

Tutorial: Open Source Tools for Data Analysis and Visualization

Katy Börner

Victor H. Yngve Professor of Information Science
Director, Cyberinfrastructure for Network Science Center
School of Informatics and Computing and
Indiana University Network Science Institute
Indiana University, USA

CDC, Atlanta, GA

8:30-11:30 AM EST
February 5, 2016



Please

- download the Sci2 Tool from <http://sci2.cns.iu.edu>
- these slides <http://cns.iu.edu/docs/presentations/2016-borner-cdc-tutorial.pdf>
- and complete the Pre-Tutorial Questionnaire

CNS Macroscopes are used by hundreds of thousands around the globe



Our mission is to advance datasets, tools, and services for the study of biomedical, social and behavioral science, physics, and other networks. A specific focus is research on the structure and evolution of science and technology (S&T) and the communication of results via static and interactive maps of science. Learn more at cishell.org.



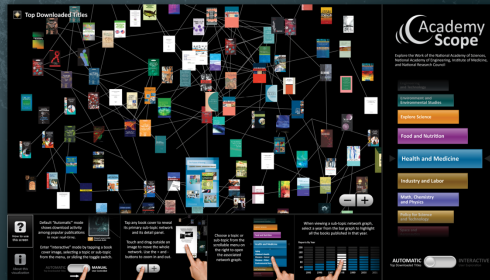
AcademyScope

AcademyScope is a state-of-the-art, interactive touch-screen visualization developed by CNS in collaboration with the National Academy of Sciences.

Using a 55-inch, multi-touch screen, viewers can explore 20 years of reports published by the National Academy of Sciences, National Academy of Engineering, Institute of Medicine, and National Research Council.

Beginning in October 2014, the *AcademyScope* web application is available to the public through the National Academies Press website. Users can access the application through the "Browse by Topic" menu on the NAP homepage (www.nap.edu), or via the "Browse Topics" button in the header of every interior page. The application can also be accessed directly at www.nap.edu/academy-scope.

Visit cns.iu.edu/interactive_displays to learn more about the design and programming.



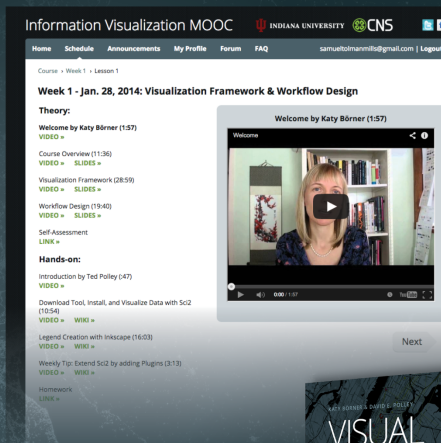
IVMOOC 2015

The Information Visualization MOOC provides an overview about the state of the art in information visualization, teaching the process of producing effective visualizations that take the needs of users into account.

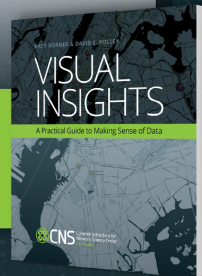
The inaugural IVMOOC, which launched in January 2013, attracted participants from more than 100 countries. It is one of the first MOOCs offered by IU and the first to offer an opportunity for students to work in teams with real clients. All registrants gain free access to the Scholarly Database and the Sci2 Tool.

The course can be taken for three Indiana University credits as part of the Online Data Science Program offered by the School of Informatics and Computing.

The course will return in January 2015. Learn more at ivmoo.cns.iu.edu.



This IVMOOC companion textbook offers a gentle introduction to the design of insightful visualizations. It seamlessly blends theory and practice, giving readers both the theoretical foundation and the practical skills necessary to render data into insights.



Tutorial Overview

8:30 Welcome and Overview of Tutorial and Attendees

9:00 The Sci2 Tool

- Download and run the Sci2 Tool
- ONE dataset, MANY analyses and visualizations

9:30 Sci2 Tool Workflows

- Temporal Analysis: Horizontal line graph of NSF projects
- Geospatial Analysis: US and world maps
- Geospatial Analysis: Geomap with network overlays
- Topical Analysis: Visualize research profiles
- Network Analysis: Co-occurrence networks and bimodal networks
- Network Analysis: Evolving collaboration networks

10:15 Networking Break

10:30 Visualization Framework

11:00 IVMOOC – MANY more Workflows

11:15 Outlook and Q&A

11:30 Adjourn

5

Tutorial Overview

8:30 Welcome and Overview of Tutorial and Attendees

9:00 The Sci2 Tool

- **Download and run the Sci2 Tool**
- **ONE dataset, MANY analyses and visualizations**

9:30 Sci2 Tool Workflows

- Temporal Analysis: Horizontal line graph of NSF projects
- Geospatial Analysis: US and world maps
- Geospatial Analysis: Geomap with network overlays
- Topical Analysis: Visualize research profiles
- Network Analysis: Co-occurrence networks and bimodal networks
- Network Analysis: Evolving collaboration networks

10:15 Networking Break

10:30 Visualization Framework

11:00 IVMOOC – MANY more Workflows

11:15 Outlook and Q&A

11:30 Adjourn

6

The Sci2 Tool: A Plug-and-Play Macroscope that implements the Visualization Framework

7

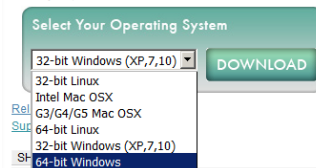


Software, Datasets, Plugins, and Documentation

- These slides
<http://cns.iu.edu/docs/presentations/2016-borner-cdc-tutorial.pdf>
- Sci2 Tool Manual
<http://sci2.wiki.cns.iu.edu>
- Sci2 Tool v1.2 beta
<http://sci2.cns.iu.edu>
- Additional Datasets
<http://sci2.wiki.cns.iu.edu/2.5+Sample+Datasets>
- Additional Plugins
<http://sci2.wiki.cns.iu.edu/3.2+Additional+Plugins>

Download

Sci2 v1.2 beta
January 7th, 2015



Make sure you have Java 1.6 (64-bit, if you selected 64-bit) or higher installed or download from <http://www.java.com/en/download>. To check your Java version, open a terminal and run 'java -version'.
Some visualizations are saved as Postscript files. A free Postscript to PDF viewer is at <http://ps2pdf.com> and a free PDF Viewer at <http://www.adobe.com/products/reader.html>.

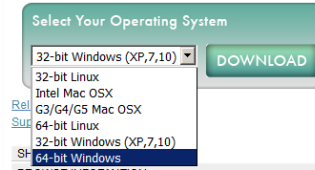
8

Install and Run Sci2

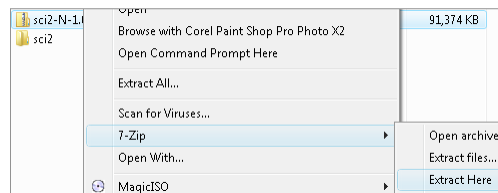
Sci2 Tool runs on Windows, Mac, and Linux.

Download

Sci2 v1.2 beta
January 7th, 2015



Unzip.



Run /sci2/sci2.exe

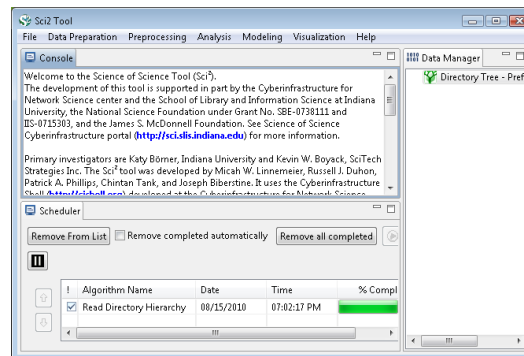
9

Sci2 Tool Interface Components

See also <http://sci2.wiki.cns.iu.edu/2.2+User+Interface>

Use

- **Menu** to read data, run algorithms.
- **Console** to see work log, references to seminal works.
- **Data Manager** to select, view, save loaded, simulated, or derived datasets.
- **Scheduler** to see status of algorithm execution.

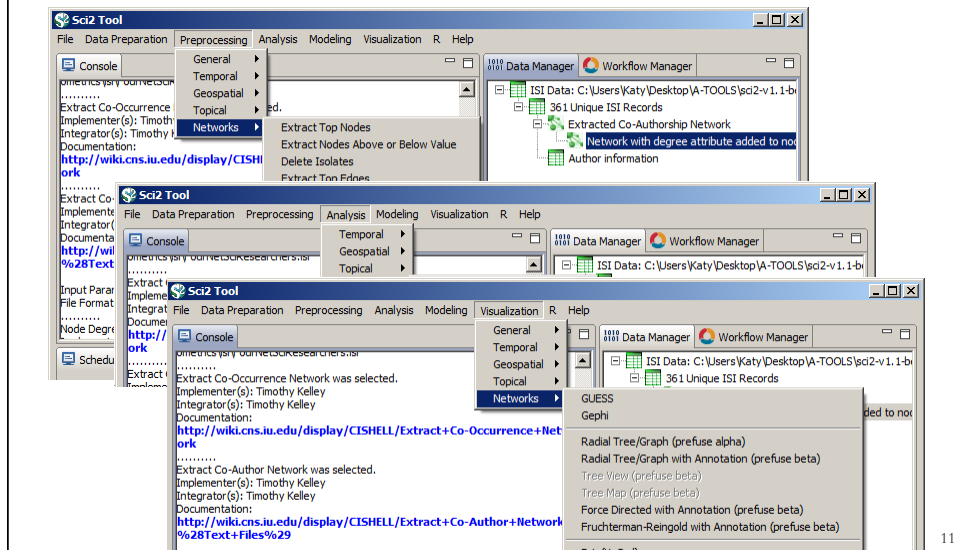


All workflows are recorded into a log file (see /sci2/logs/...), and can be re-run for easy replication. If errors occur, they are saved in a error log to ease bug reporting. All algorithms are documented online; workflows are given in Sci2 Manual at <http://sci2.wiki.cns.iu.edu>

10

Sci2 Tool Interface Components

Download for free at <http://sci2.cns.iu.edu>



11

Load **One** File and Run **Many** Analyses and Visualizations

Times Cited	Publication Year	City of Publisher	Country	Journal Title (Full)	Title	Subject Category	Authors
12	2011	NEW YORK	USA	COMMUNICATIONS OF THE ACM	Plug-and-Play Macroscopes	Computer Science	Borner, K
18	2010	MALDEN	USA	CTS-CLINICAL AND TRANSLATIONAL SCIENCE	Advancing the Science of Team Science	Research & Experimental Medicine	Falk-Krzesinski, HJ Borner, K Contractor, N Fiore, SM Hall, KL Keyton, J Spring, B Stokols, D Trochim, W Uzzi, B
13	2010	WASHINGTON	USA	SCIENCE TRANSLATIONAL MEDICINE	A Multi-Level Systems Perspective for the Science of Team Science	Cell Biology Research & Experimental Medicine	Borner, K Contractor, N Falk-Krzesinski, HJ Fiore, SM Hall, KL Keyton, J Spring, B Stokols, D Trochim, W Uzzi, B

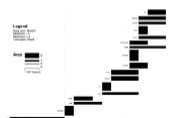
Statistical Analysis—p. 44

Temporal Burst Analysis—p. 48

Geospatial Analysis—p. 52

Geospatial Analysis—p. 52

Location	Count	# Citations
Netherlands	13	292
United States	9	318
Germany	11	36
United Kingdom	1	2

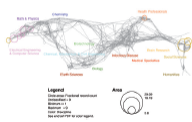


12

Load **One** File and Run **Many** Analyses and Visualizations

Times Cited	Publication Year	City of Publisher	Country	Journal Title (Full)	Title	Subject Category	Authors
12	2011	NEW YORK	USA	COMMUNICATIONS OF THE ACM	Plug-and-Play Macroscopes	Computer Science	Borner, K
18	2010	MALDEN	USA	CTS-CLINICAL AND TRANSLATIONAL SCIENCE	Advancing the Science of Team Science	Research & Experimental Medicine	Falk-Krzesinski, HJ Borner, K Contractor, N Fiore, SM Hall, KL Keyton, J Spring, B Stokols, D Trochim, W Uzzi, B
13	2010	WASHINGTON	USA	SCIENCE TRANSLATIONAL MEDICINE	A Multi-Level Systems Perspective for the Science of Team Science	Cell Biology Research & Experimental Medicine	Borner, K Contractor, N Falk-Krzesinski, HJ Fiore, SM Hall, KL Keyton, J Spring, B Stokols, D Trochim, W Uzzi, B

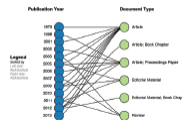
Topical Analysis—p. 56



Paper Citation Network—p. 60



Bi-Modal Network—p. 60



Co-author and many other bi-modal networks.

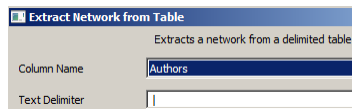
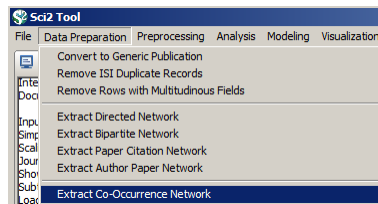
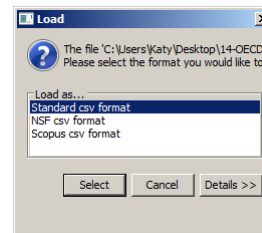
13

Load **One** File and Run **Many** Analyses and Visualizations

Download 20publications.csv from <http://wiki.cns.iu.edu/download/attachments/1245848/20publications.csv?version=1&modificationDate=1403450235951>

In Sci2, use 'File > Load' and load file as 'Standard csv format'.

Run 'Data Preparation > Extract Co-Occurrence Network' with parameters:



Co-author network will appear in **Data Manager**.

14

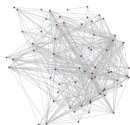
Load **One** File and Run **Many** Analyses and Visualizations

Run 'Analysis > Network Analysis Toolkit (NAT)' to get basic properties:

Nodes: 65
Isolated nodes: 0
Edges: 404
No self loops were discovered.
Average degree: 12.4308
The largest connected component consists of 65 nodes.
Density (disregarding weights): 0.1942

Select 'Extracted Network on Column Authors' network in Data Manager and run 'Visualization > GUESS' to open GUESS with file loaded.

Initial layout is random:



In GUESS, apply 'Layout > GEM':



15

Sci2 Workflows

Light, Robert, David E. Polley, and Katy Börner. 2014.
["Open Data and Open Code for Big Science of Science Studies"](#). *Scientometrics*
101 (2): 1535-1551.

16

Tutorial Overview

8:30 Welcome and Overview of Tutorial and Attendees

9:00 The Sci2 Tool

- Download and run the Sci2 Tool
- ONE dataset, MANY analyses and visualizations

9:30 Sci2 Tool Workflows

- **Temporal Analysis: Horizontal line graph of NSF projects**
- Geospatial Analysis: US and world maps
- Geospatial Analysis: Geomap with network overlays
- Topical Analysis: Visualize research profiles
- Network Analysis: Co-occurrence networks and bimodal networks
- Network Analysis: Evolving collaboration networks

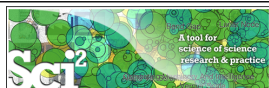
10:15 Networking Break

10:30 Visualization Framework

11:00 IVMOOC – MANY more Workflows

11:15 Outlook and Q&A

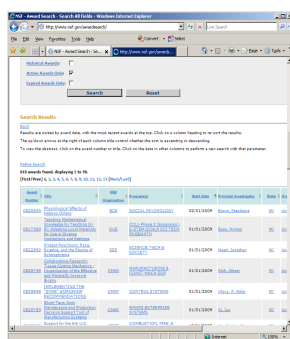
11:30 Adjourn



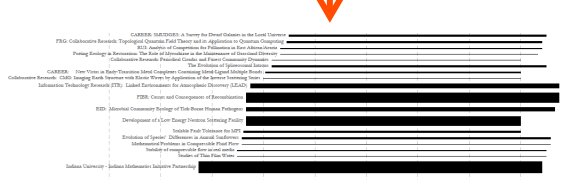
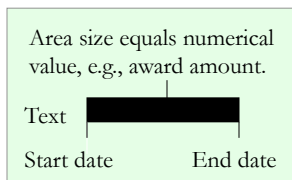
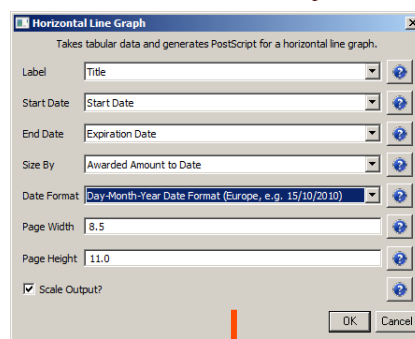
Horizontal line graph of NSF projects

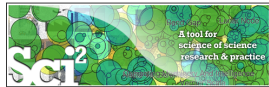
See [5.2.1 Funding Profiles of Three Universities \(NSF Data\)](#)

Download NSF data



Visualize as Horizontal Line Graph





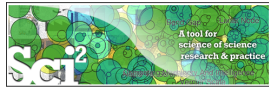
Horizontal line graph of NSF projects

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

The screenshot shows the NSF Award Search website. The search criteria are set to 'NSF Web Site'. The 'Organization' dropdown menu is highlighted with a red circle and contains the text 'University of Michigan Ann Arbor'. The search results table is visible, showing columns for Award Number, Title, PI, Start Date, End Date, and Amount. A red circle highlights the 'Save in CSV format as *institution*.nsf' option in the search results interface.

