Towards a Science of Science Cyberinfrastructure

Dr. 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

*Center for Data and Search Informatics * Spring 2009 Seminar Series*
Wednesday April 8, 4:00pm-5:00 p.m., Lindley Hall 101

Overview

What cyberinfrastructure is required to measure, model, analyze, and communicate scholarly data and ultimately scientific progress?

This talk presents our efforts to create a science of science cyberinfrastructure that supports:

- Data access and federation via the Scholarly Database, http://sdb.slis.indiana.edu,
- Data preprocessing, modeling, analysis, and visualization using plug-and-play cyberinfrastructures such as the Network Workbench, http://nwb.slis.indiana.edu, and
- Communication of science to a general audience via the *Mapping Science* exhibit at http://scimaps.org.

This talk should be particularly interesting for those interested to

- Map their very own domain of research,
- Test and compare data federation, mining, visualization algorithms on large scale datasets,
- Use advanced network science algorithms in their own research.

Overview

1. Needs Analysis
2. Conceptualizations of Science

3. Scholarly Database (SDB)
4. Network Workbench (NWB) Tool

5. **Exemplary Analyses and Visualizations using SDB/NWB**

6. Mapping Science Exhibit

1. Needs Analysis

A total of 34 science policy makers and researchers at university campus level (8), program officer level (12), and division director level at national, state, and private foundations (10) as well as science policy makers from Europe and Asia (4) were interviewed between Feb. 8th, 2008 and Oct. 2nd, 2008.

Each interview comprised a 40 min, audio-taped, informal discussion on specific information needs, datasets and tools currently used, and information on what a 'dream tool' might look and feel like. There is also a pre-interview questionnaire to acquire demographics and a post-interview questionnaire to get input on priorities.

Data compilation is in progress, should be completed in July 2009, and will be submitted as a journal paper. Some data excerpts are given here.

2. Conceptualizations of Science

See Special Issue of *Journal of Informetrics,* 3(3), Jan 2009.

Science of Science: Conceptualizations and Models of Science

Guest Editors: Katy Börner, Indiana University & Andrea Scharnhorst, Royal Netherlands Academy of Arts and Sciences

This special issue of the journal *Informetrics* aims to improve our understanding of the structure and evolution of science by reviewing and advancing existing conceptualizations and models of scholarly activity.

Existing conceptualizations and models of science have been created by scholars from very different disciplines and backgrounds. They have the form of

- philosophical concepts (Bernal, Kuhn, Popper),
- (utopian) stories (Wells, Lem),
- visual drawings (Otlet),
- empirical measurements (Price, Garfield), or
- mathematical theories (Goffman, Yablonski)

among others.

It is our belief that a theoretically grounded and practically useful shared conceptualization of science can provide the intellectual framework to interlink and puzzle together the hundreds of science models in existence today. This is analogous to how meteorologists or seismologists integrate rather different local weather models or seismic hazard predictions into a global coherent model that has higher predictive value and broader coverage. With this issue we aim to start an interdisciplinary discourse towards a science of science models.

The design of such a conceptualization requires the identification of the

- Boundaries of the system or object.
- Basic building blocks of science, e.g., units of analysis or key actors.
- Interactions of building blocks, e.g., via coupled networks.
- Basic mechanisms of growth and change.

Editorial is available at http://ivl.slis.indiana.edu/km/pub/2009-borner-scharnhorst-joi-sos-intro.pdf

3. Scholarly Database

http://sdb.slis.indiana.edu

Nianli Ma

"From Data Silos to Wind Chimes"

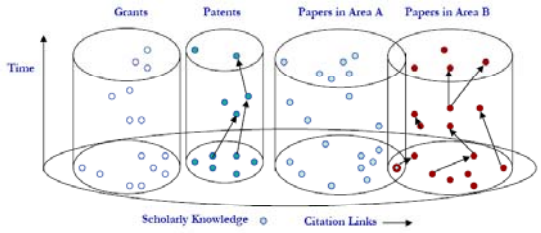
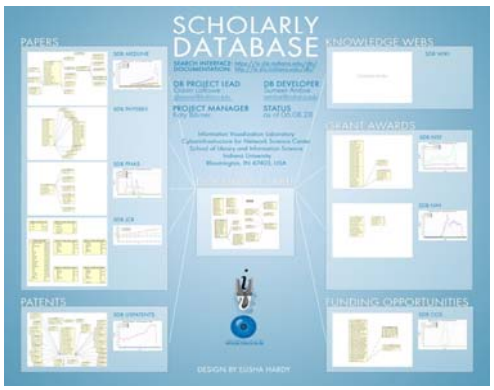


Figure 1: The interoperability and cross linkage problem. Many but not all of today's scholarly datasets, e.g., papers, patents, grants, are stored and made available so that 'vertical' citation linkages can be traversed. There are very few instances in which datasets of different origin and/or type are 'horizontally' interlinked.



- Create public databases that any scholar can use. Share the burden of data cleaning and federation.
- Interlink creators, data, software/tools, publications, patents, funding, etc.

La Rowe, Gavin, Ambre, Sumeet, Burgoon, John, Ke, Weimao and Börner, Katy. (2007) The Scholarly Database and Its Utility for Scientometrics Research. In Proceedings of the 11th International Conference on Scientometrics and Informetrics, Madrid, Spain, June 25-27, 2007, pp. 457-462. http://ella.slis.indiana.edu/~katy/paper/07-issi-sdb.pdf

3. Scholarly Database: # Records & Years Covered

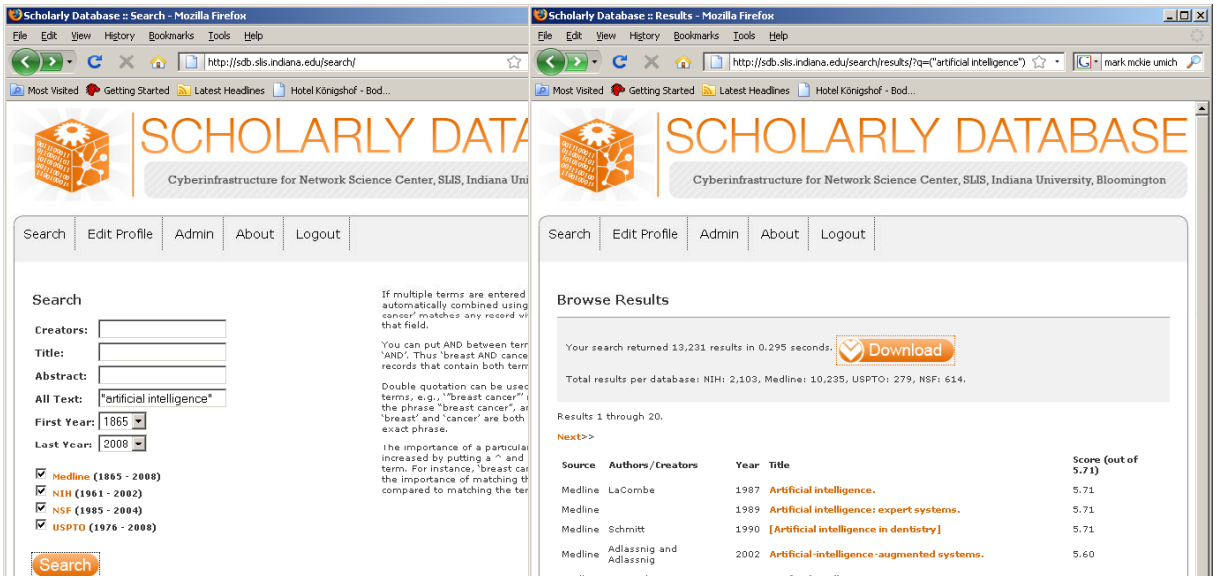
Datasets available via the Scholarly Database (* internally)

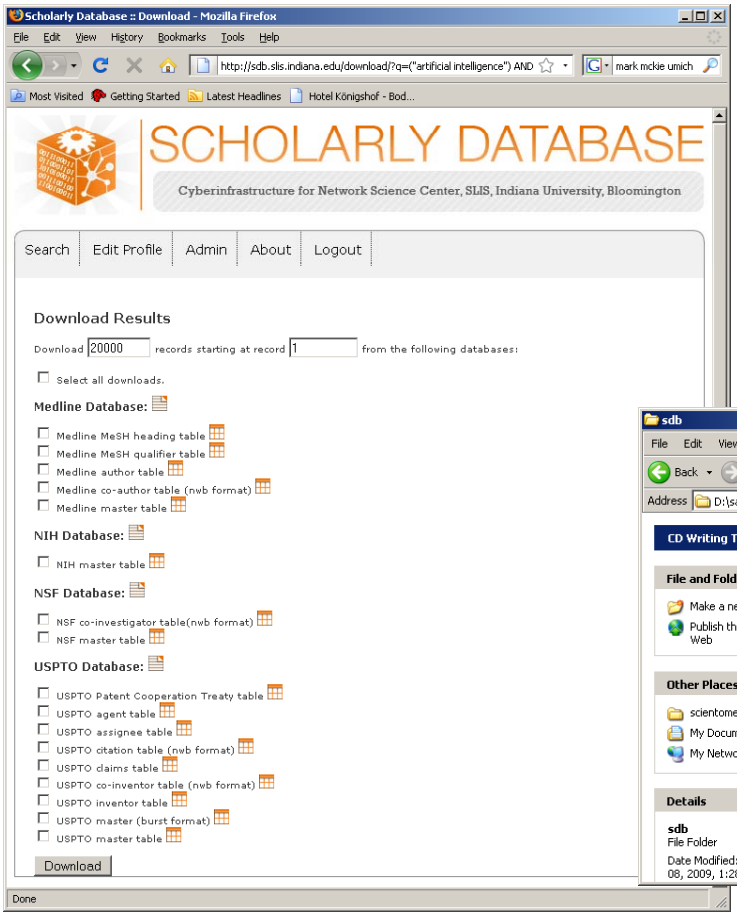
Dataset	# Records	Years Covered	Updated	Restricted Access
Medline	17,764,826	1898-2008	Yes	
PhysRev	398,005	1893-2006		Yes
PNAS	16,167	1997-2002		Yes
JCR	59,078	1974, 1979, 1984, 1989 1994-2004		Yes
USPTO	3, 710,952	1976-2008	Yes*	
NSF	174,835	1985-2002	Yes*	
NIH	1,043,804	1961-2002	Yes*	
Total	**23,167,642**	**1893-2006**	**4**	**3**

Aim for comprehensive time, geospatial, and topic coverage.

