

NETWORK AND TOPICAL ANALYSIS FOR THE HUMANITIES USING NWB AND SCI2

Scott Weingart

Cyberinfrastructure for Network Science Center
Information Visualization Laboratory
Indiana Philosophy Ontology Project

School of Library and Information Science
Department of History and Philosophy of Science

Indiana University, Bloomington, IN

<http://www.scottbot.net>

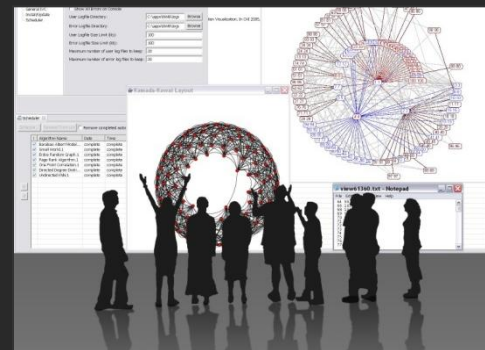
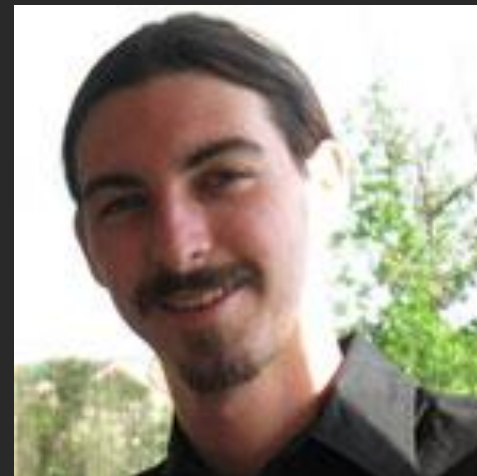
With special thanks to Katy Börner, Kevin W. Boyack, Micah Linnemeier, Russell J. Duhon, Patrick Phillips, Joseph Biberstine, Chintan Tank Nianli Ma, Hanning Guo, Mark A. Price, Angela M. Zoss, and Sean Lind

Digital Humanities 2011

Meyer Library 2080E (Language Lab)

Stanford University, Stanford, CA

13:00-16:30 on June 19, 2011





Workshop Overview

1:00-1:15 Introduction to Network Analysis

1:15-1:45 Network Analysis & Visualization in the Humanities

- Theory, Applications, and Pitfalls.
- Examples In The Wild

1:45-2:15 Collecting, Cleaning & Formatting Data

2:15-2:25 Break

2:25-3:00 Sci2 Tool Basics

- Macroscope Design and Usage.
- Download and run the tool.
- Find basic statistics and run various algorithms over the network.
- Visualize the networks as either a graph or a circular hierarchy.

3:00-3:20 Sci2 Workflow Design: Padgett's Florentine Families - Prepare, load, analyze, and visualize family and business networks from 15th century Florence.

3:20-3:35 Break

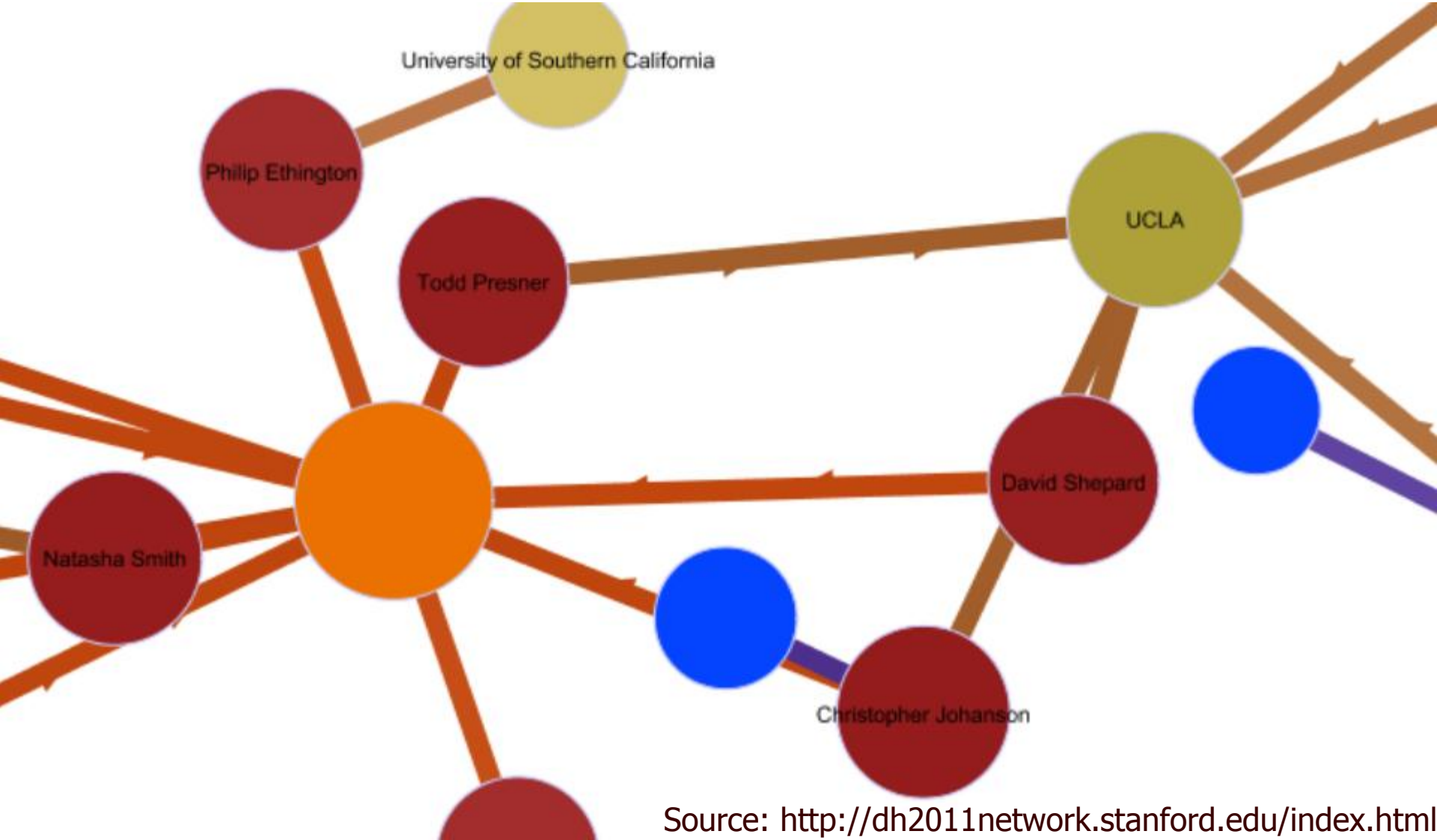
3:35-4:00 Sci2 Research Demonstration: Mapping the Republic of Letters

4:00-4:30 Q&A and Technical Assistance

Network Analysis & Visualization



What is a network?



Source: <http://dh2011network.stanford.edu/index.html>



What is a network?

- Graph

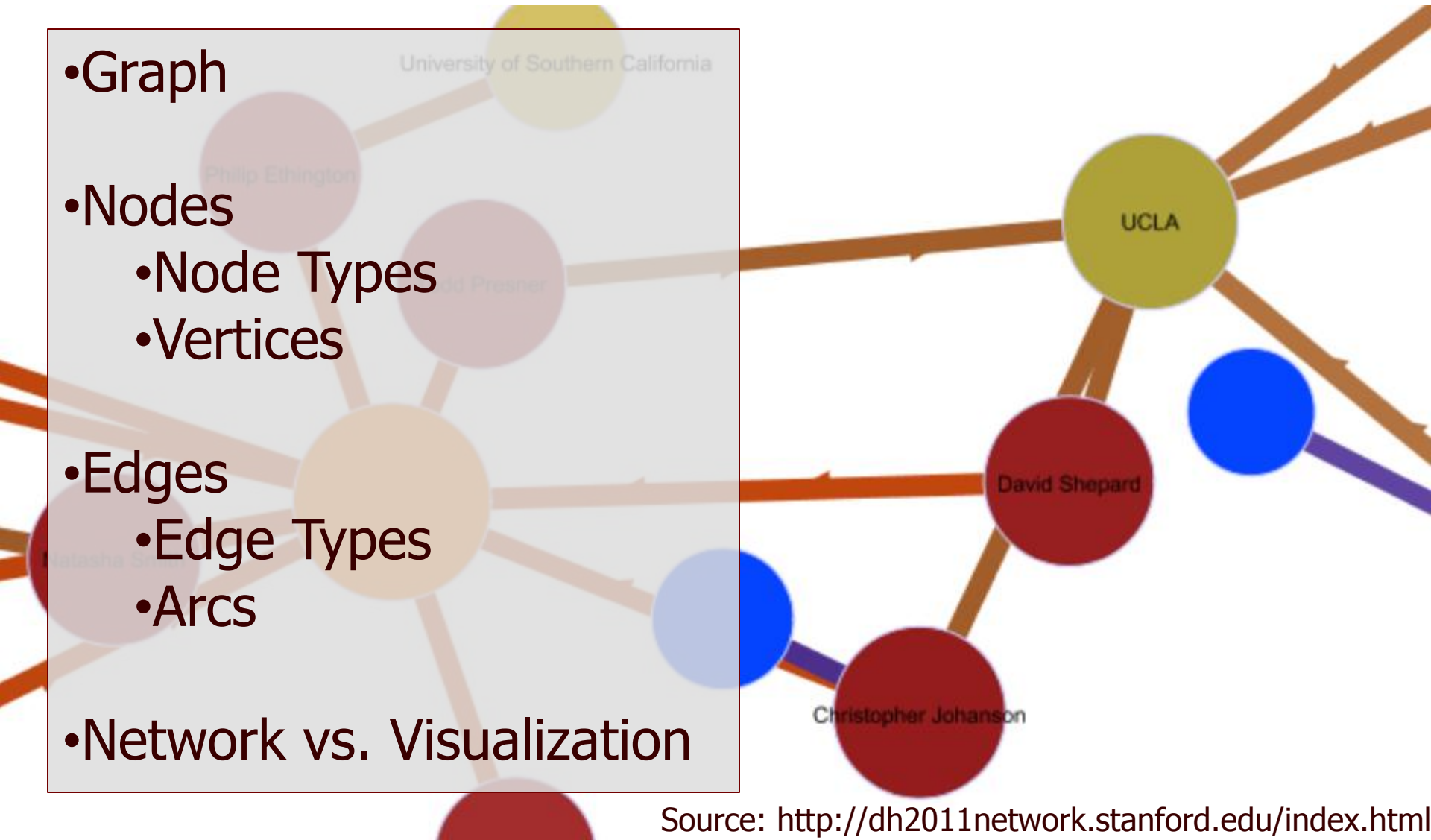
- Nodes

- Node Types
- Vertices

- Edges

- Edge Types
- Arcs

- Network vs. Visualization





What is a network?

