

DIY Data Mining, Information Visualization, and Science Maps

Katy Börner

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

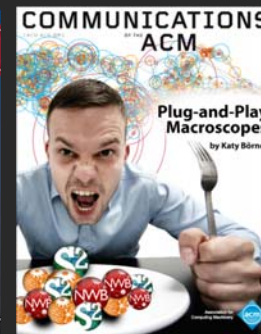


With special thanks to the members at the Cyberinfrastructure for Network Science Center; the Sci2, NWB, and EpiC team; and the VIVO Collaboration

Werklunch at DANS, KNAW
The Hague, The Netherlands

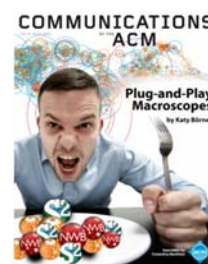
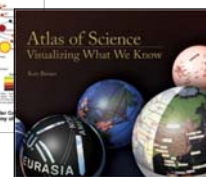


January 23, 2012



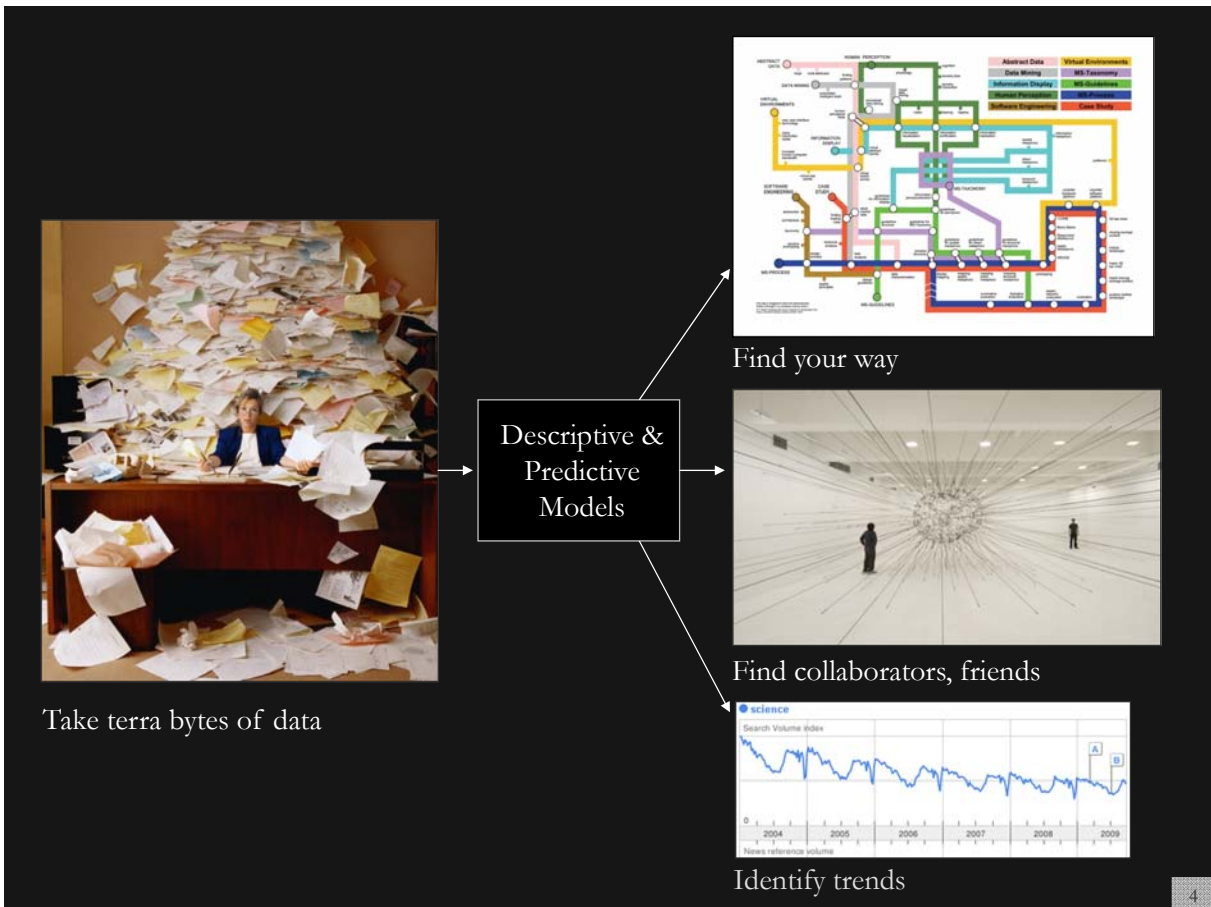
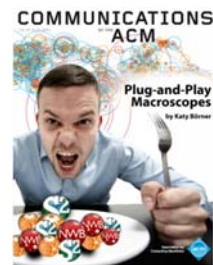
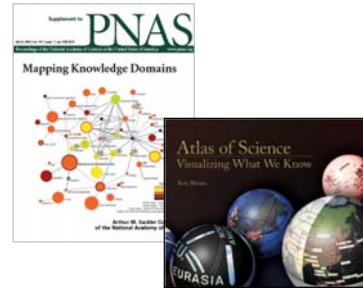
Overview

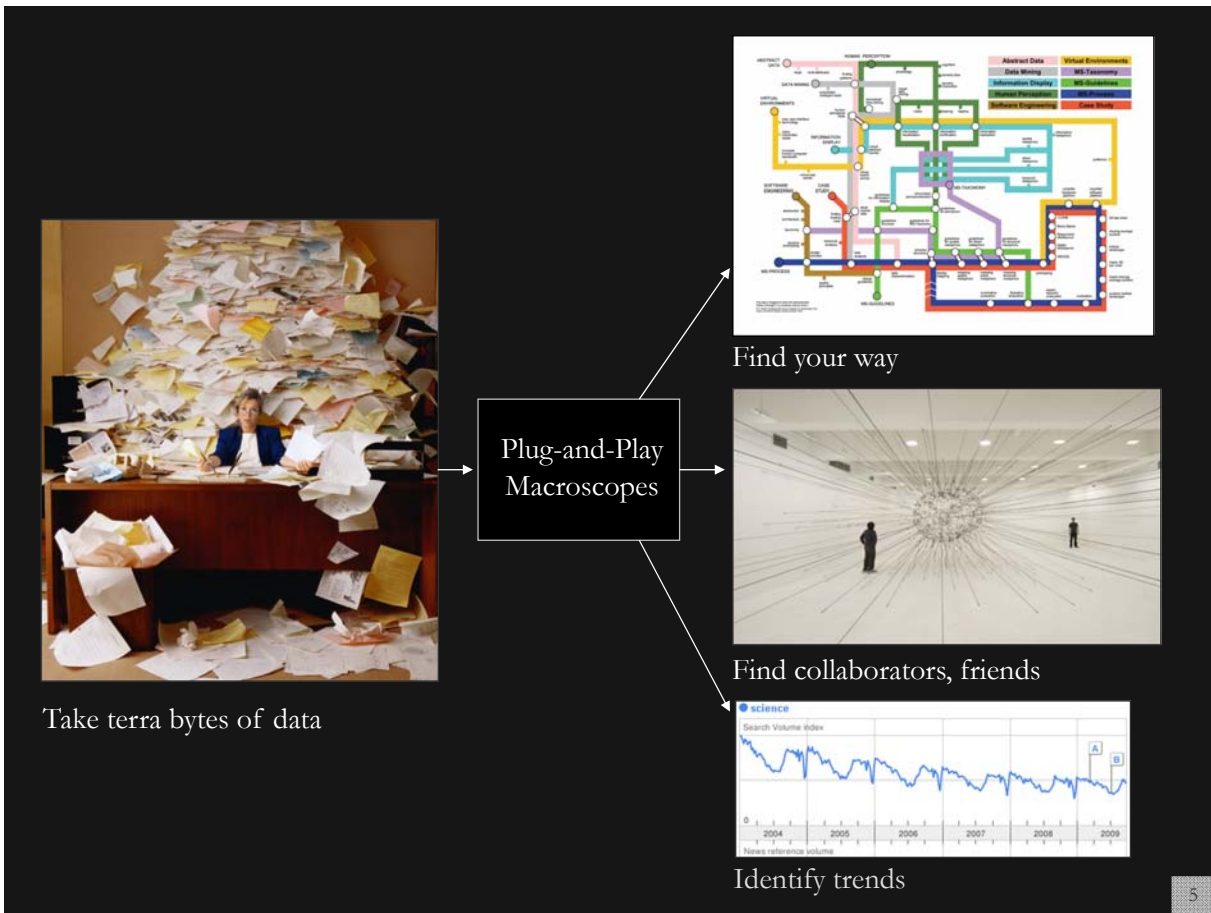
1. **Data mining and visualization research** that aims to increase our scientific understanding of the structure and dynamics of science and technology.
2. **Novel approaches and services** that improve information access, researcher networking, and research management.
3. **Data services and plug-and-play macroscope tools** that commoditize data mining and visualization.



Overview

1. **Data mining and visualization research** that aims to increase our scientific understanding of the structure and dynamics of science and technology.
2. **Novel approaches and services** that improve information access, researcher networking, and research management.
3. **Data services and plug-and-play macroscope tools** that commoditize data mining and visualization.





Type of Analysis vs. Level of Analysis

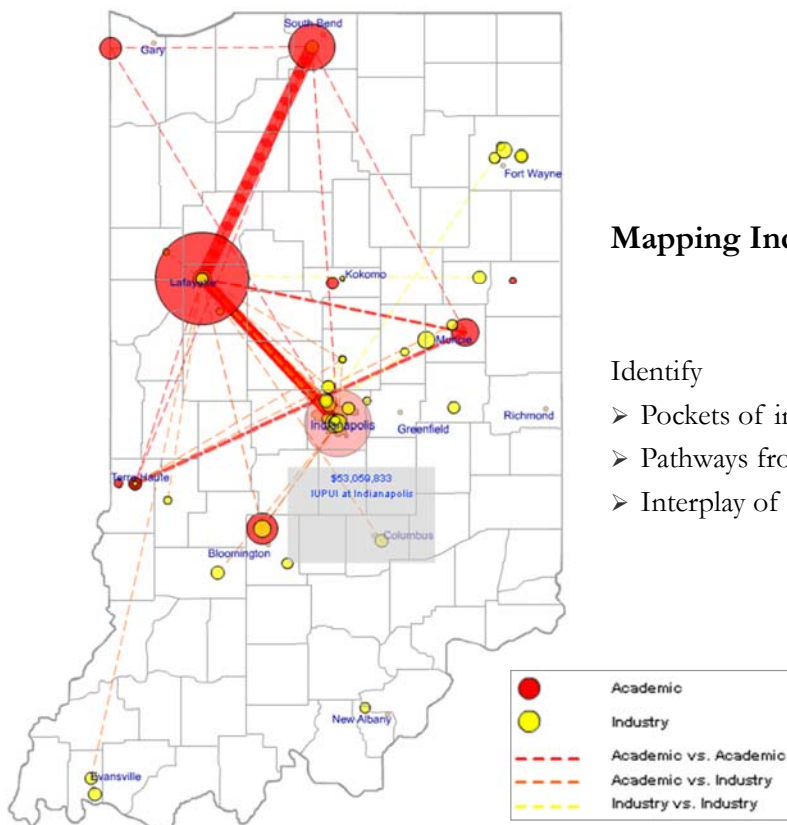
	<i>Micro/Individual</i> (1-100 records)	<i>Meso/Local</i> (101-10,000 records)	<i>Macro/Global</i> (10,000 < records)
Statistical Analysis/Profiling	Individual person and their expertise profiles	Larger labs, centers, universities, research domains, or states	All of NSF, all of USA, all of science.
Temporal Analysis (When)	Funding portfolio of one individual	Mapping topic bursts in 20-years of PNAS	113 Years of Physics Research
Geospatial Analysis (Where)	Career trajectory of one individual	Mapping a states intellectual landscape	PNAS publications
Topical Analysis (What)	Base knowledge from which one grant draws.	Knowledge flows in Chemistry research	VxOrd/Topic maps of NIH funding
Network Analysis (With Whom?)	NSF Co-PI network of one individual	Co-author network	NIH's core competency