Award Number	Title	PI	Start Date	End Date	Amount
0822020	Expanding Mathematical Knowledge for Teachers (E-MKT)	COLLEGE	09/01/2009	08/31/2010	\$1,000,000
0817265	Teaching Mathematical Knowledge for Teachers (E-MKT)	COLLEGE	09/01/2009	08/31/2010	\$1,000,000
0822832	Teaching Mathematical Knowledge for Teachers (E-MKT)	COLLEGE	09/01/2009	08/31/2010	\$1,000,000
0822735	Teaching Mathematical Knowledge for Teachers (E-MKT)	COLLEGE	09/01/2009	08/31/2010	\$1,000,000
0822838	Teaching Mathematical Knowledge for Teachers (E-MKT)	COLLEGE	09/01/2009	08/31/2010	\$1,000,000

19



Temporal bar graph of NSF projects

Download and load a dataset of your choice or load one of the sample data files, e.g., *'sampledata/scientometrics/nsf/Indiana.nsf'*

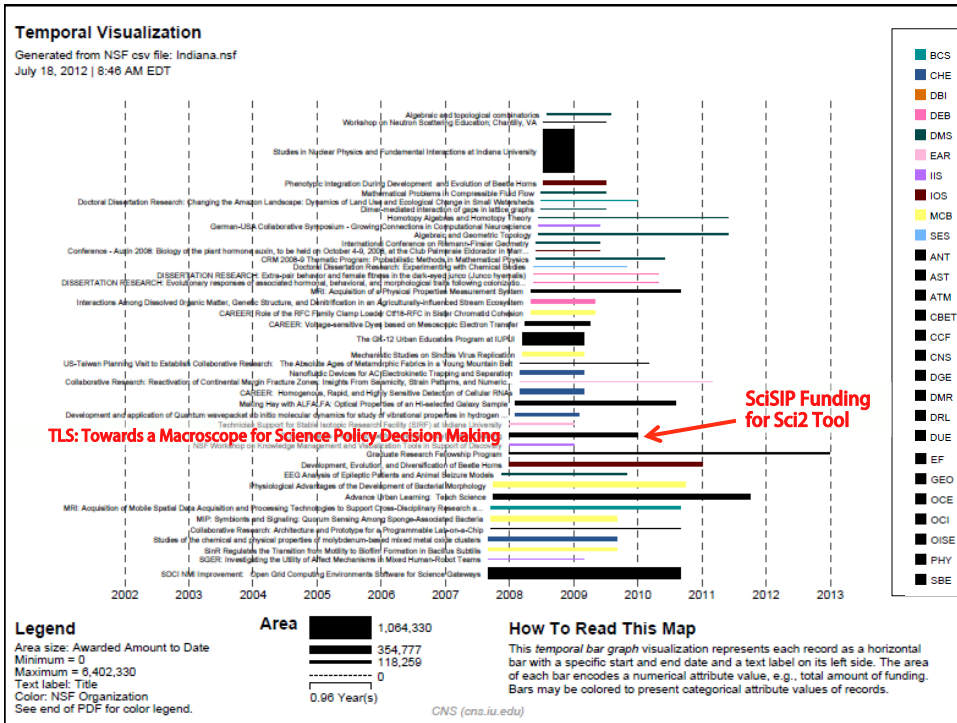
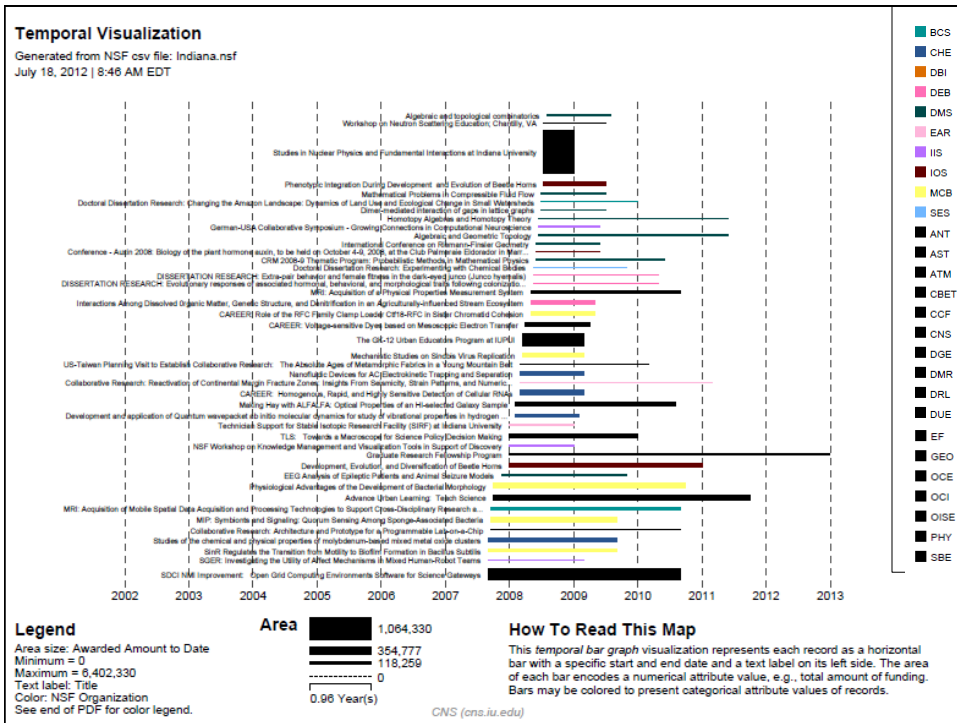
Run *'Visualization > Temporal > Temporal Bar Graph'* using parameters:

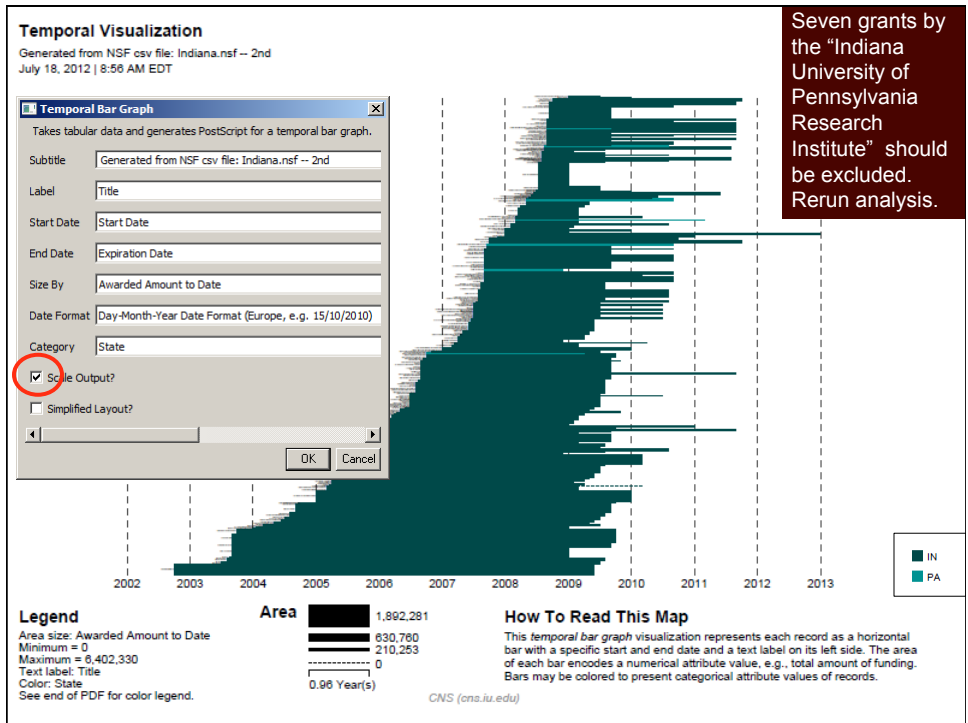
The screenshot shows the 'Temporal Bar Graph' dialog box in Sci2. The subtitle is 'Generated from NSF csv file: Indiana.nsf'. The label is 'Title'. The start date is 'Start Date' and the end date is 'Expiration Date'. The size is 'Awarded Amount to Date'. The date format is 'Day-Month-Year Date Format (Europe, e.g. 15/10/2010)'. The category is 'NSF Organization'. The 'Scale Output?' and 'Simplified Layout?' checkboxes are unchecked. The 'Save' dialog box is also visible, showing the output file name 'out_data'.

Save *'visualized with Horizontal Line Graph'* as ps or eps file. Convert into pdf and view.

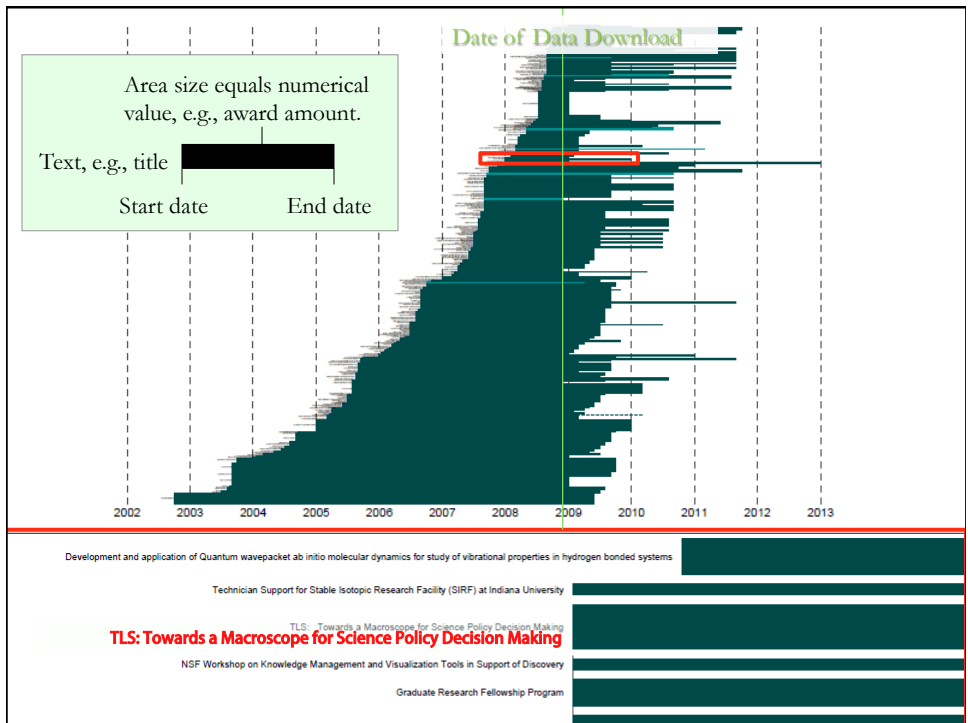
Zoom to see details in visualizations of large datasets, e.g., all NSF awards ever made.

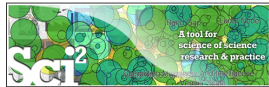
20





Seven grants by the "Indiana University of Pennsylvania Research Institute" should be excluded. Rerun analysis.



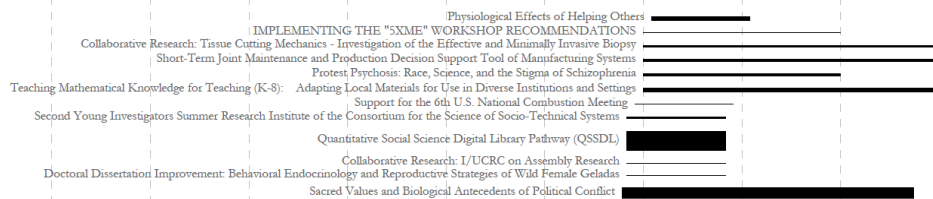
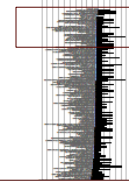


Temporal bar graph of NSF projects

Area size equals numerical value, e.g., award amount.

Text, e.g., title

Start date End date

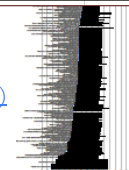


More NSF data workflows can be found in wiki tutorial:

[5.1.3 Funding Profiles of Three Researchers at Indiana University \(NSF Data\)](#)

[5.2.1 Funding Profiles of Three Universities \(NSF Data\)](#)

[5.2.3 Biomedical Funding Profile of NSF \(NSF Data\)](#)



25



Tutorial Overview

8:30 Welcome and Overview of Tutorial and Attendees

9:00 The Sci2 Tool

- Download and run the Sci2 Tool
- ONE dataset, MANY analyses and visualizations

9:30 Sci2 Tool Workflows

- Temporal Analysis: Horizontal line graph of NSF projects
- **Geospatial Analysis: US and world maps**
- Geospatial Analysis: Geomap with network overlays
- Topical Analysis: Visualize research profiles
- Network Analysis: Co-occurrence networks and bimodal networks
- Network Analysis: Evolving collaboration networks

10:15 Networking Break

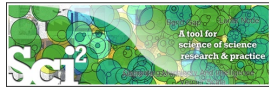
10:30 Visualization Framework

11:00 IVMOOC – MANY more Workflows

11:15 Outlook and Q&A

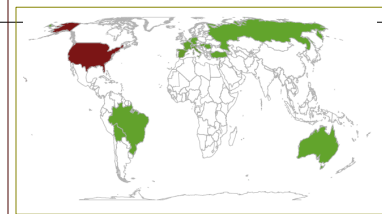
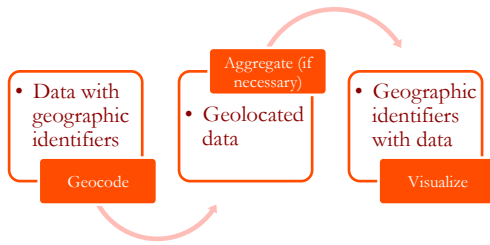
11:30 Adjourn

26

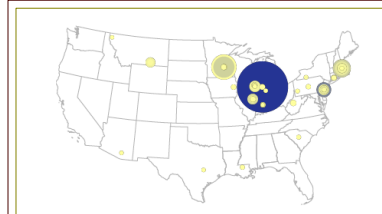


Geocoding and Geospatial Maps

<http://wiki.cns.iu.edu/display/CISHELL/Bing+Geocoder>



Region names + numeric data
(Choropleth Map)



Geocoordinates + numeric data
(Proportional Symbol Map)

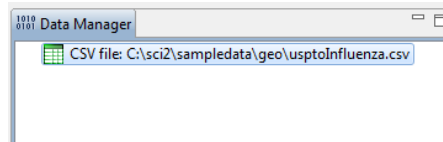
27



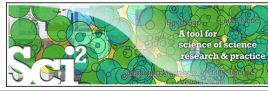
Load File with Address and Times Cited Fields

Run 'File > Load...' and select the sample data table 'sampledata/geo/usptoInfluenza.csv'
Create a map of influenza patents held by different countries.

	A	B	C	D	E
1	Country	Latitude	Longitude	Patents	Times Cited
2	Hungary	47.16116	19.504959	0.083333333	4
3	Belgium	50.500992	4.47677	3.017857143	11
4	Germany	51.090839	10.45424	4.783333333	4
5	Canada	62.35873	-96.582092	5.539285714	21
6	Russia	59.461479	108.831779	0.266666667	2
7	Austria	47.69651	13.34577	4.2	17
8	Netherlands	52.108089	5.33033	1	2
9	Switzerland	46.813091	8.22414	0.507575758	6
10	Taiwan	23.599751	121.023811	2	3
11	Australia	-24.916201	133.393112	1.617857143	23
12	United States	39.83	-98.58	73.9983889	220
13	France	46.712448	1.71832	2.201165501	9
14	South Africa	-28.483219	24.676991	0.333333333	1
15	Japan	37.487598	139.838287	15.991666667	39
16	Israel	31.389299	35.36124	3.5	3
17	United Kingdom	54.313919	-2.23218	3.85	12



28



Bing Geocoder



CIShell Manual / ... / Algorithms

Edit Watch Share Tools

Bing Geocoder

Created by Mayur Masrani, last modified by Adam Simpson on Aug 14, 2013

Description

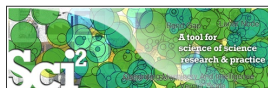
This algorithm converts place names or addresses into Latitude, Longitude co-ordinates. It accepts international addresses, countries, States of United States of America and ZIP codes of United States of America. All co-ordinates are obtained by querying Bing geocoder service. Internet access must be available during geocoding.