*Nodes

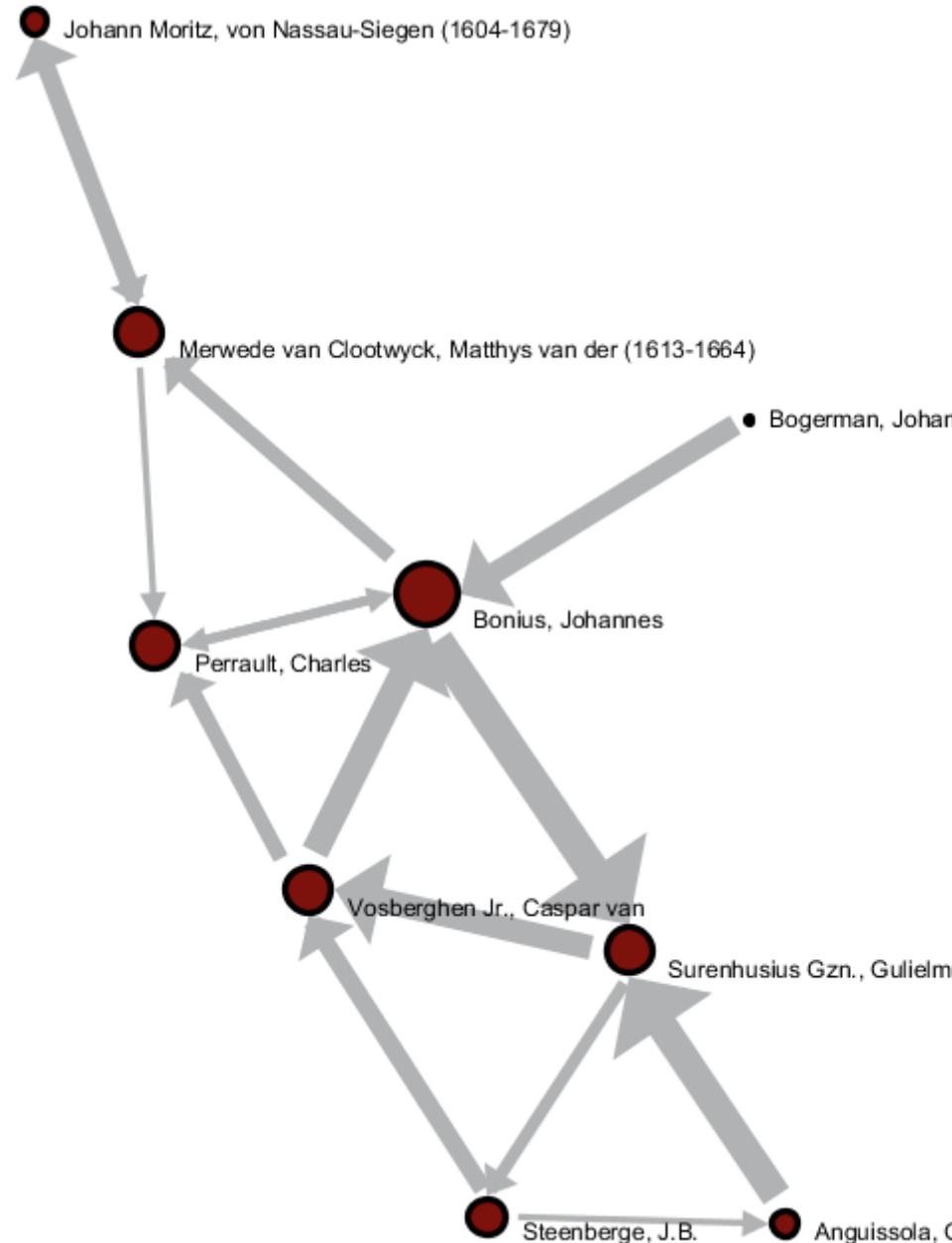
id*int label*string totaldegree*int

```
16 "Merwede van Clootwyck, Matthys van der (1613-1664)" 1
36 "Perrault, Charles" 1
48 "Bonius, Johannes" 1
67 "Surenhusius Gzn., Gulielmus" 1
99 "Anguissola, Giacomo" 1
126 "Johann Moritz, von Nassau-Siegen (1604-1679)" 6
131 "Steenberge, J.B." 1
133 "Vosberghen Jr., Caspar van" 1
151 "Bogerman, Johannes (1576-1637)" 25
```

*DirectedEdges

source*int target*int weight*float eyear*int syear*int

```
16 36 1 1640 1650
16 126 5 1641 1649
36 48 2 1630 1633
48 16 4 1637 1644
48 67 10 1645 1648
48 36 2 1632 1638
67 133 7 1644 1648
67 131 3 1642 1643
99 67 9 1640 1645
126 16 3 1641 1646
131 133 5 1630 1638
131 99 1 1637 1639
133 36 4 1645 1648
133 48 8 1632 1636
151 48 6 1644 1647
```





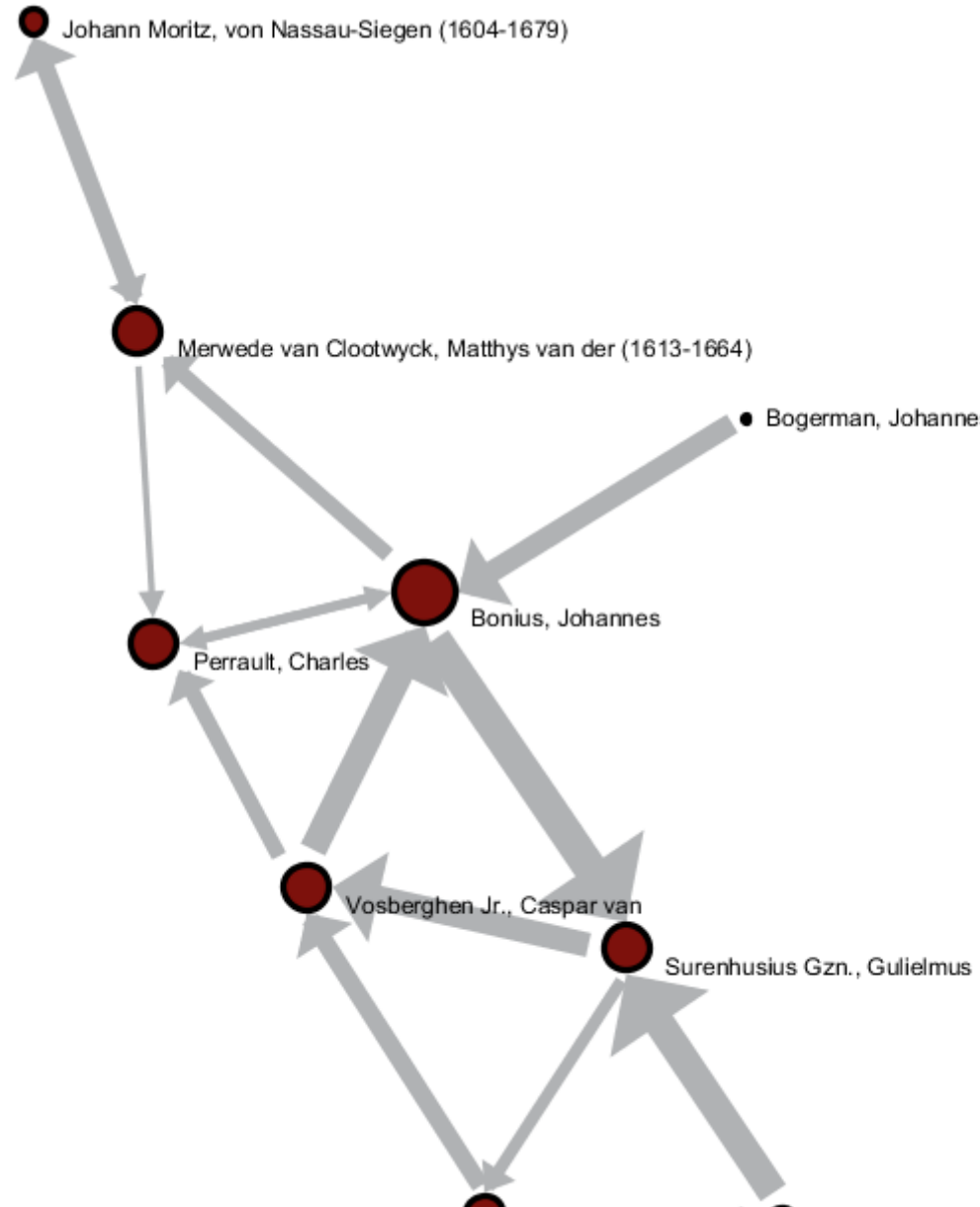
Graph Metrics – Nodes / Vertices / Entities

- Degree Centrality - direct connections
- Betweenness Centrality – # of shortest paths a node sits between
- Closeness Centrality – Distance to all nodes
- Hubs / Bridges / Isolates



Graph Metrics – Nodes / Vertices / Entities

- Degree Centrality
- Betweenness Centrality
- Closeness Centrality
- Hubs / Bridges / Isolates





