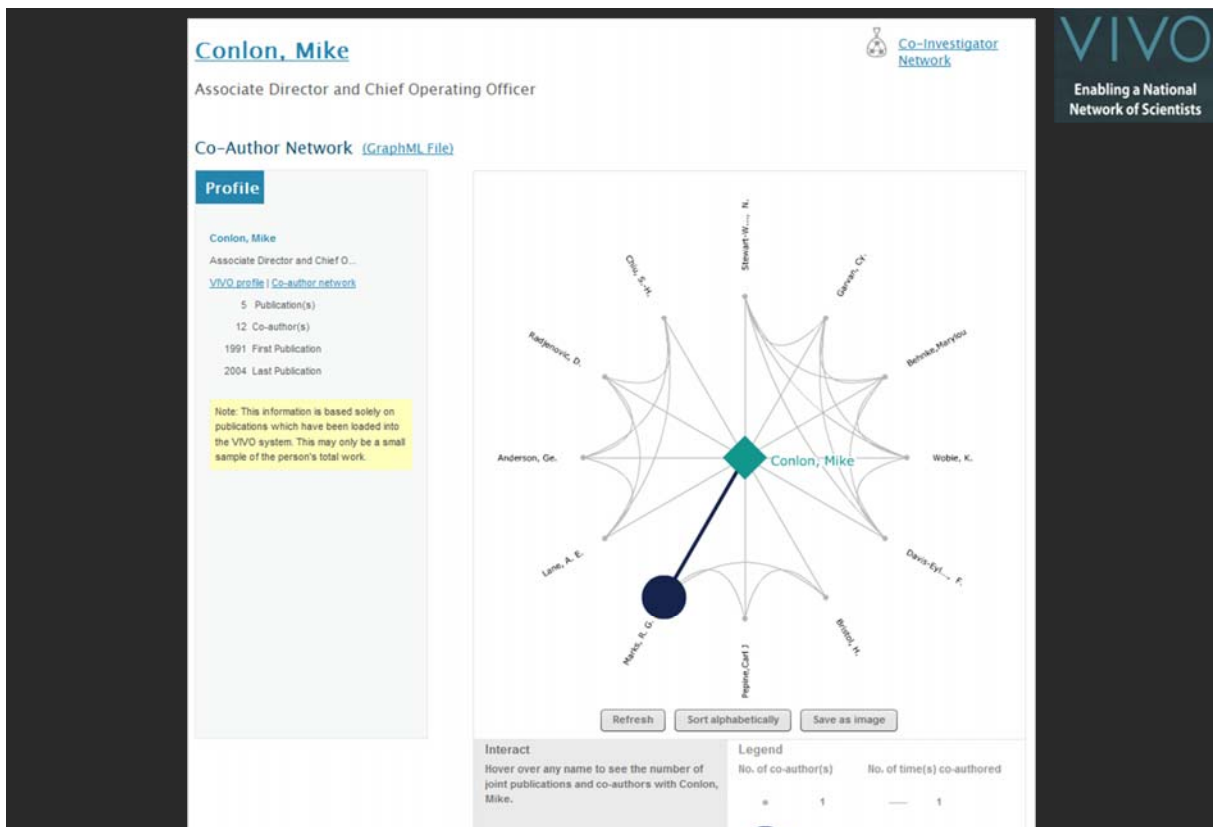
Google

Top 290 disciplines shown

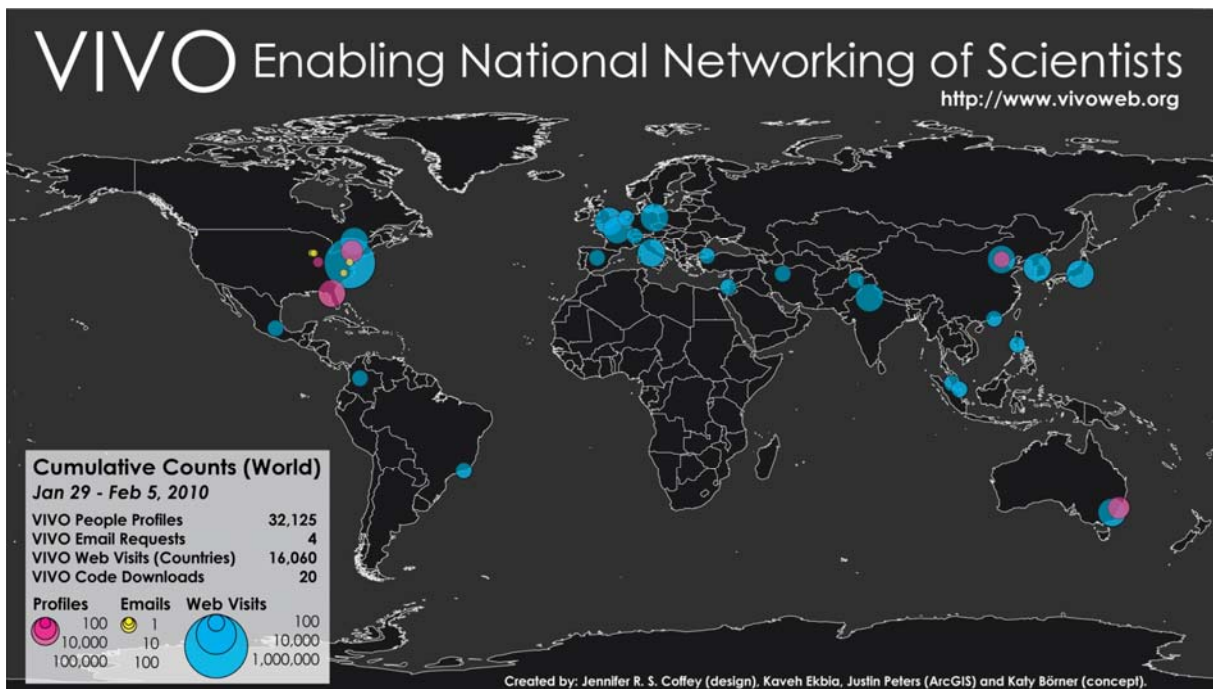
Terms of Use

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?