3. Scholarly Database: Web Interface

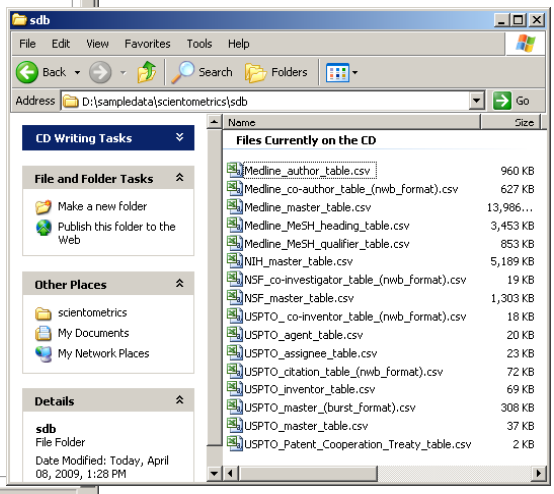
Anybody can register for free to search the about 23 million records and download results as data dumps.
Currently the system has over 120 registered users from academia, industry, and government from over 60 institutions and four continents.





Since March 2009:
Users can download networks:
- Co-author
- Co-investigator
- Co-inventor
- Patent citation

and tables for
burst analysis in NWB.



4. Scientometrics Filling of Network Workbench Tool

will ultimately be 'packaged' as a SciPolicy' tool

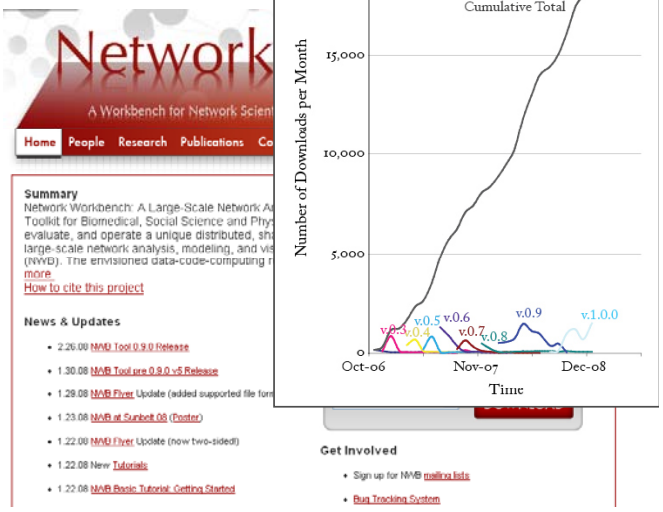
http://nwb.slis.indiana.edu/

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 Feb. 2009, the tool provides more 100 plugins that support the preprocessing, analysis, modeling, and visualization of networks.

More than 40 of these plugins can be applied or were specifically designed for S&T studies.

It has been downloaded more than 19,000 times since Dec. 2006.

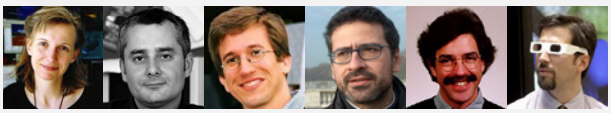


Herr II, Bruce W., Huang, Weixia (Bonnie), Penumarthy, Shashikant & Börner, Katy. (2007). Designing Highly Flexible and Usable Cyberinfrastructures for Convergence. In Bainbridge, William S. & Roco, Mihail C. (Eds.), Progress in Convergence - Technologies for Human Wellbeing (Vol. 1093, pp. 161-179),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

NetworkWorkbench
A Workbench for Network Scientists

Project Details

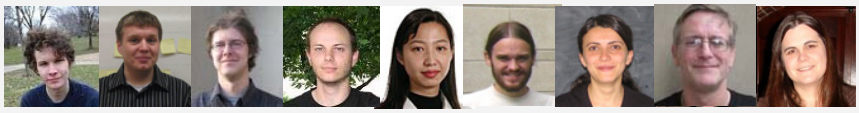
Investigators: Katy Börner, Albert-Laszlo Barabasi, Santiago Schnell, Alessandro Vespignani & Stanley Wasserman, Eric Wernert



Software Team: Lead: Micah Linnemeier

Members: Patrick Phillips, Russell Duhon, Tim Kelley & Ann McCranie

Previous Developers: Weixia (Bonnie) Huang, Bruce Herr, Heng Zhang, Duygu Balcan, Mark Price, Ben Markines, Santo Fortunato, Felix Terkhorn, Ramya Sabbineni, Vivek S. Thakre & Cesar Hidalgo



Goal: Develop a large-scale network analysis, modeling and visualization toolkit for physics, biomedical, and social science research.

Amount: $1,120,926, NSF IIS-0513650 award

Duration: Sept. 2005 - Aug. 2009

Website: http://nwb.slis.indiana.edu

4. NWB Tool: Supported Data Formats

Personal Bibliographies

- Bibtex (.bib)
- Endnote Export Format (.enw)

Data Providers

- Web of Science by Thomson Scientific/Reuters (.isi)
- Scopus by Elsevier (.scopus)
- Google Scholar (access via *Publish or Perish* save as CSV, Bibtex, EndNote)
- Awards Search by National Science Foundation (.nsf)

Scholarly Database (all text files are saved as .csv)

- Medline publications by National Library of Medicine
- NIH funding awards by the National Institutes of Health (NIH)
- NSF funding awards by the National Science Foundation (NSF)
- U.S. patents by the United States Patent and Trademark Office (USPTO)
- Medline papers – NIH Funding

Network Formats

- NWB (.nwb)
- Pajek (.net)
- GraphML (.xml or .graphml)
- XGMML (.xml)

Burst Analysis Format

- Burst (.burst)

Other Formats

- CSV (.csv)
- Edgelist (.edge)
- Pajek (.mat)
- TreeML (.xml)

4. NWB Tool: Algorithms (July 1st, 2008)

See https://nwb.slis.indiana.edu/community and handout for details.

Preprocessing Edit

- **Remove Nodes**
  - Extract Top Nodes
  - Extract Nodes Above or Below Val
  - Delete High Degree Nodes
  - Delete Random Nodes
  - Delete Isolates
- **Remove Edges**
  - Extract Top Edges
  - Extract Edges Above or Below Val
  - Remove Self Loops
  - Trim By Degree[2]
  - Pathfinder Network Scaling
- **Sampling**
  - Snowball Sampling (n nodes)
  - Node Sampling
  - Edge Sampling
- **Transformations**
  - Symmetrize
  - Dichotomize
  - Multipartite Joining

Modeling Edit

- **General**
  - Random Graph
  - Watts-Strogatz Small World
  - Barabási-Albert Scale-Free
- **Structured**
  - CAN
  - Chord
- **Unstructured**
  - Hypergrid
  - PRU
- **Other**
  - TARL
  - Discrete Network Dynamics

Analysis Edit

- **General Purpose**
  - Network Analysis Toolkit[2]
- **Unweighted & Undirected**
  - **Based on degree/**
    - Node Degree
    - Node Distribution
  - **Based on clustering**
    - k-Nearest Neighbor
    - Watts Strogatz Clustering Coefficie
    - Watts Strogatz Clustering Coefficie
  - **Based on path**
    - Diameter
    - Average Shortest Path
    - Shortest Path Distribution
    - Node Betweenness Centrality
  - **Based on components**
    - Connected Components
    - Weak Component Clustering
  - **K-Core**
    - Extract K-Core[2]
    - Annotate K-Coreness[2]
- **Unweighted & Directed**
  - **Based on degree**
    - Node Indegree
    - Node Outdegree
    - Indegree Distribution
    - Outdegree Distribution
  - **Based on local graph structure**
    - k-Nearest Neighbor
    - Single Node In-Out Degree Correla
  - **Unnamed Category?**
    - Page Rank
  - **Based on local graph structure**
    - Dyad Reciprocity[2]
    - Arc Reciprocity[2]
    - Adjacency Transitivity[2]
  - **Based on components**

Visualization Edit

- **Tools**
  - GUESS
  - GnuPlot[2]
- **Predefined Positions Layout**
  - DrL (VxOrd)
  - Pre-defined Positions (prefuse beta)[2]
- **Move**
  - Circular
- **Tree Layouts**
  - Radial Tree (prefuse alpha)
  - Radial Tree with Annotations (prefuse beta)[2]
  - Tree Map
  - Tree View
  - Balloon Graph (prefuse alpha)[2]
- **Network Layouts**
  - Force Directed with Annotation (prefuse beta)
  - Kamada-Kawai (JUNG)
  - Fruchterman-Reingold (JUNG)
  - Fruchterman-Reingold with Annotation (prefuse beta)
  - Spring (JUNG)
  - Small World (prefuse alpha)
- **Other Layouts**
  - Parallel Coordinates (demo)[2]
  - LaNet (k-Core Decomposition)

Scientometrics Edit

- **Extract Network From Table**
  - Extract Co-Authorship Network
  - Extract Co-Occurrence Network From Table[2]
  - Extract Directed Network From Table[2]
- **Extract Network From Another Network**
  - Extract Bibliographic Coupling Similarity Network
  - Extract Co-Citation Similarity Network[2]
- **Cleaning**
  - Remove ISI Duplicate Records
  - Detect Duplicate Nodes
  - Remove Rows With Multitudinous Fields[2]

4. NWB Tool: Output Formats

NWB tool can be used for data conversion. Supported output formats comprise:

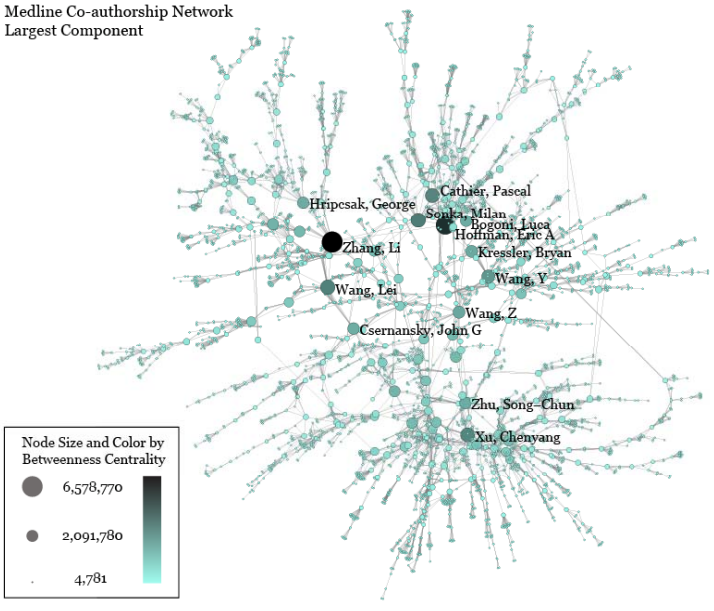
- CSV (.csv)
- NWB (.nwb)
- Pajek (.net)
- Pajek (.mat)
- GraphML (.xml or .graphml)
- XGMML (.xml)

GUESS

- Supports export of images into common image file formats.