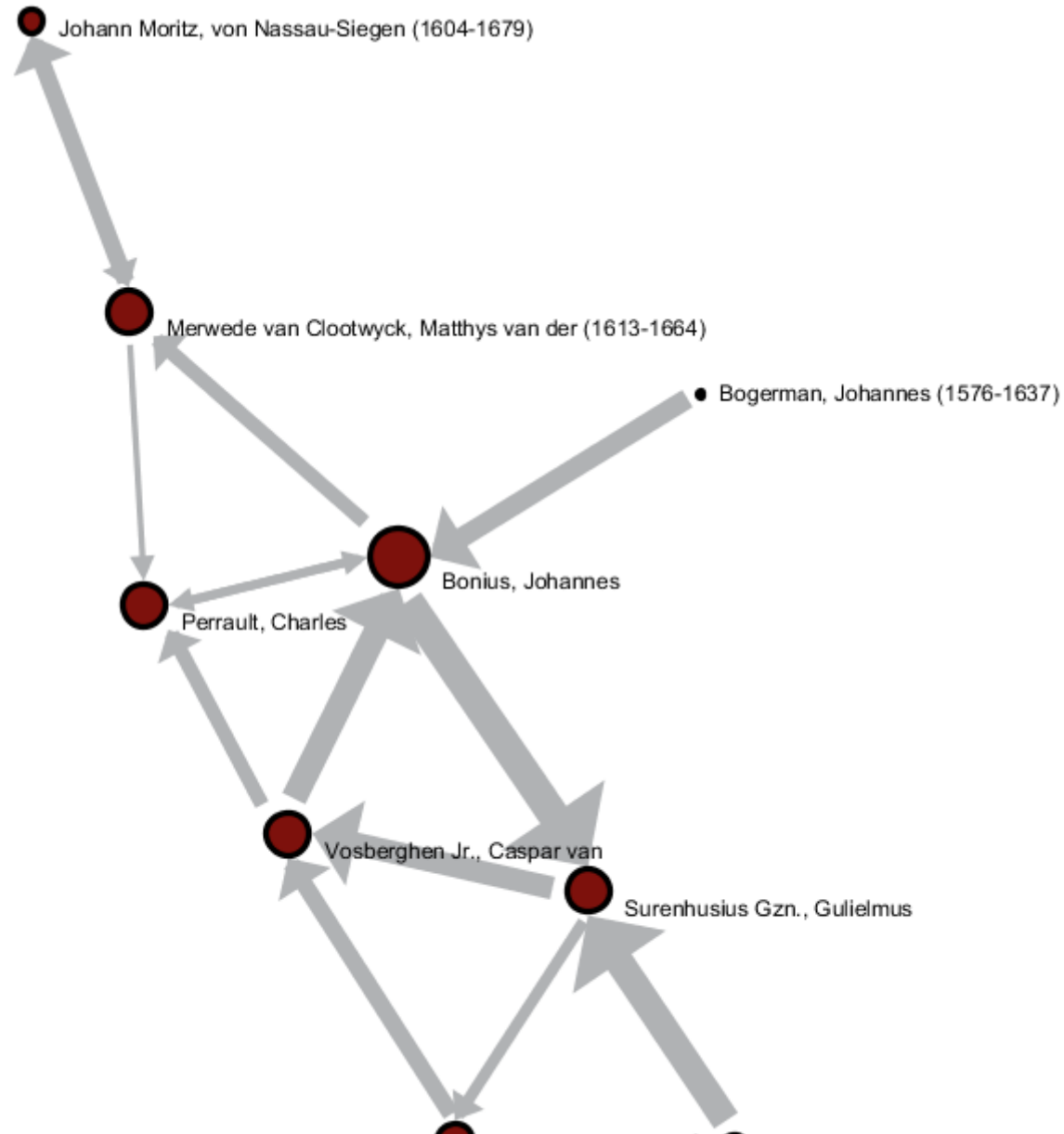
Graph Metrics – Edges / Links / Arcs / Connections / Ties

- Shortest paths – shortest distance between two nodes
- Weight – strength of tie
- Directionality – is the connection one or two-way?
- Bridge – deleting would change structure



Graph Metrics – Edges / Links / Arcs / Connections / Ties

- Shortest paths
- Weight
- Directionality
- Bridge





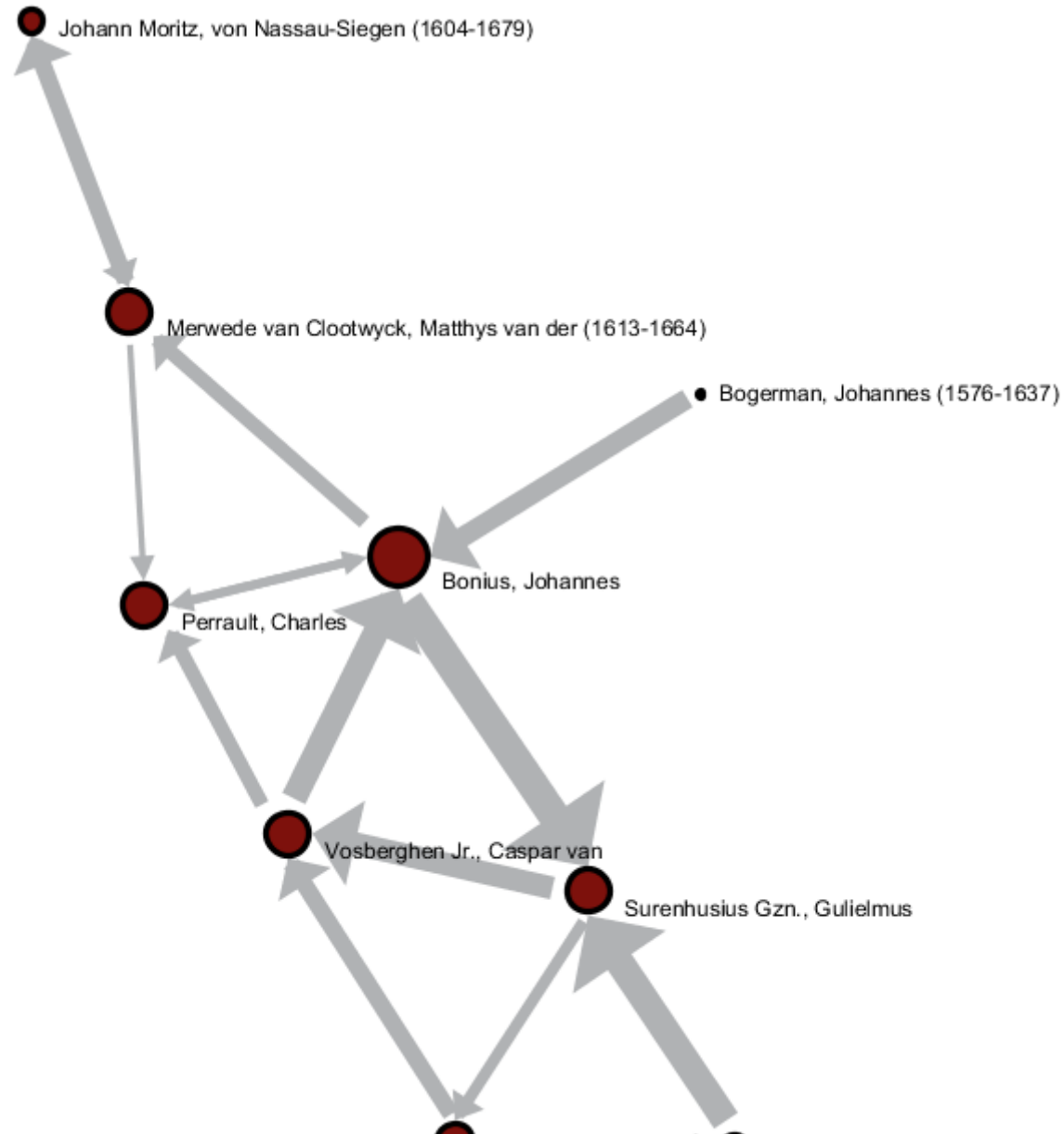
Graph Metrics – Structural

- Clusters – closely connected or similar nodes & edges
- Clustering coefficient – ‘clique-ness’
- Centralization – egocentric or decentralized
- Density – ‘spaghetti-ness’
- Average Path Length – wide or narrow



Graph Metrics – Structural

- Clusters
- Clustering coefficient
- Centralization
- Density
- Average Path Length





Graph Metrics – Types Of Networks

- Bi-Modal or Multi-Nodal
- Egocentric
- Dense/Sparse

University of Southern California

Philip Ethington

Todd Presner

UCLA

David Shepard

Katasha Smith

Christopher Johanson



Networks – Visualization

- Networks visualized as graphs
- Layouts
 - Force-based / Spring / GEM
 - Tree
 - Circular
- Distance Re-training



Workshop Overview

1:00-1:15 Introduction to Network Analysis

1:15-1:45 Network Analysis & Visualization in the Humanities

- **Theory, Applications, and Pitfalls.**
- **Examples In The Wild**

1:45-2:15 Collecting, Cleaning & Formatting Data

2:15-2:25 Break

2:25-3:00 Sci2 Tool Basics

- **Macroscope Design and Usage.**
- **Download and run the tool.**
- **Find basic statistics and run various algorithms over the network.**
- **Visualize the networks as either a graph or a circular hierarchy.**

3:00-3:20 Sci2 Workflow Design: Padgett's Florentine Families - Prepare, load, analyze, and visualize family and business networks from 15th century Florence.

3:20-3:35 Break

3:35-4:00 Sci2 Research Demonstration: Mapping the Republic of Letters

4:00-4:30 Q&A and Technical Assistance

Network Analysis & Visualization in the Humanities: Theory, Applications, and Pitfalls



Uses of Visualization

- **Solidifies objects of inquiry**
- **Exploration**
- **Discovery**
- **Trend-spotting**
- **Evidence**
- **Audience Engagement**
- **Engaging public / funding agencies**



The Importance of Visualization

[Visualization] aim at more than making the invisible visible. [It aspires] to all-at-once-ness, the condensation of laborious, step-by-step procedures in to an immediate *coup d'oeil*... What was a painstaking process of calculation and correlation—for example, in the construction of a table of variables—becomes a flash of intuition. And all-at-once intuition is traditionally the way that angels know, in contrast to the plodding demonstrations of humans.

Descartes's craving for angelic all-at-once-ness emerged forcefully in his mathematics..., compressing the steps of mathematical proof into a single bright flare of insight: "I see the whole thing at once, by intuition."

Lorraine Daston – On Scientific Observation



Warnings

[H]umanists have adopted many applications such as GIS mapping, graphs, and charts for statistical display that were developed in other disciplines... such graphical tools are a kind of intellectual Trojan horse...

Data pass themselves off as mere descriptions of a priori conditions. Rendering *observation* (the act of creating a statistical, empirical, or subjective account or image) as if it were *the same as the phenomena observed* collapses the critical distance between the phenomenal world and its interpretation, undoing the basis of interpretation on which humanistic knowledge production is based... we seem ready and eager to suspend critical judgment in a rush to visualization.