Type of Analysis vs. Level of Analysis

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



7



Mapping Indiana's Intellectual Space

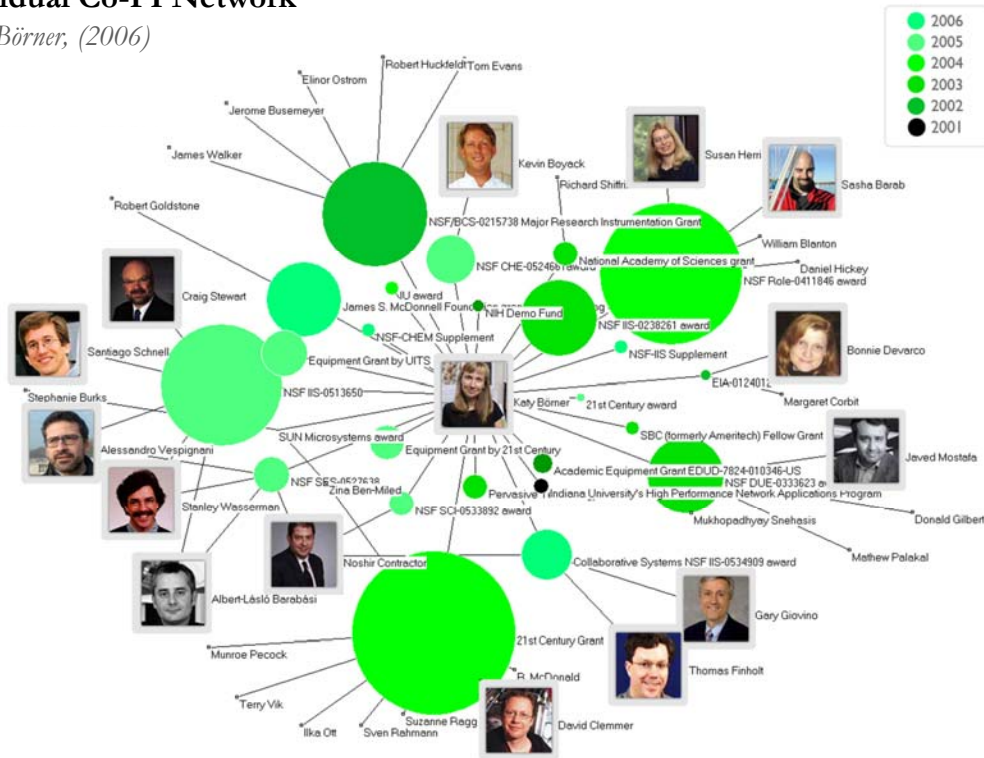
Identify

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

8

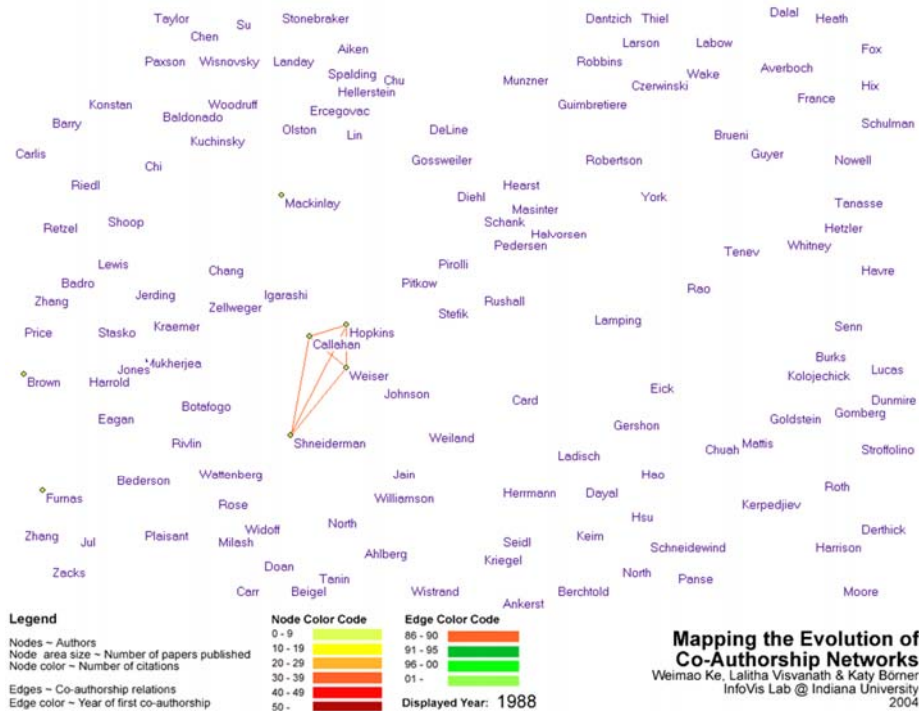
Individual Co-PI Network

Ke & Börner, (2006)



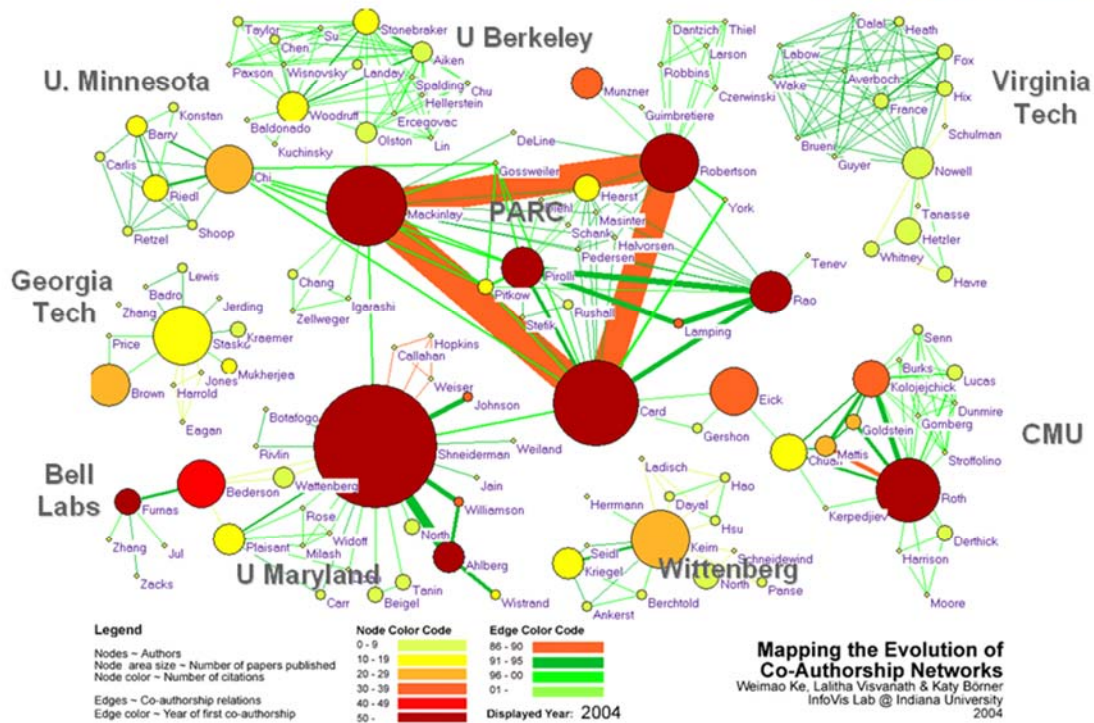
Mapping the Evolution of Co-Authorship Networks

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



Mapping the Evolution of Co-Authorship Networks

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



11

Studying the Emerging Global Brain: Analyzing and Visualizing the Impact of Co-Authorship Teams

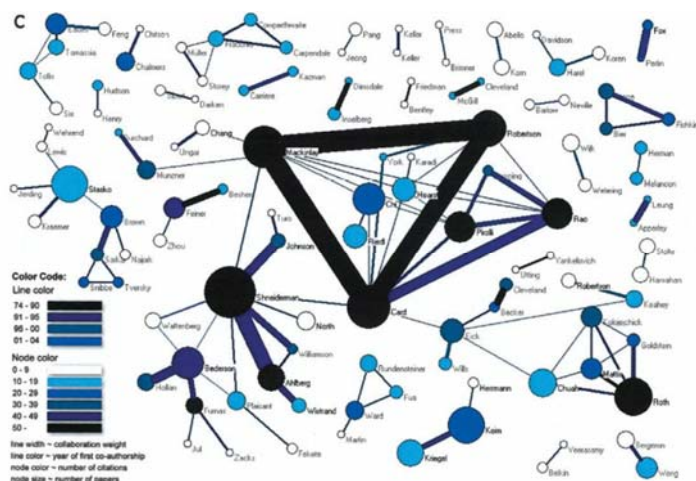
Börner, Dall'Asta, Ke & Vespignani (2005)
Complexity, 10(4):58-67.

Research question:

- Is science driven by prolific single experts or by high-impact co-authorship teams?

Contributions:

- New approach to allocate citational credit.
- Novel weighted graph representation.
- Visualization of the growth of weighted co-author network.
- Centrality measures to identify author impact.
- Global statistical analysis of paper production and citations in correlation with co-authorship team size over time.
- Local, author-centered entropy measure.



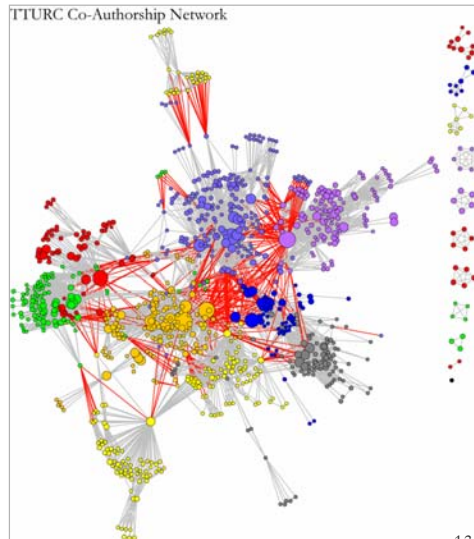
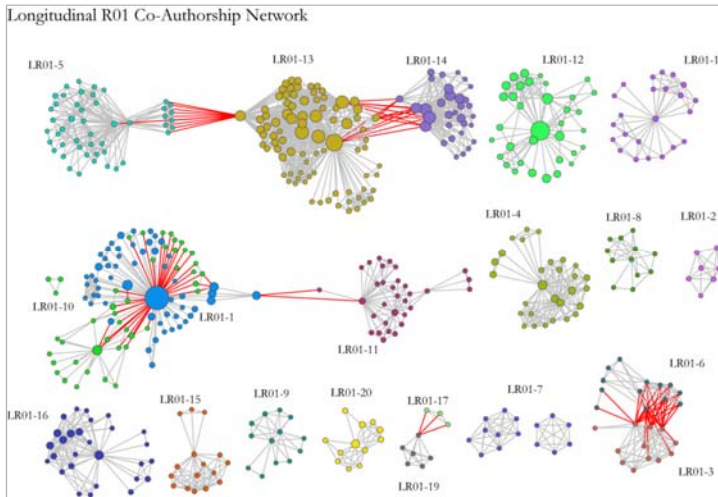
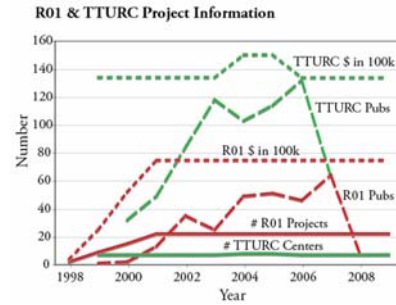
12

Mapping Transdisciplinary Tobacco Use Research Centers Publications

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

Zoss & Börner, *forthcoming*.

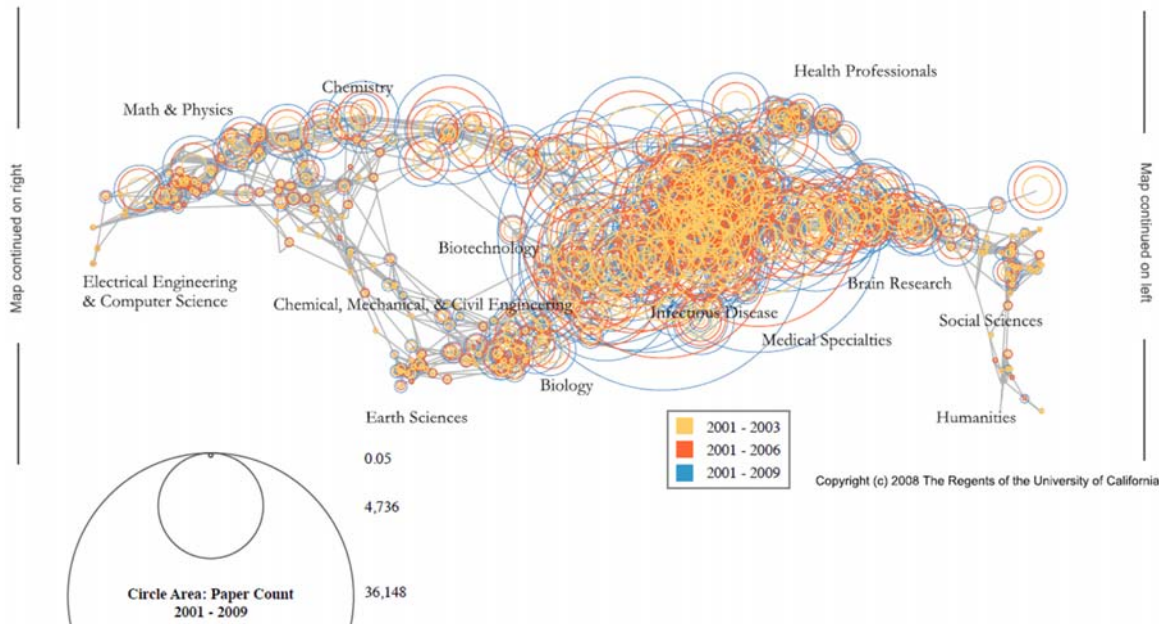
Supported by NIH/NCI Contract HHSN261200800812



13

MEDLINE Publication Output by The National Institutes of Health (NIH) Using Nine Years of ExPORTER Data

Katy Börner, Nianli Ma, Joseph R. Biberstine, Cyberinfrastructure for Network Science Center, SLIS, Indiana University, Robin M. Wagner, Rediet Berhane, Hong Jiang, Susan E. Ivey, Katrina Pearson and Carl McCabe, Reporting Branch, Division of Information Services, Office of Research Information Systems, Office of Extramural Research, Office of the Director, National Institutes of Health (NIH), Bethesda, MD.