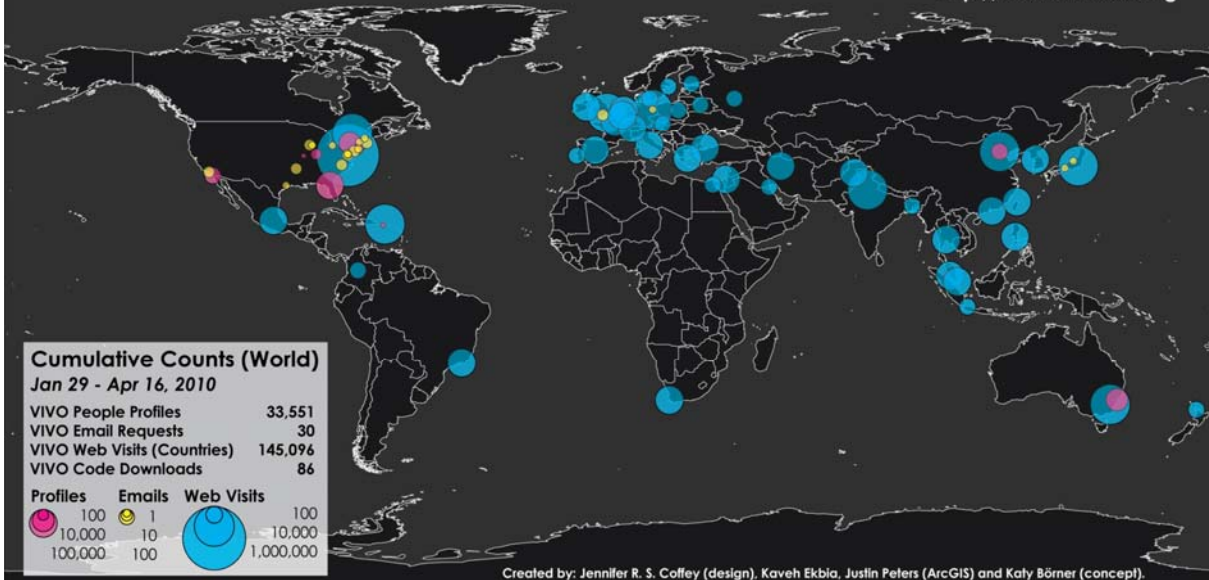


Science is global. World view of VIVO activity.
Web site visits are aggregated at the country level.