Johanna Drucker – Humanities Approaches to Graphical Display



Warnings

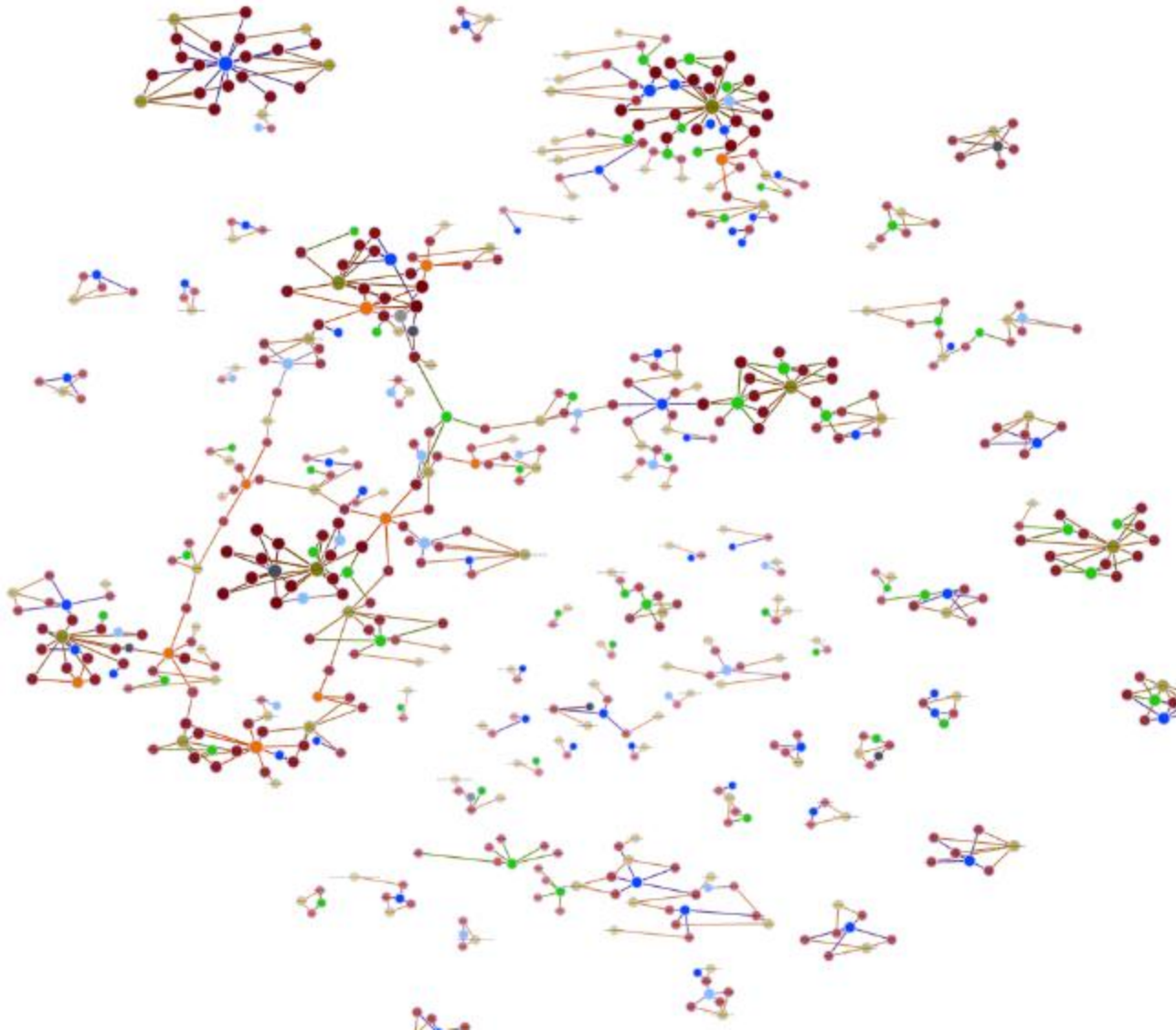
- **Data format limits use, already an act of interpretation.**
- **Statistics is often misused (wield it very carefully).**
- **Interpreting spatial distance as meaningful.**
- **Always include a legend (this presentation breaks that rule).**
- **Accidental legitimization in eyes of public.**

