The Science of Science (Sci2) Tool and Its Utility for Research

Katy Börner and the Cyberinfrastructure for Network Science Center
School of Library and Information Science
Indiana University, Bloomington, IN
katy@indiana.edu

With special thanks to Kevin W. Boyack, Micah Linnemeier, Russell J. Duhon, Patrick Phillips, Joseph Biberstine, Chintan Tank Nianli Ma, Angela M. Zoss, Hanning Guo, Mark A. Price, Scott Weingart

Networks and Complex Systems Talk Series, IUB
December 7th, 2009

Overview

1. Plug-and-Play Macroscopes

2. Sci2 Tool Introduction and Demos

3. Sample Science of Science Studies
The Changing Scientific Landscape

**Star Scientist -> Research Teams:** In former times, science was driven by key scientists. Today, science is driven by effectively collaborating co-author teams often comprising expertise from multiple disciplines and several geospatial locations (Börner, Dall’Asta, Ke, & Vespignani, 2005; Shneiderman, 2008).

**Users -> Contributors:** Web 2.0 technologies empower anybody to contribute to Wikipedia and to exchange images and videos via Fickr and YouTube. WikiSpecies, WikiProfessionals, or WikiProteins combine wiki and semantic technology in support of real time community annotation of scientific datasets (Mons et al., 2008).

**Cross-disciplinary:** The best tools frequently borrow and synergistically combine methods and techniques from different disciplines of science and empower interdisciplinary and/or international teams of researchers, practitioners, or educators to fine-tune and interpret results collectively.

**One Specimen -> Data Streams:** Microscopes and telescopes were originally used to study one specimen at a time. Today, many researchers must make sense of massive streams of multiple types of data with different formats, dynamics, and origin.

**Static Instrument -> Evolving Cyberinfrastructure (CI):** The importance of hardware instruments that are rather static and expensive decreases relative to software infrastructures that are highly flexible and continuously evolving according to the needs of different sciences. Some of the most successful services and tools are decentralized increasing scalability and fault tolerance.

**Modularity:** The design of software modules with well defined functionality that can be flexibly combined helps reduce costs, makes it possible to have many contribute, and increases flexibility in tool development, augmentation, and customization.

**Standardization:** Adoption of standards speeds up development as existing code can be leveraged. It helps pool resources, supports interoperability, but also eases the migration from research code to production code and hence the transfer of research results into industry applications and products.

**Open data and open code:** Lets anybody check, improve, or repurpose code and eases the replication of scientific studies.

Microscopes, Telescopes, and Macroscopes

Just as the **microscope** empowered our naked eyes to see cells, microbes, and viruses thereby advancing the progress of biology and medicine or the **telescope** opened our minds to the immensity of the cosmos and has prepared mankind for the conquest of space, **macroscopes** promise to help us cope with another infinite: the infinitely complex. Macroscopes give us a 'vision of the whole' and help us 'synthesize'. They let us detect patterns, trends, outliers, and access details in the landscape of science. Instead of making things larger or smaller, macroscopes let us observe what is at once too great, too slow, or too complex for our eyes.
Desirable Features of Plug-and-Play Macrosopes

Division of Labor: Ideally, labor is divided in a way that the expertise and skills of computer scientists are utilized for the design of standardized, modular, easy to maintain and extend “core architecture”. Dataset and algorithm plugins, i.e., the “filling”, are initially provided by those that care and know most about the data and developed the algorithms: the domain experts.

Ease of Use: As most plugin contributions and usage will come from non-computer scientists it must be possible to contribute, share, and use new plugins without writing one line of code. Wizard-driven integration of new algorithms and data sets by domain experts, sharing via email or online sites, deploying plugins by adding them to the ‘plugin’ directory, and running them via a Menu driven user interfaces (as used in Word processing systems or Web browsers) seems to work well.

Plugin Content and Interfaces: Should a plugin represent one algorithm or an entire tool? What about data converters needed to make the output of one algorithm compatible with the input of the next? Should those be part of the algorithm plugin or should they be packaged separately?

Supported (Central) Data Models: Some tools use a central data model to which all algorithms conform, e.g., Cytoscape, see Related Work section. Other tools support many internal data models and provide an extensive set of data converters, e.g., Network Workbench, see below. The former often speeds up execution and visual rendering while the latter eases the integration of new algorithms. In addition, most tools support an extensive set of input and output formats.