Pros & Cons

1. The performance is slower than the [Geocoder](#) and may vary due to the network latency since the queries are requested through internet service.
2. Bing Geocoder supports address geocoding with international coverage which is not supported by [Geocoder](#).
3. To use Bing Geocoder, user has to obtain an API Keys from [Bing Maps](#). Save your api keys and provide it when requested by the Bing Geocoder. Since each api key is allowed to geocode 50,000 locations per 24 hours, the user is encouraged to test on a small set of data first.

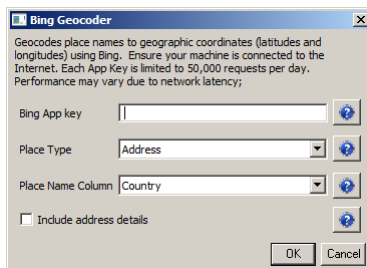
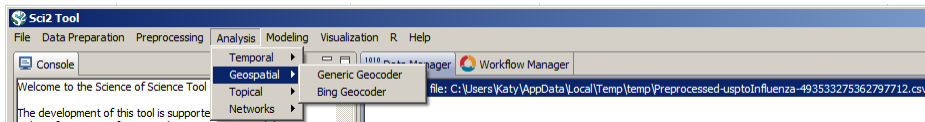
<http://wiki.cns.in.edu/display/CISHELL/Bing+Geocoder>

29



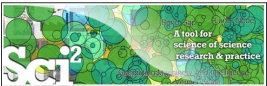
Using Bing Geocoder

Run *'Analysis > Geospatial > Bing Geocoder'*



Enter your Bing app key.
You can obtain one from [here](#)

30



Aggregate by Country

Aggregate Data was selected.
 Implementer(s): Chintan Tank
 Documentation: <http://wiki.cns.iu.edu/display/CISHELL/Aggregate+Data>
 Input Parameters:
 Aggregate on column: Country
 Delimiter for Country: |
 Longitude: AVERAGE
 Latitude: AVERAGE
 Times Cited: SUM

Aggregated by "": All rows of Latitude column were skipped due to no non-null, non-empty values.
 Aggregated by "": All rows of Longitude column were skipped due to no non-null, non-empty values.
 Frequency of unique "Country" values added to "Count" column.

Preprocessing Analysis Modeling Visualization

General ▶
 Temporal ▶
 Geospatial ▶
 Topical ▶

Extract Top N% Records
 Extract Top N Records
Aggregate Data

Aggregate Data

Aggregate data in the table based on a column.

Aggregate on column: Country

Times Cited: Sum


Latitude: Average

Longitude: Average

Delimiter for Country: |

OK Cancel

	A	B	C	D	E
1	Times Cited	Latitude	Longitude	Country	Count
2	7	42.02946091	-87.68838501	United States	194
3	0				57
4	0				14
5	2	42.34999466	-71.08765411	United States	4
6	14	41.70074844	-86.23918915	United States	1
7	15	41.70074844	-86.23918915	United States	2
8	29	41.89422607	-87.61901855	United States	2
9	32	41.70074844	-86.23918915	United States	12
10	7	41.70074844	-86.23918915	United States	5
11	5	41.70074844	-86.23918915	United States	2
12	2	41.11500168	-85.73377991	United States	13
13	10	47.50622177	19.06481934	Hungary	46
14	44	41.70074844	-86.23918915	United States	3
15	0	47.50622559	19.06481934	Hungary	2
16	19	41.70074844	-86.23918915	United States	1



Choropleth Map

Visualization R Help

General ▶
 Temporal ▶
 Geospatial ▶
 Topical ▶
 Networks ▶

Proportional Symbol Map
Choropleth Map
 Geospatial Network Layout with Base Map

Choropleth Map

Color-codes the name proportion to associate

Subtitle: Generated from CSV

Map: World

Region Name: Country

Color By: Times Cited

Color Scaling: Logarithmic

Color Range: Yellow to Blue

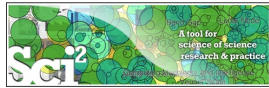
Right-click and **Save** map as PostScript file. Use PostScript Viewer or convert to pdf to view.

Data Manager

CSV file: C:\sci2\sampladata\geo\usptoInfluenza.csv

Geospatial Visualization (Choropleth Map)

Save

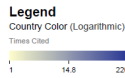
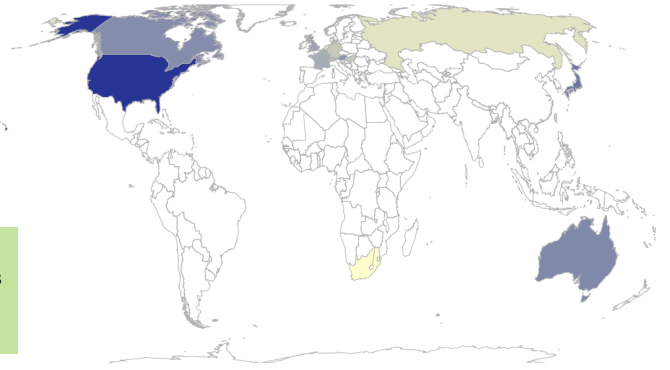


Reading the Choropleth Map

Geospatial Visualization (Choropleth Map)
 Generated from CSV file: C:\sci2\sampladata\geo\usptoInfluenza.csv
 Jun 14, 2012 | 05:33:37 PM EDT

Header shows visualization type, data description, and creation date

Legend shows how data matches up with visual representation

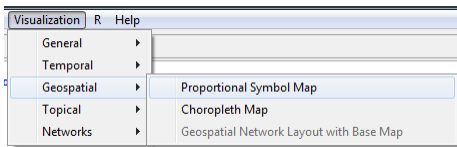


How to Read this Map
 This choropleth map shows 209 countries of the world using the equal-area Eckert IV projection. Each country may be color coded in proportion to a numerical value. Minimum and maximum data values are given in the legend.

33



Proportional Symbol Map



Proportional Symbol Map

Maps geospatial coordinates as circle color-coded in proportion to associated data values.

Subtitle: Generated from CSV file: C:\s

Map: World

Latitude: Latitude

Longitude: Longitude

Size Circles By: Times Cited

Size Scaling: Logarithmic

Color Circle Exteriors By: None (uniform coloring)

Exterior Color Scaling: Linear

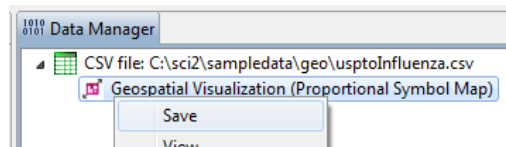
Exterior Color Range: White to Green

Color Circle Interiors By: None (no coloring)

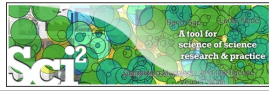
Interior Color Scaling: Linear

Interior Color Range: White to Green

Right-click and **Save** map as PostScript file. Use PostScript Viewer or convert to pdf to view.



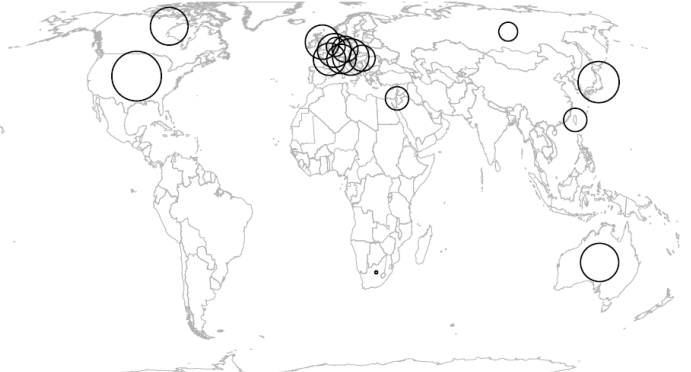
34



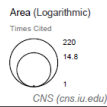
Reading the Proportional Symbol Map

Geospatial Visualization (Proportional Symbol Map)
 Generated from CSV file: C:\sci2\sampledata\geo\usptoinfluenza.csv
 Jun 14, 2012 | 05:56:39 PM EDT

Header shows visualization type, data description, and creation date



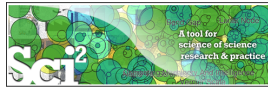
Legend shows how data matches up with visual representation



How to Read this Map

This proportional symbol map shows 209 countries of the world using the equal-area Eckert IV projection. Each dataset record is represented by a circle centered at its geolocation. The area, interior color, and exterior color of each circle may represent numeric attribute values. Minimum and maximum data values are given in the legend.

35



Relevant Sci2 Manual entry



5.2.4 Mapping Scientometrics (ISI Data)

9 Added by Ted Polley, last edited by Ted Polley on Nov 14, 2011 (View change)

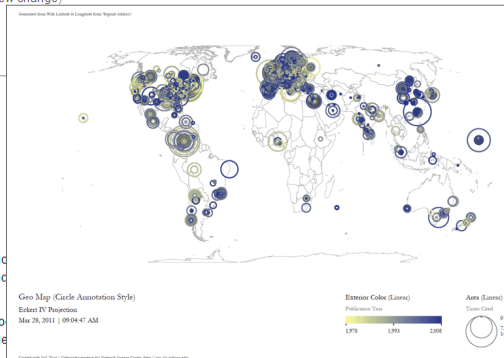
Edit Add Tools

5.2.4.1 Document Co-Citation

Scientometrics.isi	
Time frame:	1978-2008
Region(s):	Miscellaneous
Topical Area(s):	Scientometrics
Analysis Type(s):	Document Co-Citation Network

Scientometrics is a discipline which uses statistical and computational science. Here we use ISI data from the journal "Scientometrics" and Awards Search.

Download [Scientometrics.isi](#). Load the file using 'File > Load' and load document co-citation analysis, as the scale is large enough that the similarity within the domain of scientometrics.



New ISI File Format

Web of Science made a change to their output format in September, 2011. Older versions of Sci2 tool may refuse to load these new files, with an error like "Invalid ISI format file selected."

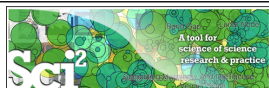
<http://wiki.cns.iu.edu/display/SCI2TUTORIAL/5.2.4+Mapping+Scientometrics+%28ISI+Data%29>

36

Tutorial Overview

- 8:30 Welcome and Overview of Tutorial and Attendees
- 9:00 The Sci2 Tool
 - Download and run the Sci2 Tool
 - ONE dataset, MANY analyses and visualizations
- 9:30 Sci2 Tool Workflows
 - Temporal Analysis: Horizontal line graph of NSF projects
 - Geospatial Analysis: US and world maps
 - **Geospatial Analysis: Geomap with network overlays**
 - Topical Analysis: Visualize research profiles
 - Network Analysis: Co-occurrence networks and bimodal networks
 - Network Analysis: Evolving collaboration networks
- 10:15 Networking Break**
- 10:30 Visualization Framework
- 11:00 IVMOOC – MANY more Workflows
- 11:15 Outlook and Q&A
- 11:30 Adjourn**

37



Geomap with Gephi Network Overlay

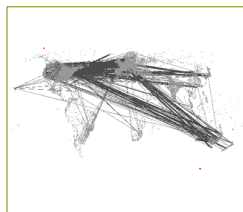
See 4.7.6 on <http://sci2.wiki.cns.iu.edu>

File with geolocations and linkage info, e.g., an isi bibliography file. → Use Bing Geocoder to identify Latitude, Longitude for each geolocation
 Extract attributes per geolocation, e.g., total times cited (TC)
 Extract linkages and their attributes, e.g., number of co-occurrences
 See sample /geo/LaszloBarabasiGeo.net with co-occurrence of “Research Addresses” and full counting of TC per geolocation.

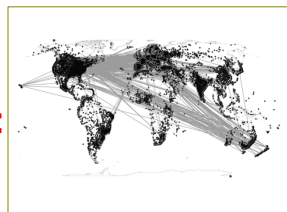
↓
 Read into Sci2 Tool to generate geomap and network file → Layout network in Gephi → Combine geomap and network in Photoshop



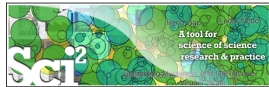
+



=



38



Relevant Sci2 Manual entry

Dashboard > Sci2 Manual > ... > 4 Workflow Design > 4.7 Geospatial Analysis (Where) Browse Log In Search Confluence

Search

- Home
- 1 Introduction
- 2 Getting Started
- 3 Algorithms, Tools, and Plugins
- 4 Workflow Design
 - 4.1 Overview
 - 4.2 Data Acquisition and Preparation
 - 4.3 Database Loading and Manipulation
 - 4.4 Summaries and Table Extractions
 - 4.5 Statistical Analysis and Profiling
 - 4.6 Temporal Analysis (When)
 - 4.7 Geospatial Analysis (Where)**
 - 4.8 Topical Analysis (What)
 - 4.9 Network Analysis (With Whom?)
 - 4.10 Modeling (Why?)
- 5 Sample Workflows
- 6 Sample Science Studies & Online Services
- 7 Extending the Sci2 Tool
- 8 Relevant Datasets and Tools
- 9 References

4.7.6 Using Gephi to Render Networks Overlaid on Geo Maps

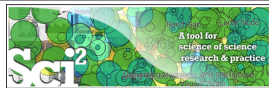
Loading and Saving Geovisualization Files in Sci2

This algorithm allows for the geospatial visualization of network data. The algorithm produces a network file and corresponding blank map. Gephi is used to edit the network produced by Sci2. Once the network has been edited in Gephi it can be exported in a format that will allow it to be overlaid on the map, facilitating visualization of the geospatial data. The following is a brief workflow explaining the process, beginning to end.

1. Load this network in Sci2.
2. Once the network had been loaded in Sci2 run "Visualization > Geospatial > Geo Map (network template, not fully rendered)" and set the following parameters:

<http://sci2.wiki.cns.iu.edu/display/SCI2TUTORIAL/4.7+Geospatial+Analysis+%28Where%29#4.7GeospatialAnalysis%28Where%29-4.7.6UsingGephiToRenderNetworksOverlaidonGeoMaps>

39



Use Sci2 Tool to Generate Geomap and Network File

Read prepared .net file and run:

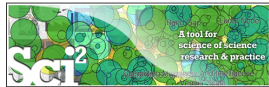
```
Loaded:  
C:\Users\katy\Desktop\sci2-12.20\sampledata\scientometrics\sci2-12.20\sampledata\geoNetwork.net
```

Save map file as Postscript file and use Adobe or other view to read. It looks like:

Save .net file as GraphML (Prefuse) and rename to .graphml so that Gephi can read it.

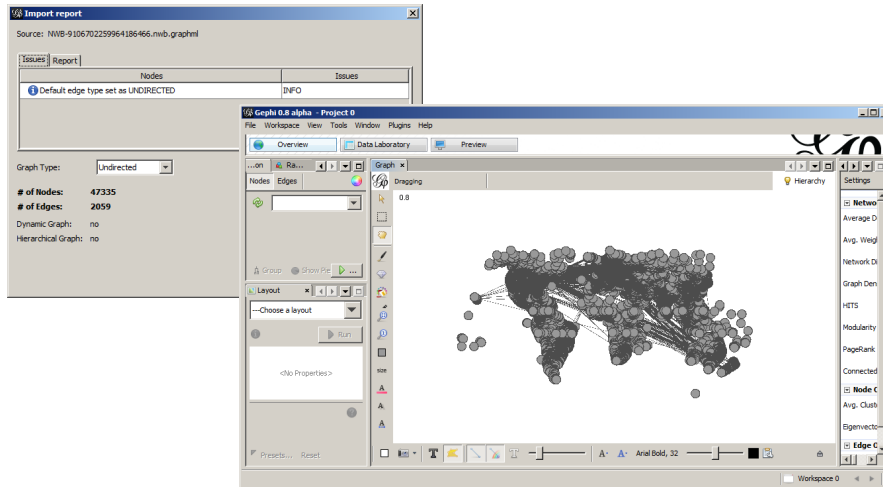


40



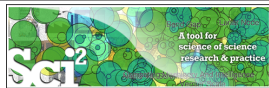
Use Gephi to Generate Network Layout

Start gephi. Use *New Project > Open a graph file* to read .graphml file that Sci2 generated.