Network Analysis & Visualization in the Humanities: Examples In The Wild



Digital Humanities 2011 – Elijah Meeks

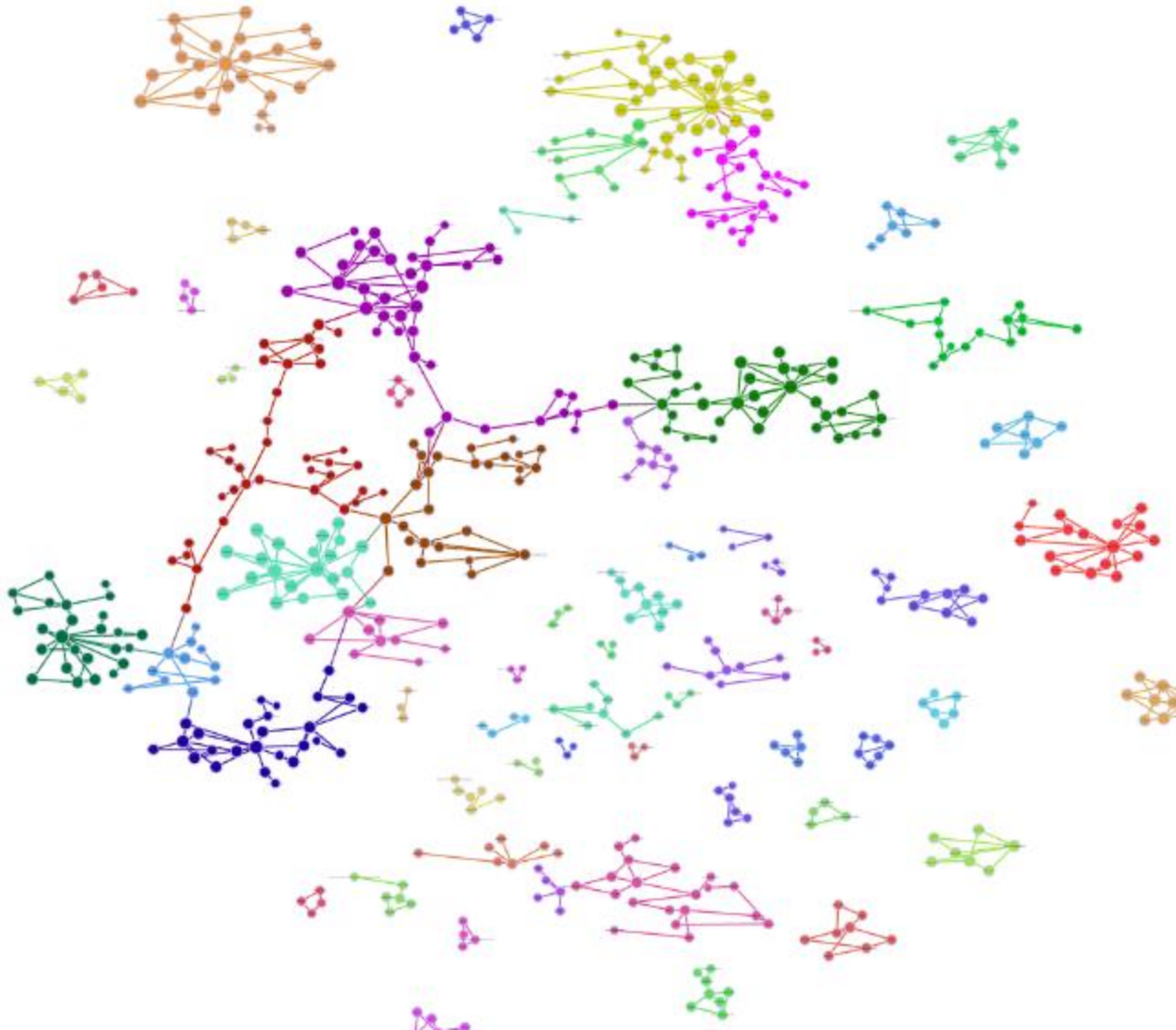
<http://dh2011network.stanford.edu/>





Digital Humanities 2011 – Elijah Meeks

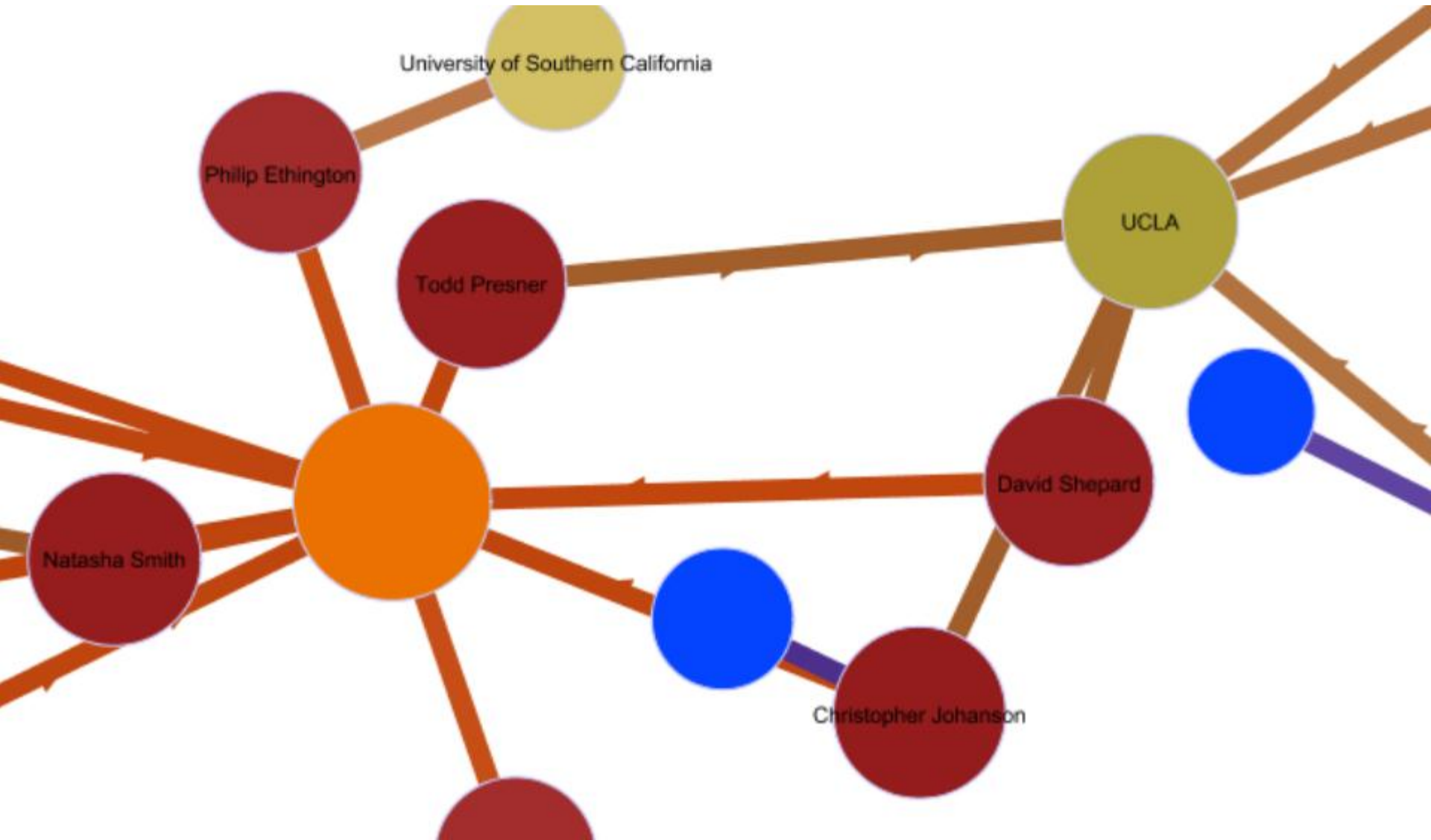
<http://dh2011network.stanford.edu/>





Digital Humanities 2011 – Elijah Meeks

<http://dh2011network.stanford.edu/>





Digital Humanities 2011 – Elijah Meeks

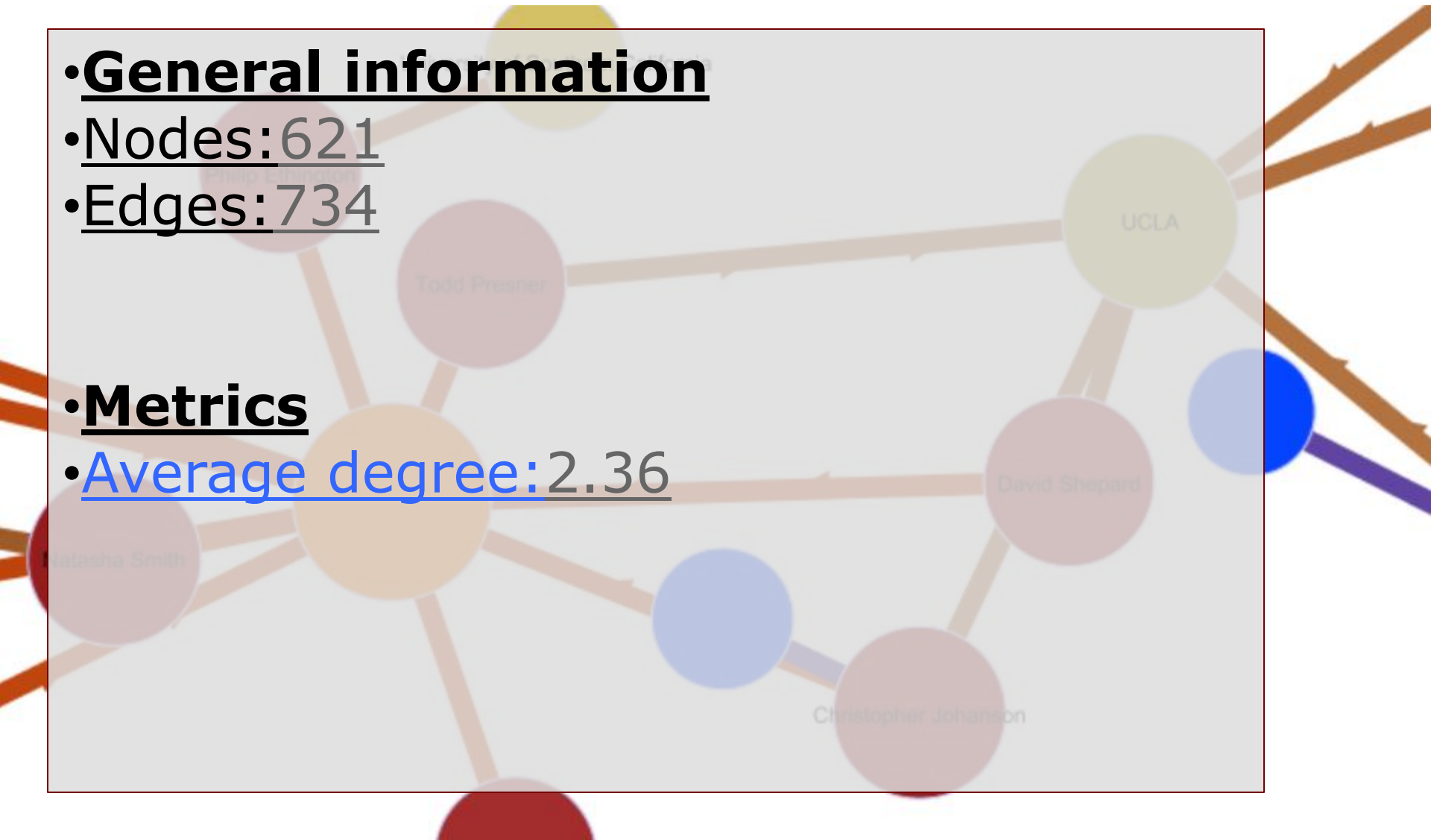
<http://dh2011network.stanford.edu/>

• General information

- Nodes: 621
- Edges: 734

• Metrics

- Average degree: 2.36





Digital Humanities 2011 – Elijah Meeks

<http://dh2011network.stanford.edu/>





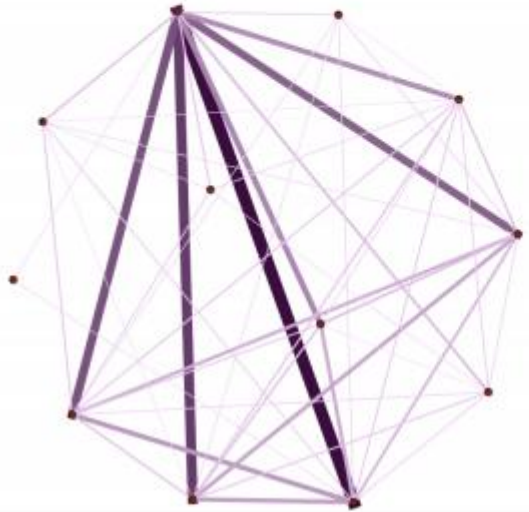
Character Networks in the 19th Century British Novel -Graham Sack

I use computational methods to count the frequency and co-occurrence of a generally ignored sub-class of common words, namely, character names. Character names are often regarded as noise and excluded from authorship and stylistics analysis because they are not consistent across texts. This study makes character names its main object of analysis because the objective is quite different: rather than style or authorship, this study attempts to make inferences about *characterization* and *social form*, two areas about which computational analysis has had comparatively little to say.

Character Network Sociograms

Figures 20 a, b, & C

The Ambassadors (James)



General Features:

- Small network (12 characters)
- No isolates
- Very high graph density (71%) and clustering coefficient (85%)
- Low average path length (1.3)
- Low degree inequality (-4.9)
- High proportion of strong ties (28%)

Conclusions

- Tightly knit social world focused on deep relationships between small set of characters
- Social interaction broadly evenly distributed

Middlemarch (Eliot)



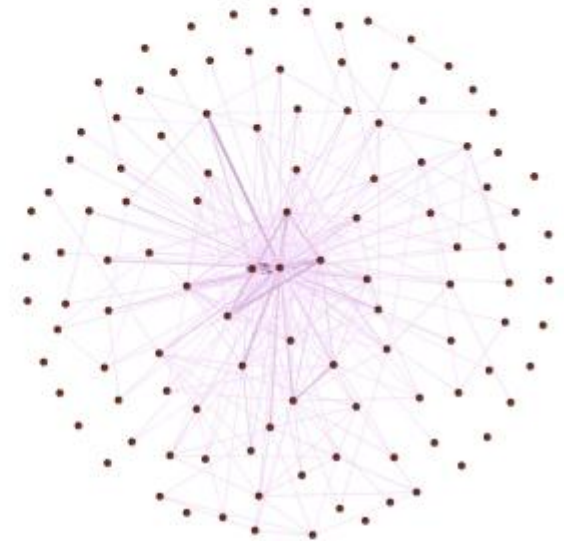
General Features:

- Large network (99 characters)
- Moderately high % of isolates (17%)
- Low graph density (7%) and clustering coefficient (73%)
- High average path length (2.4)
- Moderate degree inequality (1.9)
- Moderate proportion of strong ties (18%)

Conclusions

- Large but comparatively integrated social world with deep interaction between core characters

The Pickwick Papers (Dickens)



General Features:

- Large network (112 characters)
- High proportion of isolates (20%)
- Very low graph density (4%) and clustering coefficient (72%)
- High average path length (2.2)
- High degree inequality (3.0)
- Low proportion of strong ties (13%)