14

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/

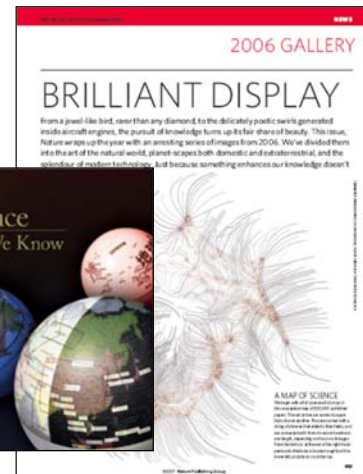
Börner, Katy, Sanyal, Soma and Vespignani, Alessandro (2007). **Network Science**. In Blaise Cronin (Ed.), *ARIST*, Information Today, Inc., Volume 41, Chapter 12, pp. 537-607.

<http://ivl.slis.indiana.edu/km/pub/2007-borner-arist.pdf>

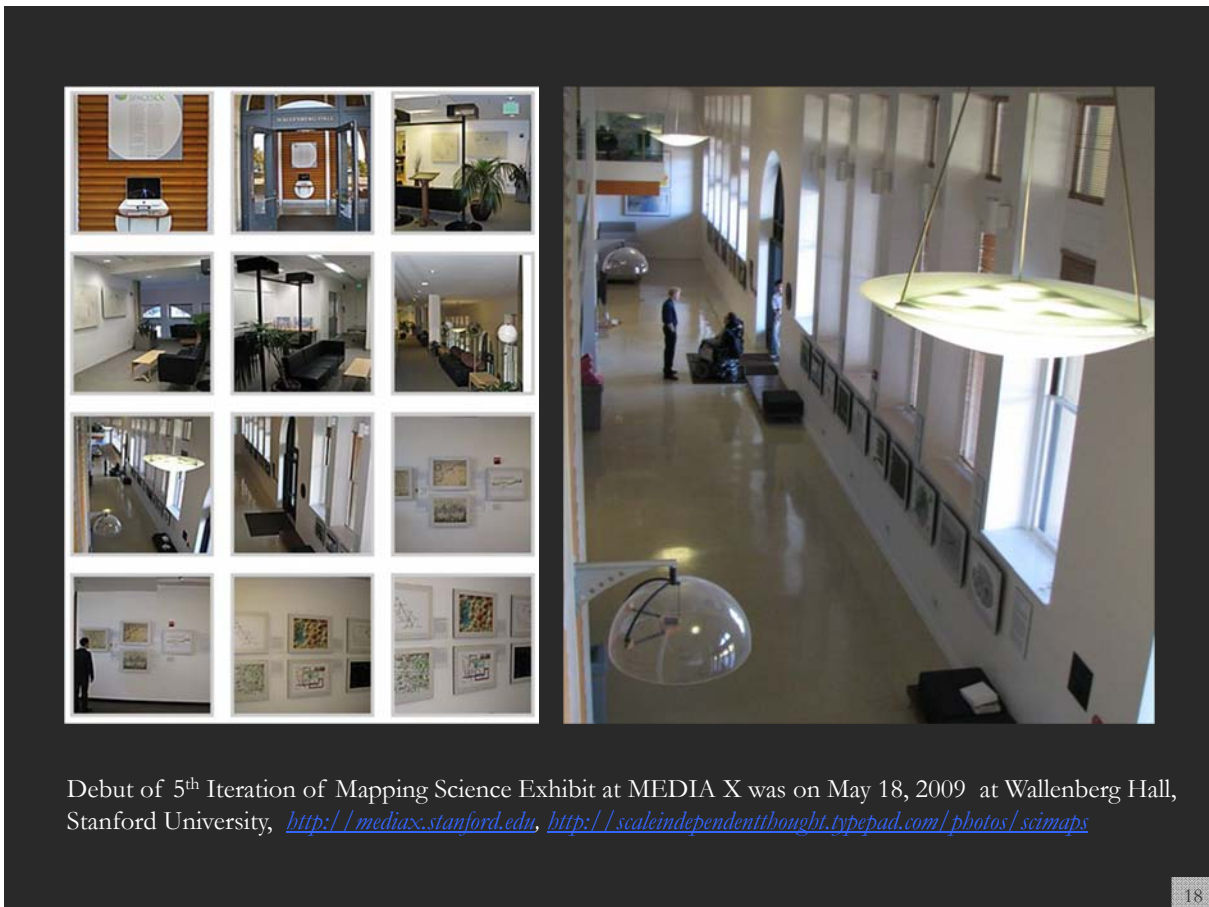
Börner, Katy (2010) **Atlas of Science**. MIT Press.

<http://scimaps.org/atlas>

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



17



Debut of 5th Iteration of 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>

18



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

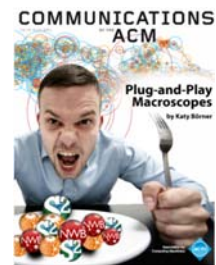
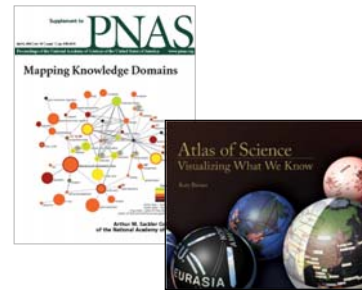
<http://www.expedition-zukunft.de>



Information on how to host the exhibit or acquire a subset of the maps is at <http://scimaps.org>

Overview

1. **Data mining and visualization research** that aims to increase our scientific understanding of the structure and dynamics of science and technology.
2. **Novel approaches and services** that improve information access, researcher networking, and research management.
3. **Data services and plug-and-play macroscope tools** that commoditize data mining and visualization.



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.

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/Mentors—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 need 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.



Mapping Sustainability Research

Geographic Map Science Map



Funding
 NIH
 NSF
 USDA

Publications
 DOE
 ISI
 Medline

Patents
 USPTO

Citations Count

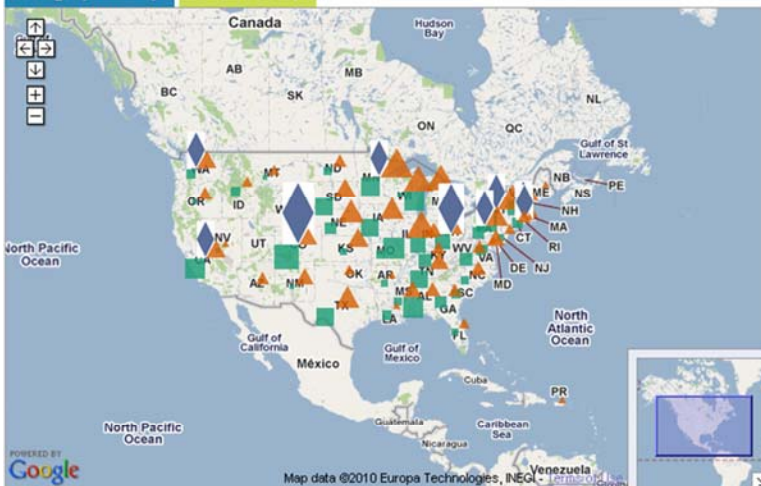
Amount Count

From year 1901 to year 2009

Search by keyword

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

Geographic Map Science Map



Funding Patents
 NIH USPTO
 NSF ISI
 USDA Medline

Citations Count

Amount Count

From year 1901 to year 2009

Search by keyword

Maps Detail Data About

Datasets

The dataset covers 13,528 records on "biomass" and "biofuel" research and technology from seven different publication, patent, funding datasets for the years 1901 to 2010.

Funding

National Institutes of Health (NIH) awards retrieved from the Scholarly Database (<http://sdb.slis.indiana.edu>) at Indiana University on 11/20/2010. Search query used was biomass OR biofuel OR "bio mass" OR "bio fuel" in the 'All Text' field.

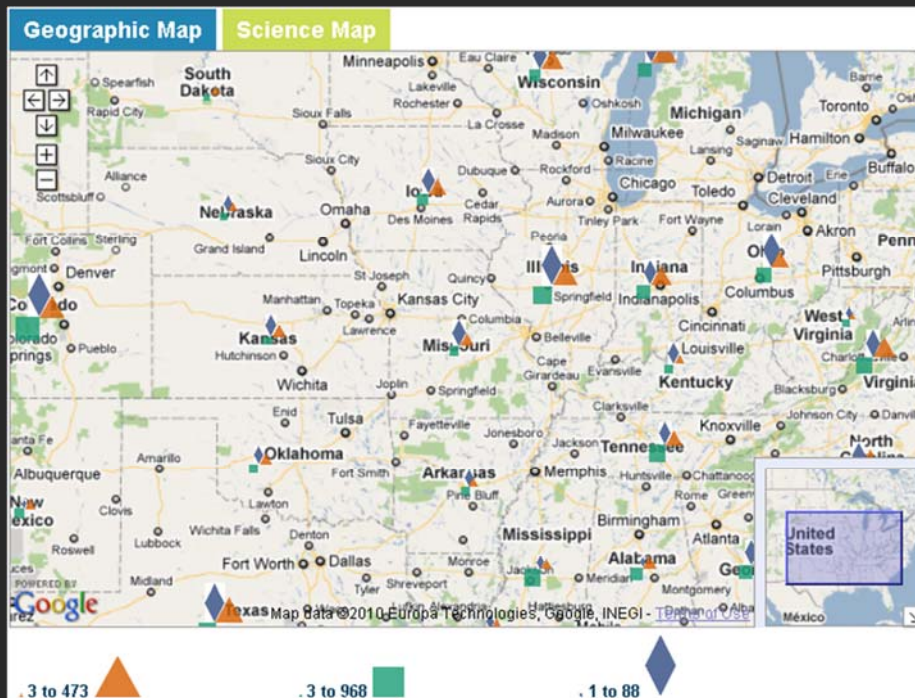
National Science Foundation (NSF) awards retrieved from the Scholarly Database (<http://sdb.slis.indiana.edu>) at Indiana University on 11/20/2010. Search query used was biomass OR biofuel OR "bio mass" OR "bio fuel" in the 'All Text' field.

US Department of Agriculture (USDA) awards made available by a staff member of the Office of Scientific and Technical Information from the US Department of Energy (DOE).