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>



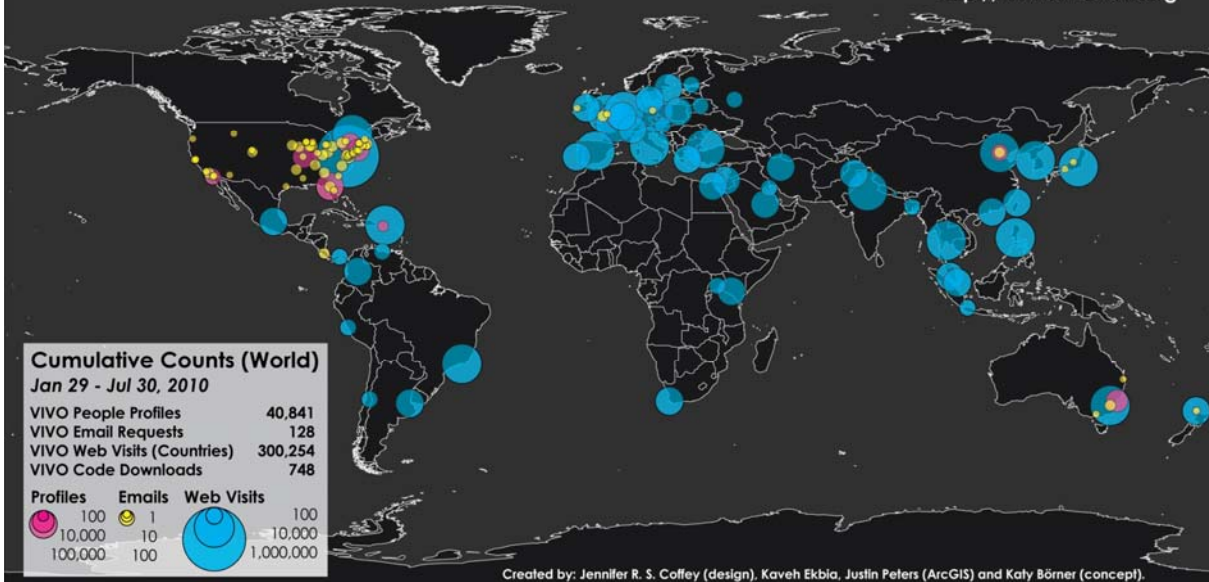
Shown are the

- 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.

25

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.

26

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](#)

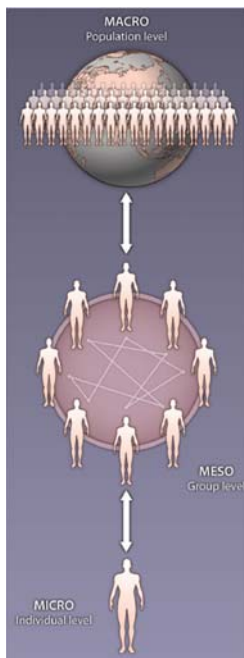
at Indiana University. The site was created in support of a NSF IGERT grant application. A major intent is to cross-fertilize between research done in the social and behavioral sciences, research in natural sciences such as biology or physics, but also research on Internet technologies.

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>

27



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.

28

DIGGING INTO DATA CHALLENGE

Wednesday, April 13, 2011 Home > Repositories >>> Search

Site Navigation

- Home
- Repositories
- DiD Logos
- Award Recipients - 2009
- Conference

List of Data Repositories

Last Updated: 31 March, 2011

Below is a list of digital libraries, data archives, and data repositories that are inviting Digging into Data researchers to use their collections. For each repository, you'll find a description of their contents, contact information, and other details.

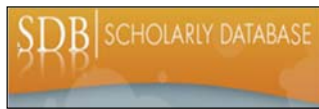
This list is being frequently updated, so check back often! If you are a digital repository and would like to be included on this list, please [get in touch with us](#).

The Archaeology Data Service (ADS)

ads.ahds.ac.uk

About: The ADS catalogue holds the digital archives of a huge number of archaeological interventions from the UK and beyond in around 400 collections, these range from the outputs of single excavations to large scale developer funded projects encompassing hundreds of individual archaeological interventions. As well as digital archives and fieldwork outputs the catalogue contains a number of scholarly resources intended specifically as reference sources for further research on topics such as lithics, ceramics and animal bone. The catalogue also contains digitised (or born digital) versions of various significant journals and series running to many thousands

<http://www.diggingintodata.org/Repositories/tabid/167/Default.aspx>



Scholarly Database at Indiana University

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

Supports federated search of 25 million publication, patent, grant records.
 Results can be downloaded as data dump and (evolving) co-author, paper-citation networks.

Cyberinfrastructure for Network Science Center, SLIS, Indiana University, Bloomington

IU User

IU Users must login using the Central Authentication Service (CAS), the standard IU authentication system. Please click the button below to proceed to the IU login page.

[Go to IU Login](#)

Non-IU User

Email:

Password:

[Login](#)

[Not Registered Yet?](#)

[Register as a New IU User](#)

In the News

Whitfield, John, 2006, *Group Theory*, Nature, 435, 9: 720-723.

Please Cite As

La Rosa, Sean, Ambler, Sumner, Burgess, John, Ye, Weimes and Böhm, Katy. (2007) The Scholarly Database and Its Utility for Scientometric Research. In Proceedings of the 13th International Conference on Scientometrics and Informetrics, Madrid, Spain, June 29-31, 2007, pp. 487-492.