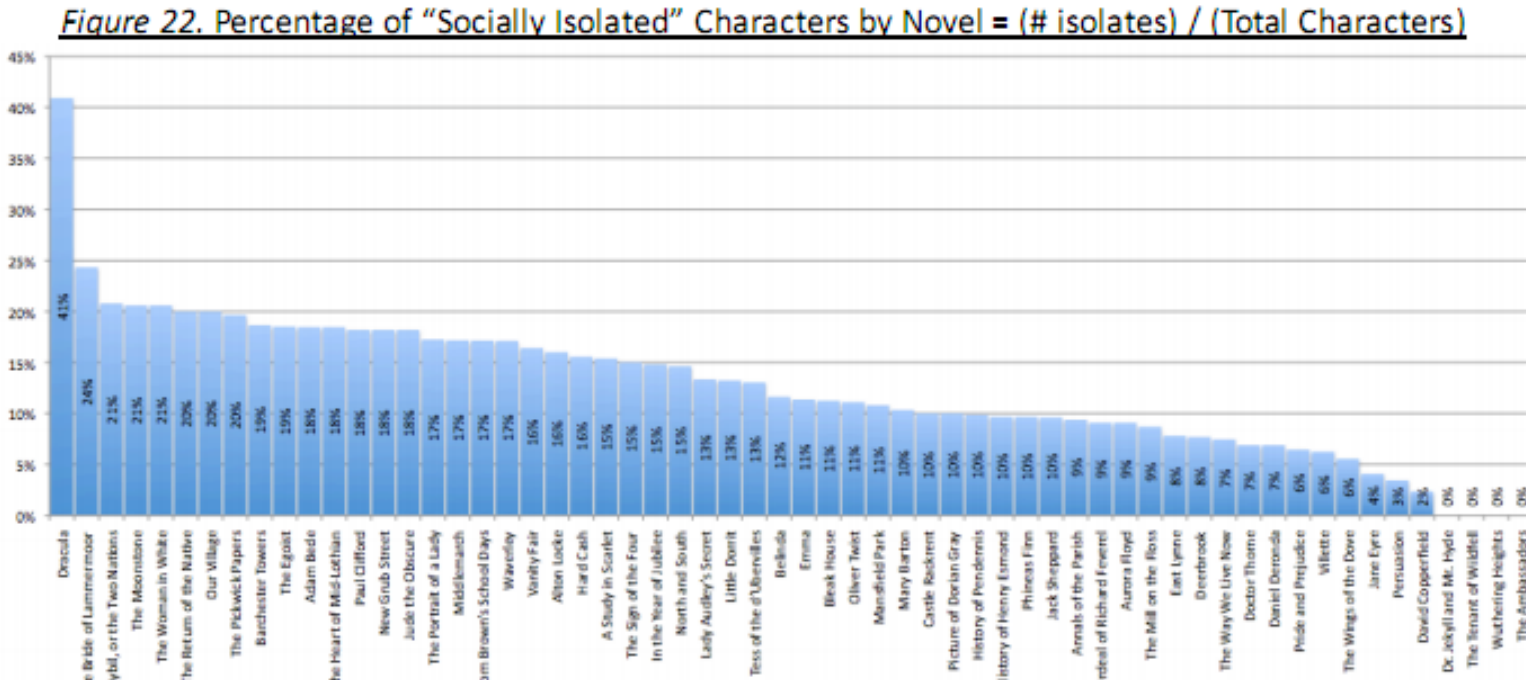
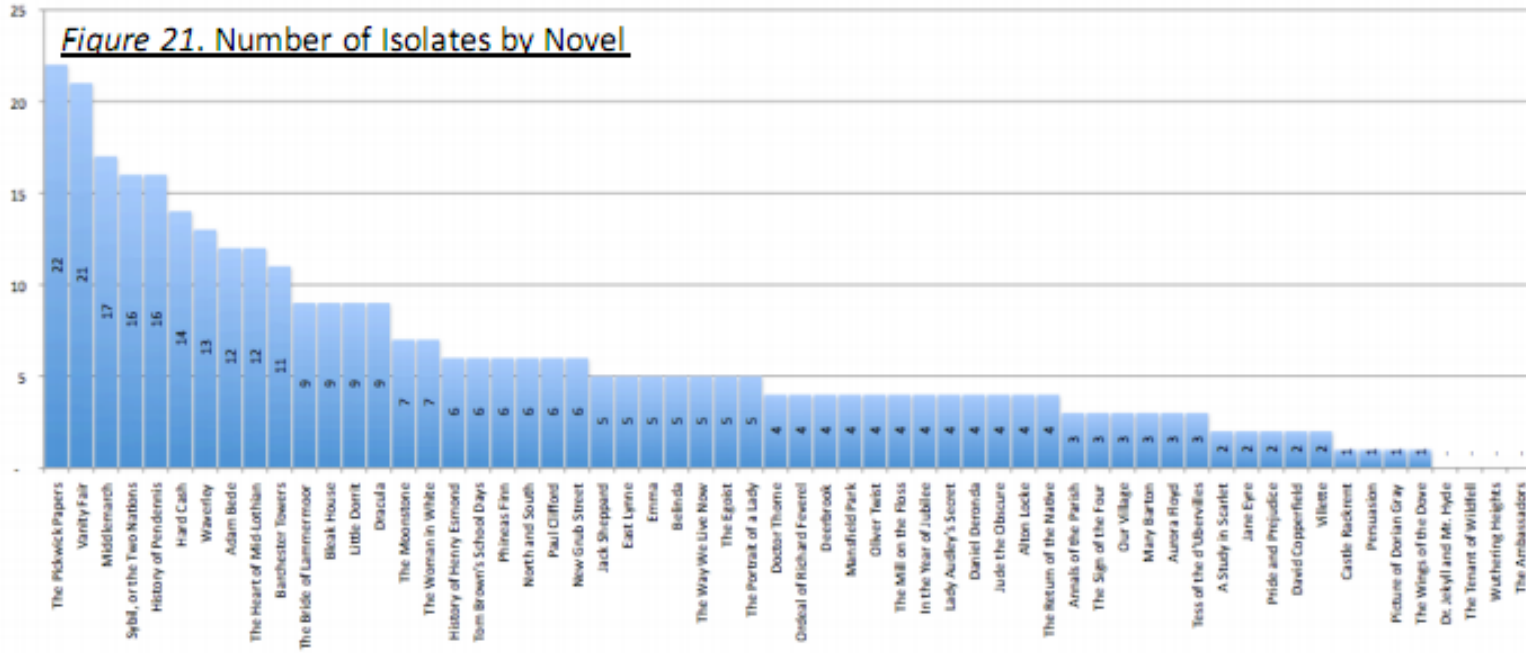
Conclusions

- Expansive but diffuse social world with passing social interactions and many isolated characters



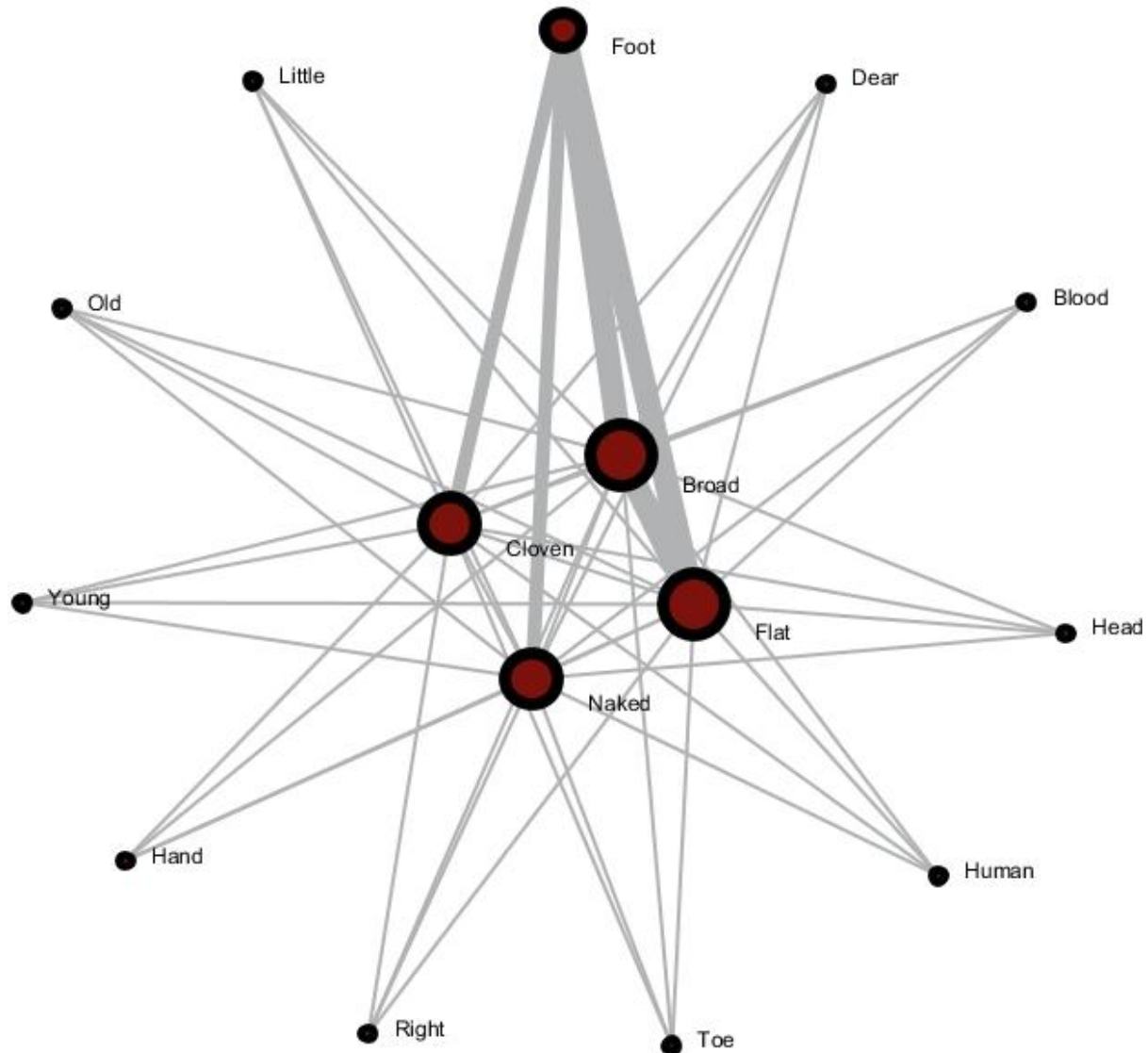
Character Networks in the 19th Century British Novel - Graham Sack

Social Metrics – By Novel (1/3)



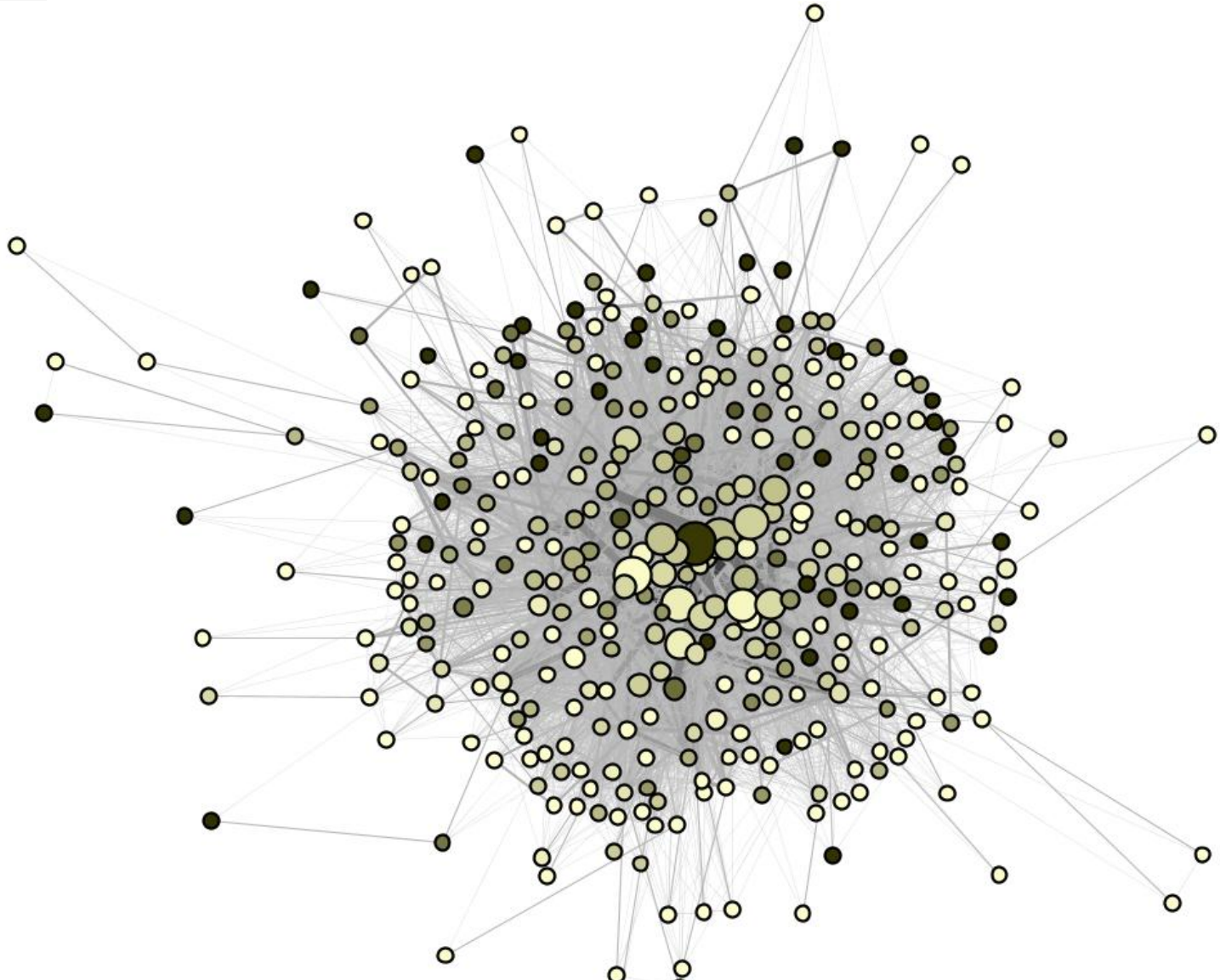


Word Co-Occurrences in European Fairytales -Jorgensen & Weingart





Word Co-Occurrences in European Fairytales -Jorgensen & Weingart





Mapping the Republic of Letters – Chang et al.