Core vs. Plugins: As will be shown, the “core architecture” and the “plugin filling” can be implemented as sets of plugin bundles. Answers to questions such as: “Should the graphical user interface (GUI), interface menu, scheduler, or data manager be part of the core or its filling?” will depend on the type of tools and services to be delivered.

Supported Platforms: If the software is to be used via Web interfaces then Web services need to be implemented. If a majority of domain experts prefers a stand-alone tool running on a specific operating system then a different deployment is necessary.

Hubble telescope has cost a total of about $10 billion over its 20 years of operation. Ed Weiler, HST space science chief, computed that this "...equates to about two cents per week per American taxpayer over 20 years" – a true bargain.
Illuminated Diagram Display


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.

A single person’s spreading influence is shown as a series of four snapshots. First, we light up topics and places relating to that person’s papers—papers that are still highly cited today. The second light everything that cites that original work; note that this first-generation impact extends to far more topics than did the original work. The third snapshot lights scientific work that cites the second, and the fourth lights science that cites the third.
图例：科学学科是怎样相互关联的

探索科学学科的相互关联

所有科学学科

纳米技术

纳米技术和科学学科的相互关联

世界地图：科学研究在哪里进行着

图例：科学学科是怎样相互关联的
TOPIC MAP: HOW SCIENTIFIC PARADIGMS RELATE
Overview

1. Plug-and-Play Macrosopes

2. Sci2 Tool Introduction and Demos

3. Sample Science of Science Studies
Design comprehensive databases that capture relevant data and cyberinfrastructures that can be used to make sense of this data stream.

Science studies can be conducted at different levels:
- **local** (individual),
- **meso** (local, e.g., one institute, one funding agency), or
- **global** level (all of science or world wide).

Using:
- Statistical Analysis/Profiling
- Temporal Analysis (When)
- Geospatial Analysis (Where)
- Topical Analysis (What)
- Network Analysis (With Whom?)

---

**Open Data and Open Code**

Studying Individual, Local, and Global Flows and Activity Patterns

---

**CI for a Science of Science Studies**

- **Scholarly Database**: 23 million scholarly records
  - [http://sdb.slis.indiana.edu](http://sdb.slis.indiana.edu)
- **Information Visualization Cyberinfrastructure**
  - [http://iv.slis.indiana.edu](http://iv.slis.indiana.edu)
- **Network Workbench Tool + Community Wiki**
  - [http://nwb.slis.indiana.edu](http://nwb.slis.indiana.edu)
- **Sci² Tool and Science of Science CI Portal**
  - [http://sci.slis.indiana.edu](http://sci.slis.indiana.edu)
- **Epidemics Cyberinfrastructure**
  - [http://epic.slis.indiana.edu/](http://epic.slis.indiana.edu/)
Network Workbench Tool
A large-scale network analysis, modeling and visualization toolkit for biomedical, social science and physics research.

http://nwb.slis.indiana.edu

Science of Science Tool
for science of science studies at the individual, local, or global level for temporal, geospatial, semantic, or network analysis and visualization.

http://sci.slis.indiana.edu

Epidemics Research Tool
An open computational infrastructure for epidemics research.

http://epic.slis.indiana.edu

Sci2 Tool
Branded OSGi/CIShell based tool with NWB plugins and many new new plugins.

http://sci.slis.indiana.edu

“Open Code for S&T Assessment”

Sci² Tool

Welcome to the Science of Science Tool (Sci² Tool)!
The development of this tool is supported by the National Science Foundation, the James S. McDonnell Foundation, and the Indiana University Network Science Center (http://iub.divine.indiana.edu).

The primary investigators are Katy Bauer, Joakim Malmqvist, and Dan Malarz. The Sci² Tool was developed by D. Shery, Patrick A. Phillips, Christian Hark, and D. M. Lindeberg. Many algorithm plugins were derived from the Network Science Center (http://iub.divine.indiana.edu).

Please cite as follows:

Circular Hierarchy

Geo Maps

Topic Mapping: UCSD Science Map

Journal locations for FourNetSciResearchers.txt

342 journal references matched out of 361 found.
These 342 references are associated with 3 of 15 disciplines of science and 238 of 534 research specialties in the UCSD Map of Science.
**Sci² 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)

---

**Sci² Tool: Algorithms**

See [https://nwb.slis.indiana.edu/community](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
- Barabasi-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
- 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

---

24
See https://nwb.slis.indiana.edu/community

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)

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

NWB=Sci² 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.
Sample Study – NSF Funding of STEM

download relevant NSF awards that have “stem” and “education” in title, abstract, and awards.
Active awards only.

