Towards a Macroscope for Science Policy Decision Making

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

Kevin Boyack
SciTech Strategies Inc.
kboyack@mapofscience.com

With special thanks to the members at the Cyberinfrastructure for Network Science Center, Mapping Science exhibit map makers and advisory board members, and the VIVO team.

October 19, 2010
AAAS

Computational Scientometrics
Cyberinfrastructures

Scholarly Database: 25 million scholarly records
http://sdb.slis.indiana.edu

VIVO Research Networking
http://vivoweb.org

Information Visualization Cyberinfrastructure
http://iv.slis.indiana.edu

Network Workbench Tool & Community Wiki
http://nwb.slis.indiana.edu

Science of Science (Sci²) Tool and CI Portal
http://sci.slis.indiana.edu

Epidemics Cyberinfrastructure
http://epic.slis.indiana.edu/
Sci² Tool – “Open Code for S&T Assessment”

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

Sci² Tool

Sci² Tool: Algorithms
See https://nwb.slis.indiana.edu/community

Preprocessing
Extract Top N% Records
Extract Top N Records
Normalize Text
Slice Table by Line
---
Extract Top Nodes
Extract Nodes Above or Below Value
Delete Isolates
---
Extract top Edges
Extract Edges Above or Below Value
Remove Self Loops
Trim by Degree
MST-Pathfinder Network Scaling
Fast Pathfinder Network Scaling
---
Snowball Sampling (in nodes)
Node Sampling
Edge Sampling
---
Symmetrize
Dichotomize
Multipartite Joining
---
GeoCoder
---
Extract ZIP Code

Modeling
Random Graph
Watts-Strogatz
Small World
Barabási-Albert Scale-Free
TARL

Analysis
Network Analysis Toolkit (NAT)
---
Unweighted & Undirected
Node Degree
Degree Distribution
---
K-Nearest Neighbor (Java)
Watts-Strogatz Clustering Coefficient
Watts Strogatz Clustering Coefficient over K
Diameter
Average Shortest Path
Shortest Path Distribution
Node Betweenness Centrality
Weak Component Clustering
Global Connected Components
---
Extract K-Cores
Annotate K-Coreness
HITS

Weighted & Undirected
Clustering Coefficient
Nearest Neighbor Degree
Strength vs Degree
Degree & Strength
Average Weight vs End-point Degree
Strength Distribution
Weight Distribution
Randomize Weights
---
Blondel Community Detection
HITS
---
Unweighted & Directed
Node Indegree
Node Outdegree
Indegree Distribution
Outdegree Distribution
---
K-Nearest Neighbor
Single Node in-Out Degree Correlations
---
Dyad Reciprocity
Arc Reciprocity
Adjacency Transitivity
---
Weak Component Clustering
Strong Component Clustering
See https://nwb.slis.indiana.edu/community

----------------------------------
Extract K-Core
Annotate K-Coreness

----------------------------------
HITS
PageRank
Weighted & Directed
HITS
Weighted PageRank

Textual
Burst Detection

Visualization
GnuPlot
GUESS
Image Viewer

_________________________________
Radial Tree/Graph (prefuse alpha)
Radial Tree/Graph with Annotation
(prefuse beta)
Tree View (prefuse beta)
Tree Map (prefuse beta)
Force Directed with Annotation
(prefuse beta)
Fruchterman-Reingold with Annotation
(prefuse beta)

DrL (VxOrd)
Specified (prefuse beta)

----------------------------------
Horizontal Line Graph
Circular Hierarchy
Geo Map (Circle Annotation Style)
Geo Map (Colored-Region Annotation Style)
Science Map (Circle Annotation)

Scientometrics
Remove ISI Duplicate Records
Remove Rows with Multitudinous Fields
Detect Duplicate Nodes
Update Network by Merging Nodes

----------------------------------
Extract Directed Network
Extract Paper Citation Network
Extract Author Paper Network

----------------------------------
Extract Co-Occurrence Network
Extract Word Co-Occurrence Network
Extract Co-Author Network
Extract Reference Co-Occurrence
(Bibliographic Coupling) Network

----------------------------------
Extract Document Co-Citation Network

NEW:
Database support for ISI and NSF data.

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.
Zass & Börner, forthcoming.
Interactive Science Map of NIH Funding


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.
Where Are the Academic Jobs? Interactive Exploration of Job Advertisements in Geospatial and Topical Space

Angela Zoss, Michael Connover, Katy Börner (2010)

http://cns-nd3.slis.indiana.edu/mapjobs/geo

Type of Analysis vs. Level of Analysis

<table>
<thead>
<tr>
<th>Level of Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Micro/Individual (1-100 records)</td>
</tr>
<tr>
<td>Meso/Local (101-10,000 records)</td>
</tr>
<tr>
<td>Macro/Global (10,000 &lt; records)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Analysis/Profiling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistical Analysis/Profiling</td>
</tr>
<tr>
<td>Temporal Analysis (When)</td>
</tr>
<tr>
<td>Geospatial Analysis (Where)</td>
</tr>
<tr>
<td>Topical Analysis (What)</td>
</tr>
<tr>
<td>Network Analysis (With Whom?)</td>
</tr>
</tbody>
</table>

- Individual person and their expertise profiles
- Funding portfolio of one individual
- Career trajectory of one individual
- NIH's core competency
- NSF Co-PI network of one individual

- Larger labs, centers, universities, research domains or states
- Mapping topic bursts in 20-years of PNAS
- Mapping intellectual landscape of PNAS publications
- Knowledge flows in Chemistry research
- NIH funding

- All of NSF, all of USA, all of science
- 113 Years of Physics Research
- VxOrd/Topic maps of NIH funding
- NIH's core competency

Copyright © 2008 The Regents of the University of California | General Use
Computational Scientometrics
Cyberinfrastructures

Scholarly Database: 25 million scholarly records
http://sdb.slis.indiana.edu

VIVO Research Networking
http://vivoweb.org

Information Visualization Cyberinfrastructure
http://iv.slis.indiana.edu

Network Workbench Tool & Community Wiki
http://nwb.slis.indiana.edu

Science of Science (Sci²) Tool and CI Portal
http://sci.slis.indiana.edu

Epidemics Cyberinfrastructure
http://epic.slis.indiana.edu/

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

Register for free access at http://sdb.slis.indiana.edu
Since March 2009:
Users can download networks:
- Co-author
- Co-investigator
- Co-inventor
- Patent citation
and tables for
burst analysis in NWB.
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.

Soon:
- Simplify reporting tasks, e.g., generate biosketch, department report.


Indiana University: Katy Borner (IU PI), William Barnett, Shamsan Chen, Ying Ding, Russell Duhon, Jon Dunn, Micah Linneemeier, Nianli Ma, Robert McDonald, Barbara Ann O'Leary, Mark Prince, 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, Michael 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 Klar, Peter Michelini, Grace Migliorini, John Ruffing, Jason Specland, Tru Tran, Jesse Turner, Vinay Varughese.
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)

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)
Publications per year (CSV File), see top file.
Co-authors (CSV File)

**Computational Scientometrics**

**References**


Katy Börner, Network Workbench: A CI-Marketplace for Network Scientists

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