<https://republicofletters.stanford.edu/#maps>





Workshop Overview

1:00-1:15 Introduction to Network Analysis

1:15-1:45 Network Analysis & Visualization in the Humanities

- Theory, Applications, and Pitfalls.
- Examples In The Wild

1:45-2:15 Collecting, Cleaning & Formatting Data

2:15-2:25 Break

2:25-3:00 Sci2 Tool Basics

- Macroscope Design and Usage.
- Download and run the tool.
- Find basic statistics and run various algorithms over the network.
- Visualize the networks as either a graph or a circular hierarchy.

3:00-3:20 Sci2 Workflow Design: Padgett's Florentine Families - Prepare, load, analyze, and visualize family and business networks from 15th century Florence.

3:20-3:35 Break

3:35-4:00 Sci2 Research Demonstration: Mapping the Republic of Letters

4:00-4:30 Q&A and Technical Assistance

Collecting, Cleaning & Formatting Data



Data Collection

- **65% collecting, 25% cleaning, 5% formatting, 5% analysis.**
- **Keep analysis in mind before first data collected**
- **Remember not just current use but future use as well; nobody wants to do the same work five times.**
- **Make sure system for data entry is easy, transferrable, and as granular as possible.**



Data Cleaning & Validation

- **Multiple coders? Undergraduates? Check for intercoder reliability. Understand biases going in.**
- **Microsoft Excel is your friend (ubiquitous, easy).**
- **Consistency in word use and coding.**
- **Sanity checks at the extremes – order by size, etc.**
- **Check against existing databases if available.**
- **Make sure basic statistics make sense.**



File-types

➤ **Excel**

➤ **Database**

➤ **Text**

➤ **CSV**



Network Formats

	Newton	Oldenburg	Flamsteed
Newton	0	13	38
Oldenburg	24	0	45
Flamsteed	62	7	0

- **Matrix**
- **Adjacency List**
- **Node & Edge List**

Newton	Oldenburg	13
Newton	Flamsteed	38
Oldenburg	Newton	24
Oldenburg	Flamsteed	45
Flamsteed	Newton	62
Flamsteed	Oldenburg	7

Nodes		
1	Newton	
2	Oldenburg	
3	Flamsteed	
Edges		
1	2	13
1	3	38
2	1	24
2	3	45
3	1	62
3	2	7



NWB Format

*Nodes

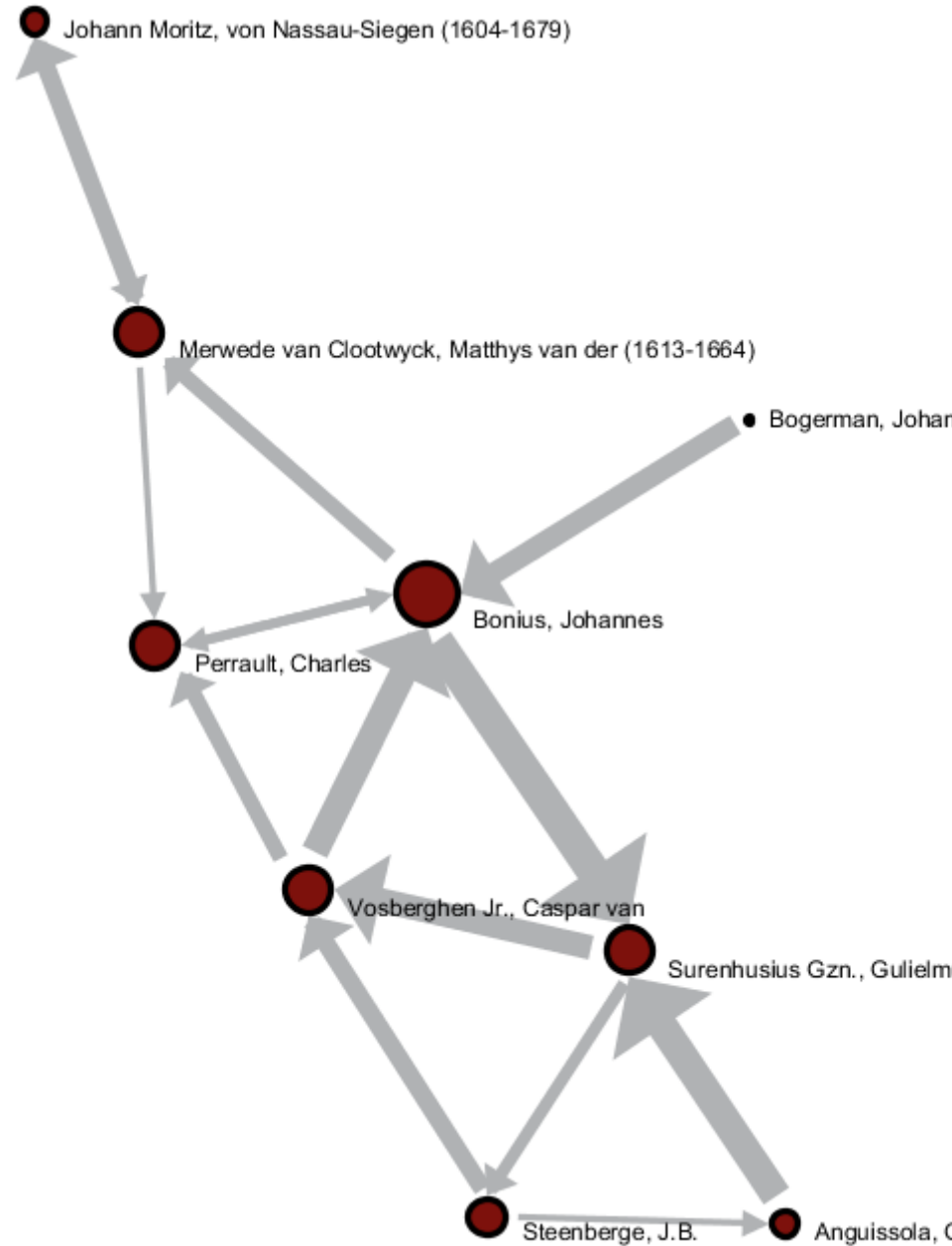
id*int label*string totaldegree*int

```
16 "Merwede van Clootwyck, Matthys van der (1613-1664)" 1
36 "Perrault, Charles" 1
48 "Bonius, Johannes" 1
67 "Surenhusius Gzn., Gulielmus" 1
99 "Anguissola, Giacomo" 1
126 "Johann Moritz, von Nassau-Siegen (1604-1679)" 6
131 "Steenberge, J.B." 1
133 "Vosberghen Jr., Caspar van" 1
151 "Bogerman, Johannes (1576-1637)" 25
```

*DirectedEdges

source*int target*int weight*float eyear*int syear*int

```
16 36 1 1640 1650
16 126 5 1641 1649
36 48 2 1630 1633
48 16 4 1637 1644
48 67 10 1645 1648
48 36 2 1632 1638
67 133 7 1644 1648
67 131 3 1642 1643
99 67 9 1640 1645
126 16 3 1641 1646
131 133 5 1630 1638
131 99 1 1637 1639
133 36 4 1645 1648
133 48 8 1632 1636
151 48 6 1644 1647
```





Workshop Overview

1:00-1:15 Introduction to Network Analysis

1:15-1:45 Network Analysis & Visualization in the Humanities

- Theory, Applications, and Pitfalls.
- Examples In The Wild

1:45-2:15 Collecting, Cleaning & Formatting Data

2:15-2:25 Break

2:25-3:00 Sci2 Tool Basics

- Macroscope Design and Usage.
- Download and run the tool.
- Find basic statistics and run various algorithms over the network.
- Visualize the networks as either a graph or a circular hierarchy.

3:00-3:20 Sci2 Workflow Design: Padgett's Florentine Families - Prepare, load, analyze, and visualize family and business networks from 15th century Florence.

3:20-3:35 Break

3:35-4:00 Sci2 Research Demonstration: Mapping the Republic of Letters

4:00-4:30 Q&A and Technical Assistance



Workshop Overview

1:00-1:15 Introduction to Network Analysis

1:15-1:45 Network Analysis & Visualization in the Humanities

- Theory, Applications, and Pitfalls.
- Examples In The Wild

1:45-2:15 Collecting, Cleaning & Formatting Data

2:15-2:25 Break

2:25-3:00 Sci2 Tool Basics

- **Macroscope Design and Usage.**
- **Download and run the tool.**
- **Find basic statistics and run various algorithms over the network.**
- **Visualize the networks as either a graph or a circular hierarchy.**

3:00-3:20 Sci2 Workflow Design: Padgett's Florentine Families - Prepare, load, analyze, and visualize family and business networks from 15th century Florence.

3:20-3:35 Break

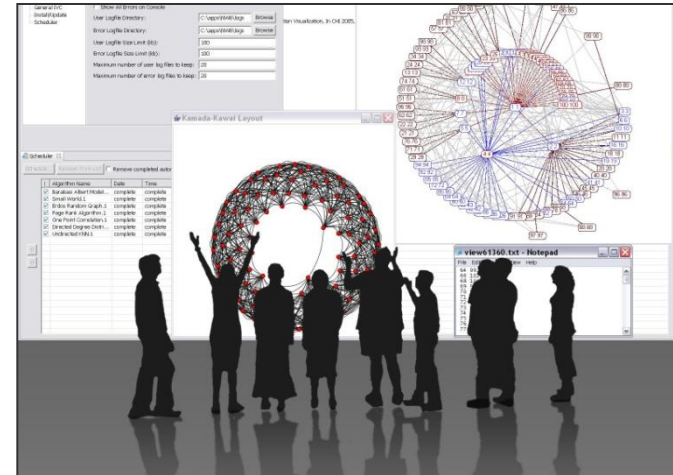
3:35-4:00 Sci2 Research Demonstration: Mapping the Republic of Letters

4:00-4:30 Q&A and Technical Assistance

Sci2 Tool Basics: Macroscope Design and Usage



Microscopes, Telescopes, and Macroscopes



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



Desirable Features of Macroscopes

Core Architecture & Plugins/Division of Labor: Computer scientists need to design the standardized, modular, easy to maintain and extend “core architecture”. Dataset and algorithm plugins, i.e., the “filling”, are provided by those that care and know most about the data and developed the algorithms: the domain experts.