Number of awards: 1,340
Total awarded amount to date: $1,347,802,833

Retrieved on Oct 18, 2009

Federal K-12 STEM Education Program Funding in 2006

Search for awards that have “stem” and “education” in title, abstract, and awards.
Active awards only. Query run on 10/18/2009.
### Top-10 Projects with highest Award Amount to Date

<table>
<thead>
<tr>
<th>Title</th>
<th>NSF Org</th>
<th>Program(s)</th>
<th>PI</th>
<th>State</th>
<th>Organization</th>
<th>Awarded To Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Next Generation Biometrics: Achieving Strength in Molecular EPS</td>
<td>DUE</td>
<td>RESEARCH/INFRASTRUCTURE/Ped HI</td>
<td>WV</td>
<td>NJ</td>
<td>Higher Education Policy Commission</td>
<td>10,799,940</td>
</tr>
<tr>
<td>UMB-EPDS STEM Project</td>
<td>DUE</td>
<td>RESEARCH/INFRASTRUCTURE/Ped HI</td>
<td>WV</td>
<td>NJ</td>
<td>Higher Education Policy Commission</td>
<td>10,799,940</td>
</tr>
<tr>
<td>MRSER: MRSEC on Nanstructured Interfaces</td>
<td>DMR</td>
<td>MATERIALS PHYS SCI &amp; EN &amp; Juan De Pablo</td>
<td>WI</td>
<td>Wisconsin-Madison</td>
<td></td>
<td>11,664,193</td>
</tr>
<tr>
<td>Spatial Intelligence and Learning Center (SILC)</td>
<td>JSEP</td>
<td>NSERC ACTIVITIES/RESOURCES/James Perriot</td>
<td>ON</td>
<td>Nova Scotia</td>
<td></td>
<td>12,578,414</td>
</tr>
<tr>
<td>Project Pathways: Opening Routes to Math &amp; Science Suc</td>
<td>DUE</td>
<td>Teaching &amp; Ment Tch Fellow/William Emmons</td>
<td>PA</td>
<td>Temple University</td>
<td></td>
<td>12,792,641</td>
</tr>
<tr>
<td>ACL AT A Center to Develop Nanoscale Science and Engine (CALL)</td>
<td>NANO CTB FOR LRN &amp; TEACH P &amp; Chang</td>
<td></td>
<td>IL</td>
<td>Northwestern University</td>
<td></td>
<td>15,980,654</td>
</tr>
<tr>
<td>Enabling a Giant Segmented Mirror Telescope for the Unit ACT (ENSM)</td>
<td>INDIAN (ST) INSTRUMENTATION/Mark Cavaliere</td>
<td></td>
<td>DC</td>
<td>National Optical Astronomy Ob</td>
<td></td>
<td>16,482,000</td>
</tr>
<tr>
<td>Support of Synchrotron Radiation Center Operations (SREC)</td>
<td>DMR</td>
<td>MPS/DMR INSTRUMENTATION/Mark Cavaliere</td>
<td>WI</td>
<td>Wisconsin-Madison</td>
<td></td>
<td>19,526,500</td>
</tr>
<tr>
<td>Testbed Extension: Bridging to 10</td>
<td>DUE</td>
<td>Teaching &amp; Ment Tch Fellow/Telma Millar</td>
<td>WI</td>
<td>Wisconsin-Madison</td>
<td></td>
<td>36,900,000</td>
</tr>
<tr>
<td>System-Wide Changes for All Learners and Educators</td>
<td>DUE</td>
<td>Teaching &amp; Ment Tch Fellow/Telma Millar</td>
<td>WI</td>
<td>Wisconsin-Madison</td>
<td></td>
<td>36,900,000</td>
</tr>
</tbody>
</table>

### Search Results

Search Results:
Results are sorted by award date, with the most recent awards at the top. Click on a column heading to re-sort the results.
The up/down arrows at the right of each column title control whether the sort is ascending or descending.
To view the abstract, click on the award number or title. Click on the data in other columns to perform a new search with that parameter.