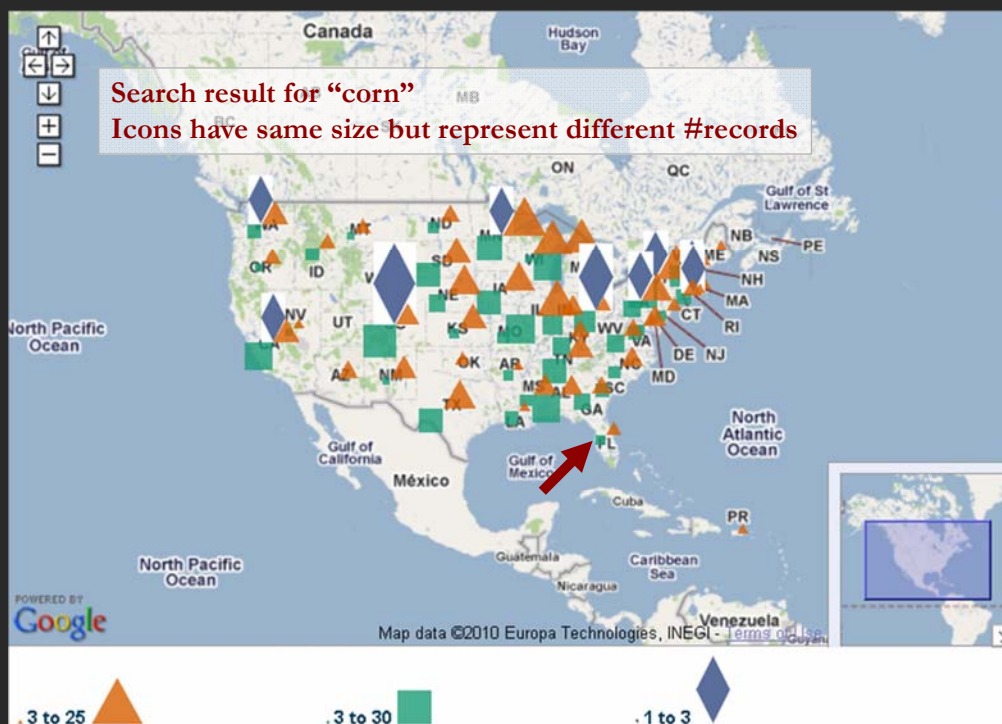
Publications

MEDLINE papers by the National Library of Medicine retrieved from the Scholarly Database (<http://sdb.slis.indiana.edu>) at Indiana University on 11/20/2010. Search query used was biomass OR biofuel OR "bio mass" OR "bio fuel" in the 'All Text' field.

The geographic map at state level.

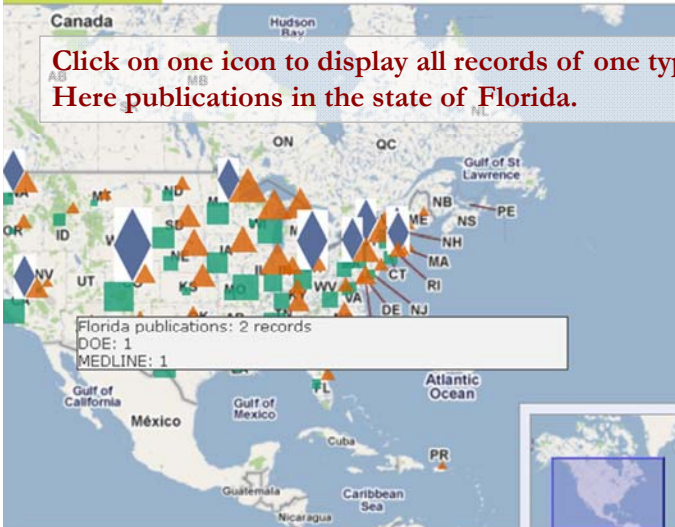


The geographic map at city level.



Science Map

Click on one icon to display all records of one type.
Here publications in the state of Florida.



Maps

Detail

Data

About

> Florida

MEDLINE

2002

- [Recovery Of Dairy Manure Nutrients By Benthic Freshwater Algae](#)

DOE

1985

- [Enzymatic Hydrolysis And Fermentation Of Corn For Fuel Alcohol](#)

29

Information Bridge: DOE Scientific and Technical Information - - Document #5789929 - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.osti.gov/bridge/product.biblio.jsp?osti_id=5789929

MapSustain x Information Bridge: DOE Scientifi... x

DOE Scientific and Technical Information

DOE • OSTI

INFORMATION BRIDGE

Home • Basic Search • Fielded Search • Alerts • Help

FAQ • Widget • Site Map

SHARE

Bibliographic Citation

[See/Add Document Discussions](#) [Return to Search Results](#) [Return to Original Search Page](#) [Download as EndNote](#)

Full Text Availability information may be found in the Availability, Publisher, Research Organization, Resource Relation and/or Author (affiliation information) fields and/or via the "Full-text Availability" link. For a journal article, please see the Resource Relation field.

Title Enzymatic hydrolysis and fermentation of corn for fuel alcohol
[Word Cloud](#) | [More Like This](#)

Creator/Author [Mullins, J.T.](#)

Publication Date 1985 Jan 01

OSTI Identifier OSTI ID: 5789929

Other Number(s) Journal ID: CODEN: BIBA

Resource Type Journal Article

Resource Relation Journal Name: Biotechnol. Bioeng.; (United States); Journal Volume: 27:3

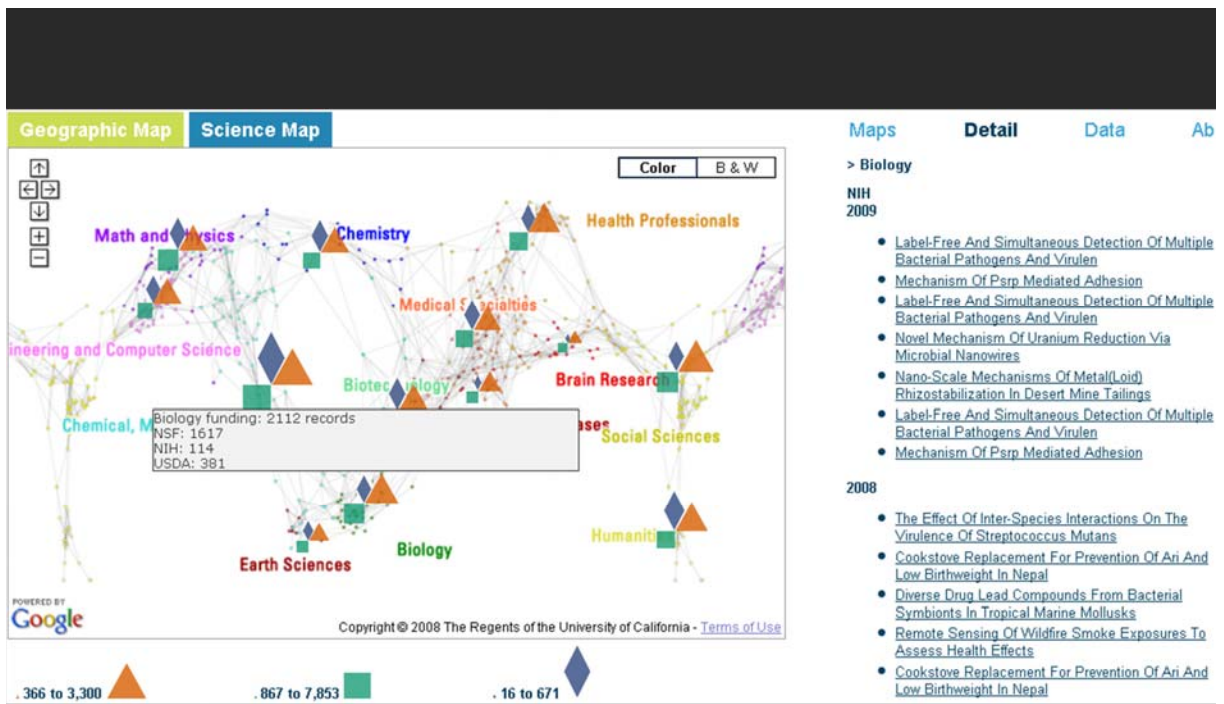
Research Org Univ. of Florida, Gainesville

Subject 09 BIOMASS FUELS; 32 ENERGY CONSERVATION, CONSUMPTION, AND UTILIZATION; ETHANOL FUELS; BIOSYNTHESIS; MAIZE; ENZYMATIC HYDROLYSIS; FERMENTATION; PRODUCTIVITY; COST; ENERGY EFFICIENCY; EXPERIMENTAL DATA; WASTE PRODUCT UTILIZATION; ALCOHOL FUELS; BIOCONVERSION; CEREALES; CHEMICAL REACTIONS; DATA; DECOMPOSITION; EFFICIENCY; FUEL; GLASS; HYDROLYSIS; INFORMATION; LYSIS; NUMERICAL DATA; PLANTS;

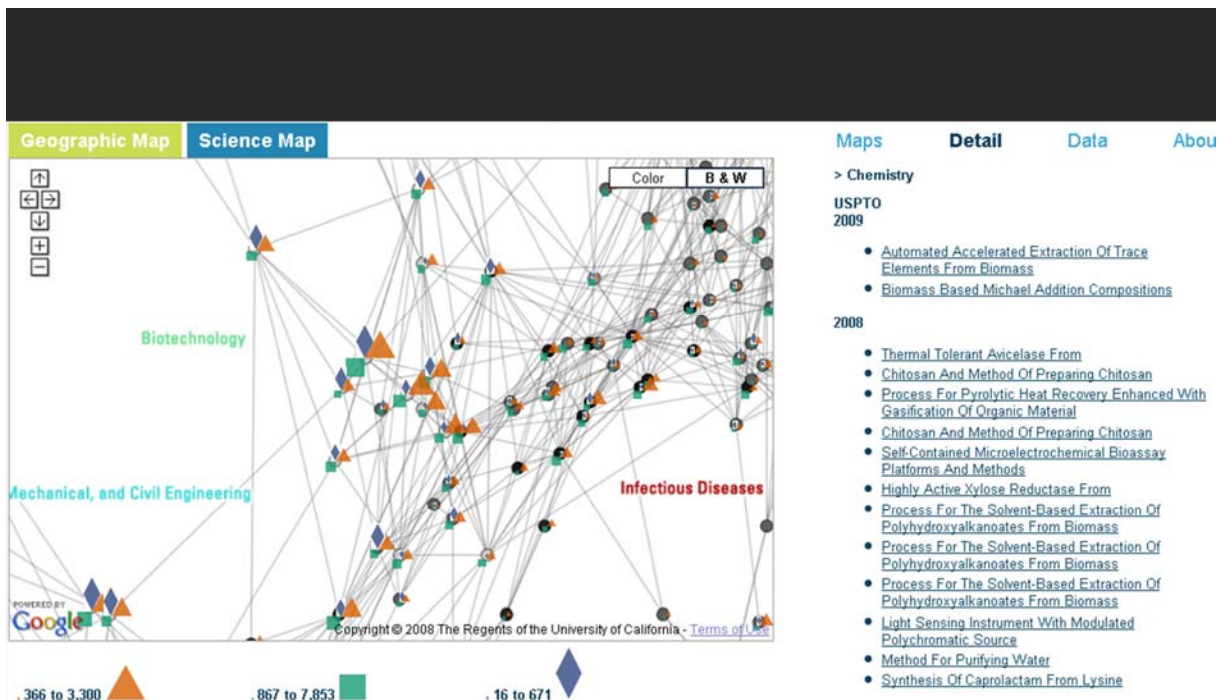
Done

Detailed information on demand
via original source site for exploration and study.

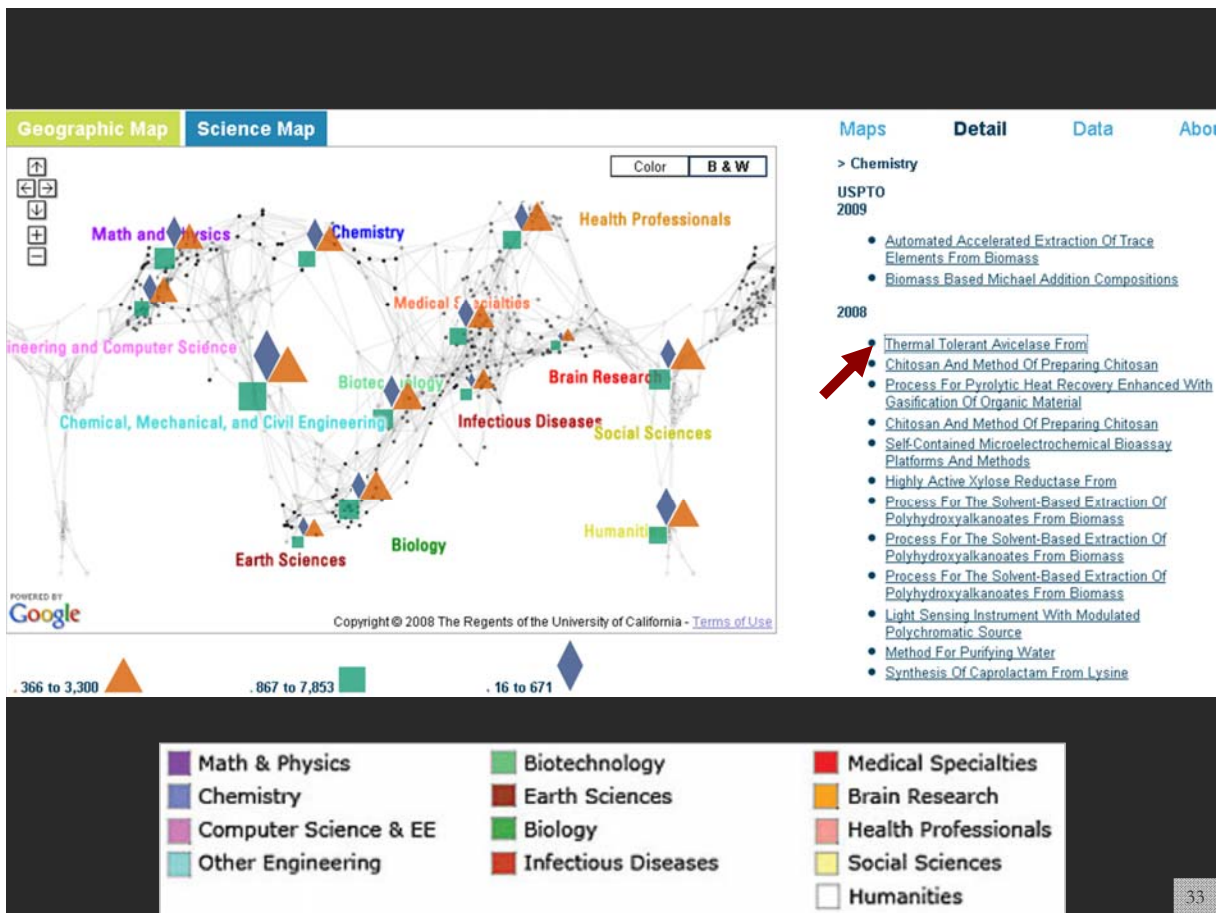
30



The science map at 13 top-level scientific disciplines level.