<http://sdb.wiki.cns.iu.edu/~katy/paper/87%20sdb.pdf>

Acknowledgements

The Scholarly Database is funded by the School of Library and Information Science and the Cyberinfrastructure for Network Science center at Indiana University, the National Science Foundation under Grants No. IRI-0239243 and IRI-0519450, and a James S. McDonnell Foundation grant in area *Sharing Complex Systems*.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Cyberinfrastructure for Network Science Center, SLIS, Indiana University, Bloomington

Search Edit Profile Admin About Logout

Search

Creators:

Title:

Abstract:

Full Text:

First Year: 1898

Last Year: 2008

Headline (1898 - 2008)
 NIH (1961 - 2002)
 NSF (1985 - 2004)
 USPTO (1976 - 2007)

[Search](#)

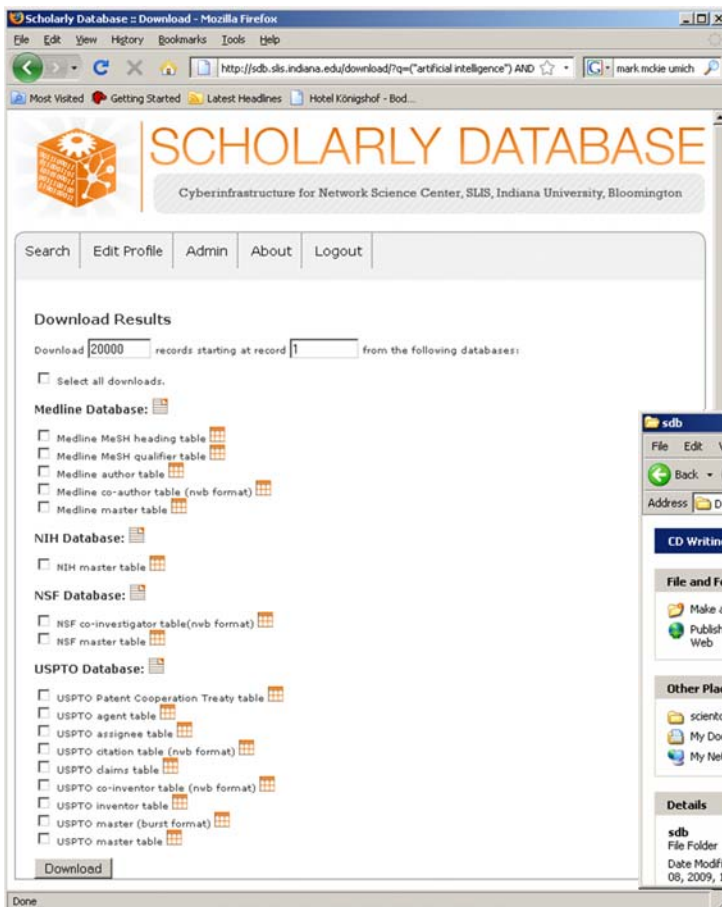
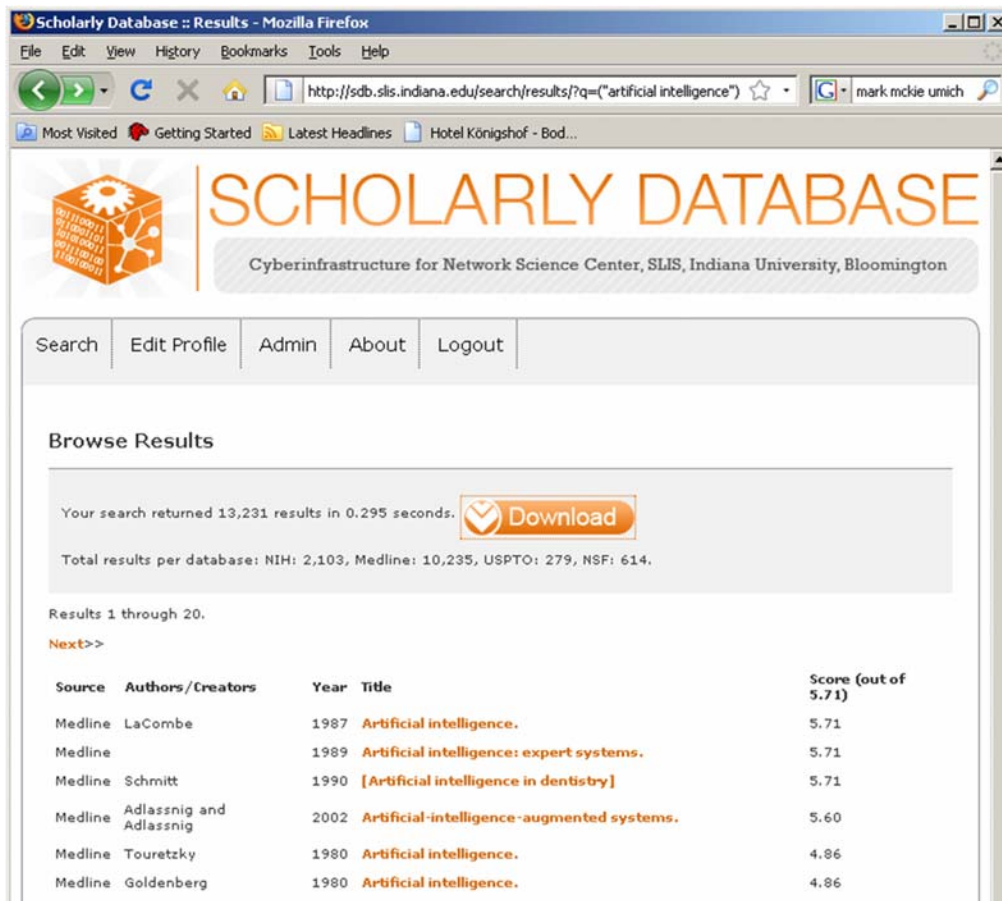
If multiple terms are entered in a field, they are automatically combined using "OR". So, "breast cancer" matches any record with "breast" or "cancer" in that field.