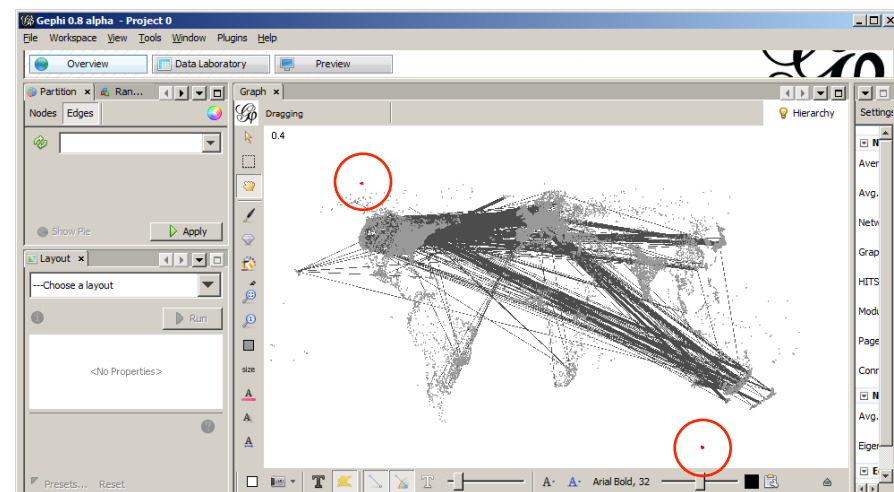
Follow instructions in online tutorial on **Manipulating the Network File in Gephi**

41

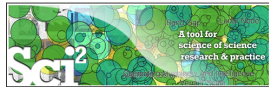


Use Gephi to Generate Network Layout

Color or size code the “Near Alaska” and “Near Antarctica” anchor nodes to ease alignment of geomap and network overlay, see instructions in online tutorial on **Manipulating the Network File in Gephi**. Save result using *File > Export > SVG/PDF file*.

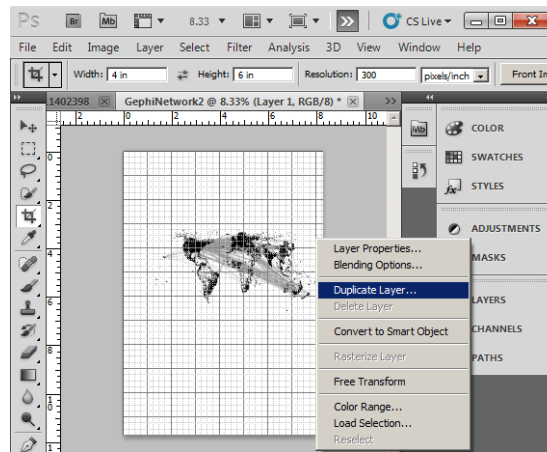


42

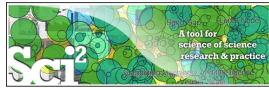


Use Photoshop to Overlay Network on Geomap

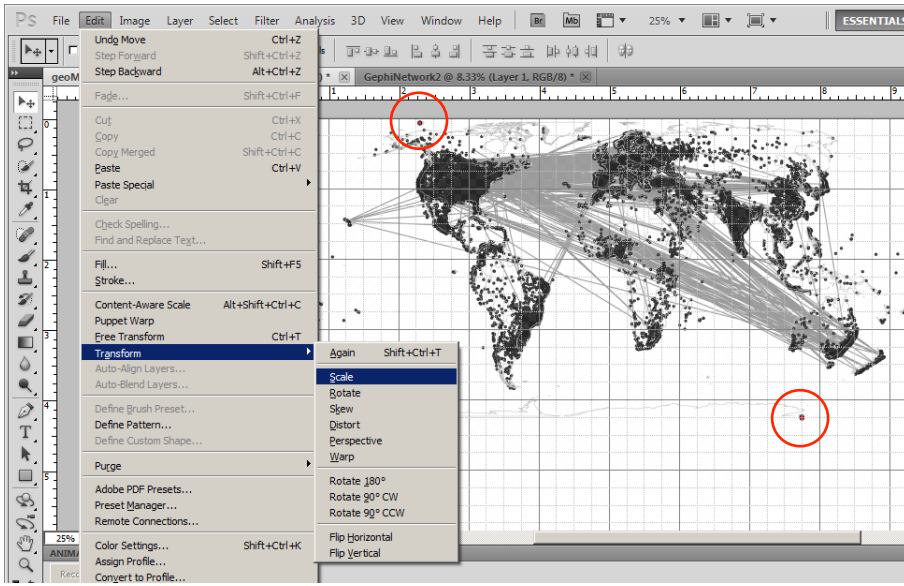
Load geomap and network files into Photoshop. Select 'network' layer and use 'Right click, Duplicate Layer' to copy network over to 'geomap' file as a second layer. Use Edit > Transform > Scale' and align using the "Near Antarctica" anchor nodes, see instructions in online tutorial on **Creating the Visualization in Photoshop**.



43

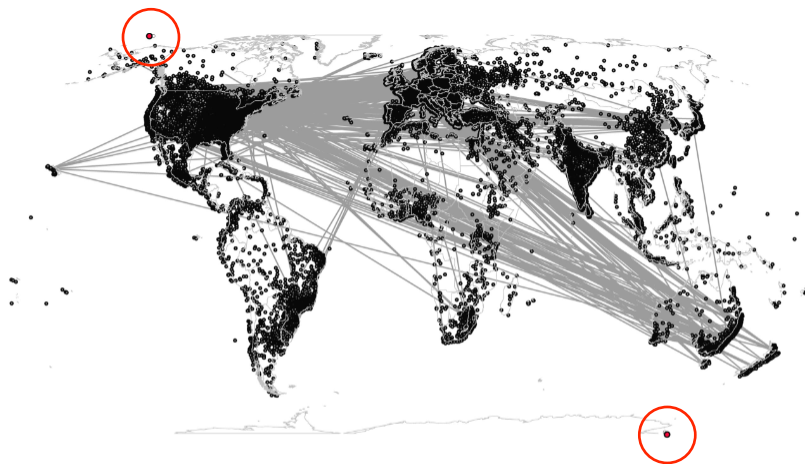


Use Photoshop to Overlay Network on Geomap



44

Delete anchor nodes and save in preferred format.

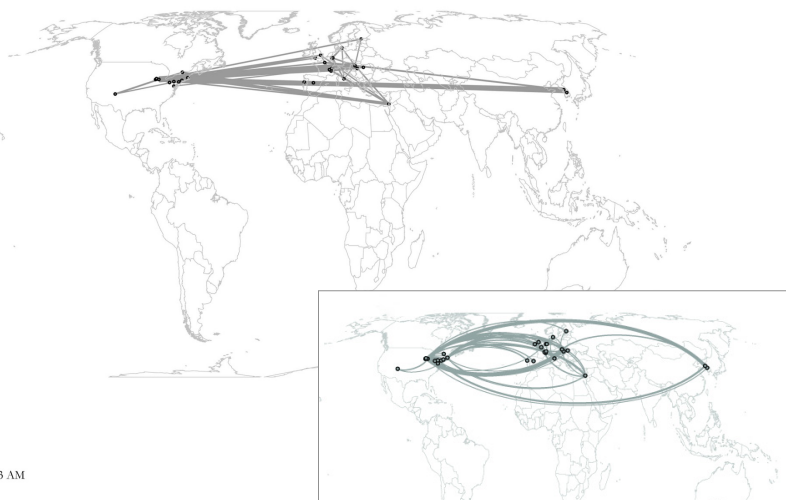


Geo Map ()
Eckert IV Projection
Apr 06, 2012 | 03:19:51 AM

Created with Sci2 Tool | Cyberinfrastructure for Network Science Center (<http://cims.nyu.edu>)

45

Practice these steps using “LaszloBarabasi-collaborations.net” linked from Sci2 wiki:
4.7.6 Using Gephi to Render Networks Overlaid on GeoMaps



Geo Map ()
Eckert IV Projection
Apr 11, 2012 | 06:20:13 AM

Created with Sci2 Tool | Cyberinfrastructure for Network Science Center (<http://cims.nyu.edu>)

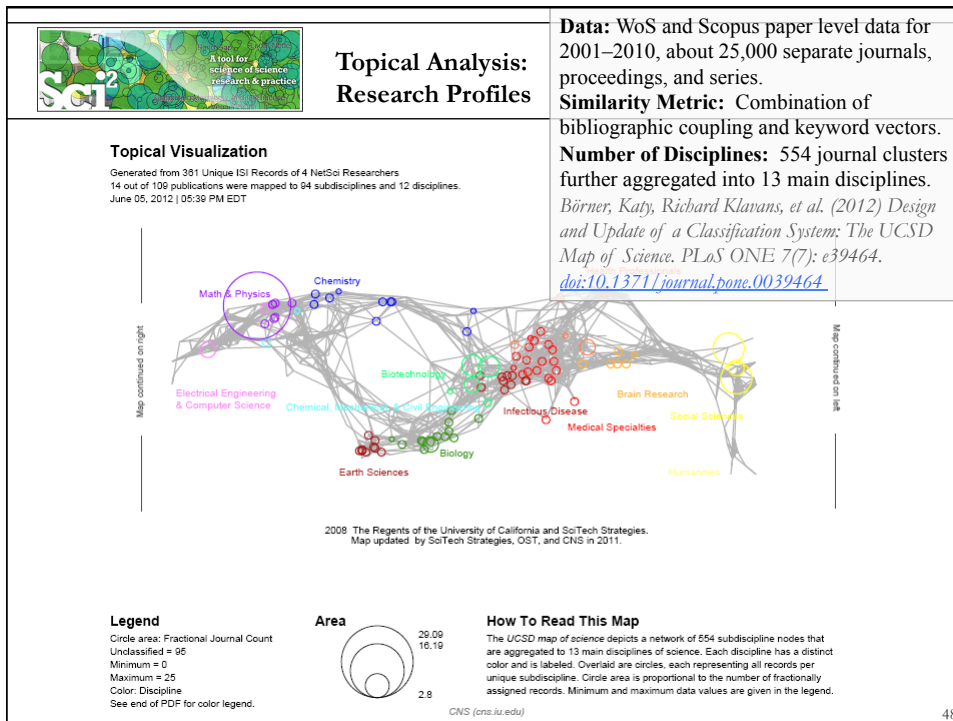
Rounded edges might increase legibility of overlapping lines.

46

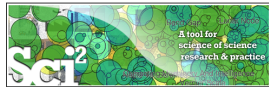
Tutorial Overview

- 8:30 Welcome and Overview of Tutorial and Attendees
- 9:00 The Sci2 Tool
 - Download and run the Sci2 Tool
 - ONE dataset, MANY analyses and visualizations
- 9:30 Sci2 Tool Workflows
 - Temporal Analysis: Horizontal line graph of NSF projects
 - Geospatial Analysis: US and world maps
 - Geospatial Analysis: Geomap with network overlays
 - **Topical Analysis: Visualize research profiles**
 - Network Analysis: Co-occurrence networks and bimodal networks
 - Network Analysis: Evolving collaboration networks
- 10:15 Networking Break**
- 10:30 Visualization Framework
- 11:00 IVMOOC – MANY more Workflows
- 11:15 Outlook and Q&A
- 11:30 Adjourn**

47

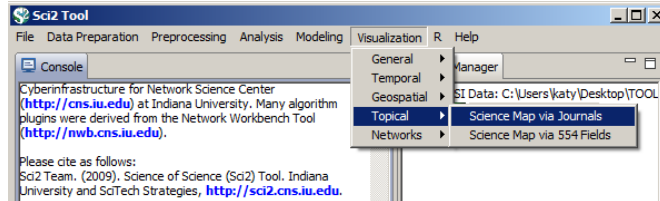


48



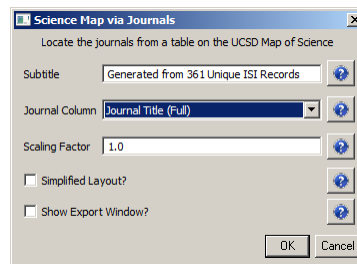
Research Profiles—Publication Data

Load an ISI (*.isi), Bibtex (*.bib), Endnote Export Format (*.enw), Scopus csv (*.scopus) file such as /sci2/sampledata/scientometrics/isi/FourNetSciResearchers.isi



Run 'Visualization > Topical > Science Map via Journals' using parameters given to the right.

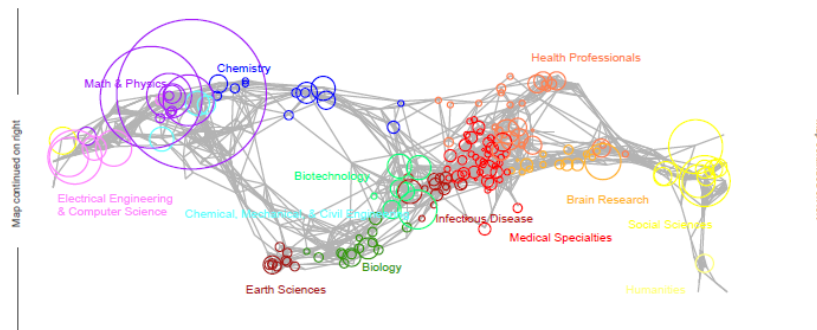
Postscript file will appear in Data Manager.
Save and open with a Postscript Viewer.



49

Topical Visualization

Generated from 361 Unique ISI Records
90 out of 112 publications were mapped to 182 subdisciplines and 13 disciplines.
June 24, 2012 | 04:04 PM EDT



© 2008 The Regents of the University of California and SoTech Strategies.
Map updated by SoTech Strategies, OST, and CNS in 2011.

Legend

Circle area: Fractional Journal Count
Unclassified = 22
Minimum = 0
Maximum = 98
Color: Discipline
See end of PDF for color legend.

Area



How To Read This Map

The UCSD map of science depicts a network of 554 subdiscipline nodes that are aggregated to 13 main disciplines of science. Each discipline has a distinct color and is labeled. Overlaid are circles, each representing all records per unique subdiscipline. Circle area is proportional to the number of fractionally assigned records. Minimum and maximum data values are given in the legend.

CNS (cns.iu.edu)

Topical Visualization

Generated from 361 Unique ISI Records
90 out of 112 publications were mapped to 182 subdisciplines and 13 disciplines.
June 24, 2012 | 04:04 PM EDT

Biology

- 1 BMC EVOLUTIONARY BIOLOGY
- 1 NATURWISSENSCHAFTEN

Biotechnology

- 1 BMC BIOINFORMATICS
- 2 FEBS JOURNAL
- 1 GENOME RESEARCH
- 1 INTERNATIONAL MICROBIOLOGY
- 1 NATURE BIOTECHNOLOGY
- 3 NATURE GENETICS
- 1 NATURE REVIEWS GENETICS
- 1 NUCLEIC ACIDS RESEARCH
- 2 PROTEOMICS

Brain Research

- 5 JOURNAL OF MATHEMATICAL PSYCHOLOGY

Chemical, Mechanical, & Civil Engineering

- 1 JOURNAL OF CERAMIC PROCESSING RESEARCH
- 2 MATERIALS SCIENCE AND ENGINEERING A--STRUCTURAL MATERIA...
- 1 PHYSICS WORLD
- 1 SCIENTIFIC AMERICAN

Chemistry

- 1 COMPUTER PHYSICS COMMUNICATIONS
- 2 JOURNAL OF CHEMICAL INFORMATION AND COMPUTER SCIENCES
- 1 JOURNAL OF THE INDIAN INSTITUTE OF SCIENCE
- 1 PURE AND APPLIED CHEMISTRY

Earth Sciences

- 1 CURRENT SCIENCE

Electrical Engineering & Computer Science

- 1 ASIST 2003: PROCEEDINGS OF THE 66TH ASIST ANNUAL MEETING...
- 1 CANADIAN JOURNAL OF INFORMATION AND LIBRARY SCIENCE-REV...
- 5 IEEE TRANSACTIONS ON PROFESSIONAL AND LIBRARY COMMUNICATION
- 1 INFORMATION TECHNOLOGY AND LIBRARIES
- 5 JOURNAL OF INFORMATION SCIENCE
- 3 JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE
- 5 JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENC...
- 2 LIBRARY QUARTERLY
- 1 LIBRI
- 1 PROCEEDINGS OF THE AMERICAN SOCIETY FOR INFORMATION SC...