*Input Parameters:*
Start Date: Start Date
Size By: Awarded Amount to Date
Label: Title
End Date: Expiration Date

---

**1340 Funded Projects**

**Horizontal Bar Graphs**

Horizontal Line Graph was selected.
Input Parameters:
Start Date: Start Date
Size By: Awarded Amount to Date
Label: Title
End Date: Expiration Date
1,340 Funded Projects

Geographic Maps

Geocoder was selected.
Input Parameters:
Place Name Column: Organization State
Place Type: STATE

Geo Map (Circle Annotation Style) was selected.
Input Parameters:
Longitude: Longitude
Size Circles By: Awarded Amount to Date
Color Circle Exteriors By: Awarded Amount to Date
Color Circle Interiors By: None (no inner color)
Exterior Color Scaling: Linear
Exterior Color Range: Green to Red
Interior Color Range: Green to Red
Size Scaling: Linear
Projection: Albers Equal-Area Conic
Map: US States
Author Name: K. Borner
Interior Color Scaling: Linear
Latitude: Latitude

What Co-PI Networks Exist?

Extract Directed Network was selected.
Input Parameters:
Source Column: Principal Investigator
Text Delimiter: |
Target Column: Co-PI Name(s)

Network Analysis Toolkit (NAT) was selected.
Nodes: 3225
Isolated nodes: 276
Edges: 2265
Average total degree: 1.4047
Average in degree: 0.7023
Average out degree: 0.7023

Delete Isolates was selected.

Node Degree was selected.

Weak Component Clustering was selected.
Number of top clusters: 10
722 clusters found, generating graphs for the top 10 clusters.

Giant component has 39 nodes
Next largest networks have 35, 17, 16 nodes
Co-PI Networks – Giant Component

Nodes = investigators
Size and color coded by number of collaborators (degree)

Directed edges from PI to Co-PI

What Projects Fund Which PIs?

Extract Directed Network was selected.
Input Parameters:
Source Column: Title
Text Delimiter: |
Target Column: Principal Investigator

Network Analysis Toolkit (NAT) was selected.
Nodes: 2478
Isolated nodes: 0
Edges: 1337
Average total degree: 1.0791
Average in degree: 0.5395
Average out degree: 0.5395
This graph is not weakly connected.
There are 1144 weakly connected components.
The largest connected component consists of 14 nodes.
Density (disregarding weights): 0.0002

Node Indegree was selected.
Node Outdegree was selected.
GUESS
What Projects Fund Which PIs - Zoom

What Programs at NSF are Co-Funding STEM?

Extract Co-Occurrence Network was selected.
Input Parameters:
Text Delimiter: |
Column Name: Program(s)

Node Degree was selected.

Network Analysis Toolkit (NAT) was selected.
Nodes: 226
Isolated nodes: 71
Edges: 483
No self loops were discovered.
Average degree: 4.2743
Density (disregarding weights): 0.019

GUESS

Weak Component Clustering was selected.
79 clusters found

Network Analysis Toolkit (NAT) was selected.
Nodes: 135
Isolated nodes: 0
Edges: 467
No self loops were discovered.
Average degree: 6.9185
Density (disregarding weights): 0.0516

GUESS
What Programs at NSF are Co-Funding STEM – Giant Component

What Organizations are funded by what NSF Programs?

Extract Directed Network was selected.
Input Parameters:
Source Column: Organization
Text Delimiter: |
Target Column: Program(s)

Network Analysis Toolkit (NAT) was selected.
Nodes: 794
Isolated nodes: 1
Edges: 1592
Average total degree: 4.0101
Average in degree: 2.005
Average out degree: 2.005
The largest connected component consists of 777 nodes.
Density (disregarding weights): 0.0025

Node Indegree was selected.

Node Outdegree was selected.

GUESS
What Organizations are funded by what NSF Programs?

Color and size coding by number of NSF programs that fund these organizations. Institutions which are funded by 10 or more programs are labeled.

What NSF Programs fund how many Organizations?

Color and size coding by number of organizations that are funded by these programs. NSF programs which fund 10 or more organizations are labeled.
Overview

1. Plug-and-Play Macroscopes

2. Sci2 Tool Introduction and Demos

3. Sample Science of Science Studies
### Type of Analysis vs. Level of Analysis

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

Geospatial/Network Analysis
2001-2006, BioMed, IN Scope
Academic-Industry collaborations and knowledge diffusion

Mapping Topic Bursts

Co-word space of the top 50 highly frequent and bursty words used in the top 10% most highly cited PNAS publications in 1982-2001.