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 not the exact phrase.

The importance of a particular term in a query can be increased by putting a "*" and a number after the term. For instance, "breast cancer*10" would increase the importance of matching the term "cancer" by ten compared to matching the term "breast".

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

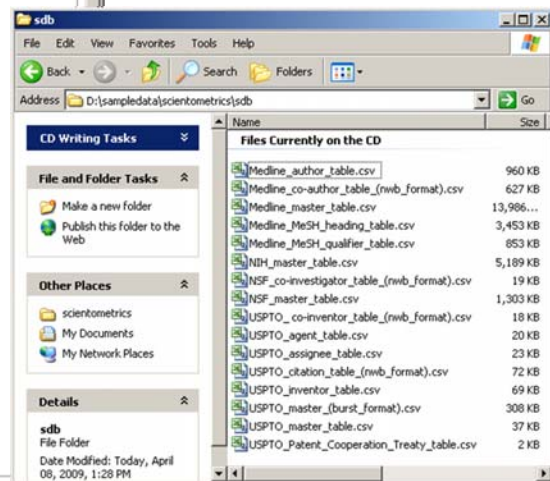


Since March 2009:

Users can download networks:

- Co-author
- Co-investigator
- Co-inventor
- Patent citation

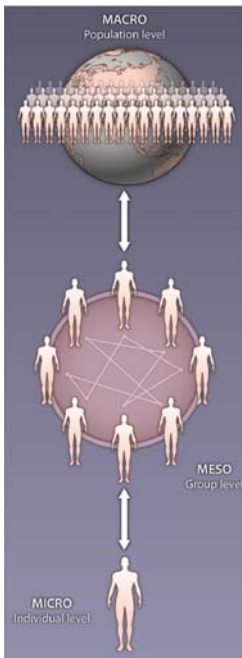
and tables for burst analysis in NWB.



National Researcher Networking Visualization 1.0

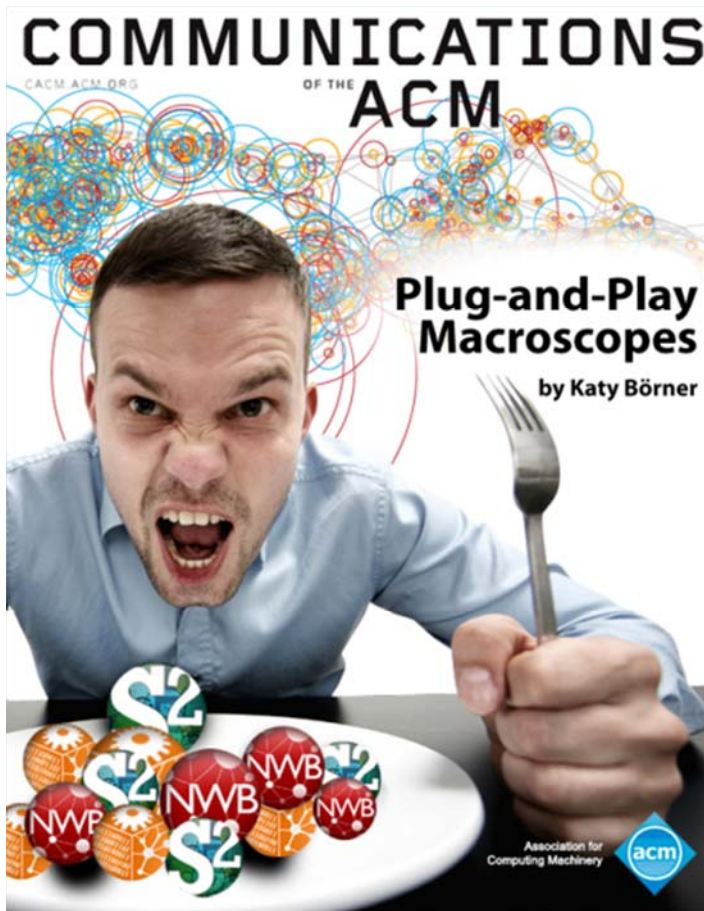


<http://nrn.cns.iu.edu>



Tools

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



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>



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.

Plug-and-play macroscopes **empower** me, my students, colleagues, and 100,000 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.**



Microscopes



Telescopes



Macroscopes

41



Macroscopes cont.