Horizontal Bar Graphs

- saves out raster and ps files.



5. Exemplary Analyses and Visualizations

Individual Level

A. Loading ISI files of major network science researchers, extracting, analyzing and visualizing paper-citation networks and co-author networks.

B. Loading NSF datasets with currently active NSF funding for 3 researchers at Indiana U

Institution Level

C. Indiana U, Cornell U, and Michigan U, extracting, and comparing Co-PI networks.

Scientific Field Level

D. Extracting co-author networks, patent-citation networks, and detecting bursts in SDB data.

5. Exemplary Analyses and Visualizations

Individual Level

A. Loading ISI files of major network science researchers, extracting, analyzing and visualizing paper-citation networks and co-author networks.

B. Loading NSF datasets with currently active NSF funding for 3 researchers at Indiana U

Institution Level

C. Indiana U, Cornell U, and Michigan U, extracting, and comparing Co-PI networks.

Scientific Field Level

D. Extracting co-author networks, patent-citation networks, and detecting bursts in SDB data.

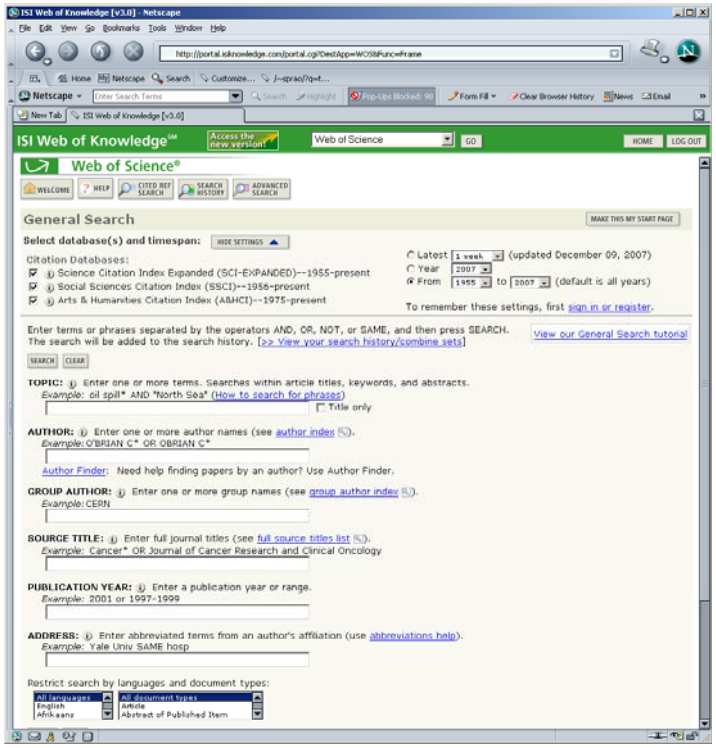
Data Acquisition from Web of Science

Download all papers by

- Eugene Garfield
- Stanley Wasserman
- Alessandro Vespignani
- Albert-László Barabási

from

- Science Citation Index Expanded (SCI-EXPANDED) --1955-present
- Social Sciences Citation Index (SSCI)--1956-present
- Arts & Humanities Citation Index (A&HCI)--1975-present



Comparison of Counts

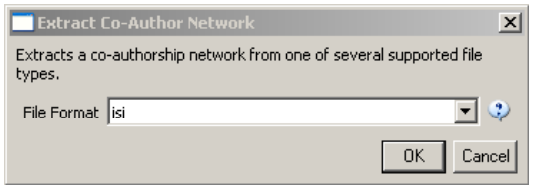
No books and other non-WoS publications are covered.

	Age	Total # Cites	Total # Papers	H-Index	
Eugene Garfield	82	1,525	672	31	
Stanley Wasserman		122	35	17	
Alessandro Vespignani	42	451	101	33	
Albert-László Barabási	40	2,218	126	47	*(Dec 2007)*
	41	16,920	159	52	*(Dec 2008)*

Extract Co-Author Network

Load **yournwbdirectory*/sampledata/scientometrics/isi/FourNetSciResearchers.isi'* using *'File > Load and Clean ISI File'*.

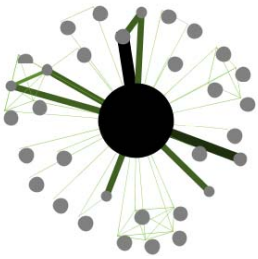
To extract the co-author network, select the *'361 Unique ISI Records'* table and run *'Scientometrics > Extract Co-Author Network'* using isi file format:



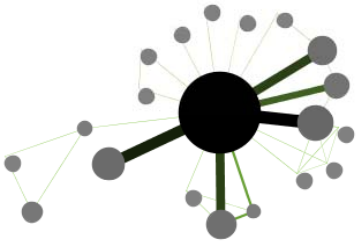
The result is an undirected network of co-authors in the Data Manager. It has 247 nodes and 891 edges.

To view the complete network, select the network and run *'Visualization > GUESS > GEM'*. Run *Script > Run Script… . And select Script folder > GUESS > co-author-nw.py*.

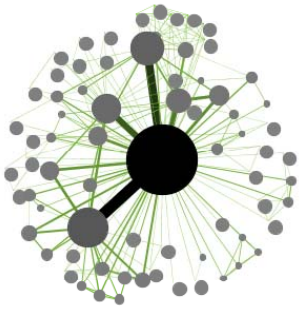
Comparison of Co-Author Networks



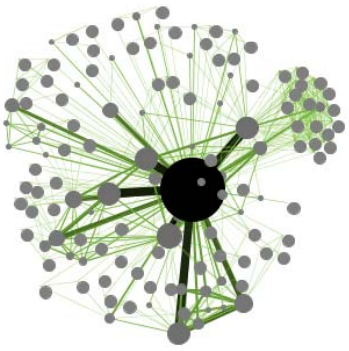
Eugene Garfield



Stanley Wasserman

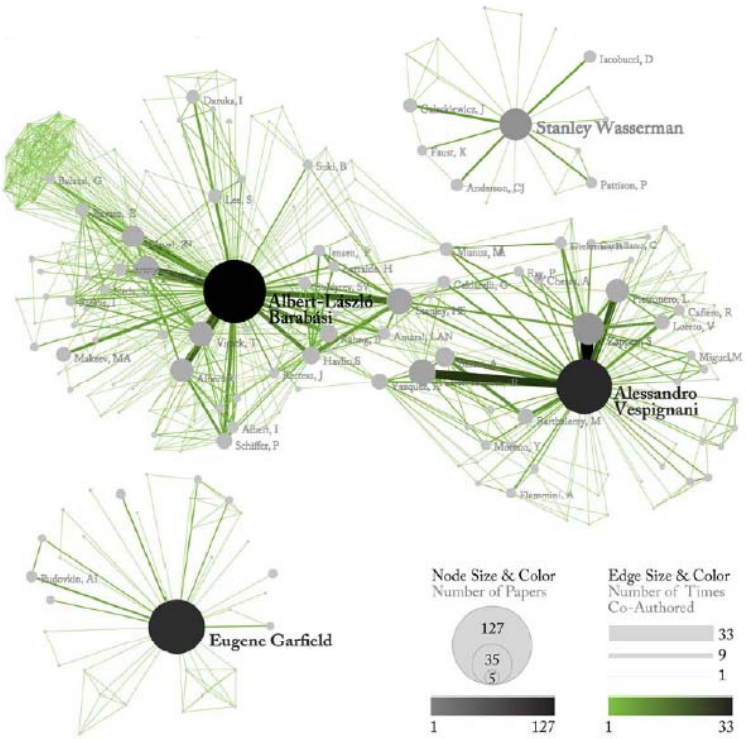


Alessandro Vespignani



Albert-László Barabási

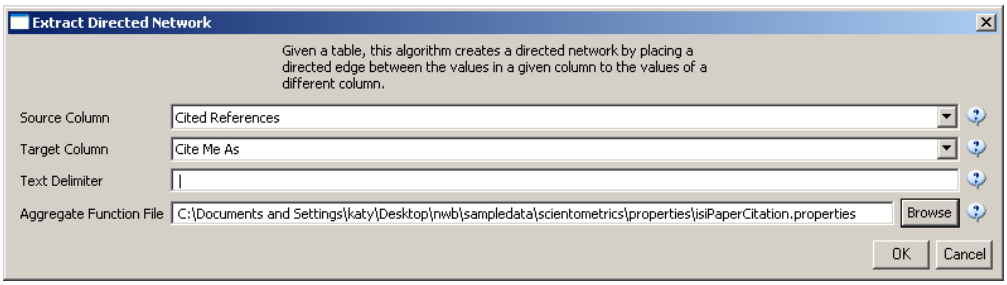
Joint Co-Author Network of all Four NetsSci Researchers