PNAS, 101(Suppl. 1): 5287-5290.

Temporal/Topical Analysis
Topic evolution and bursts
Research questions:
1. Does space still matter in the Internet age?
2. Does one still have to study and work at major research institutions in order to have access to high quality data and expertise and produce high quality research?
3. Does the Internet lead to more global citation patterns, i.e., more citation links between papers produced at geographically distant research institutions?

Contributions:
- Answer to Qs 1 + 2 is YES.
- Answer to Qs 3 is NO.
- Novel approach to analyzing the dual role of institutions as information producers and consumers and to study and visualize the diffusion of information among them.

Research Collaborations by the Chinese Academy of Sciences
By Weixia (Bonnie) Huang, Russell J. Duhon, Elisha F. Hardy, Katy Börner, Indiana University, USA

This map highlights the collaborations of the Chinese Academy of Sciences with locations in China and countries around the world. The large geographic map shows the research collaborations of all CAS institutes. Each smaller geographic map shows the research collaborations by the CAS researchers in one province-level administrative division. Collaborations between CAS researchers are not included in the data. On each map, locations are colored on a logarithmic scale by the number of collaborations from red to yellow. The darkest red is 3,395 collaborations by all of CAS with researchers in Beijing. Also, flow lines are drawn from the location of focus to all locations collaborated with. The width of the flow line is linearly proportional to the number of collaborations with the locations it goes to, with the smallest flow lines representing one collaboration and the largest representing differing amounts on each geographic map.
Individual Co-PI Network
Ke & Börner, (2006)

Temporal/Network Analysis
2001-2006, US, InfoVis Scope
Evolving project-PI networks

Mapping the Evolution of Co-Authorship Networks
**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.
113 Years of Physical Review

Topical/Network Analysis
1893-2006, World, Physics Scope
Topic evolution, knowledge diffusion, Nobel predictions

Topical Composition and Knowledge Flow Patterns in Chemistry Research for 1974 and 2004
Kevin W. Boyack, Katy Börner, & Richard Klavans (2007)

Chemistry - Biology Interface
1974

Temporal/Network Analysis
1974-2004, US, NSF Chemistry Funding Scope
Mapping the main structure, topic evolution, and knowledge diffusion
Interactive Maps of Science – NIH Funding

Google maps with charts and tables

Topical/Network Analysis
2007, US, NIH Funding Scope
Mapping the main structure color coded by funding institute


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.

Temporal/Network Analysis
Comparing co-author networks with different funding
Reference Mapper

Duhon & Börner, forthcoming.

Topical/Network Analysis
2009, US, NSF Funding
Grouping interdisciplinary funding proposals for review

Mapping S&T Job Market Data in Real Time – GeoMap
Angela Zoss, Michael Conover

Data
Thousands of full-text, location-specific, time stamped job postings from Nature Jobs and Science Careers RSS feeds. The posts have been parsed and stored in a relational MySQL database.

Jobs have been geolocated on a Google map.

Visualization of Job Postings

Geographic Visualization
Here we have a more traditional view of job postings – a geographic overlay. Positions are color-coded by state, city, and state in the United States. You can click on a job to see its details.
The UCSD Map of Science used here is the product of a large study by researchers at the University of California - San Diego using 7.2 million papers and over 16,000 separate journals, proceedings, and series from Thomson Scientific and Scopus over the five year period from 2001 to 2005.

Jobs were associated with nodes in the Map of Science by way of keyword extraction.

Mapping S&T Job Market Data in Real Time – SciMap
Angela Zoss, Michael Conover
Geospatial and Topical/Network Analysis
2008-2009, English speaking, Job RSS feeds
Support interactive search for job postings in geo and topic space.

The Power of Maps (2005)
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)
Telling Lies With Science Maps (2014)

Exhibit has been shown in 72 venues on four continents. Currently at
- NSF, 10th Floor, 4201 Wilson Boulevard, Arlington, VA
- Wallenberg Hall, Stanford University, CA
- Center of Advanced European Studies and Research, Bonn, Germany
- Science Train, Germany.

http://scimaps.org

This is the only mockup in this slide show. Everything else is available today.
All papers, maps, cyberinfrastructures, talks, press are linked from http://cns.slis.indiana.edu