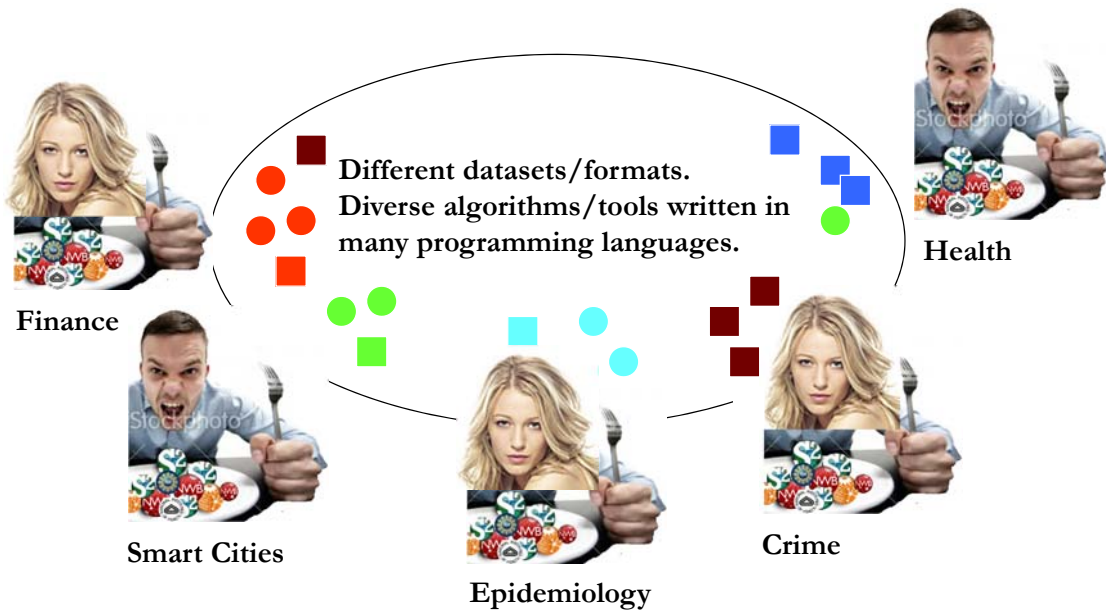
While microscopes and telescopes are physical instruments, **macroscopes resemble continuously changing bundles of software plug-ins.** Macroscopes make it easy to select and combine algorithm and tool plug-ins but also interface plug-ins, workflow support, logging, scheduling, and other plug-ins needed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

42



Macrosopes Serve the Changing Scientific Landscape



43



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

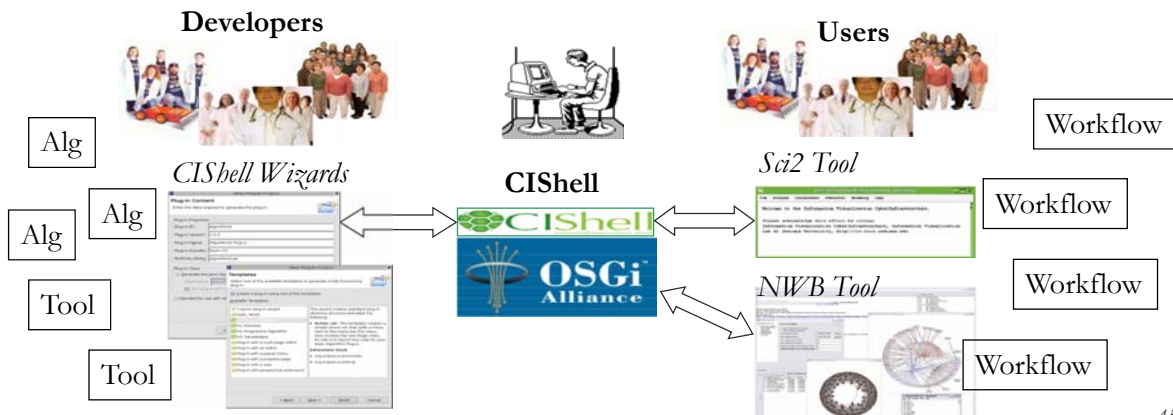
- Cancer Biomedical Informatics Grid (caBIG) (<http://cabig.nci.nih.gov>)
- Biomedical Informatics Research Network (BIRN) (<http://nbirn.net>)
- Informatics for Integrating Biology and the Bedside (i2b2) (<https://www.i2b2.org>)
- HUBzero (<http://hubzero.org>) platform for scientific collaboration 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**.

44

- 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.