Paper-Citation Network Layout

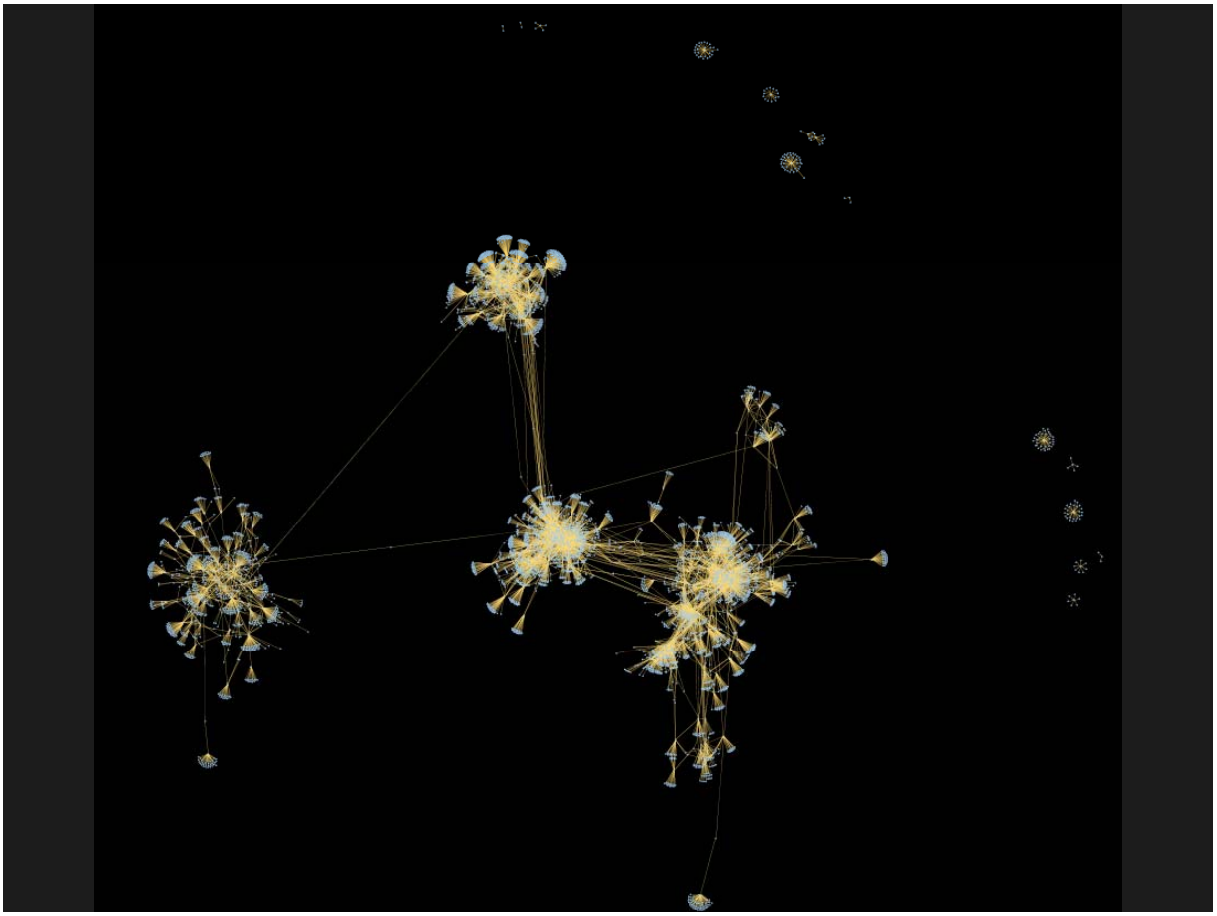
Load *'*yournwbdirectory*/sampledata/scientometrics/isi/FourNetSciResearchers.isi'* using *'File > Load and Clean ISI File'*.

To extract the paper-citation network, select the *'361 Unique ISI Records'* table and run *'Scientometrics > Extract Directed Network'* using the parameters:



The result is a directed network of paper citations in the Data Manager. It has 5,335 nodes and 9,595 edges.

To view the complete network, select the network and run *'Visualization > GUESS'*. Run *'Script > Run Script …' and select 'yournwbdirectory*/script/GUESS/paper-citation-nw.py'*.



5. Exemplary Analyses and Visualizations

Individual Level

A. Loading ISI files of major network science researchers, extracting, analyzing and visualizing paper-citation networks and co-author networks.

B. Loading NSF datasets with currently active NSF funding for 3 researchers at Indiana U

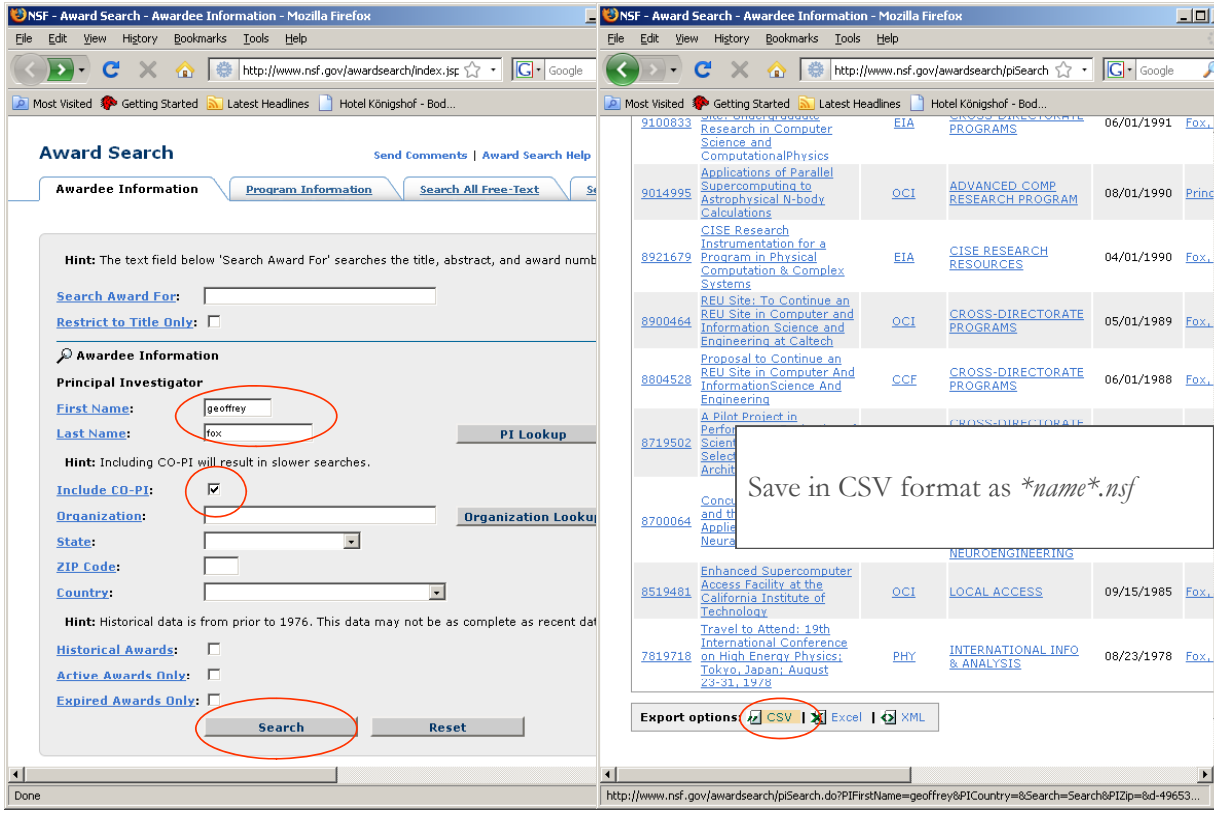
Institution Level

C. Indiana U, Cornell U, and Michigan U, extracting, and comparing Co-PI networks.

Scientific Field Level

D. Extracting co-author networks, patent-citation networks, and detecting bursts in SDB data.

NSF Awards Search via http://www.nsf.gov/awardsearch



NSF Awards Search Results

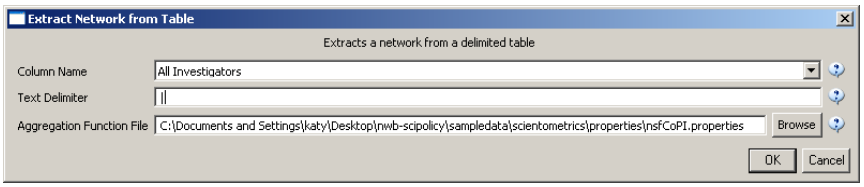
Name	# Awards	First A. Starts	Total Amount to Date
Geoffrey Fox	27	Aug 1978	12,196,260
Michael McRobbie	8	July 1997	19,611,178
Beth Plale	10	Aug 2005	7,224,522

Disclaimer:

Only NSF funding, no funding in which they were senior personnel, only as good as NSF's internal record keeping and unique person ID. If there are 'collaborative' awards then only their portion of the project (award) will be included.