The science map at 554 sub-disciplines level.



- Maps Detail Data About
- > Chemistry
- USPTO
- 2009
- Automated Accelerated Extraction Of Trace Elements From Biomass
 - Biomass Based Michael Addition Compositions
- 2008
- Thermal Tolerant Avicelase From Chitosan And Method Of Preparing Chitosan
 - Process For Pyrolytic Heat Recovery Enhanced With Gasification Of Organic Material
 - Chitosan And Method Of Preparing Chitosan
 - Self-Contained Microelectrochemical Bioassay Platforms And Methods
 - Highly Active Xylose Reductase From Process For The Solvent-Based Extraction Of Polyhydroxyalkanoates From Biomass
 - Process For The Solvent-Based Extraction Of Polyhydroxyalkanoates From Biomass
 - Process For The Solvent-Based Extraction Of Polyhydroxyalkanoates From Biomass
 - Light Sensing Instrument With Modulated Polychromatic Source
 - Method For Purifying Water
 - Synthesis Of Caprolactam From Lysine

United States Patent: 7364890 - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://patft.uspto.gov/netacgi/nph-Parser?Sect1=PTO1&Sect2=HITOFF&d=PALL&p=1&u=/netacgi/PTO/srch

MapSustain United States Patent: 7364890 United States Patent: 7364890 Information Bridge: DOE Scientific a...

USPTO PATENT FULL-TEXT AND IMAGE DATABASE

Home Quick Advanced Pat Num Help

Bottom

View Cart Add to Cart

Images

(1 of 1)

United States Patent 7,364,890
Ding, et al. April 29, 2008

Thermal tolerant avicelase from *Acidothermus cellulolyticus*

Abstract

The invention provides a thermal tolerant (thermostable) cellulase, AvIII, that is a member of the glycoside hydrolase (GH) family. AvIII was isolated and characterized from *Acidothermus cellulolyticus* and, like many cellulases, the disclosed polypeptide and/or its derivatives may be useful for the conversion of biomass into biofuels and chemicals.

Inventors: Ding; Shi-You (Golden, CO), Adney; William S. (Golden, CO), Vinzant; Todd B. (Golden, CO), Himmel; Michael E. (Littleton, CO)
Assignee: Midwest Research Institute (Kansas City, MO)
Appl. No.: 09/017,276

Done

34

NIH Topic Maps

NIH TOPIC MAPS

A Topic Database of NIH-Funded Grants

NIH Map Viewer [Show Topic Browser](#) [Export Data](#) [Methods](#) [Feedback](#)

2009 [?](#) [add](#) [delete](#) [AND](#) [Topic Words](#) 20 0/0 [?](#) [Search](#) [Clear Search](#)

Institutes (9)

NIH Inst	# Grants	Count	+
NCI		116	
NCRR		10	
NIEHS		5	
NCMHD		1	
NIA		-	

Topics

%	Title Words	+
25.9	breast, cancer, cancer_risk, women, cancer_sui	
3.86	risk, risk_factors, cancer, prospective, women,	
3.76	genome_wide_association, loci, genome_wide,	
3.70	genetic, genetics, genes, gene_environment, i	

Grants (137)

NIH Inst	Grant	+
NCRR	2P20RR011792-10S2-6914 OBESITY, INSULIN RESISTANCE, IGF'S, AND BREAST CANCER RISK IN AFRICAN AMERICANS PI: CUI, YONG	
NCI	3R01CA120562-03S1 Commonly Used Medications and Breast Cancer Recurrence PI: BOUDREAU, DENISE M	
NCI	5R01CA120562-03 Commonly Used Medications and Breast Cancer Recurrence PI: BOUDREAU, DENISE M	
NCI	5R01CA093772-06 Long-term Survivorship in Older Women with Early Stage Breast	

Powered by ChalkLabs

<https://app.nihmaps.org>

36

NIH TOPIC MAPS

A Topic Database of NIH-Funded Grants

NIH Topic Browser

Show Map Viewer ?

Export Data

Methods

Feedback

Topics by NIH Institute

Topics by Category

2009

?

add

delete

AND

Exact Text

cancer

Search

Clear Search

2009 Grants (137)

Institutes (9)

Col	NIH Inst	Project/Subproj#	Title	Investigator(s)	# 1 Topic	# 1 Topic Work	NIH Inst	# Grants	Count
	NCCR	3P20RR011792-10S2 6914	OBESITY, INSULIN RESISTANCE, IGF'S, AND BREAST CANCER RISK IN AFRICAN AMERICANS	CUI, YONG	686 (50%)	cancer brea...	NCI	116	
	NCI	3R01CA120562-03S1	Commonly Used Medications and Breast Cancer Recurrence	BOUDREAU, DENISE M	686 (42%)	cancer brea...	NCRB	10	
	NCI	5R01CA120562-03	Commonly Used Medications and Breast Cancer Recurrence	BOUDREAU, DENISE M	686 (42%)	cancer brea...	NIHHS	5	
	NCI	5R01CA093772-06	Long-term Survivorship in Older Women with Early Stage Breast Cancer	SILLIMAN, REBECCA A	686 (42%)	cancer brea...	NCMHD	1	
	NCI	5R01CA064277-11	Shanghai Breast Cancer Study	ZHENG, WEI	686 (41%)	cancer brea...	NIA	1	
							NCCAM	1	
							NICHD	1	
							NIHR	1	
							NHGRI	1	

Topics

Similar Grants

Show Top 100 on Map

%	Topic	Topic Words	Title Words	Similar	C	NIH Inst	Grant
25.91	686	cancer breast	cancers cancer_risk cancer_patients	6.51	NCI	1R01CA129639-01A2	Genome-Wide Association Study of Radiation Exposure and Bilateral Breast Cancer PI: BERNSTEIN, JONINE LISA
3.86	437	risk risk_factors cases cohort prospective high_ris	risk, risk_factors, v	6.46	NCI	1K07CA136758-01A1	Genetic variants in the PI3K pathway in mammographic density and breast cancer PI: THOMPSON, CHERYL L.
3.76	544	snps snp genome_wide_association cases genes	genome_wide_ass	6.31	NCI	5P50CA116199-05	UTMDACC SPORE in Breast Cancer PI: HORTOBAGYI, GABRIEL N
3.70	173	genetic genes risk susceptibility polymorphisms	genetic, genetics,	6.02	NCI	2R01CA050385-21A1	Risk Factors for Breast Cancer in Younger Nurses PI: WILLETT, WALTER C.
2.62	252	treatment patients management patient outcom	management, tre	4.6	NCI	5R01CA127617-02	Who Cares For Older Breast Cancer Survivors And How Does It Affect Quality? PI: MANDELBLATT, JEANNE
1.64	235	conference meeting workshop symposium scienti	th, conference, sy				
1.63	351	community implementation community_based he	community, preve				
1.54	325	million disease treatment united_states public_h	disease, treatmen				
1.51	580	training candidate career skills applicant program	treatment, depres				

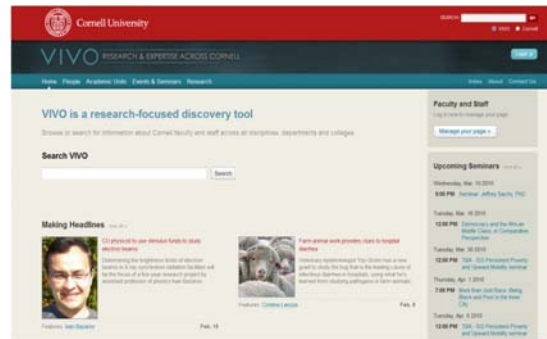
<https://app.nihmaps.org>

37

VIVO International Researcher Network

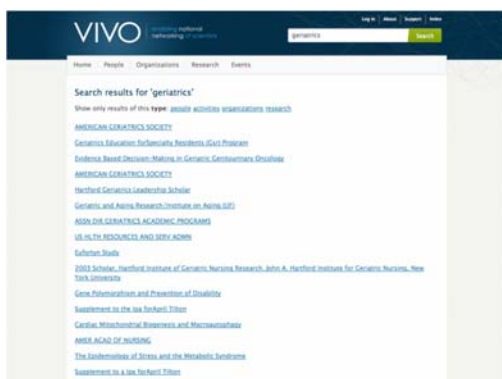
VIVO: A Semantic Approach to Creating a National Network of Researchers (<http://vivoweb.org>)

- 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 Ping, 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.