Health Professionals

- 1 ANNALS OF BIOMEDICAL ENGINEERING
- 1 BULLETIN OF THE MEDICAL LIBRARY ASSOCIATION
- 1 CROATIAN MEDICAL JOURNAL
- 2 JOURNAL OF APPLIED PHYSIOLOGY
- 1 JOURNAL OF PUBLIC HEALTH DENTISTRY
- 1 METHODS OF INFORMATION IN MEDICINE
- 1 PLASTIC AND RECONSTRUCTIVE SURGERY
- 1 TEXAS MEDICINE
- 1 UNFALLCHIRURG
- 1 WIENER KLINISCHE WOCHENSCHRIFT

Humanities

- 1 BULLETIN OF THE ATOMIC SCIENTISTS

Infectious Diseases

- 1 FEMS MICROBIOLOGY LETTERS
- 1 JOURNAL OF BACTERIOLOGY

Math & Physics

- 1 ADVANCES IN APPLIED PROBABILITY

CNS (cns.iu.edu)

Topical Visualization

Generated from 361 Unique ISI Records
90 out of 112 publications were mapped to 182 subdisciplines and 13 disciplines.
June 24, 2012 | 04:04 PM EDT

Math & Physics

- 10 APPLIED PHYSICS LETTERS
- 1 BRAZILIAN JOURNAL OF PHYSICS
- 3 CHAOS SOLITONS & FRACTALS
- 1 COMPLEXITY
- 1 COMPUTATIONAL MATERIALS SCIENCE
- 11 EUROPEAN PHYSICAL JOURNAL B
- 12 EUROPHYSICS LETTERS
- 2 INTERNATIONAL JOURNAL OF MODERN PHYSICS B
- 6 JOURNAL OF PHYSICS A--MATHEMATICAL AND GENERAL
- 1 JOURNAL OF STATISTICAL MECHANICS--THEORY AND EXPERIMENT
- 1 JOURNAL OF STATISTICAL PHYSICS
- 1 JOURNAL OF THE KOREAN PHYSICAL SOCIETY
- 1 MATERIALS SCIENCE AND ENGINEERING B--SOLID STATE MATERIA...
- 3 NATURE PHYSICS
- 3 NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH SEC...
- 12 PHYSICA A
- 5 PHYSICAL REVIEW A
- 2 PHYSICAL REVIEW B
- 45 PHYSICAL REVIEW LETTERS
- 2 REVIEWS OF MODERN PHYSICS

Medical Specialties

- 1 ANNALS OF INTERNAL MEDICINE
- 1 REVISTA DE INVESTIGACION CLINICA

Social Sciences

- 1 ADMINISTRATIVE SCIENCE QUARTERLY
- 1 AMERICAN BEHAVIORAL SCIENTIST
- 1 AMERICAN SOCIOLOGICAL REVIEW
- 1 ANNALS OF THE AMERICAN ACADEMY OF POLITICAL AND SOCIAL S...
- 1 ARBOR--CIENCIA PENSAMIENTO Y CULTURA
- 3 BRITISH JOURNAL OF MATHEMATICAL & STATISTICAL PSYCHOLOGY
- 1 JOURNAL OF CLASSIFICATION

Social Sciences

- 2 JOURNAL OF MATHEMATICAL SOCIOLOGY
- 3 JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION
- 2 PSYCHOLOGICAL BULLETIN
- 5 PSYCHOMETRIKA
- 1 RECHERCHE
- 5 SCIENTOMETRICS
- 1 SOCIAL FORCES
- 6 SOCIAL NETWORKS
- 3 SOCIOLOGICAL METHODS & RESEARCH

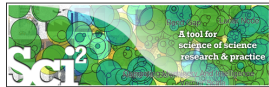
Multiple Categories

- 1 BRITISH MEDICAL JOURNAL
- 2 JAMA--JOURNAL OF THE AMERICAN MEDICAL ASSOCIATION
- 1 JOURNAL OF THEORETICAL BIOLOGY
- 18 NATURE
- 44 PHYSICAL REVIEW E
- 5 PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE ...
- 6 SCIENCE

Unclassified

- 1 ALGORITHMS AND MODELS FOR THE WEB--GRAPHS, PROCEEDINGS
- 2 AMERICAN DOCUMENTATION
- 2 ASIST 2002: PROCEEDINGS OF THE 65TH ASIST ANNUAL MEETING, ...
- 1 BIOLOGIYA MORYA--MARINE BIOLOGY
- 1 BULLETIN OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE
- 1 CHEMIKER-ZEITUNG
- 3 CHEMTECH
- 1 COMBINATORIAL AND ALGORITHMIC ASPECTS OF NETWORKING
- 7 CURRENT COMMENTS
- 3 CURRENT CONTENTS/LIFE SCIENCES
- 1 FEDERATION PROCEEDINGS
- 5 FRACTALS--AN INTERDISCIPLINARY JOURNAL ON THE COMPLEX GE...
- 1 FRONTIERS OF LIBRARIANSHIP--SYRACUSE UNIVERSITY

CNS (cns.iu.edu)



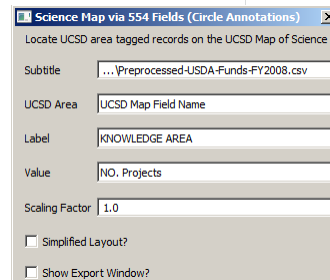
Research Profiles—Existing Classifications

In addition to using [journal names](#) to

- Map career trajectories
- Identify evolving expertise areas
- Compare expertise profiles

[Existing classifications](#) can be aligned and used to generate science map overlays.

	B	C	D	E	F	G
KNOWLEDGE AREA		NO. Projects	USDA Staff Years	STATE APPR	TOTAL FUNDS	UCSD Map Field Name
101 Appraisal of Soil Resources						315
102 Soil, Plant, Water, Nutrient Relationships						227
103 Management of Saline and Sodic Soils and Salinity						158
104 Protect Soil from Harmful Effects of Natural Elements						120
111 Conservation and Efficient Use of Water						245
112 Watershed Protection and Management						245
121 Management of Range Resources						520
122 Management and Control of Forest and Range Fires						520
123 Management and Sustainability of Forest Resources						231
124 Urban Forestry						231
125 Agroforestry						231

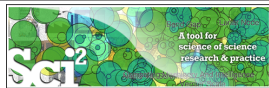


Run *Visualization > Topical > Science Map via 554 Fields* using parameters given to the right.

Postscript file will appear in *Data Manager*.

Save and open with a Postscript Viewer.

53



Align Science Basemaps using the Sci2 Tool

Topical Visualization

Legend

Area

How To Read This Map

UCSD Map

Loet et al science maps ISI categories

<http://vosviewer.com>

NIH Map
<https://app.nihmaps.org>

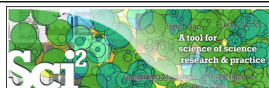
Elsevier's SciVal Map

Science-Matrix.com

54

Tutorial Overview

- 8:30 Welcome and Overview of Tutorial and Attendees
- 9:00 The Sci2 Tool
 - Download and run the Sci2 Tool
 - ONE dataset, MANY analyses and visualizations
- 9:30 Sci2 Tool Workflows
 - Temporal Analysis: Horizontal line graph of NSF projects
 - Geospatial Analysis: US and world maps
 - Geospatial Analysis: Geomap with network overlays
 - Topical Analysis: Visualize research profiles
 - **Network Analysis: Co-occurrence networks and bimodal networks**
 - Network Analysis: Evolving collaboration networks
- 10:15 Networking Break**
- 10:30 Visualization Framework
- 11:00 IVMOOC – MANY more Workflows
- 11:15 Outlook and Q&A
- 11:30 Adjourn**



General Network Extraction: Weighted, Undirected Co-Occurrence Network

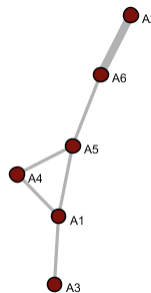
	A	B	C	D
1	Paper	Authors	References	Year
2	P1	A1		1970
3	P2	A2;A6	P1	1980
4	P3	A1;A3	P1;P2	1990
5	P4	A1;A4;A5	P2	1995
6	P5	A5;A6	P1;P2;P3;P4	1995
7	P6	A2;A6	P5	2000

Author co-occurrence network

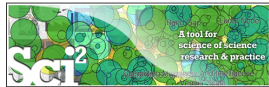
Extract Network from Table
Extracts a network from a delimited table

Column Name:

Text Delimiter:



- *Vertices 6
- 1 A1
- 2 A6
- 3 A2
- 4 A3
- 5 A5
- 6 A4
- *Edges 6
- 2 3 2
- 1 4 1
- 1 5 1
- 5 6 1
- 1 6 1
- 2 5 1



General Network Extraction: Unweighted, Directed Bipartite Network

	A	B	C	D
1	Paper	Authors	References	Year
2	P1	A1		1970
3	P2	A2;A6	P1	1980
4	P3	A1;A3	P1;P2	1990
5	P4	A1;A4;A5	P2	1995
6	P5	A5;A6	P1;P2;P3;P4	1995
7	P6	A2;A6	P5	2000

Paper-author bipartite (2-mode) network

Extract Bipartite Network

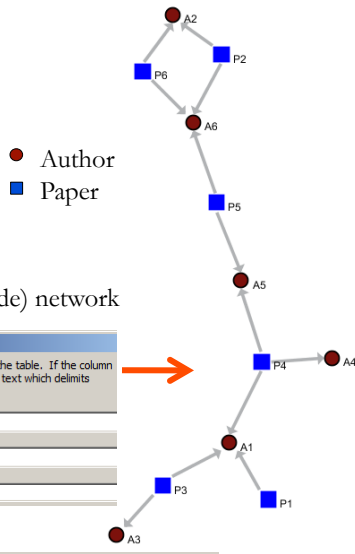
Extract a bipartite network from two columns in the table. If the column values may list multiple entries, enter the special text which delimits them.

First column:

Second column:

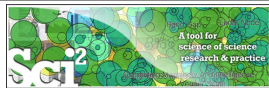
Text Delimiter:

Object: nodes based on -> Property: bipartitetype Operator: == Value: Paper



- *Vertices 12
- 1 P1 bipartitetype "Paper"
 - 2 A1 bipartitetype "Authors"
 - 3 P2 bipartitetype "Paper"
 - 4 A2 bipartitetype "Authors"
 - 5 A6 bipartitetype "Authors"
 - 6 P3 bipartitetype "Paper"
 - 7 A3 bipartitetype "Authors"
 - 8 P4 bipartitetype "Paper"
 - 9 A4 bipartitetype "Authors"
 - 10 A5 bipartitetype "Authors"
 - 11 P5 bipartitetype "Paper"
 - 12 P6 bipartitetype "Paper"
- *Arcs
- 1 2
 - 3 4
 - 3 5
 - 6 2
 - 6 7
 - 8 2
 - 8 10
 - 8 9
 - 11 5
 - 11 10
 - 12 4
 - 12 5

57



General Network Extraction: Unweighted, Directed Network

	A	B	C	D
1	Paper	Authors	References	Year
2	P1	A1		1970
3	P2	A2;A6	P1	1980
4	P3	A1;A3	P1;P2	1990
5	P4	A1;A4;A5	P2	1995
6	P5	A5;A6	P1;P2;P3;P4	1995
7	P6	A2;A6	P5	2000

Extract Directed Network

Given a table, this algorithm extracts a directed edge that starts at a column node.

Source Column:

Target Column:

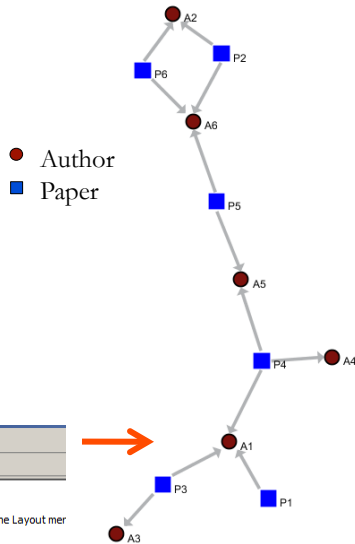
Text Delimiter:

Analysis Modeling Visualization R Help

Temporal Geospatial Topical

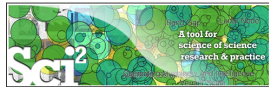
Networks

- Network Analysis Toolkit (NAT)
- Unweighted & Undirected
- Weighted & Undirected
- Unweighted & Directed**



- *Vertices 12
- 1 P1 indegree 0
 - 2 A1 indegree 3
 - 3 P2 indegree 0
 - 4 A2 indegree 2
 - 5 A6 indegree 3**
 - 6 P3 indegree 0
 - 7 A3 indegree 1
 - 8 P4 indegree 0
 - 9 A4 indegree 1
 - 10 A5 indegree 2
 - 11 P5 indegree 0
 - 12 P6 indegree 0
- *Arcs
- 1 2
 - 3 4
 - 3 5
 - 6 2
 - 6 7
 - 8 10
 - 8 2
 - 8 9
 - 11 10
 - 11 5
 - 12 4
 - 12 5

58



General Network Extraction: Unweighted, Directed Paper-Citation Network

	A	B	C	D
1	Paper	Authors	References	Year
2	P1	A1		1970
3	P2	A2;A6	P1	1980
4	P3	A1;A3	P1;P2	1990
5	P4	A1;A4;A5	P2	1995
6	P5	A5;A6	P1;P2;P3;P4	1995
7	P6	A2;A6	P5	2000

- *Vertices 6
- 1 P1
- 2 P2
- 3 P3
- 4 P4
- 5 P5
- 6 P6
- *Arcs
- 2 1
- 3 1
- 3 2
- 4 2
- 5 4
- 5 3
- 5 1
- 5 2
- 6 5

Extract Directed Network

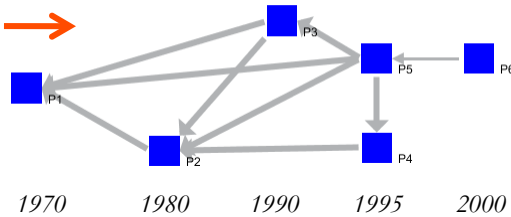
Given a table, this algorithm directed edge that starts a column node.

Source Column:

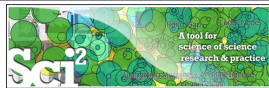
Target Column:

Text Delimiter:

Arcs from papers to references



59



General Network Extraction: Unweighted, Directed Bi-Partite Network

	A	B	C	D
1	Paper	Authors	References	Year
2	P1	A1		1970
3	P2	A2;A6	P1	1980
4	P3	A1;A3	P1;P2	1990
5	P4	A1;A4;A5	P2	1995
6	P5	A5;A6	P1;P2;P3;P4	1995
7	P6	A2;A6	P5	2000

- *Vertices 11
- 1 P1 bipartitetype "Paper"
- 2 P2 bipartitetype "Paper"
- 3 P1 bipartitetype "References"
- 4 P3 bipartitetype "Paper"
- 5 P2 bipartitetype "References"
- 6 P4 bipartitetype "Paper"
- 7 P5 bipartitetype "Paper"
- 8 P4 bipartitetype "References"
- 9 P3 bipartitetype "References"
- 10 P6 bipartitetype "Paper"
- 11 P5 bipartitetype "References"
- *Arcs
- 2 3
- 4 3
- 4 5
- 6 5
- 7 3
- 7 9
- 7 5
- 7 8
- 10 11

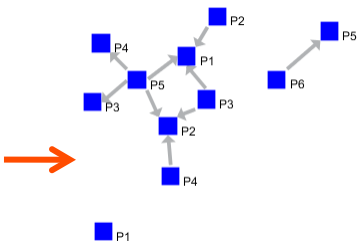
Extract Bipartite Network

Extract a bipartite network values may list multiple ent them.

First column:

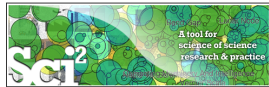
Second column:

Text Delimiter:



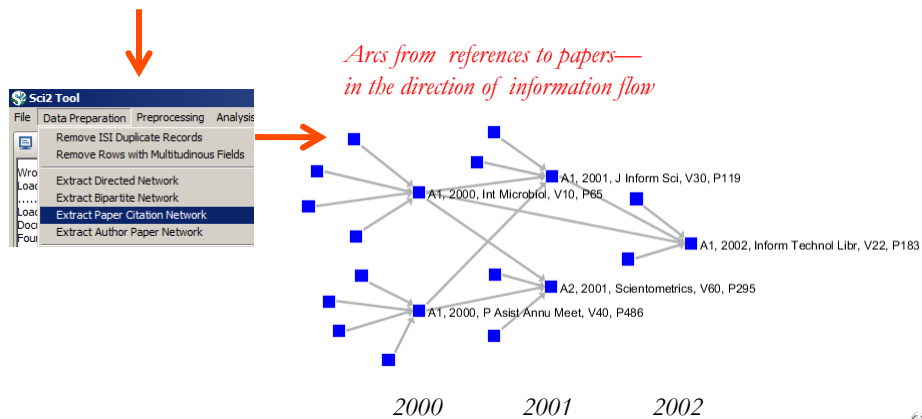
WRONG!!!