45

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](#)

46

47



Network Workbench Tool

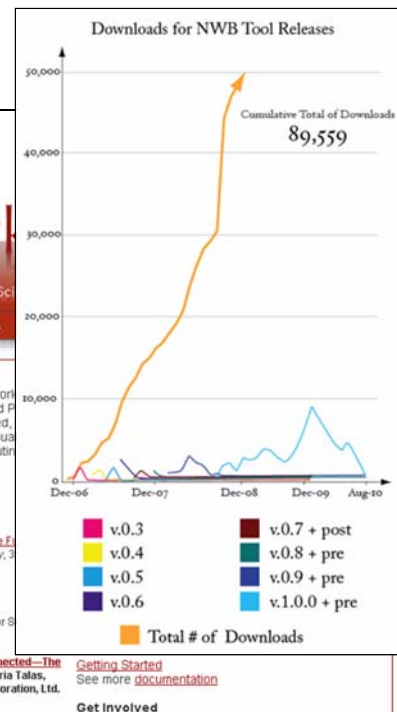
<http://nwb.cns.edu>

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 than 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.



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, Mibail C. (Eds.), *Progress in Convergence - Technologies for Human Wellbeing* (Vol. 1093, pp. 161-179), *Annals of the New York Academy of Sciences*, Boston, MA.

48

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](#)

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

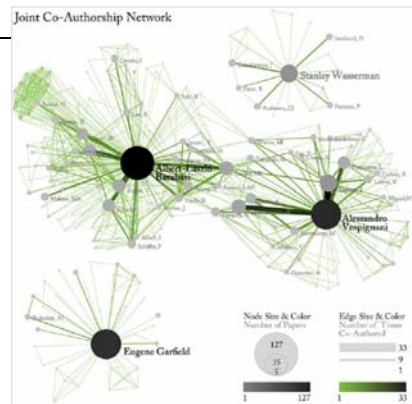
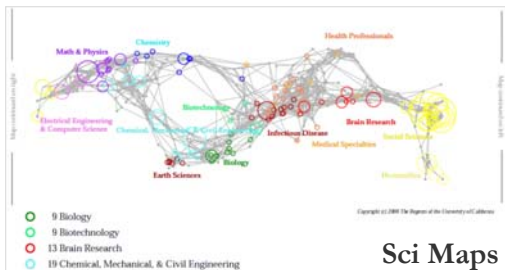
Geetha Senthil (2010). [Multidisciplinary Nature of Work With Reference to PIs and ICs Within a Portfolio](#). PA Group at NIH.

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



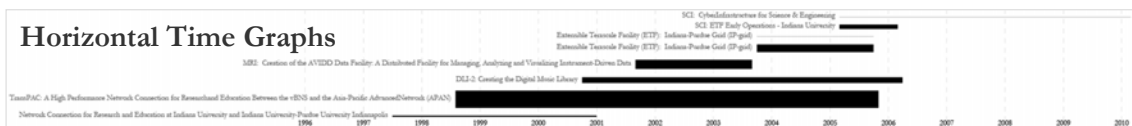
Sci² Tool – “Open Code for S&T Assessment”

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



GUESS Network Vis

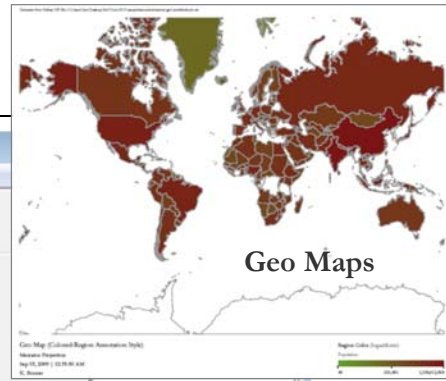
Horizontal Time Graphs



Börner, Katy, Huang, Weixia (Bonnie), Linnemeier, Micah, Dubon, 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 (Sci²). The development of this tool is supported in part by the National Science Foundation (NSF) under grant numbers IRI-0715303 and IRI-0715303, and the James S. McDonnell Cyberinfrastructure portal (<http://sci.slis.indiana.edu>).

The primary investigators are Katy Börner, INdiana University; 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>).

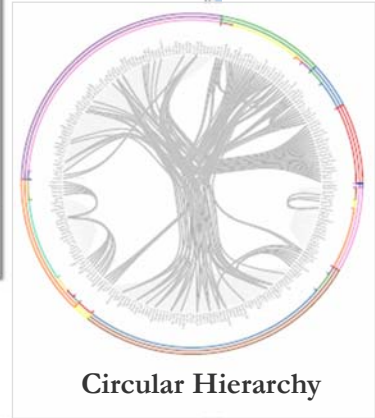
Please cite as follows:
Sci² Team. (2009). Science of Science Tool. Indiana University. <http://sci.slis.indiana.edu>.

Scheduler

Remove From List Remove completed

!	Algorithm Name	Date	Time	% Complete
<input checked="" type="checkbox"/>	Extract Co-Author Netw...	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 National Science Foundation (NSF) under grant numbers NIH-RM-07-004 award. The primary investigators are Chintan Tank, Joseph Biberstine, and Dr. Jim Sherman.

The EpiC tool was developed by Chintan Tank, Joseph Biberstine, and Chin Hu.

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.