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

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

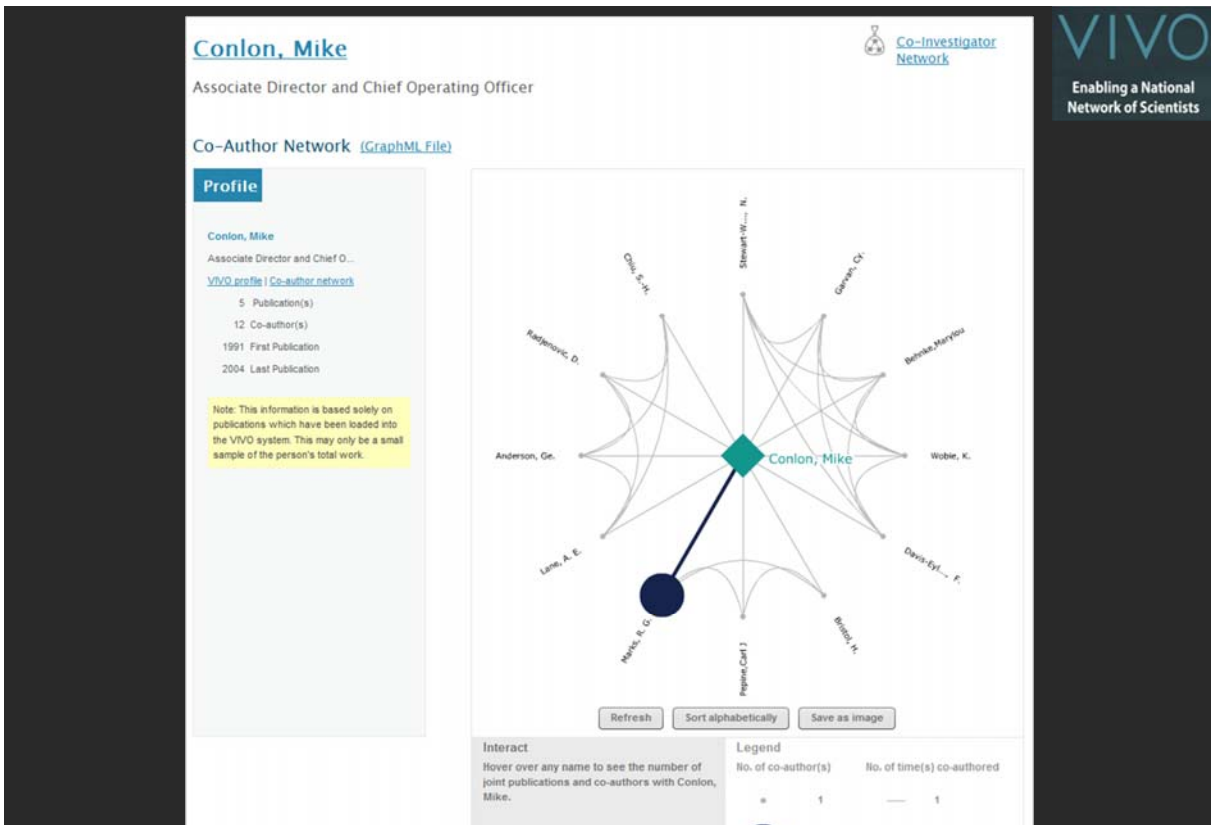
1 - 13 of 554

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

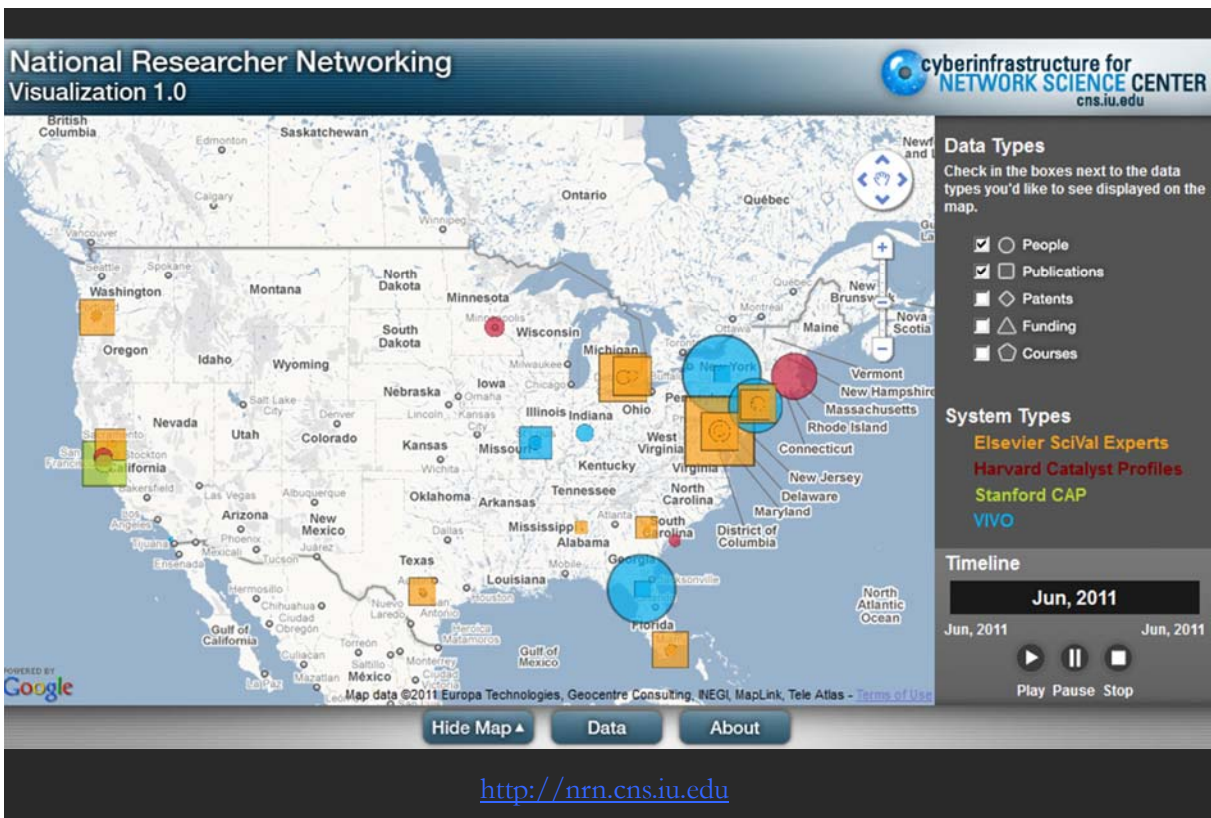
Top 290 disciplines shown

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? 43



Geospatial Analysis (Where) Where is what science performed by whom? Science is global and needs to be studied globally. 44

VIVO On-The-Go

Overview, Interactivity, Details on Demand

come to commonly used devices and environments



Borner, Katy
Person

This information is based solely on publications which have been loaded into the VIVO system. This may only be a small sample of the person's total work.

General Statistics

- 36 publications from 2001 to 2010 [\(.CSV File\)](#)
- 87 co-author(s) from 2001 to 2010 [\(.CSV File\)](#)

Co-Author Network [\(GraphML File\)](#)

13 co-author(s)
26 co-author(s)

Borner, Katy
Person

- 36 Publications
- 13 Co-author(s)
- 2001 First Publication
- 2007 Last Publication

Legend

- No. of publication(s)
- No. of BMC(s) co-authored

Interact

Hover over any name to see the number of past publications and co-author(s) with Borner, Katy. Click on a name to see details on the right.

Thresholding

Only people that co-authored more than 1 paper(s) with Borner, Katy are shown.
13 out of 87 co-author(s) are shown.
Borner and co-author(s) are placed near one another if they regularly collaborate with each other and each other's co-author(s) in the graph.

[Change to log scale](#) [Refresh](#) [Sort alphabetically](#) [Save as Image](#)

Tables

Year	Publications per Year	Co-author(s)	Publications with Borner, Katy
2001	2		
2002	4	Chen C	6
2003	2	Stropick K W	4
2004	7	Mine K K	4
2005	7	Va V K	3
2006	3	Panumathy S	3
2007	10	Vespiarian, Alessandro	2
		Herr B	2
		Hards E	2
		Holloway T	2
		Herr D P W	2
		Thurston S	2
		Feng Y	2
		Mine K	2

Download Data

General Statistics

- 36 publication(s) from 2001 to 2010 [\(.CSV File\)](#)
- 80 co-author(s) from 2001 to 2010 [\(.CSV File\)](#)

Co-Author Network

[\(GraphML File\)](#)

Save as Image (.PNG file)

Tables

- Publications per year [\(.CSV File\)](#)
- Co-authors [\(.CSV File\)](#)

<http://vivo.iu.edu/vis/author-network/person25557>

36 publication(s) from 2001 to 2010 ([.CSV File](#))

Year	Publications
2001	2
2002	4
2003	2
2004	7
2005	7
2006	3
2007	10
2010	1

80 co-author(s) from 2001 to 2010 ([.CSV File](#))

Year	Count	Co-Author(s)
2001	1	Chen C.
2002	3	Chen C.; McMahon T.; Feng Y.
2003	2	Chen C.; Boyack K.W.
2004	17	Sengupta A.; Penumarthy S.; Thakur S.; Sooriamurthy R.; Maru J.T.; Shiffrin R.M.; Mane K.; Moor K.A.;

Co-author network ([GraphML File](#))

```

1 <?xml version="1.0" encoding="UTF-8" ?>
2 <graphml xmlns="http://graphml.graphdrawing.org/xmlns"
3 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
4 xsi:schemaLocation="http://graphml.graphdrawing.org/xmlns
5 http://graphml.graphdrawing.org/xmlns/1.0/graphml.xsd">
6 <key id="label" for="node" attr.name="label" attr.type="string" />
7 <key id="number_of_authored_works" for="node" attr.name="number_of_authored_works" attr.type="int" />
8 <key id="num_unknown_publication" for="node" attr.name="num_unknown_publication" attr.type="int" />
9 <key id="num_latest_publication" for="node" attr.name="num_latest_publication" attr.type="int" />
10 <key id="latest_publication" for="node" attr.name="latest_publication" attr.type="int" />
11 <key id="profile_url" for="node" attr.name="profile_url" attr.type="string" />

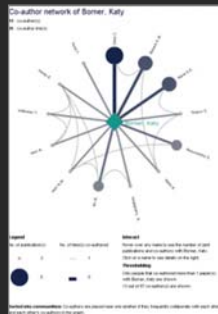
```

Save as Image (.PNG file)

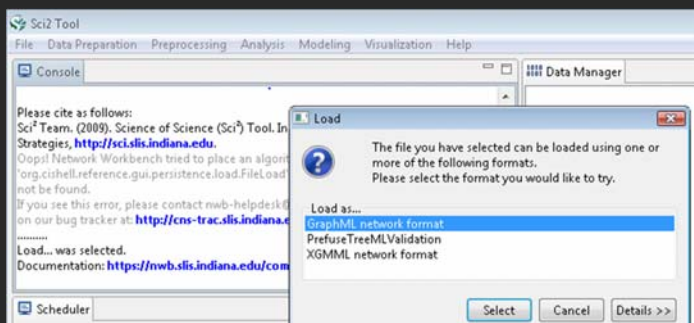
Publications per year ([.CSV File](#)), see top file.

Co-authors ([.CSV File](#))

Co-Author	Count
Andrienko G.	1
Andrienko N.	1
Ben-Miled Z.	1
Blackwell A.	1
Boyack K.W.	4
Bozicevic M.	1
Brodbeck D.	1
Burkhard R.A.	1
Chen C.	5



Run Sci2 Tool and Load Co-Author Network ([GraphML File](#))



Network Analysis Toolkit
Nodes: 81
Edges: 390

Visualize the file using Radial Graph layout.



Click on node to focus on it.

Hover over a node to highlight its co-authors.

Code and tutorials are linked from <http://sci2.wiki.cns.edu>

Network Analysis Toolkit (NAT) was selected.

Nodes: 109

Isolated nodes: 0

Node attributes present: label, number_of_authored_works, num_unknown_publication, num_latest_publication, latest_publication, profile_url, num_earliest_publication, earliest_publication, url

Edges: 731

No self loops were discovered.

No parallel edges were discovered.

Average degree: 13.4128

This graph is weakly connected.

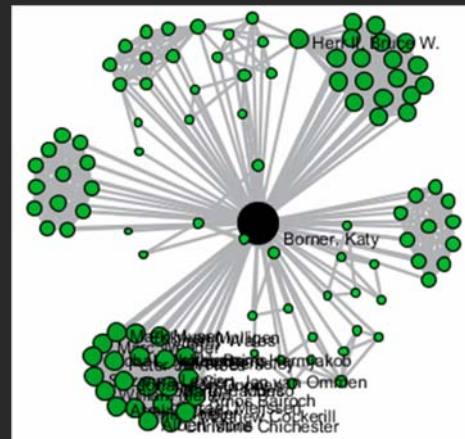
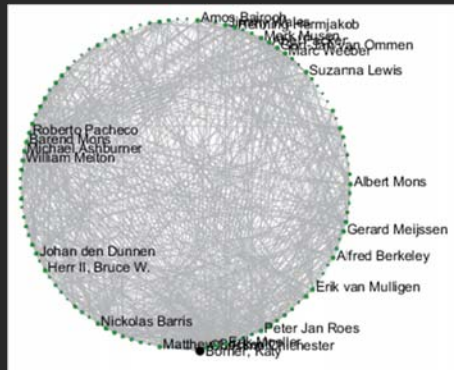
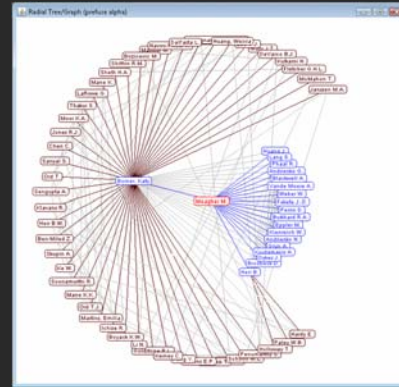
There are 1 weakly connected components. (0 isolates)

The largest connected component consists of 109 nodes.

Density (disregarding weights): 0.1242

.....

Node Degree was selected.