60



ISI Paper-Citation Network Extraction

	A	B	C	D	E
1	Authors	Cited References	Publication Year	Title	Cite Me As
2	A1 A2	BENSMAN SJ, 1998, LIBR RESOUR TECH SER, V42, P147 BRO	2000 T1	A1, 2000, INT MICROBIOL, V10, P65	
3	A1	BENSMAN SJ, 1999, LIBR RESOUR TECH SER, V42, P147 BRO	2000 T2	A1, 2000, P ASIST ANNU MEET, V40, P486	
4	A2 A3	GARFIELD E, 1985, ESSAYS INFORMATION S, V8, P403 GILBE	2001 T3	A2, 2001, SCIENTOMETRICS, V60, P295	
5	A1	ASIMOV A, 1963, GENETIC CODE LEDERBERG J, 1972, NATU	2001 T4	A1, 2001, J INFORM SCI, V30, P119	
6	A1 A2	AVERY OT, 1944, J EXP MED, V79, P137 SMALL H, 1985, J INF	2002 T5	A1, 2002, INFORM TECHNOL LIBR, V22, P183	



61



Tutorial Overview

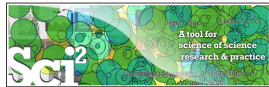
- 8:30 Welcome and Overview of Tutorial and Attendees
- 9:00 The Sci2 Tool
 - Download and run the Sci2 Tool
 - ONE dataset, MANY analyses and visualizations
- 9:30 Sci2 Tool Workflows
 - Temporal Analysis: Horizontal line graph of NSF projects
 - Geospatial Analysis: US and world maps
 - Geospatial Analysis: Geomap with network overlays
 - Topical Analysis: Visualize research profiles
 - Network Analysis: Co-occurrence networks and bimodal networks
 - **Network Analysis: Evolving collaboration networks**

10:15 Networking Break

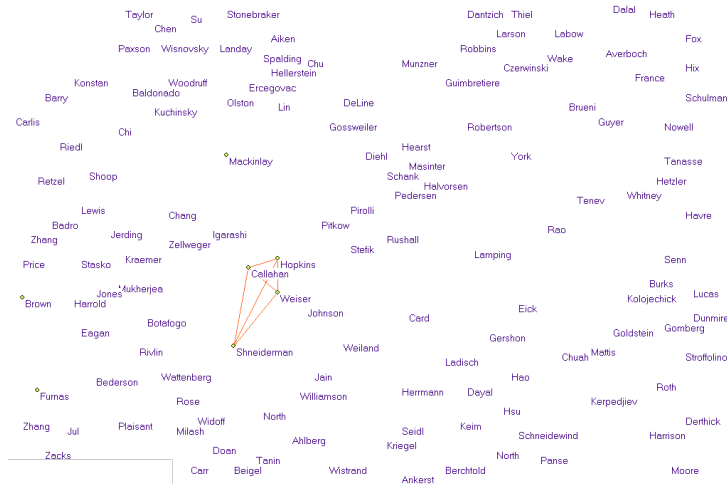
- 10:30 Visualization Framework
- 11:00 IVMOOC – MANY more Workflows
- 11:15 Outlook and Q&A

11:30 Adjourn

62

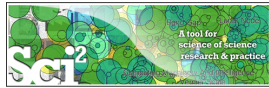


Evolving collaboration networks



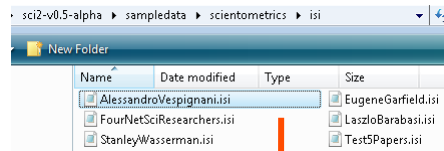
Mapping the Evolution of Co-Authorship Networks
Weimao Ke, Lalitha Visvanath & Katy Börner
InfoVis Lab @ Indiana University
2004

63



Evolving Collaboration Networks

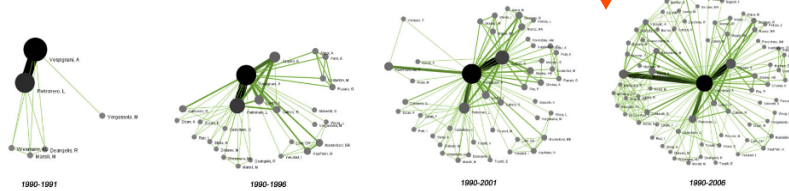
Load isi formatted file



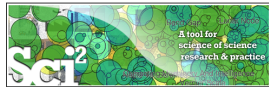
As csv, file looks like:

	A	B	C	D	E	F	G
1	Abstract	Authors	Authors (Full Names)	Beginning	Book Serie	Book Serie	Cited Pate
2	The systematic study of Colizza, V Barrat, A Barthelemy, M Vespignani, A			2015			
3	Uncovering the hidden ri Colizza, V Flammini, A Serrano, M Vespignani, A			110			
4	Computer viruses can s Vespignani, A			135			
5	Mapping the Internet gei Dall'Asta, L Alvarez-Hamelin, I Barrat, A Yaquez, A Vespignani, A			140			LECTURE NOTES IN

Visualize each time slide separately:



64



Relevant Sci2 Manual entry

- Home
- 1 Introduction
- 2 Getting Started
- 3 Algorithms, Tools, and Plugins
- 4 Workflow Design
- 5 Sample Workflows
 - 5.1 Individual Level Studies - Micro
 - 5.1.1 Mapping Collaboration, Publication, and Funding Profiles of One Researcher (EndNote and NSF Data)
 - 5.1.2 Time Slicing of Co-Authorship Networks (ISI Data)
 - 5.1.3 Funding Profiles of Three Researchers at Indiana University (NSF Data)
 - 5.1.4 Studying Four Major Theoretical Researchers (ISI Data)
 - 5.2 Institution Level Studies - Meso
 - 5.3 Global Level Studies - Macro
- 6 Sample Science Studies & Online Services
- 7 Extending the Sci2 Tool
- 8 Relevant Datasets and Tools
- 9 References



5.1.2 Time Slicing of Co-Authorship Networks (ISI Data)

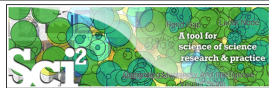
Added by Ted Polley, last edited by Scott Weingart on Mar 16, 2011 (view change)

AlessandroVespignani.isi	
Time frame:	1990-2006
Region(s):	Indiana University, University of Rome, Yale University, Leiden University, International Center for Theoretical Physics, University of Paris-Sud
Topical Area(s):	Informatics, Complex Network Science and System Research, Physics, Statistics, Epidemics
Analysis Type(s):	Co-Authorship Network

The Sci2 Tool supports the analysis of evolving networks. For this study, load Alessandro Vespignani's publication history from ISI, which can be downloaded from Thomson's Web of Science or loaded using 'File > Load' and following this path: 'yoursci2directory/sampledata/scientometrics/Asu/AlessandroVespignani/asi'. Slice the data into five year intervals from 1990-2006 using 'Preprocessing > Temporal > Slice Table by Time' and the following parameters:

[http://sci2.wiki.cns.inu.edu/5.1.2+Time+Slicing+of+Co-Authorship+Networks+\(ISI+Data\)](http://sci2.wiki.cns.inu.edu/5.1.2+Time+Slicing+of+Co-Authorship+Networks+(ISI+Data))

65



Slice Table by Time

"Slice Into" allows the user to slice the table by days, weeks, months, quarters, years, decades, and centuries. There are two additional parameters for time slicing: cumulative and align with calendar. The former produces tables containing all data from the beginning to the end of each table's time interval, which can be seen in the Data Manager and below.

- 101 Unique ISI Records
- slice from beginning of 1990 to end of 2006 (101 records)
- slice from beginning of 1990 to end of 2001 (65 records)
- slice from beginning of 1990 to end of 1996 (26 records)
- slice from beginning of 1990 to end of 1991 (4 records)

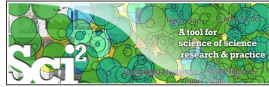
The latter option aligns the output tables according to calendar intervals:

- 101 Unique ISI Records
- slice from beginning of 2002 to end of 2006 (36 records)
- slice from beginning of 1997 to end of 2001 (39 records)
- slice from beginning of 1992 to end of 1996 (22 records)
- slice from beginning of 1990 to end of 1991 (4 records)

Choosing "Years" under "Slice Into" creates multiple tables beginning from January 1st of the first year. If "Months" is chosen, it will start from the first day of the earliest month in the chosen time interval.

[http://sci2.wiki.cns.inu.edu/5.1.2+Time+Slicing+of+Co-Authorship+Networks+\(ISI+Data\)](http://sci2.wiki.cns.inu.edu/5.1.2+Time+Slicing+of+Co-Authorship+Networks+(ISI+Data))

66

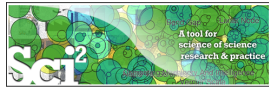


Visualize Each Network, Keep Node Positions

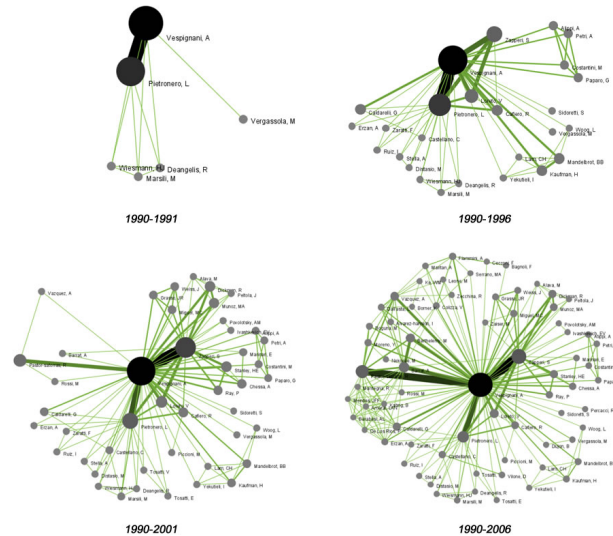
1. To see the evolution of Vespignani's co-authorship network over time, check 'cumulative'.
2. Extract co-authorship networks one at a time for each sliced time table using 'Data Preparation > Extract Co-Author Network', making sure to select "ISI" from the pop-up window during the extraction.
3. To view each of the Co-Authorship Networks over time using the same graph layout, begin by clicking on longest slice network (the 'Extracted Co-Authorship Network' under 'slice from beginning of 1990 to end of 2006 (101 records)') in the data manager. Visualize it in GUESS using 'Visualization > Networks > GUESS'.
4. From here, run 'Layout > GEM' followed by 'Layout > Bin Pack'. Run 'Script > Run Script ...' and select 'yoursci2directory/scripts/GUESS/co-author-nw.py'.
5. In order to save the x, y coordinates of each node and to apply them to the other time slices in GUESS, select 'File > Export Node Positions' and save the result as 'yoursci2directory/NodePositions.csv'. Load the remaining three networks in GUESS using the steps described above and for each network visualization, run 'File > Import Node Positions' and open 'yoursci2directory/NodePositions.csv'.
6. To match the resulting networks stylistically with the original visualization, run 'Script > Run Script ...' and select 'yoursci2directory/scripts/GUESS/co-author-nw.py', followed by 'Layout > Bin Pack', for each.

[http://sci2.wiki.cns.inu.edu/5.1.2+Time+Slicing+of+Co-Authorship+Networks+\(ISI+Data\)](http://sci2.wiki.cns.inu.edu/5.1.2+Time+Slicing+of+Co-Authorship+Networks+(ISI+Data))

67

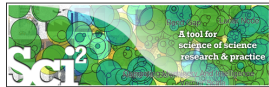


Visualize Each Network, Keep Node Positions



[http://sci2.wiki.cns.inu.edu/5.1.2+Time+Slicing+of+Co-Authorship+Networks+\(ISI+Data\)](http://sci2.wiki.cns.inu.edu/5.1.2+Time+Slicing+of+Co-Authorship+Networks+(ISI+Data))

68



Relevant CShell plugin



Slice Table by Time

Tools ▾

Added by [Aurelia Alencar](#), last edited by [Ted Polley](#) on Jan 12, 2011 ([view changes](#))

Description

Slice Table By Time is an algorithm to chop a table up into new tables, based on a date/time column. It takes the column with the date/time data, a string describing the format of that column, the intervals that the data should be sliced into, whether or not the slices are cumulative, whether or not the slices should be aligned with the calendar, and what day the week is considered to start on (which only matters if the slices are aligned with the calendar) as parameters.

The column to use for date/time values should have a single value for each row of data. It is used by the algorithm to choose which slice(s) the row should end up in. In order to determine what date/time is represented by that row, you must provide the algorithm with a descriptive format, in the second parameter. For instance, a four digit year would be represented by yyyy (the default value). See <http://joda-time.sourceforge.net/api-release/org/joda/time/format/DateTimeFormat.html> for details of all the various formatting options.

The next dropdown has the available intervals to slice the table into. These include milliseconds, seconds, minutes, hours, days, weeks, fortnights, months, quarters, years, decades, and centuries. A future version of the algorithm may include the ability to select how many of these intervals should be grouped together at once.

The checkbox that follows determines if the slices will be cumulative. If the slices are not cumulative, every row in the original table is in one and only one resulting slice. However, if the slices are cumulative, every row in the original table is in the slice it is for and every slice for a period after that.

The checkbox that follows determines if the slices will be aligned with the calendar. For instance, if the first row is for June 7th, 2006 and yearly slices are chosen, then the default behavior will be to have the first slice be from June 7th, 2006 to June 6th, 2007. However, if the slices are aligned with the calendar, the first slice will be from January 1st, 2006 to December 31st, 2006. Alignment does not affect the output for intervals of fortnights, quarters, decades, or milliseconds.

If the slices are aligned with the calendar and are weekly, then the day the week starts is used to determine how they are aligned.

Pros & Cons

The output of the slice algorithm is in separate tables, so a longitudinal analysis will require working with each slice separately, which can be awkward. There will likely be future versions of the time slice algorithm that annotate the original table with the slice the rows belong to.

Applications

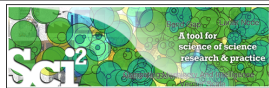
When doing longitudinal analysis of data, it can be useful to consider it in chunks, such as to calculate how statistics have changed over time. Alternatively, only a particular time period might be of interest, and this algorithm can extract it from data for a larger time range.

Implementation Details

This algorithm uses the Joda Time library extensively, which provides significantly improved capabilities compared to the default Java algorithms for dates and times.

<http://cishell.wiki.cns.iu.edu/Slice+Table+by+Time>

69




Network Visualization with GUESS

Pan:


“grab” the background by holding left-click and moving your mouse.

Zoom:

Using scroll wheel, press the “+” and “-” buttons in the upper-left hand corner, or right-click and move the mouse left or right. Center graph by selecting ‘View -> Center’.

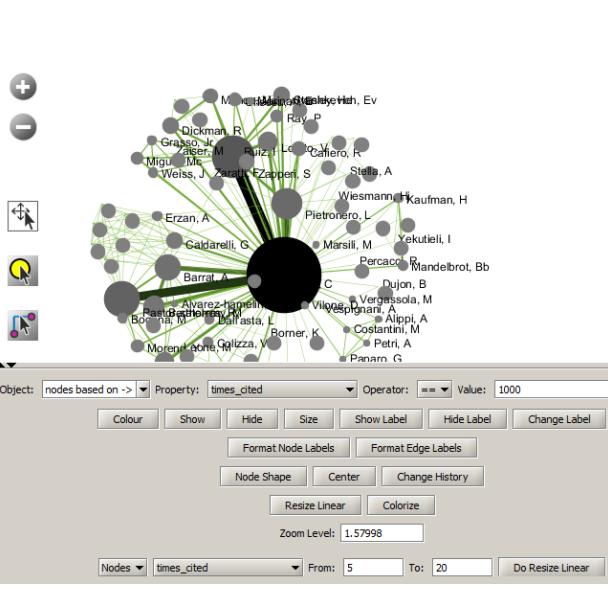
Select  to select/move single nodes. Hold down ‘Shift’ to select multiple.

Right click node/edge to modify Color, Shape, etc.



A tool for science of science research & practice

Network Visualization with GUESS



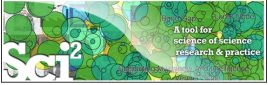
Graph Modifier:

Select “all nodes” in the Object drop-down menu and click ‘Show Label’ button.

Select ‘Resize Linear > Nodes > times_cited’ drop-down menu, then type “5” and “20” into the From” and To” Value box separately. Then select ‘Do Resize Linear’.