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



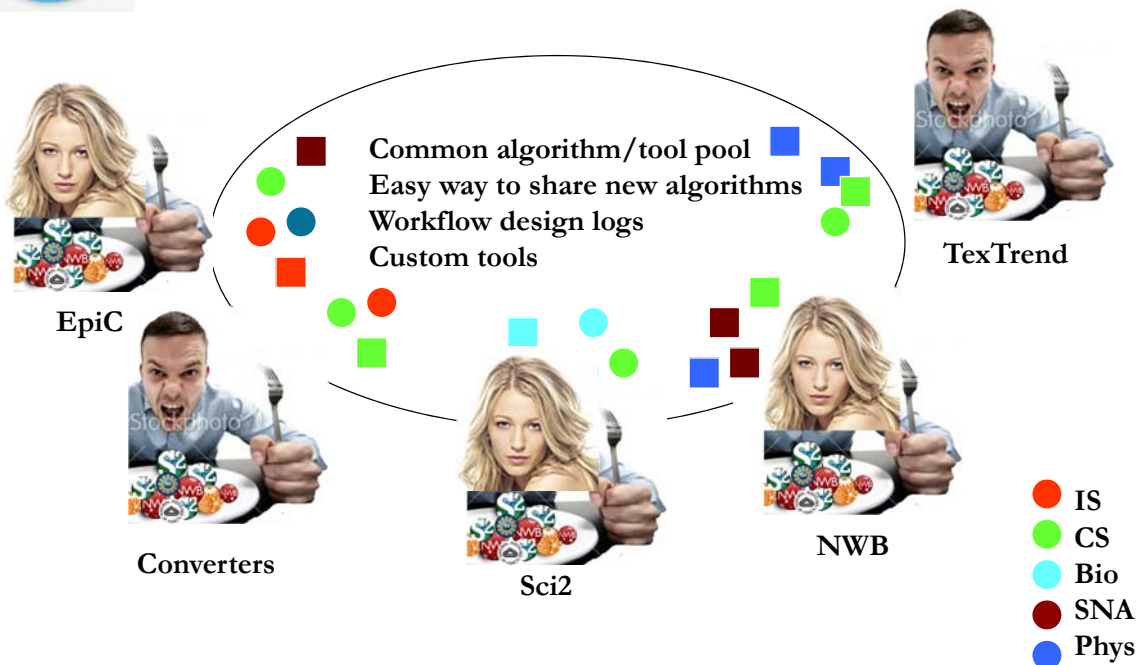
OSGi/CIShell Adoption

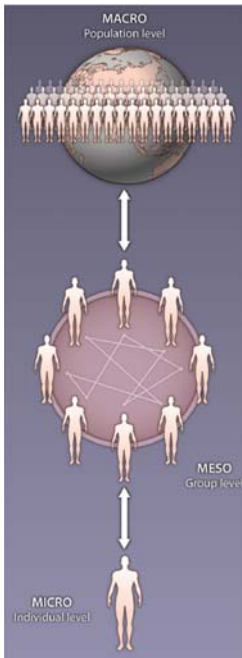
A number of other projects recently adopted OSGi and/or CIShell:

- 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





Mixed-methods, multi-level SciTS needs:

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.

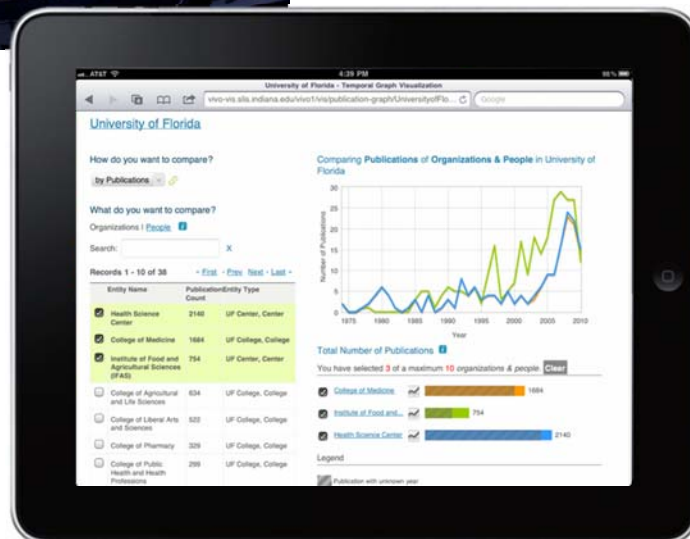
Data – find, interlink, unify, merge, reformat, share them, e.g., using web sites analogous to <http://www.diggingintodata.org/Repositories/tabid/167/Default.aspx>, SDB, or LOD.

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



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

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





Debut of 5th Iteration of the Mapping Science Exhibit at MEDIA X was in 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. <http://www.expedition-zukunft.de>

Science & Technology Forecasts @ Times Square in 2016



References

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

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

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

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

Scharnhorst, Andrea, Börner, Katy, van den Besselaar, Peter (2011) **Models of Science Dynamics**. Springer Verlag.