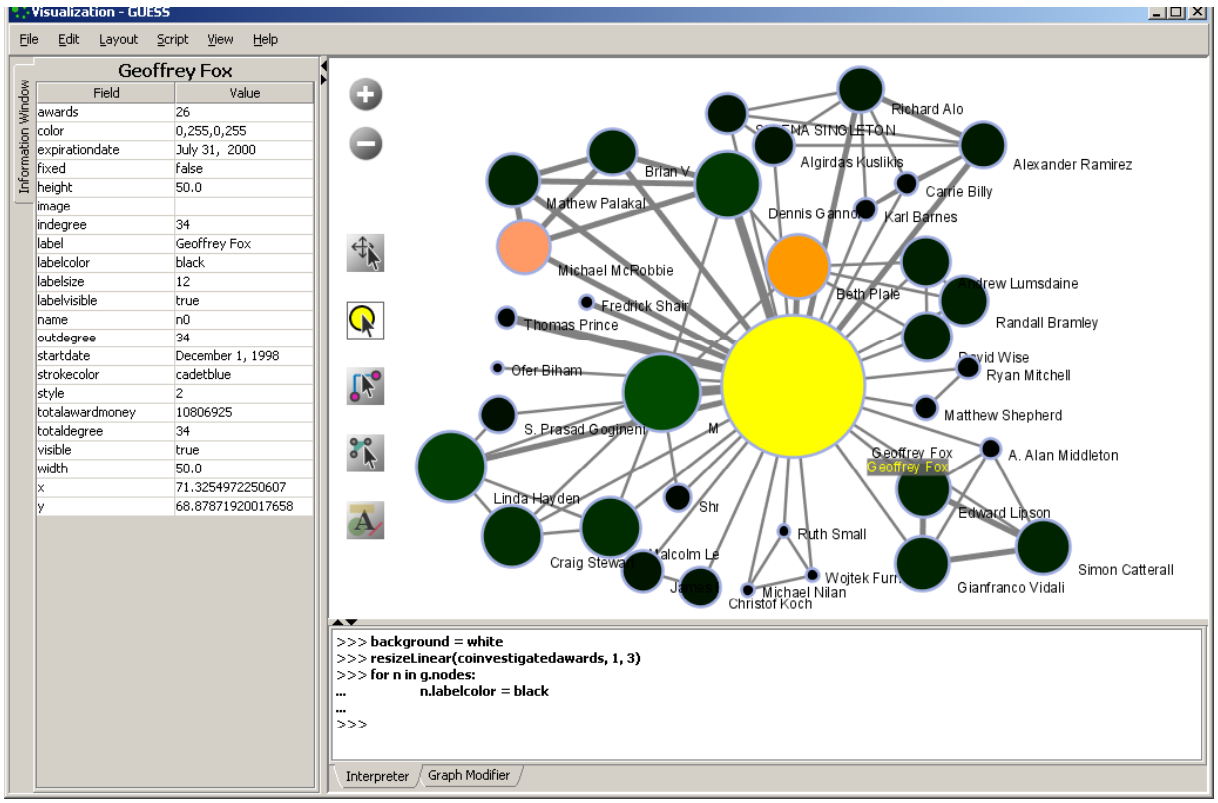
Using NWB to Extract Co-PI Networks

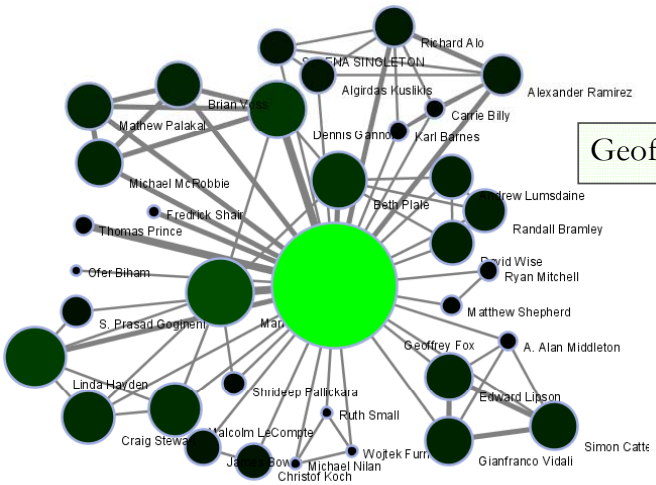
- Load into NWB, open file to count records, compute total award amount.
- Run '*Scientometrics* > *Extract Co-Occurrence Network*' using parameters:



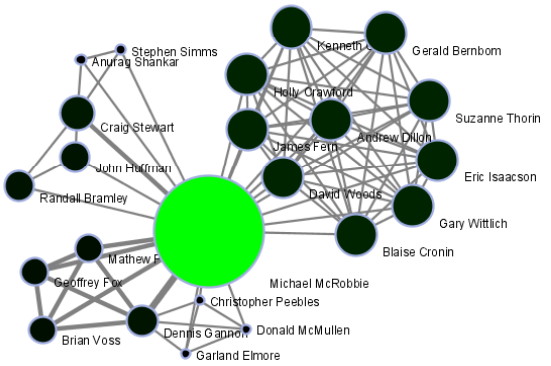
- Select *"Extracted Network .."* and run *'Analysis* > *Network Analysis Toolkit (NAT)'*
- Remove unconnected nodes via *'Preprocessing* > *Delete Isolates'*.
- *'Visualization* > *GUESS'*, layout with GEM
- Run 'co-PI-nw.py' GUESS script to color/size code.

Totalawardmoney is the total of all funding associated with an investigator.

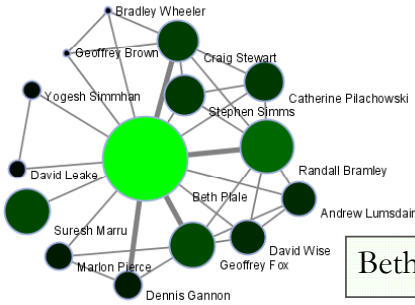




Geoffrey Fox



Michael McRobbie



Beth Plale

Geoffrey Fox

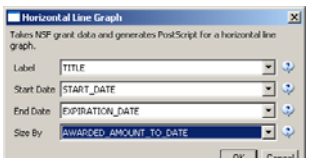
Last Expiration date

July 10

Michael McRobbie

Feb 10

Beth Plale



Sept 09

5. Exemplary Analyses and Visualizations

Individual Level

A. Loading ISI files of major network science researchers, extracting, analyzing and visualizing paper-citation networks and co-author networks.

B. Loading NSF datasets with currently active NSF funding for 3 researchers at Indiana U

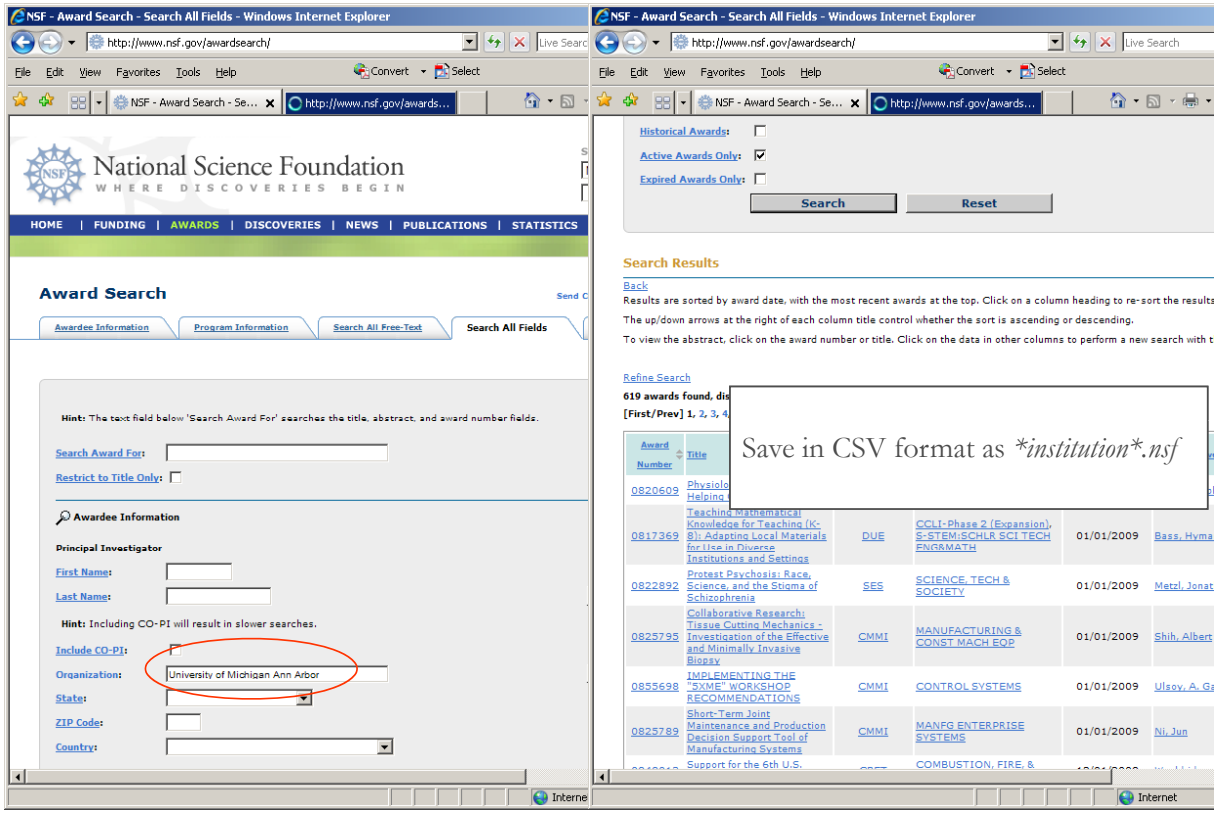
Institution Level

C. Indiana U, Cornell U, and Michigan U, extracting, and comparing Co-PI networks.

Scientific Field Level

D. Extracting co-author networks, patent-citation networks, and detecting bursts in SDB data.

NSF Awards Search via http://www.nsf.gov/awardsearch



Active NSF Awards on 11/07/2008:

- Indiana University 257

  (there is also Indiana University at South Bend Indiana University Foundation, Indiana University Northwest, Indiana University-Purdue University at Fort Wayne, Indiana University-Purdue University at Indianapolis, Indiana University-Purdue University School of Medicine)

- Cornell University 501

  (there is also Cornell University – State, Joan and Sanford I. Weill Medical College of Cornell University)

- University of Michigan Ann Arbor 619

  (there is also University of Michigan Central Office, University of Michigan Dearborn, University of Michigan Flint, University of Michigan Medical School)

Save files as csv but rename into .nsf.

Or simply use the files saved in *'*yournwbdirectory*/sampledata/scientometrics/nsf/'*.

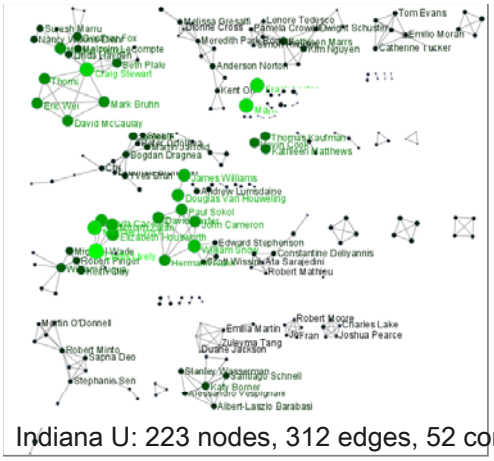
Extracting Co-PI Networks

Load NSF data, selecting the loaded dataset in the Data Manager window, run *'Scientometrics > Extract Co-Occurrence Network'* using parameters:

Two derived files will appear in the Data Manager window: the co-PI network and a merge table. In the network, nodes represent investigators and edges denote their co-PI relationships. The merge table can be used to further clean PI names.