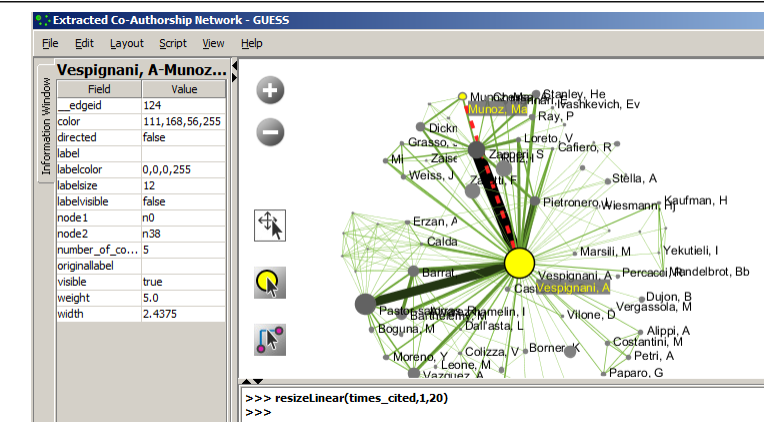
Select ‘Colorize> Nodes>totalities’, then select white and enter (204,0,51) in the pop-up color boxes on in the “From” and “To” buttons.

Select “Format Node Labels”, replace default text {originallabel} with your own label in the pop-up box ‘Enter a formatting string for node labels.’



A tool for science of science research & practice

Network Visualization with GUESS



Interpreter uses Jython a combination of Java and Python.

Try

```
resizeLinear(times_cited,1,20)
colorize(times_cited, white, red)
```

BREAK

73



Tutorial Overview

8:30 Welcome and Overview of Tutorial and Attendees

9:00 The Sci2 Tool

- Download and run the Sci2 Tool
- ONE dataset, MANY analyses and visualizations

9:30 Sci2 Tool Workflows

- Temporal Analysis: Horizontal line graph of NSF projects
- Geospatial Analysis: US and world maps
- Geospatial Analysis: Geomap with network overlays
- Topical Analysis: Visualize research profiles
- Network Analysis: Co-occurrence networks and bimodal networks
- Network Analysis: Evolving collaboration networks

10:15 Networking Break

10:30 Visualization Framework

11:00 IVMOOC – MANY more Workflows

11:15 Outlook and Q&A

11:30 Adjourn

74

Visualization Framework

75

Theoretically Grounded and Practically Useful Visualization Framework

developed to empower the broadest spectrum of users to read and make data visualizations that are useful and meaningful to them.

The visualization framework was used to

- design the aforementioned study and
- develop plug-and-play macroscope tools that improve the data visualization literacy of researchers, practitioners, IVMOOC students, museum visitors, and others.

Börner, Katy. 2015. *Atlas of Knowledge: Anyone Can Map*. The MIT Press.
<http://scimaps.org/atlas2>

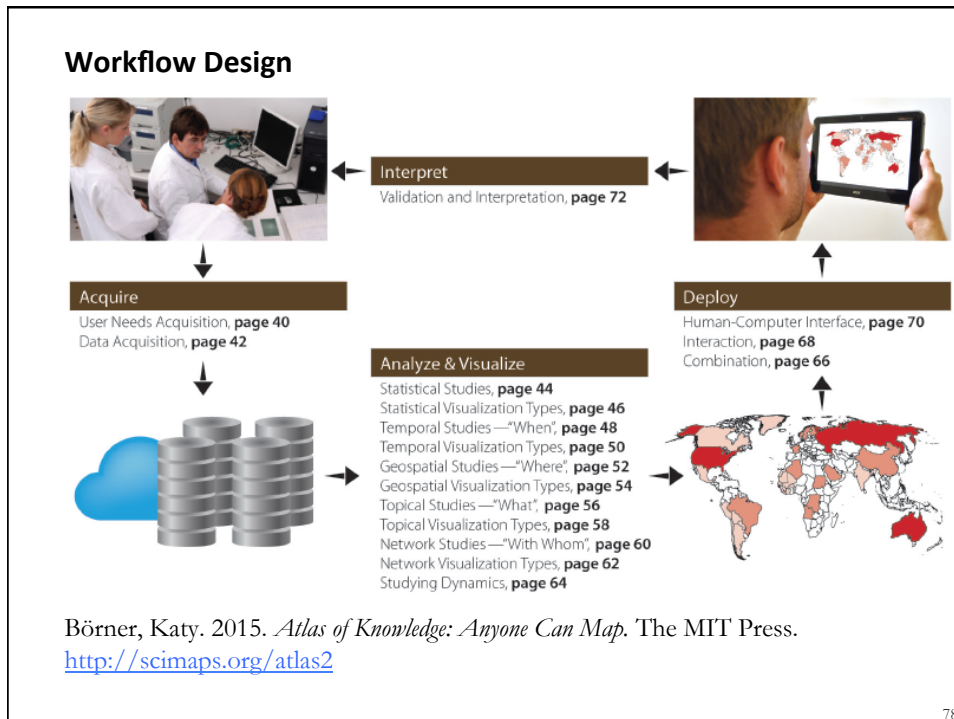
76

Tasks	LEVELS		
	MICRO: Individual Level about 1–1,000 records page 6	MESO: Local Level about 1,001–100,000 records page 8	MACRO: Global Level more than 100,000 records page 10
TYPES			
Statistical Analysis page 44			
WHEN: Temporal Analysis page 48			
WHERE: Geospatial Analysis page 52			
WHAT: Topical Analysis page 56			
WITH WHOM: Network Analysis page 60			

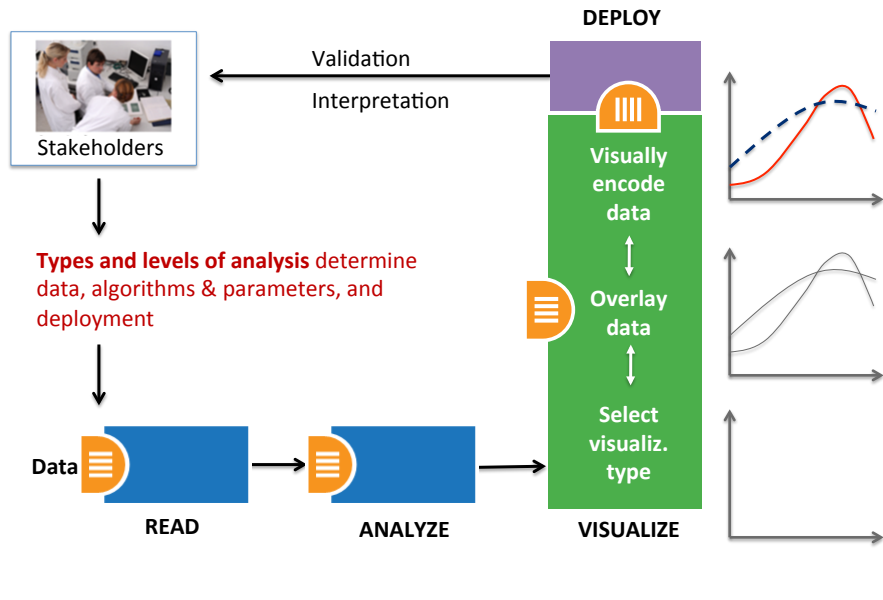
Atlas of Knowledge
Anyone Can Map

See page 5

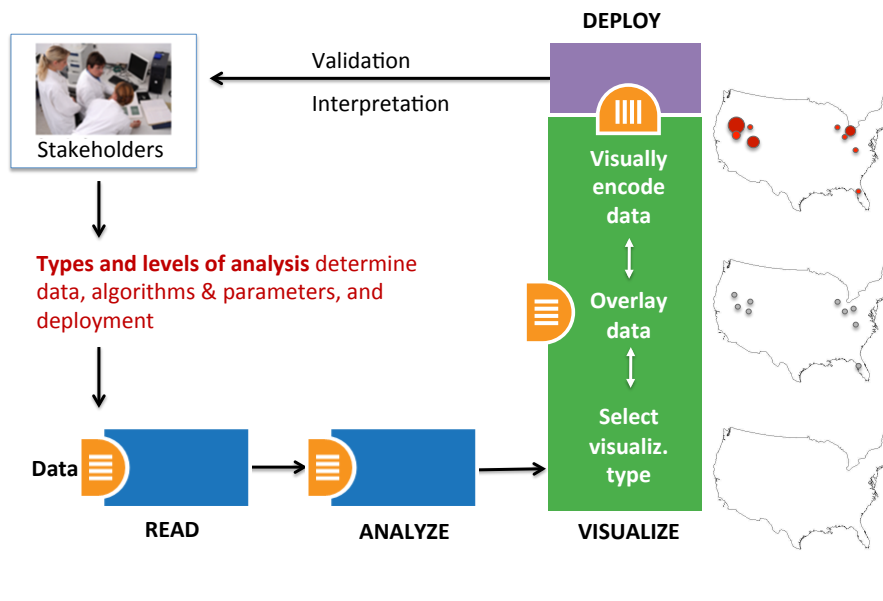
77



Needs-Driven Workflow Design



Needs-Driven Workflow Design



Types

Insight Need Types page 26	Data Scale Types page 28	Visualization Types page 30	Graphic Symbol Types page 32	Graphic Variable Types page 34	Interaction Types page 26
<ul style="list-style-type: none"> • categorize/cluster • order/rank/sort • distributions (also outliers, gaps) • comparisons • trends (process and time) • geospatial • compositions (also of text) • correlations/relationships 	<ul style="list-style-type: none"> • nominal • ordinal • interval • ratio 	<ul style="list-style-type: none"> • table • chart • graph • map • network layout 	<ul style="list-style-type: none"> • geometric symbols <ul style="list-style-type: none"> point line area surface volume • linguistic symbols <ul style="list-style-type: none"> text numerals punctuation marks • pictorial symbols <ul style="list-style-type: none"> images icons statistical glyphs 	<ul style="list-style-type: none"> • spatial <ul style="list-style-type: none"> position • retinal <ul style="list-style-type: none"> form color optics motion 	<ul style="list-style-type: none"> • overview • zoom • search and locate • filter • details-on-demand • history • extract • link and brush • projection • distortion

Börner, Katy. 2015. *Atlas of Knowledge: Anyone Can Map*. The MIT Press.
<http://scimaps.org/atlas2>


81

Types

Basic Task Types								
Bertin, 1967	Wehrend & Lewis, 1996	Few, 2004	Yau, 2011	Rendgen & Wiedemann, 2012	Frankel, 2012	Tool: Many Eyes	Tool: Chart Chooser	Börner, 2014
selection	categorize			category				categorize/cluster
order	rank	ranking					table	order/rank/sort
	distribution	distribution					distribution	distributions (also outliers, gaps)
	compare	nominal comparison & deviation	differences		compare and contrast	compare data values	comparison	comparisons
		time series	patterns over time	time	process and time	track rises and falls over time	trend	trends (process and time)
		geospatial	spatial relations	location		generate maps		geospatial
quantity		part-to-whole	proportions		form and structure	see parts of whole, analyze text	composition	compositions (also of text)
association	correlate	correlation	relationships	hierarchy		relations between data points	relationship	correlations/relationships

82

Insight Need Types page 26	Data Scale Types page 28	Visualization Types page 30	Graphic Symbol Types page 32	Graphic Variable Types page 34	Interaction Types page 26
<ul style="list-style-type: none"> • categorize/cluster • order/rank/sort • distributions (also outliers, gaps) • comparisons • trends (process and time) • geospatial • compositions (also of text) • correlations/relationships 	<ul style="list-style-type: none"> • nominal • ordinal • interval • ratio 	<ul style="list-style-type: none"> • table • chart • graph • map • network layout 	<ul style="list-style-type: none"> • geometric symbols <ul style="list-style-type: none"> point line area surface volume • linguistic symbols <ul style="list-style-type: none"> text numerals punctuation marks • pictorial symbols <ul style="list-style-type: none"> images icons statistical glyphs 	<ul style="list-style-type: none"> • spatial <ul style="list-style-type: none"> position • retinal <ul style="list-style-type: none"> form color optics motion 	<ul style="list-style-type: none"> • overview • zoom • search and locate • filter • details-on-demand • history • extract • link and brush • projection • distortion


See page 24
83

Visualization Types (Reference Systems)

- 1. Charts:** No reference system—e.g., Wordle.com, pie charts
- 2. Tables:** Categorical axes that can be selected, reordered; cells can be color coded and might contain proportional symbols. Special kind of graph.
- 3. Graphs:** Quantitative or qualitative (categorical) axes. Timelines, bar graphs, scatter plots.
- 4. Geospatial maps:** Use latitude and longitude reference system. World or city maps.
- 5. Network layouts:** Node position might depends on node attributes or node similarity. **Trees:** hierarchies, taxonomies, genealogies. **Networks:** social networks, migration flows.

Types

Insight Need Types page 26	Data Scale Types page 28	Visualization Types page 30	Graphic Symbol Types page 32	Graphic Variable Types page 34	Interaction Types page 26
<ul style="list-style-type: none"> categorize/cluster order/rank/sort distributions (also outliers, gaps) comparisons trends (process and time) geospatial compositions (also of text) correlations/relationships 	<ul style="list-style-type: none"> nominal ordinal interval ratio 	<ul style="list-style-type: none"> table chart graph map network layout 	<ul style="list-style-type: none"> geometric symbols <ul style="list-style-type: none"> point line area surface volume linguistic symbols <ul style="list-style-type: none"> text numerals punctuation marks pictorial symbols <ul style="list-style-type: none"> images icons statistical glyphs 	<ul style="list-style-type: none"> spatial position retinal <ul style="list-style-type: none"> form color optics motion 	<ul style="list-style-type: none"> overview zoom search and locate filter details-on-demand history extract link and brush projection distortion

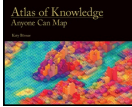
Börner, Katy. 2015. *Atlas of Knowledge: Anyone Can Map*. The MIT Press.
<http://scimaps.org/atlas2>

Graphic Variable Types Versus Graphic Symbol Types

			Point	Line	Geometric Symbols Area
Spatial	x	quantitative			
	y	quantitative			
	z	quantitative			
Retinal	Form	Size	quantitative	NA (Not Applicable)	
		Shape	qualitative	NA	
		Rotation	quantitative	NA	
		Curvature	quantitative	NA	
		Angle	quantitative	NA	
		Closure	quantitative	NA	
Color	Value	quantitative			
	Hue	qualitative			
	Saturation	quantitative			

Graphic Variable Types Versus Graphic Symbol Types

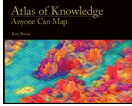
		Geometric Symbols			
		Point	Line	Area	
Spatial	x	quantitative			
	y	quantitative			
	z	quantitative			
Form	Size	quantitative	NA (Not Applicable)		
	Shape	qualitative	NA		
	Rotation	quantitative	NA		
	Curvature	quantitative	NA		
	Angle	quantitative	NA		
	Closure	quantitative	NA		
Color	Value	quantitative			
	Hue	qualitative			
	Saturation	quantitative			



See page 36

Graphic Variable Types Versus Graphic Symbol Types

		Geometric Symbols				Linguistic Symbols				Pictorial Symbols						
		Point	Line	Area	Surface	Volume	Text, Numbers, Punctuation Marks	Text	Text	Text	Text	Images, Icons, Qualitative Graphs	Text	Text	Text	Text
Spatial	x	quantitative														
	y	quantitative														
	z	quantitative														
Form	Size	quantitative	NA (Not Applicable)													
	Shape	qualitative	NA													
	Rotation	quantitative	NA													
	Curvature	quantitative	NA													
	Angle	quantitative	NA													
	Closure	quantitative	NA													
Color	Value	quantitative														
	Hue	qualitative														
	Saturation	quantitative														



See pages 36-39

Tutorial Overview

8:30 Welcome and Overview of Tutorial and Attendees

9:00 The Sci2 Tool

- Download and run the Sci2 Tool
- ONE dataset, MANY analyses and visualizations

9:30 Sci2 Tool Workflows

- Temporal Analysis: Horizontal line graph of NSF projects
- Geospatial Analysis: US and world maps
- Geospatial Analysis: Geomap with network overlays
- Topical Analysis: Visualize research profiles
- Network Analysis: Co-occurrence networks and bimodal networks
- Network Analysis: Evolving collaboration networks

10:15 Networking Break

10:30 Visualization Framework

11:00 IVMOOC – MANY more Workflows

11:15 Outlook and Q&A

11:30 Adjourn

89

Information Visualization MOOC (IVMOOC)
Teaches the
Visualization Framework and the Sci2 Tool

90



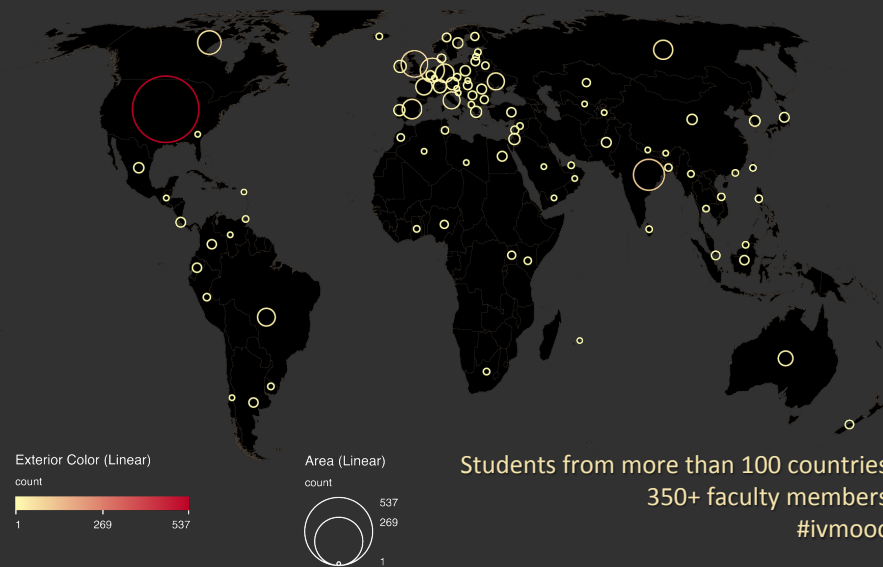
IVMOOC 2016

MENU



Register for free: <http://ivmooc.cns.iu.edu>. Class started Jan 12, 2016.