49

Develop VIVO Visualizations

See also *Visualization in VIVO Workshop on Aug 24, 2011*

<http://wiki.cns.in.edu/display/PRES/VIVO+Presentation>



VIVO Presentation

4 Added by Chin Hua Kong, last edited by Chintan Tank on Aug 24, 2011 (view change)

August, 2011 Workshop

Material

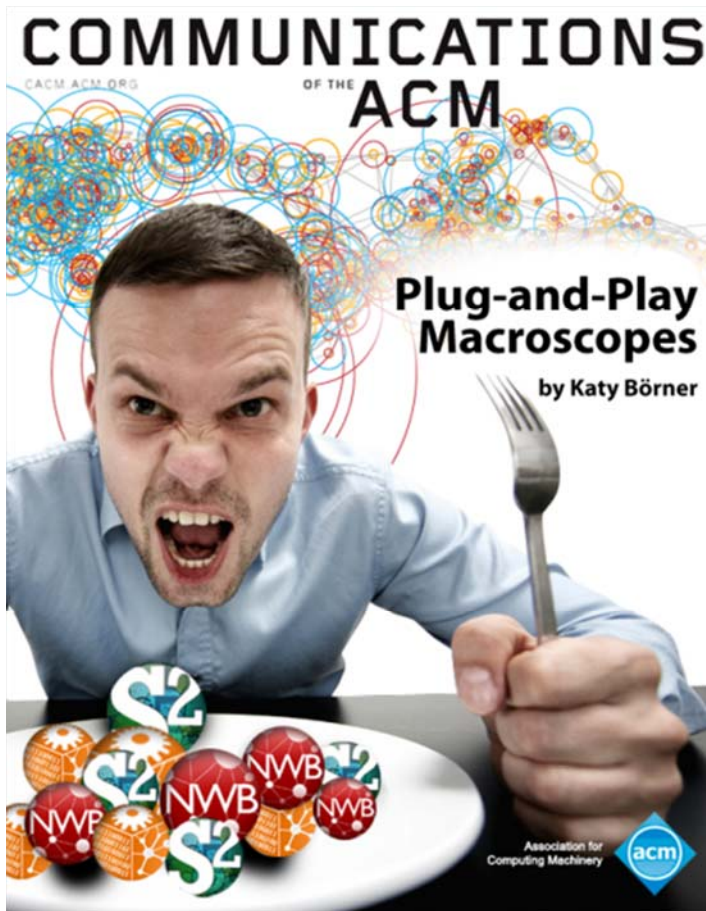
- [Java 1.5 or higher](#) - A programming language and computing platform for developing cross OS softwares.
- [Science of Science tool \(Sci2\)](#) - An desktop application for information analysis and visualization.
- [Gephi](#) - An interactive visualization tool for networks and complex systems, dynamic and hierarchical graphs.
- [VIVO August 2011 workshop data.zip](#) - Hands on workshop data package

Slides

- [Tutorial Slides](#) presented at the VIVO Conference 2011
- [Pre-Questionnaire and Post-Questionnaire](#)

Demo Links

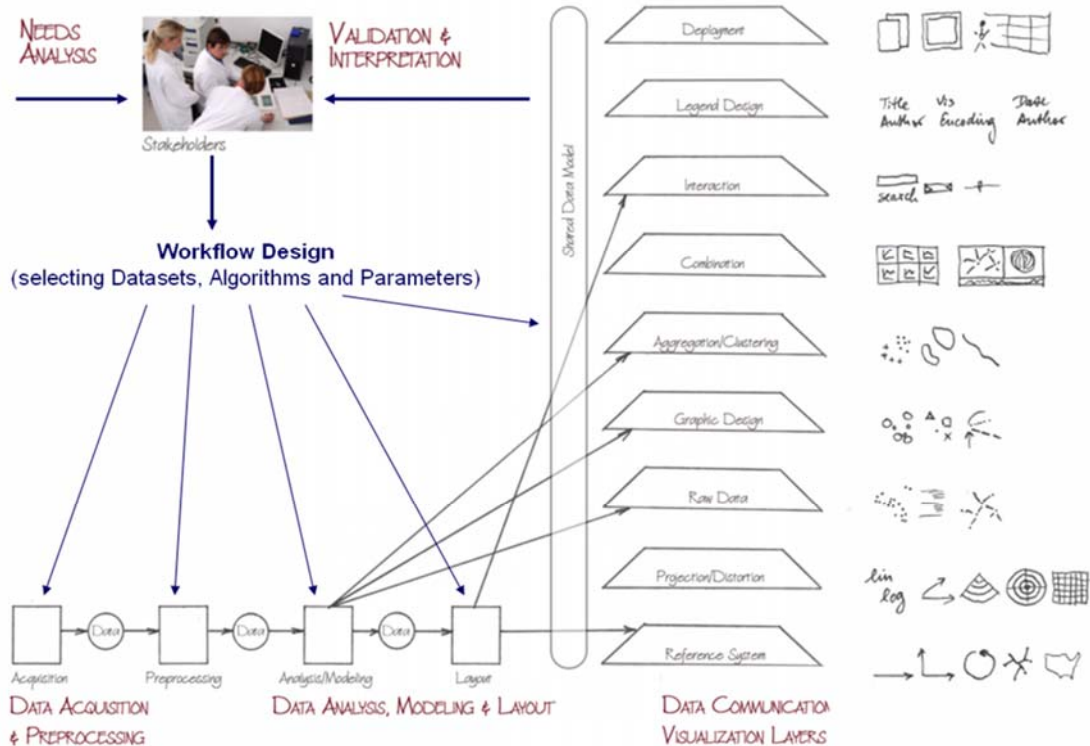
- [Map of Science Visualization \(dev link\)](#)
- [Temporal Graph Visualization \(dev link\)](#)
- [National Researcher Networking Visualization](#)
- [Word Cloud Visualization dev link](#)



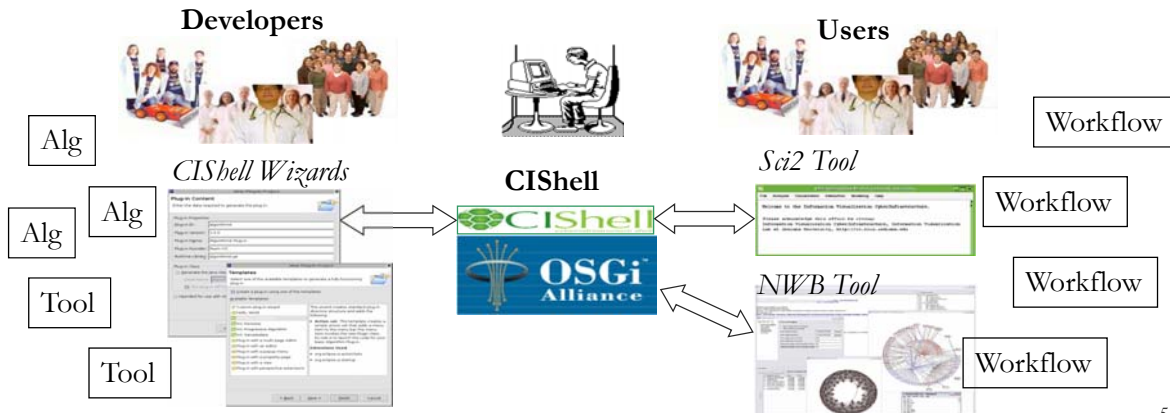
Börner, Katy. (March 2011). Plug-and-Play Macroscopes. *Communications of the ACM*, 54(3), 60-69.

Video and paper are at <http://www.scivee.tv/node/27704>


Needs-Driven Workflow Design using a modular data acquisition/analysis/ modeling/ visualization pipeline as well as modular visualization layers.



- 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.



55



[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](#)

56

57



Network Workbench Tool

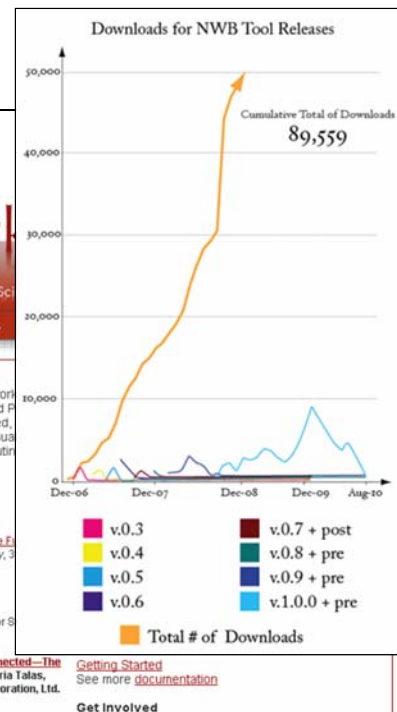
<http://nwb.cns.edu>

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.



Computational Proteomics

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

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



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

59

Computational Economics

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

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

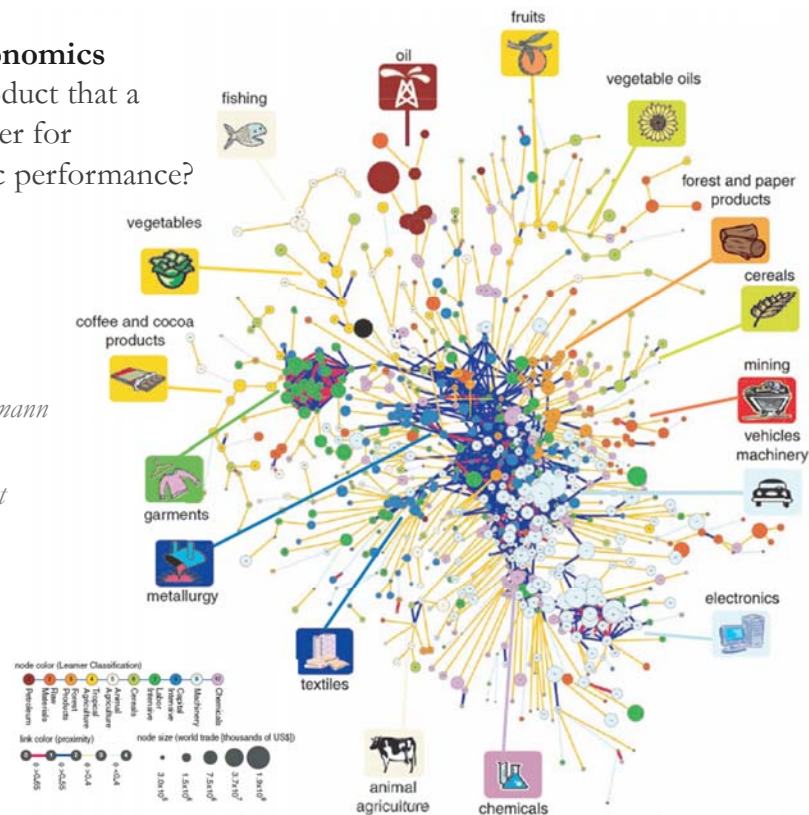


Fig. 1. The product space. (A) Hierarchically clustered proximity matrix representing the 775 SITC-4 product classes exported in the 1998–2000 period. (B) Network representation of the product space. Links are color coded with their proximity value. The sizes of the nodes are proportional to world trade,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

60

Computational Social Science

Studying large scale social networks such as Wikipedia

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

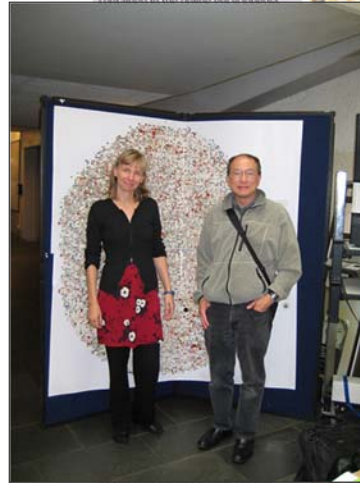


Second sight

Image: Bruce W. Hest and Todd M. Holloway

Power struggle

How do you keep track of the bobbling mass of information that is Wikipedia? This chaotic-looking mosaic is one attempt to show which topics are contained in the online encyclopedia.



...which were the most viewed pages at the time of writing include entries on Sheffield Wednesday football club, Mikhail Gorbachev and pigs). The mosaic has been commended in a competition for images that visualise network dynamics, coinciding with this week's International Workshop and Conference on Network Science in Bloomington.

www.newscientist.com



19 May 2007 | NewScientist | 55

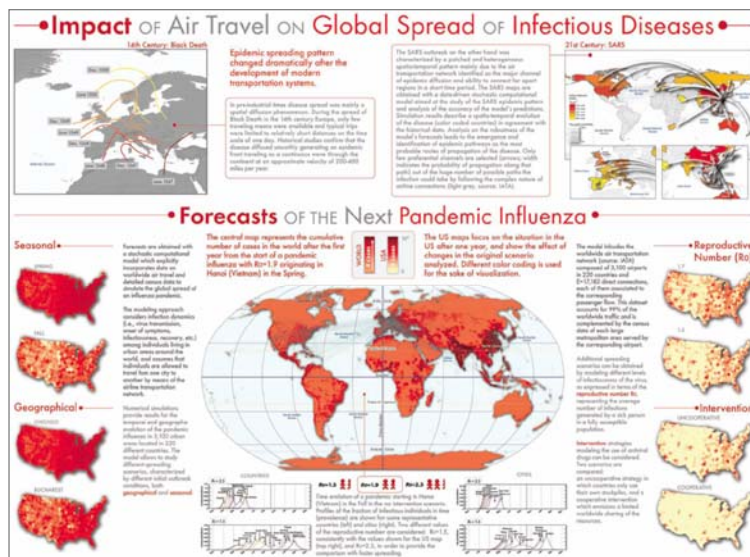
Computational Epidemics

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

Epidemic Modeling in Complex realities, V. Colizza, A. Barrat, M. Barthelemy, A. Vespignani, Comptes Rendus Biologie, 330, 364-374 (2007).

Reaction-diffusion processes and metapopulation models in heterogeneous networks, V. Colizza, R. Pastor-Satorras, A. Vespignani, Nature Physics 3, 276-282 (2007).