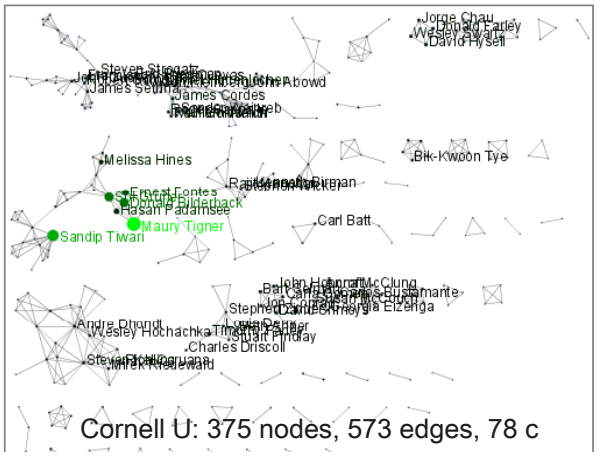
Running the *'Analysis > Network Analysis Toolkit (NAT)'* reveals that the number of nodes and edges but also of isolate nodes that can be removed running *'Preprocessing > Delete Isolates'*.

Select *'Visualization > GUESS'* to visualize. Run 'co-PI-nw.py' script.



Indiana U: 223 nodes, 312 edges, 52 components

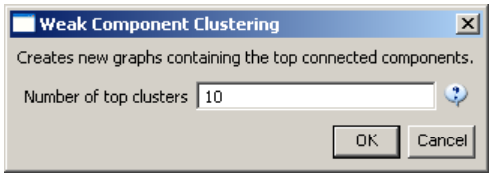
U of Michigan: 497 nodes, 672 edges, 117 c



Cornell U: 375 nodes, 573 edges, 78 c

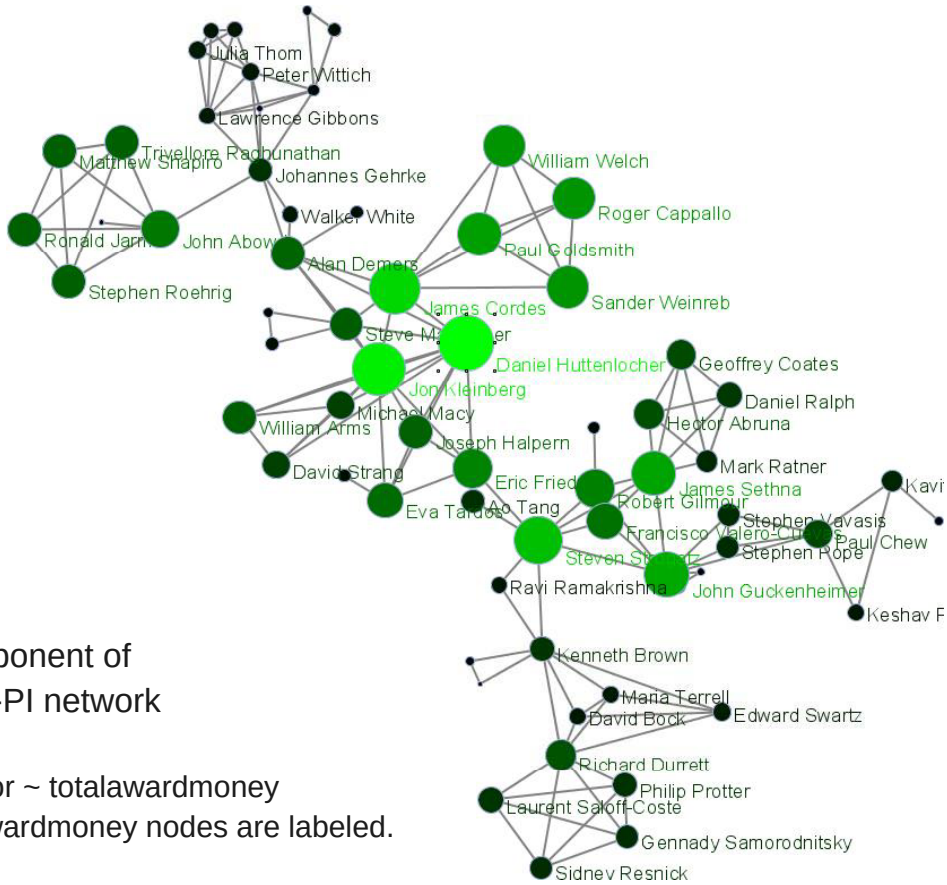
Extract Giant Component

Select network after removing isolates and run *'Analysis > Unweighted and Undirected > Weak Component Clustering' with parameter*



Indiana's largest component has 19 nodes, Cornell's has 67 nodes, Michigan's has 55 nodes.

Visualize Cornell network in GUESS using same .py script and save via *'File > Export Image'* as jpg.



Largest component of
Cornell U co-PI network

Node size/color ~ totalawardmoney
Top-50 totalawardmoney nodes are labeled.

Top-10 Investigators by Total Award Money

```
for i in range(0, 10):
    print str(nodesbytotalawardmoney[i].label) + ": " +
    str(nodesbytotalawardmoney[i].totalawardmoney)
```

Totalawardmoney is the total of all funding associated with an investigator.
Funding retrieved for one university might have investigators from other universities.

Indiana University		**Cornell University**		**Michigan University**	
Curtis Lively:	7,436,828	Maury Tigner:	107,216,976	Khalil Najafi:	32,541,158
Frank Lester:	6,402,330	Sandip Tiwari:	72,094,578	Kensall Wise:	32,164,404
Maynard Thompson:	6,402,330	Sol Gruner:	48,469,991	Jacquelynne Eccles:	25,890,711
Michael Lynch:	6,361,796	Donald Bilderback:	47,360,053	Georg Raithel:	23,832,421
Craig Stewart:	6,216,352	Ernest Fontes:	29,380,053	Roseanne Sension:	23,812,921
William Snow:	5,434,796	Hasan Padamsee:	18,292,000	Theodore Norris:	23,35,0921
Douglas V. Houweling:	5,068,122	Melissa Hines:	13,099,545	Paul Berman:	23,350,921
James Williams:	5,068,122	Daniel Huttenlocher:	7,614,326	Roberto Merlin:	23,350,921
Miriam Zolan:	5,000,627	Timothy Fahey:	7,223,112	Robert Schoeni:	21,991,140
Carla Caceres:	5,000,627	Jon Kleinberg:	7,165,507	Wei-Jun Jean Yeung:	21,991,140

5. Exemplary Analyses and Visualizations

Individual Level

A. Loading ISI files of major network science researchers, extracting, analyzing and visualizing paper-citation networks and co-author networks.

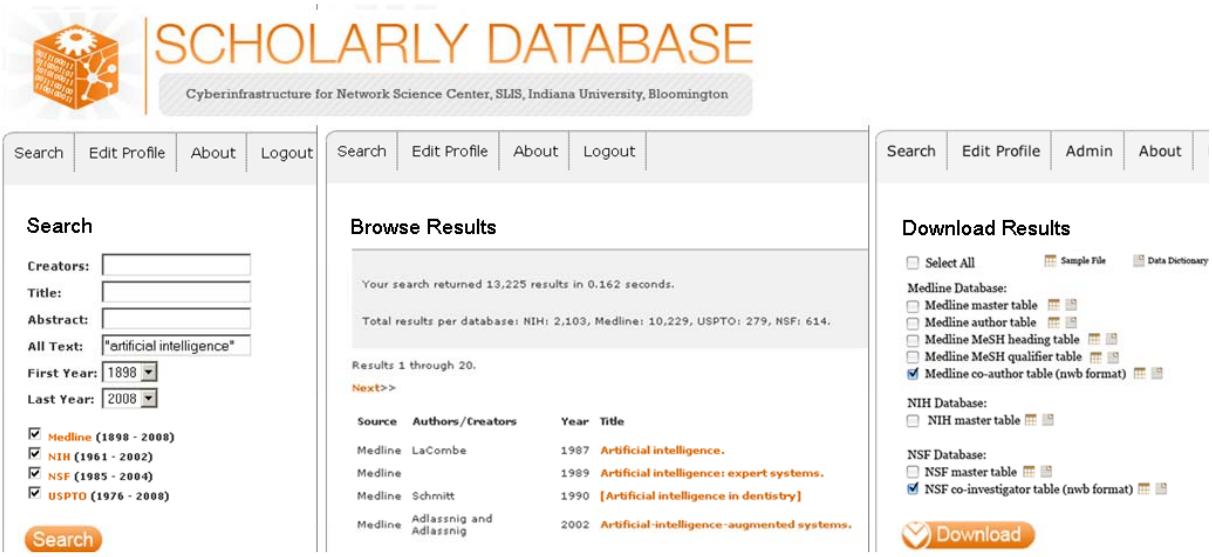
B. Loading NSF datasets with currently active NSF funding for 3 researchers at Indiana U

Institution Level

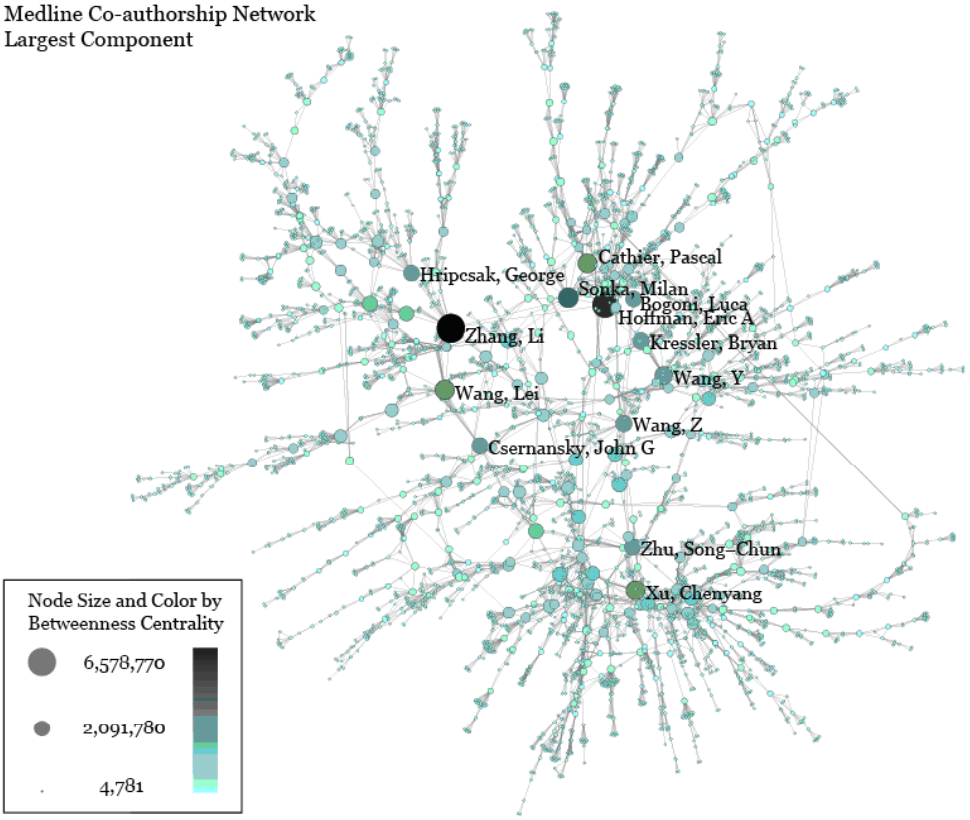
C. Indiana U, Cornell U, and Michigan U, extracting, and comparing Co-PI networks.