The Information Visualization MOOC
ivmooc.cns.iu.edu



92

Course Schedule

Part 1: Theory and Hands-On

- **Session 1** – Workflow Design and Visualization Framework
- **Session 2** – “When:” Temporal Data
- **Session 3** – “Where:” Geospatial Data
- **Session 4** – “What:” Topical Data

Mid-Term

- **Session 5** – “With Whom:” Trees
- **Session 6** – “With Whom:” Networks
- **Session 7** – Dynamic Visualizations and Deployment

Final Exam

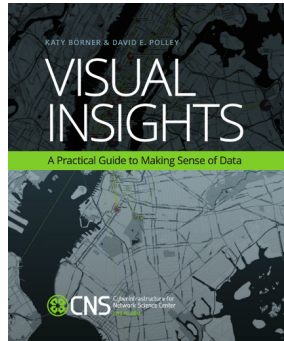
Part 2: Students work in teams on client projects.

Final grade is based on Class Participation (10%), Midterm (30%), Final Exam (30%), and Client Project(30%).



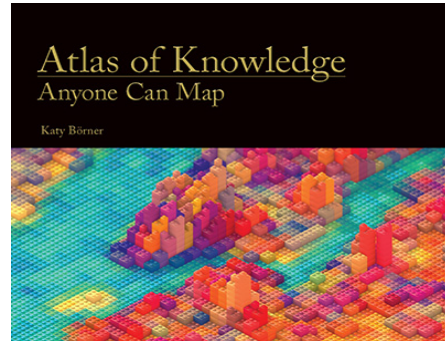
93

Books Used in the IVMOOC



Teaches timely knowledge:

Advanced algorithms, tools, and hands-on workflows.

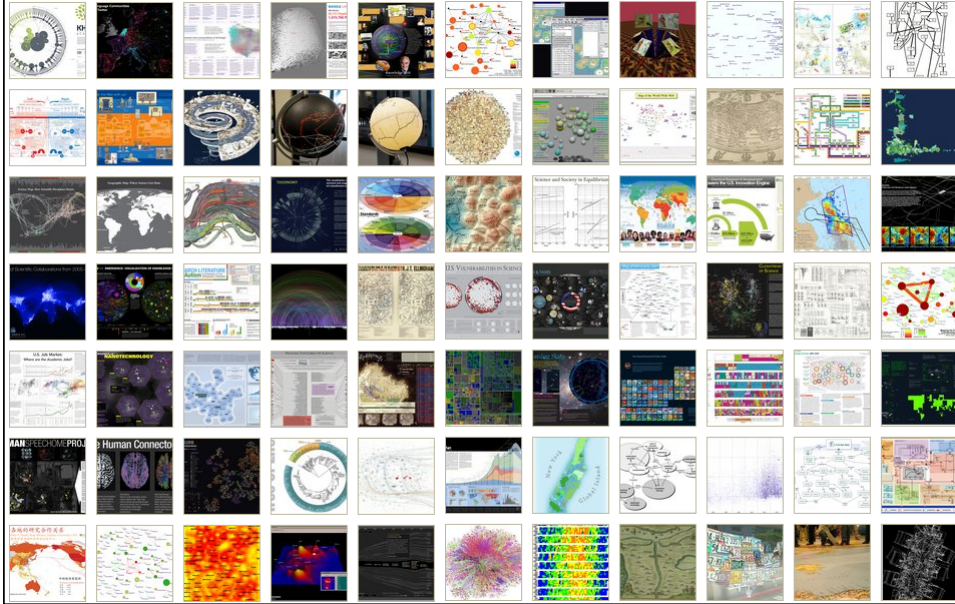


Teaches timeless knowledge:

Visualization framework—exemplified using generic visualization examples and pioneering visualizations.

94

Visualization Frameworks



How to Classify Different Visualizations?

By

- User insight needs?
- User task types?
- Data to be visualized?
- Data transformation?
- Visualization technique?
- Visual mapping transformation?
- Interaction techniques?
- Or ?

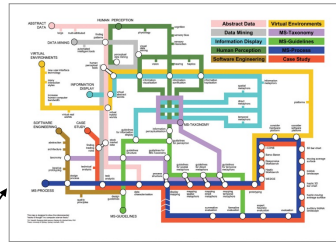


Different Question Types



Terabytes of data

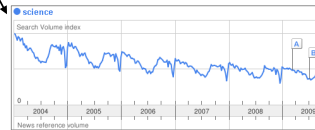
Descriptive & Predictive Models



Find your way



Find collaborators, friends

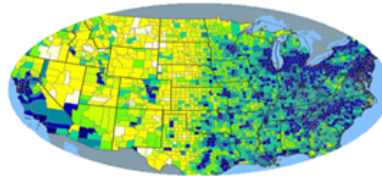


Identify trends

97

Different Levels of Abstraction/Analysis

Macro/Global
Population Level



Meso/Local
Group Level



Micro
Individual Level





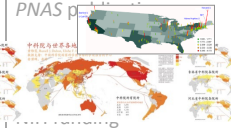




98

Type of Analysis vs. Level of Analysis

	Micro/Individual (1-100 records)	Meso/Local (101–10,000 records)	Macro/Global (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 <i>PNAS</i>	113 years of physics research
Geospatial Analysis (Where)	Career trajectory of one individual	Mapping a state's intellectual landscape	<i>PNAS</i> 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

99

Type of Analysis vs. Level of Analysis

	Micro/Individual (1-100 records)	Meso/Local (101–10,000 records)	Macro/Global (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 <i>PNAS</i>	113 years of physics research 
Geospatial Analysis (Where)	Career trajectory of one individual	Mapping a state's intellectual landscape	<i>PNAS</i> 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 

100

Clients

Information Visualization MOOC



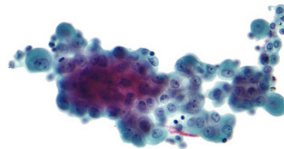
List of Clients

ISIS
Project Title: Isis: 100 Years
Client Name: Jay Malone
Project goal/scientific or practical value: A visual representation Isis' contributors and locales over the past 100 years. Isis is the journal of the History of Science Society. This representation will provide a dynamic picture of how scholarship in the history of science has shifted over the past century.
Information on dataset(s) to be used: Citation information, author locale, and issue number for Isis publications.
Relevant publications, websites, etc: <http://www.press.uchicago.edu/ucp/journals/journal/isis.html>
Conditions under which students can publish results and/or add project results to their resume: Client would like to approve results.

oycib
Project Title: e-Xploration
Client Name: Luyi
Project goal/scientific or practical value: e-Xploration is an agent-based model for the ethnographic observation and the registry, analysis, and interpretation of social practices in virtual communities for intervention in the development of collaboration and cooperation. This project will analyze the interactions between subjects and objects in a platform collaborative community called OYCIB, a project based on e-Xploration (e-crick.net).
Information on dataset(s) to be used: I can provide a data base in .graphml format for the students. The file .graphml contains the interactions between subjects and objects in a platform collaborative community called OYCIB. In the level of practice, it is not necessary that students know agent-based models for using the database. But, in another level, for example: the collaborate level for the OYCIB development, it is necessary to have basic knowledge in AMS or MAS and another competences like PHP and MySQL.
Relevant publications, websites, etc: <http://www.e-crick.net/logs>
Conditions under which students can publish results and/or add project results to their resume: If any person or institution use my dataset or another info about eXploration (e-crick.net, oycib.net), I need to approve the results and appear as co-author.

<http://ivmooc.cns.iu.edu/clients.html>

[Diogo Carmo](#)



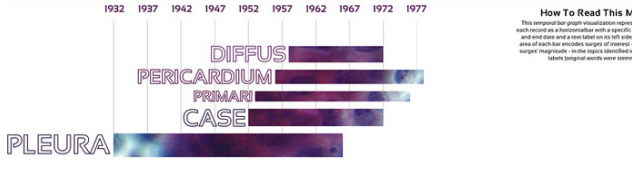
Mesothelioma

Main title topics in Medline papers

Mesothelioma (ie, more precisely, malignant mesothelioma) is a rare form of cancer that develops from an abnormal cell, originating in the mesothelium, the protective lining that covers many of the internal organs of the body. It is usually caused by exposure to asbestos. The most common anatomical site for the development of mesothelioma is the chest (the lining of the lungs, the pleural cavity), but it can also arise in the peritoneum (the lining of the abdominal cavity), and the pericardium (the sac that surrounds the heart), or the tunica vaginalis (the sac that surrounds the testis).

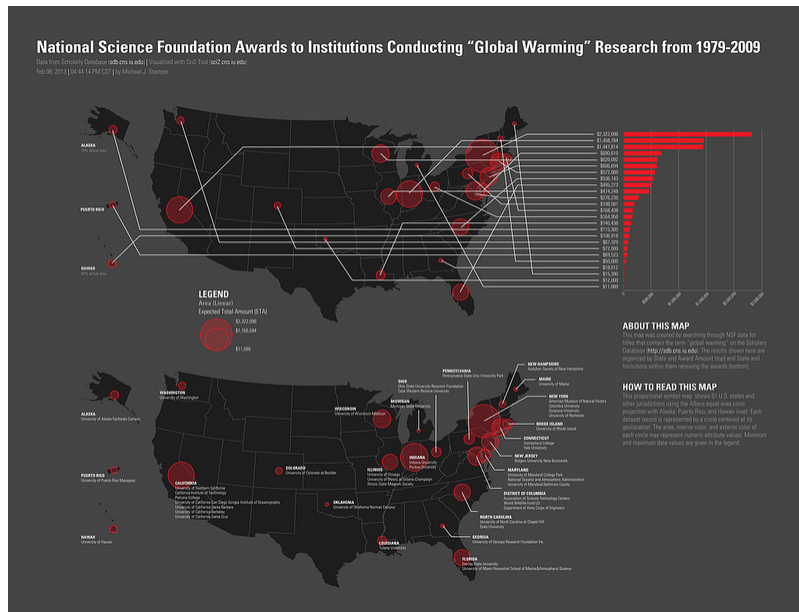
Most people who develop mesothelioma have worked in jobs where they inhaled asbestos, or were exposed to asbestos dust and fibers in other ways. It has also been suggested that washing clothes of a family member who worked with asbestos increases their risk for developing mesothelioma. Unlike lung cancer, there seems to be no association between mesothelioma and tobacco smoking, but smoking greatly increases the risk of other asbestos-related cancers. Some people who were exposed to asbestos have suffered damage from asbestos-related disease, including mesothelioma. Compensation for asbestos-related disease and/or lawsuits is an important issue in law practices regarding mesothelioma.

MALIGNANT
 PLEURAL
 CYSTIC
 BENIGN
 DIAGNOSIS



How To Read This Map
 Each bar in this visualization represents a specific year and one date and a red line on its left side. The area of each bar indicates the number of abstracts and papers published in that year identified in the library keyword search were returned.

Author: Diogo Carmo. <http://ivmooc.cns.iu.edu/clients.html>. Visualization software: Sci2 Tools (2009). Science of Science (Sci2) Tools, Indiana University and Sci Tech Strategies, <http://vis2.usf.edu/>. [Bioscience Resource Project] Medline Papers, as available in Scholarly Database: <http://pubs.ncbi.nlm.nih.gov/>. These and other visualization software, available at <http://ivmooc.cns.iu.edu/clients.html>. From: Sarason, Leo. *Born to Move*. © 2011. All Rights Reserved. This text has been licensed under a Creative Commons Attribution-NonCommercial-ShareAlike license.



[mistamper_ivmooc](#)

103

CNS Cyberinfrastructure for
Network Science Center

CDC
CENTERS FOR DISEASE
CONTROL AND PREVENTION

Tutorial Overview

- 8:30 Welcome and Overview of Tutorial and Attendees
- 9:00 The Sci2 Tool
 - Download and run the Sci2 Tool
 - ONE dataset, MANY analyses and visualizations
- 9:30 Sci2 Tool Workflows
 - Temporal Analysis: Horizontal line graph of NSF projects
 - Geospatial Analysis: US and world maps
 - Geospatial Analysis: Geomap with network overlays
 - Topical Analysis: Visualize research profiles
 - Network Analysis: Co-occurrence networks and bimodal networks
 - Network Analysis: Evolving collaboration networks
- 10:15 Networking Break**
- 10:30 Visualization Framework
- 11:00 IVMOOC – MANY more Workflows
- 11:15 Outlook and Q&A**
- 11:30 Adjourn**

104

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/

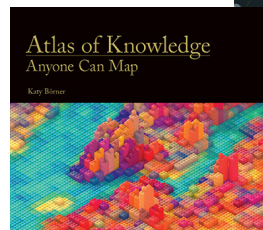
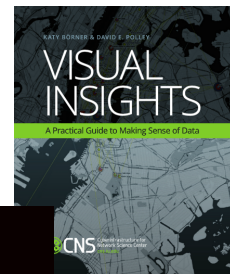
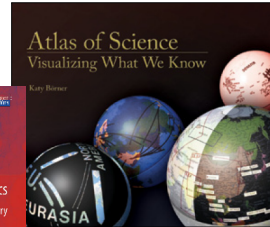
Börner, Katy (2010) **Atlas of Science: Visualizing What We Know**. The MIT Press. <http://scimaps.org/atlas>

Scharnhorst, Andrea, Börner, Katy, van den Besselaar, Peter (2012) **Models of Science Dynamics**. Springer Verlag.

Katy Börner, Michael Conlon, Jon Corson-Rikert, Cornell, Ying Ding (2012) **VIVO: A Semantic Approach to Scholarly Networking and Discovery**. Morgan & Claypool.

Katy Börner and David E Polley (2014) **Visual Insights: A Practical Guide to Making Sense of Data**. The MIT Press.

Börner, Katy (2015) **Atlas of Knowledge: Anyone Can Map**. The MIT Press. <http://scimaps.org/atlas2>



105

The screenshot shows the CNS website with the following content:

- Navigation:** About Us, Research, Development, Teaching, Outreach, Videos, News & Events, Connect With Us.
- Main Banner:** "We work closely with clients to provide custom-made data, visualization, and software solutions" with a photo of a meeting.
- Research:** Open Data and Open Code for Big Science of Science Studies.
- Latest News:** Put your money where your citations are: a proposal for a new funding system (website accessed 9/05/13).
- Upcoming Events:**
 - OCT 1: Katy Börner attends PIUG 2013 Northeast Conference
 - 10.13: Katy Börner presents Mapping Science Exhibit at VSSF
 - 10.15: Ted Polley & Google Team present IMMOOC at EDUCAUSE
 - 10.22: Katy Börner presents at the SciELO 15 Years Conference
- Development:** Behind the scenes of the design and development of AcademyScope.
- Outreach:** See some of the most fascinating data visualizations in the world.
- Videos:** Watch Katy Börner's full presentation from TEDxBloomington.
- Teaching:** Successful IMMOOC will be offered again in January of 2014.
- Our Products:** We work closely with clients to provide custom-made data, visualization, and software solutions.

All papers, maps, tools, talks, press are linked from <http://cns.iu.edu>

These slides will soon be at <http://cns.iu.edu/docs/presentations>

CNS Facebook: <http://www.facebook.com/cnscenter>

Mapping Science Exhibit Facebook: <http://www.facebook.com/mappingscience>

106