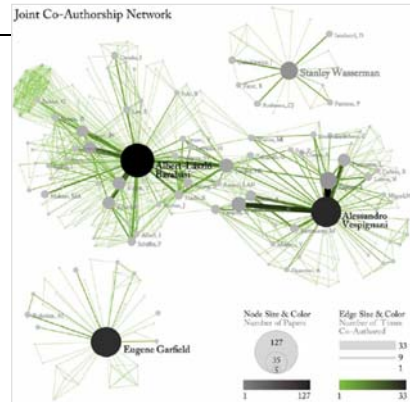
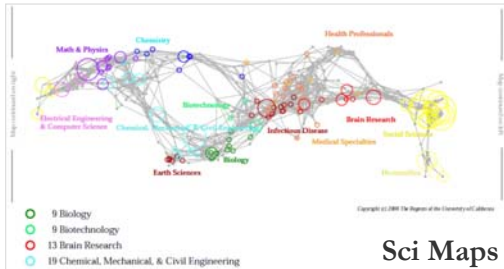
Modeling the Worldwide Spread of Pandemic Influenza: Baseline Case and Containment Interventions, V. Colizza, A. Barrat, M. Barthelemy, A.-J. Valleron, A. Vespignani, PLoS-Medicine 4, e13, 95-110 (2007).





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). *Reti-Netzwerk-Red: Analyzing and Visualizing Scholarly Networks Using the Scholarly Database and the Network Workbench Tool*. *Proceedings of ISSI 2009: 12th International Conference on Scientometrics and Informetrics, Rio de Janeiro, Brazil, July 14-17*. Vol. 2, pp. 619-630.



Sci² Tool Vis cont.

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>).

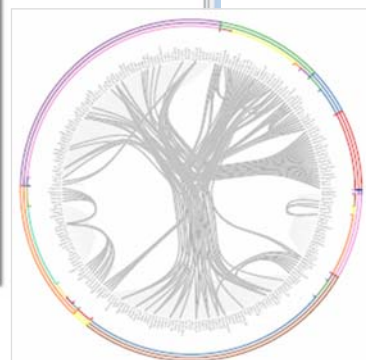
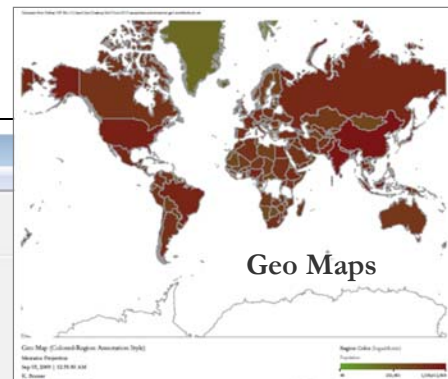
The primary investigators are Katy Börner, In SciTech Strategies Inc. The Sci² tool was developed by J. Duhon, Patrick A. Phillips, Chintan Tank, a Cyberinfrastructure Shell (<http://cishell.org>) for Network Science Center (<http://cns.slis.indiana.edu>). Many algorithm plugins were derived from the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

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

Scheduler

Remove From List Remove completed

!	Algorithm Name	Date	Time	% Con
<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%



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

EpiC Tool
File | Compartmental Modeling | Networks | Simulation | Visualization | R | Help

Welcome to the EpiC tool, which supports the modeling, analysis, and visualization of epidemic processes.

The EpiC project (<http://epic.sls.indiana.edu>) is supported in part by the NIH RM-07-004 award. The primary investigators are Dr. Katy Börner, Dr. Alessandro Vespignani, and Dr. Jim Sherman.

Please cite as follows:
EpiC Team. (2009). EpiC Tool. Indiana University.

File
Create a compartmental model
Edit compartmental model

Simulation
Single-Population
Exact
Network

Visualization
Line Graph

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

EpiC
cyberinfrastructure for NETWORK SCIENCE CENTER
CIShell Powered

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.

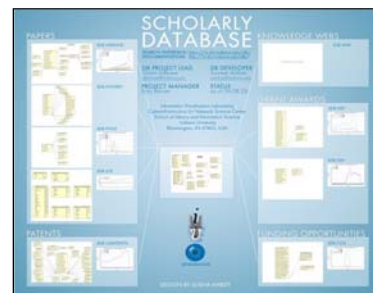
Computational Scientometrics Cyberinfrastructures



Scholarly Database: 25 million scholarly records
<http://sdb.cns.iu.edu>



VIVO Research Networking
<http://vivoweb.org>



Information Visualization Cyberinfrastructure
<http://iv.cns.iu.edu>



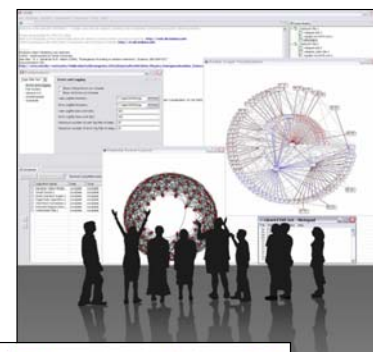
Network Workbench Tool & Community Wiki
<http://nwb.cns.iu.edu>



Science of Science (Sci²) Tool
<http://sci2.cns.iu.edu>



Epidemics Tool & Marketplace
Forthcoming



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

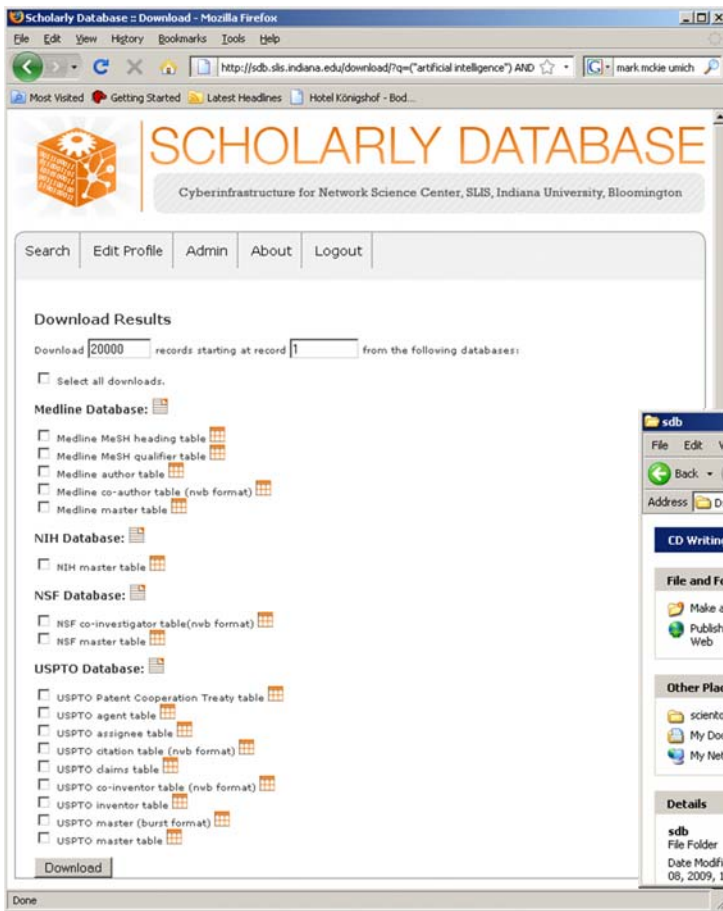
Results can be downloaded as data dump and (evolving) co-author, paper-citation networks.

The image shows two screenshots of the Scholarly Database website. The left screenshot displays the login page with options for 'IU User' and 'Non-IU User'. The right screenshot shows the search interface with a search bar, filters for 'First Year' and 'Last Year', and a list of database sources: Medline (1998 - 2008), NIH (1961 - 2002), NSF (1985 - 2004), and USPTO (1976 - 2007). A 'Search' button is visible at the bottom of the search interface.

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

The image is a screenshot of a Mozilla Firefox browser window displaying the Scholarly Database search results for the query "artificial intelligence". 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 is a navigation menu with "Search", "Edit Profile", "Admin", "About", and "Logout". The main content area is titled "Browse Results" and shows the search results summary: "Your search returned 13,231 results in 0.295 seconds." with a "Download" button. 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

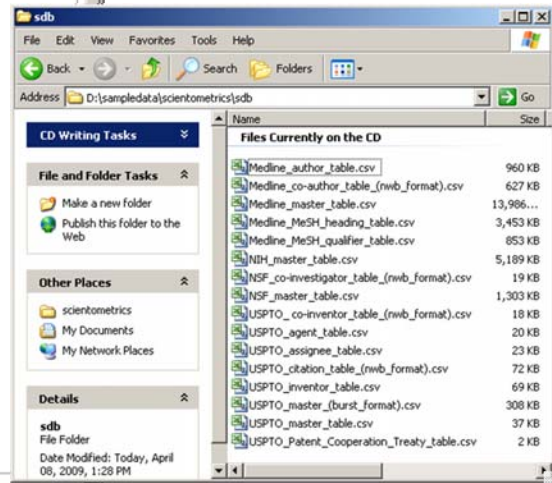


Since March 2009:

Users can download networks:

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

and tables for burst analysis in NWB.



"Sci2 Tool: Temporal, Geospatial, Topical, and Network Analysis and Visualization" Tutorial

- Instructor:** Dr. Katy Börner
- Time/Date:** 12:30-16:30 on Feb 16, 2012
- Place:** Meertens Institute, Joan Muyskenweg 25, 1096 CJ Amsterdam
- Format:** Lecture and “hands-on” training. Please bring your laptop.
- Audience:** This tutorial is designed for researchers and practitioners interested to use advanced data mining algorithms and visualizations in their research and daily decision making.
- Cost:** Free but register via <http://www.surveymonkey.com/s/TB2R7RL>

Abstract:

The Science of Science Tool (Sci²) (<http://sci2.cns.iu.edu>) was designed for researchers and practitioners interested to study and understand the structure and dynamics of science. Today is used by major federal agencies in the US but also by researchers from more than 40 countries and from many different areas of research -- including arts and humanities scholars.

"Sci2 Tool: Temporal, Geospatial, Topical, and Network Analysis and Visualization" Tutorial cont.

Abstract cont.

Sci2 is a standalone desktop application that installs and runs on Windows, Linux x86 and Mac OSX and supports:

- Reading and writing of 20 major file formats (e.g., ISI, Scopus, bibtex, nsf, EndNote, CSV, Pajek .net, XGMML, GraphML),
- Easy access to more than 180 algorithms for the temporal, geospatial, topical, and network analysis and visualization of scholarly datasets at the micro (individual), meso (local), and macro (global) levels, and
- Professional visualization of analysis results by means of large-format charts and maps.

The first hour of the tutorial provides a basic introduction of the tool. Remaining time will be spent discussing sample workflows featured in the Sci2 Tutorial at (<http://sci2.wiki.cns.iu.edu>) and new functionality such as the Yahoo! geocoder, network clustering and backbone identification algorithms, and the analysis and visualization of evolving networks.

Reference

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

73

"Sci2 Tool: Temporal, Geospatial, Topical, and Network Analysis and Visualization" Tutorial cont.

Agenda

12:30p Welcome and Overview of Tutorial and Attendees

12:45p Plug-and-Play Macroscopes, OSGi/CIShell Powered Tools

1:00p Sci2 Tool Basics

Download and run the Sci2 Tool

Load, analyze, and visualize family and business networks

Studying four major network science researchers

- Load and clean a dataset; process raw data into networks

- Find basic statistics and run various algorithms over the network

- Visualize the network using different layouts

2:30p Break

3:00p Sci2 Tool Novel Functionality

- Yahoo! geocoder

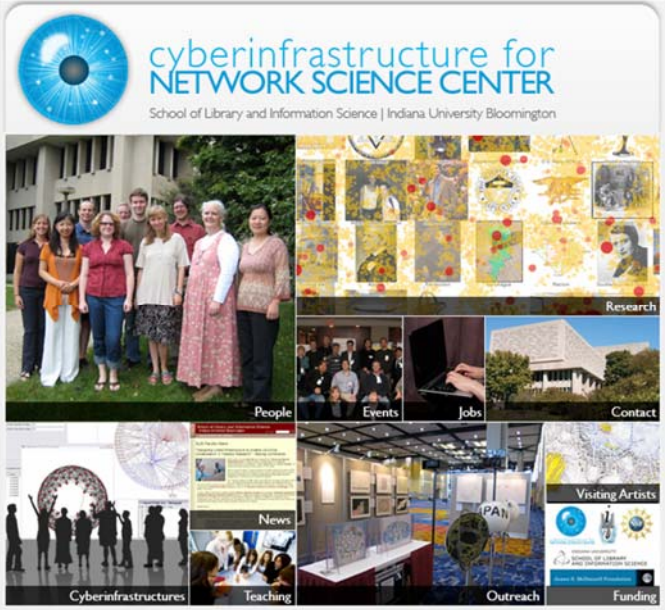
- Evolving collaboration networks

- R-Bridge

4:00p Outlook and Q&A

4:30p Adjourn

74



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>



Science & Technology Forecasts @ Times Square in 2016

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