Scientific Field Level

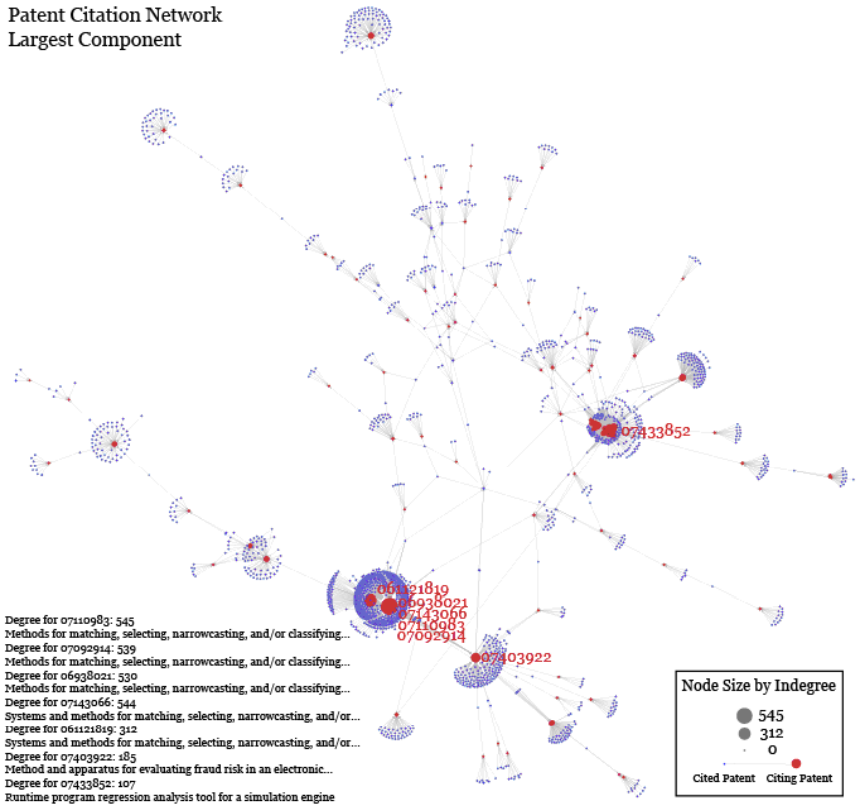
D. Extracting co-author networks, patent-citation networks, and detecting bursts in SDB data.



Medline Co-authorship Network
Largest Component



Patent Citation Network
Largest Component



Top-10 burst terms from abstracts of the AI search results.

Medline				
Word	**Length**	**Weight**	**Start**	**End**
medical	17	299.7924	1983	1999
knowledge	5	293.9375	1991	1995
knowledge	6	215.2407	1997	2002
expert	13	171.0443	1985	1997
systems	15	170.3306	1985	1999
intelligence	21	123.9794	1981	2001
patient	21	123.9297	1982	2002
care	12	106.5522	1990	2001
registration	5	104.8139	2005	
knowledge-based	16	98.83778	1987	2002

NIH				
Word	**Length**	**Weight**	**Start**	**End**
Phase	8	117.2205	1993	2000
commercial	9	87.57158	1995	
proposed	9	87.57158	1995	
mass	3	83.36952	1978	1980
protein	1	72.15788	1988	1988
networks	4	71.252	1993	1996
patterns	3	66.44826	1977	1979
being	8	66.29254	1971	1978
reasoning	2	65.68178	1984	1985
expert	4	60.49935	1987	1990

NSF				
Word	**Length**	**Weight**	**Start**	**End**
their	6	47.05097	1999	
gray	2	28.19808	2000	2001
learning	2	27.40728	1997	1998
human	5	25.4525	2000	
control	2	24.07877	1992	1993
knowledge	1	21.48756	1998	1998
students	1	21.07674	1997	1997
problems	2	20.77133	1998	1999
more	2	19.96109	2000	2001
use	1	19.38503	2001	2001

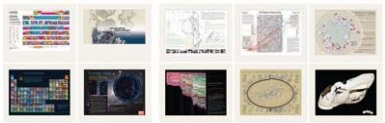
USPTO				
Word	**Length**	**Weight**	**Start**	**End**
human	3	19.03937321	2004	2006
video	3	15.32736425	1998	2000
disclosed	2	14.06694671	1999	2000
neural	3	13.30105906	2004	2006
"correct"	2	12.4336047	1999	2000
unit	2	12.35745838	2002	2003
material	1	12.08487035	2000	2000
feedback	1	12.07730195	2000	2000
rule	1	12.07730195	2000	2000
elevator	4	11.83351857	1991	1994

5. Mapping Science Exhibit – 10 Iterations in 10 years

http://scimaps.org/

The Power of Maps (2005)

The Power of Reference Systems (2006)



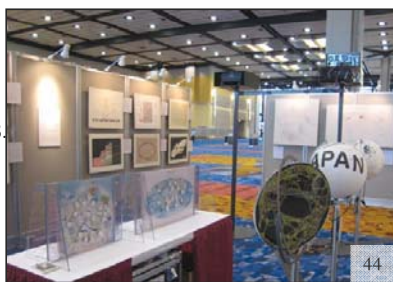
The Power of Forecasts (2007)

Science Maps for Economic Decision Makers (2008)

Science Maps for Science Policy Makers (2009)
Science Maps for Scholars (2010)
Science Maps as Visual Interfaces to Digital Libraries (2011)
Science Maps for Kids (2012)
Science Forecasts (2013)

How to Lie with Science Maps (2014)

Exhibit has been shown in 49 venues on four continents. Also at
- NSF, 10th Floor, 4201 Wilson Boulevard, Arlington, VA, permanent display.
- National Science Library of the Chinese Academy of Sciences, Beijing, China, 2008.
- University of Alberta, Edmonton, Alberta, Canada, Nov 10-Feb 31, 2009.
- The Institute for Research Information and Quality Assurance, Bonn, Germany, permanent display.



Places & Spaces: Mapping Science

a science exhibit that introduces people to maps of sciences, their makers and users.

http://scimaps.org.

Exhibit Curators: Dr. Katy Börner & Elisha F. Hardy

Illuminated Diagram Display

W. Bradford Paley, Kevin W. Boyack, Richard Kalvans, and Katy Börner (2007) Mapping, Illuminating, and Interacting with Science. SIGGRAPH 2007.

Large-scale, high resolution prints illuminated via projector or screen.

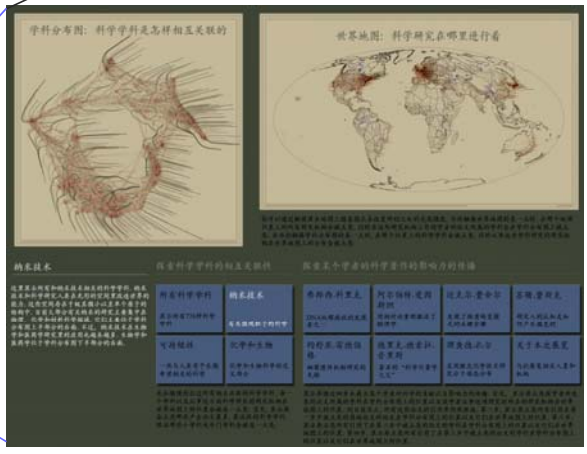
Questions:

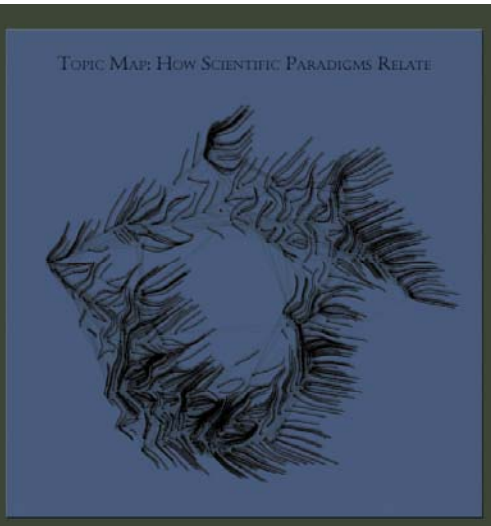
- Who is doing research on what topic and where?
- What is the 'footprint' of interdisciplinary research fields?
- What impact have scientists?

Interactive touch panel.

Contributions:

- Interactive, high resolution interface to access and make sense of data about scholarly activity.





You may run your finger over each of these maps to control the lighting on the other: touching a place on the world map will light up topics studied in that place; touching a paradigm on the topic map will light up the places that study that topic.

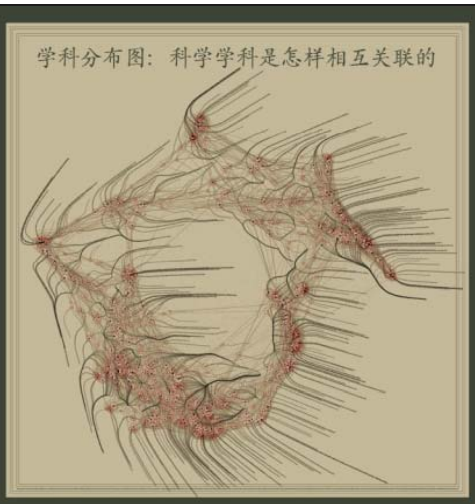
Nanotechnology

This overlay shows the distribution of nanotechnology within the paradigms of science. The majority of current work in nanotechnology takes places in physics, chemistry, and materials science, at the upper right portion of the map. However, an increasing amount of nanotechnology is being applied in the biological and medical sciences, at the lower right.

All Topics *Sweep through all 776 scientific paradigms*	**Nanotechnology** *Science on the tiny scale of molecules*	**Francis H. C. CRICK** *Co-discovered DNA's double helix*	**Albert EINSTEIN** *Revitalized physics with Relativity theories*	**Michael E. FISHER** *Models critical phase transitions of matter*	**Susan T. FISKE** *Connects perception and stereotypes*
Sustainability *The science behind our long-term hopes*	**Biology & Chemistry** *The interface between these two vital fields*	**Joshua LEDERBERG** *Pioneer in bacterial genetic mechanisms*	**Derek J. de Solla PRICE** *Known as the "Father of Scientometrics"*	**Richard N. ZARE** *Uses laser chemistry in molecular dynamics*	**About this display** *People & organizations that helped create it*

We sweep slowly through adjoining related topics, lighting up the places in the world that study each topic. You may select a subset of the topics that deal with these three interesting subjects by touching it.

A single person's spreading influence is shown as a series of four snapshots. First, we light only topics and places relating to that person's papers—papers that are still highly cited today. The second lights everything that cites that original work. Note that this first-generation impact extends to far more topics than did the original work. The third shapshot lights science that cites the second; and the fourth lights science that cites the third.



你可以通过触摸屏在地图上随意指点来改变所到之处的光亮强度。当你触摸世界地图的某一点时，在那个地理位置上的所有研究机构会被点亮。同时在这些研究机构工作的学者的论文所属的学科会在学科分布图上被点亮。而当你触摸学科分布图的某一点时，在那个位置上的科学学科会被点亮。同时从事这些学科研究的研究机构在世界地图上的分布会被点亮。

纳米技术

这里显示所有和纳米技术相关的科学学科。纳米技术和科学研究人类在无形的空间里改造世界的能力。这些空间寿在于极其微小以至单个原子的结构中。目前大部分有关纳米的研究主要集中在物理、化学和材料科学领域，它们主要位于学科分布图上半部分的右面。不过，纳米技术在生物学和医药学研究里的应用也越来越多。生物学和医药学位于学科分布图下半部分的右面。

探索科学学科的相互关联性

所有科学学科 显示所有776种科学学科	**纳米技术** 有关微观粒子的科学
可持续性 一些与人类寄予长期希望相关的科学	**化学和生物** 化学和生物科学的交叉部分

光柱缓慢的扫过所有相互关联的科学学科，每一个学科以及从事这方面科学研究的研究机构在世界地图上的位置会被逐一点亮。首先，显示屏会点亮那些产出论文最多、最活跃的科学学科，然后那些小学科或冷门学科会被逐一点亮。

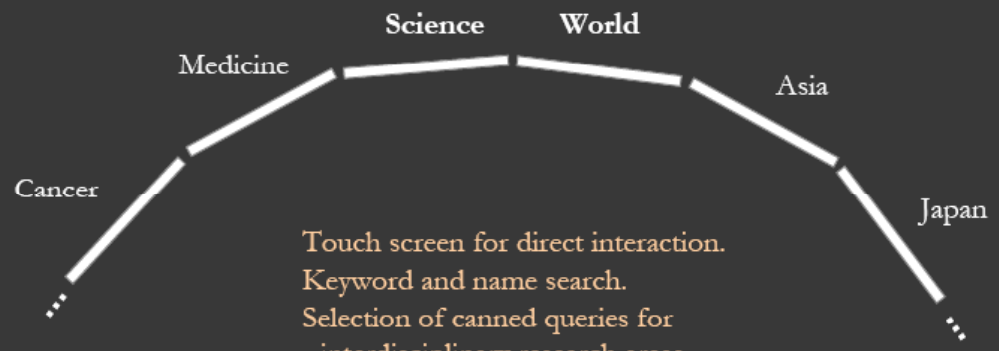
探索某个学者的科学著作的影响力的传播

弗郎西.科里克 DNA双螺旋状的发现者之一	**阿尔伯特.爱因斯坦** 用相对论重新激活了物理学	**迈克尔.费舍尔** 发现了物质转变模式的关键步骤	**苏珊.费斯克** 研究人的认知是如何产生偏见的
约舒亚.雷德伯格 细菌遗传机制研究的先驱	**德里克.德索拉.普里斯** 著名的“科学计量学之父”	**理查德.扎尔** 采用激光化学技术研究分子动态分布	**关于本次展览** 与此展览相关人员和机构

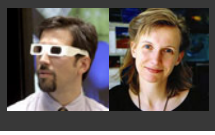
显示屏通过四步来展示某个学者对科学的贡献以及影响力的传播。首先，显示屏点亮该学者所发表的论文所属的学科在学科分布图上的位置以及该学者从事这项研究时所在的研究机构在世界地图上的位置。到目前为止，所有这些论文的引用率仍然很高。第二步，显示屏点亮所有引用在第一步中被点亮的原始论文的论文在学科分布图上的位置以及它们在世界地图上的位置。第三步，显示屏点亮所有引用了在第二步中被点亮的论文的学科在学科分布图上的位置以及它们在世界地图上的位置。第四步，显示屏点亮所有引用了在第三步中被点亮的论文的学科在学科分布图上的位置以及它们在世界地图上的位置。

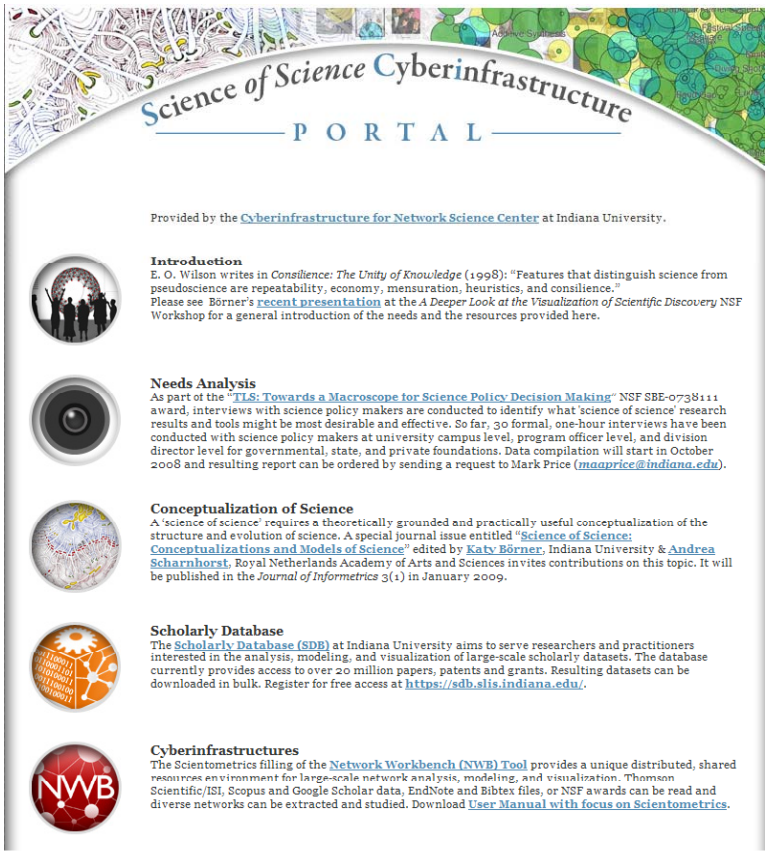
Re-implementation of Illuminated Diagram Software
by Advanced Visualization Lab, Indiana University

Drives unlimited number of ID screens.

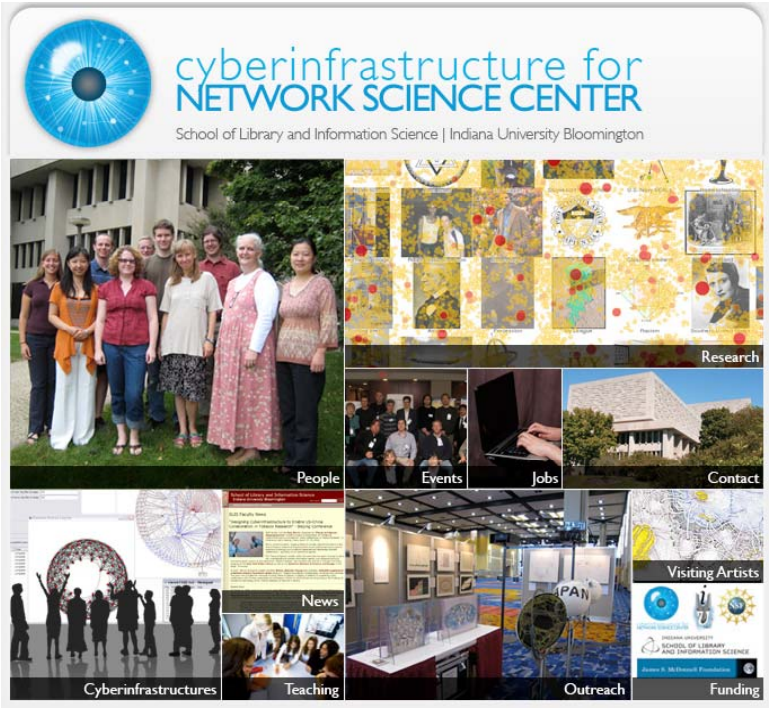


Touch screen for direct interaction.
Keyword and name search.
Selection of canned queries for
- interdisciplinary research areas
- famous people
- activity patterns, e.g., bursts, trends, etc.





http://sci.slis.indiana.edu



All papers, maps, cyberinfrastructures, talks, press are linked from http://cns.slis.indiana.edu