Ease of Use: As most plugin contributions and usage will come from non-computer scientists it must be possible to contribute, share, and use new plugins without writing one line of code. Users need guidance for constructing effective workflows from 100+ continuously changing plugins.

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

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

Open Data and Open Code: Lets anybody check, improve, or repurpose code and eases the replication of scientific studies.

Macroscopes are similar to Flickr and YouTube and but instead of sharing images or videos, you freely share datasets and algorithms with scholars around the globe.



Macroscopic Design

Custom Tools for Different Scientific Communities

Information Visualization Cyberinfrastructure

<http://iv.slis.indiana.edu>

Network Workbench Tool + Community Wiki

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

Science of Science (Sci²) Tool and Portal

<http://sci.slis.indiana.edu>

Epidemics Cyberinfrastructure

<http://epic.slis.indiana.edu/>



180+ Algorithm Plugins and Branded GUIs

+

Core Architecture

Open Services Gateway Initiative (OSGi) Framework.

<http://orgi.org>

Cyberinfrastructure Shell (CIShell)

<http://cishell.org>



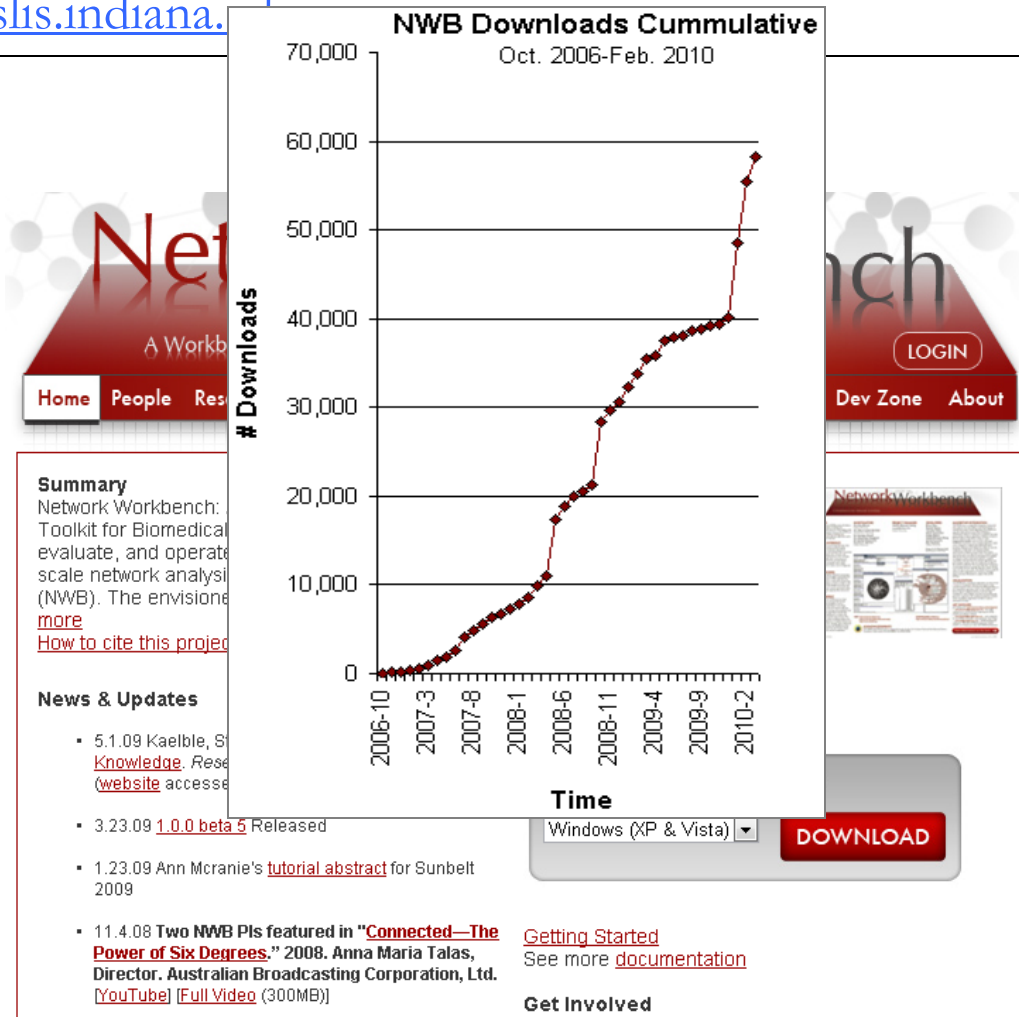
**CIShell Powered Tools:
Network Workbench (NWB)**

The Network Workbench (NWB) tool supports researchers, educators, and practitioners interested in the study of biomedical, social and behavioral science, physics, and other networks.

In February 2009, the tool provides more 169 plugins that support the preprocessing, analysis, modeling, and visualization of networks.

More than 50 of these plugins can be applied or were specifically designed for S&T studies.

It has been downloaded more than 65,000 times since December 2006.



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



Software Team: Lead: Micah Linnemeier
Members: Patrick Phillips, Russell Duhon, Tim Kelley & Ann McCranie
Previous Developers: Weixia (Bonnie) Huang, Bruce Herr, Heng Zhang, Duygu Balcan, Bryan Hook, Ben Markines, Santo Fortunato, Felix Terkhorn, Ramya Sabbineni, Vivek S. Thakre & Cesar Hidalgo



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

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

Duration: Sept. 2005 - Aug. 2009

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

NWB Advisory Board:

James Hendler (Semantic Web) <http://www.cs.umd.edu/~hendler/>

Jason Leigh (CI) <http://www.evl.uic.edu/spiff/>

Neo Martinez (Biology) <http://online.sfsu.edu/~webhead/>

Michael Macy, Cornell University (Sociology) <http://www.soc.cornell.edu/faculty/macy.shtml>

Ulrik Brandes (Graph Theory) <http://www.inf.uni-konstanz.de/~brandes/>

Mark Gerstein, Yale University (Bioinformatics) <http://bioinfo.mbb.yale.edu/>

Stephen North (AT&T) <http://public.research.att.com/viewPage.cfm?PageID=81>

Tom Snijders, University of Groningen <http://stat.gamma.rug.nl/snijders/>

Noshir Contractor, Northwestern University <http://www.spcomm.uiuc.edu/nosh/>



Computational Proteomics

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

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

© 2007 Nature Publishing Group

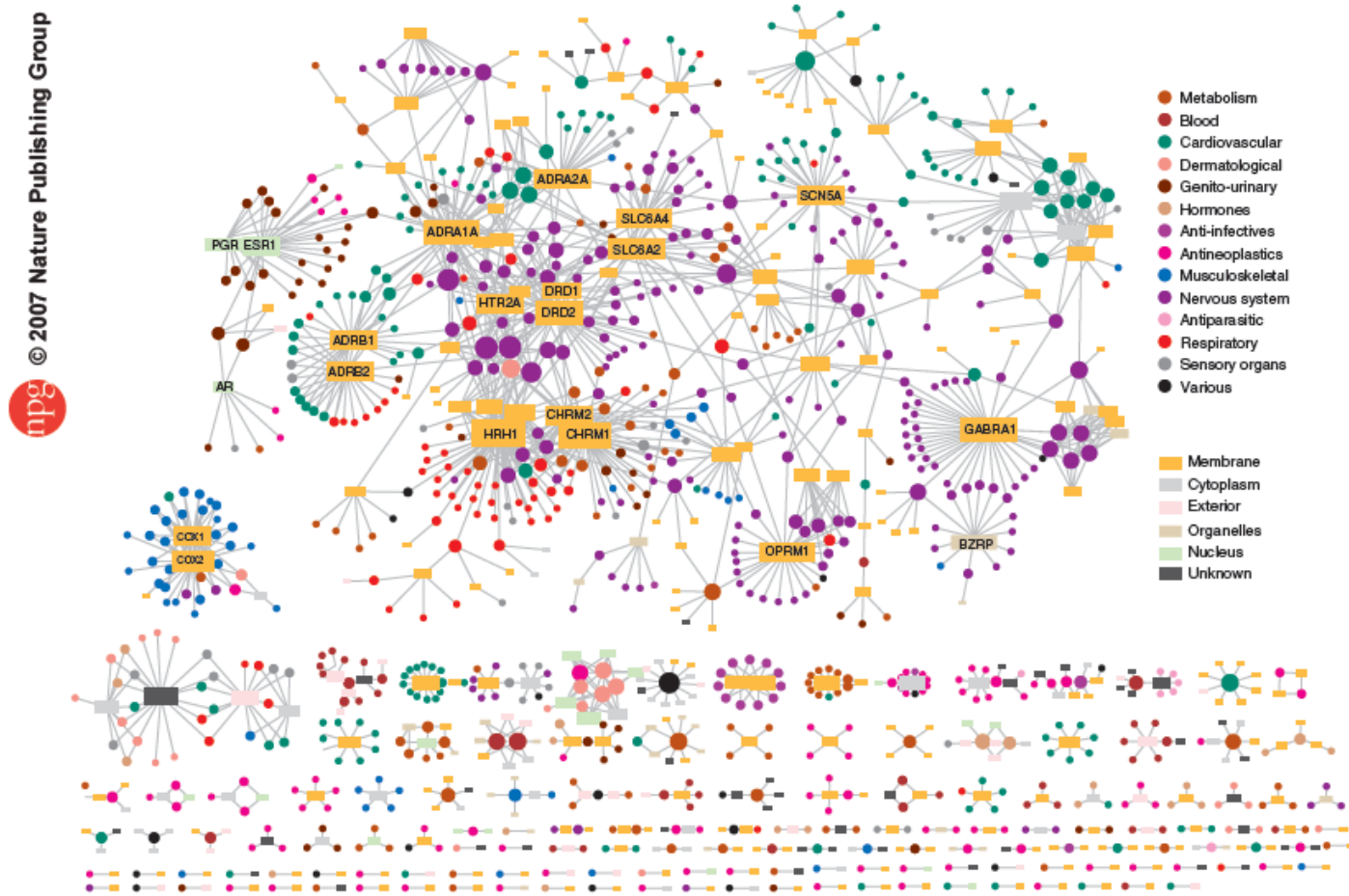


Figure 2 Drug–target network (DT network). The DT network is generated by using the known associations between FDA-approved drugs and their target proteins. Circles and rectangles correspond to drugs and target proteins, respectively. A link is placed between a drug node and a target node if the protein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



Computational Economics

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

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

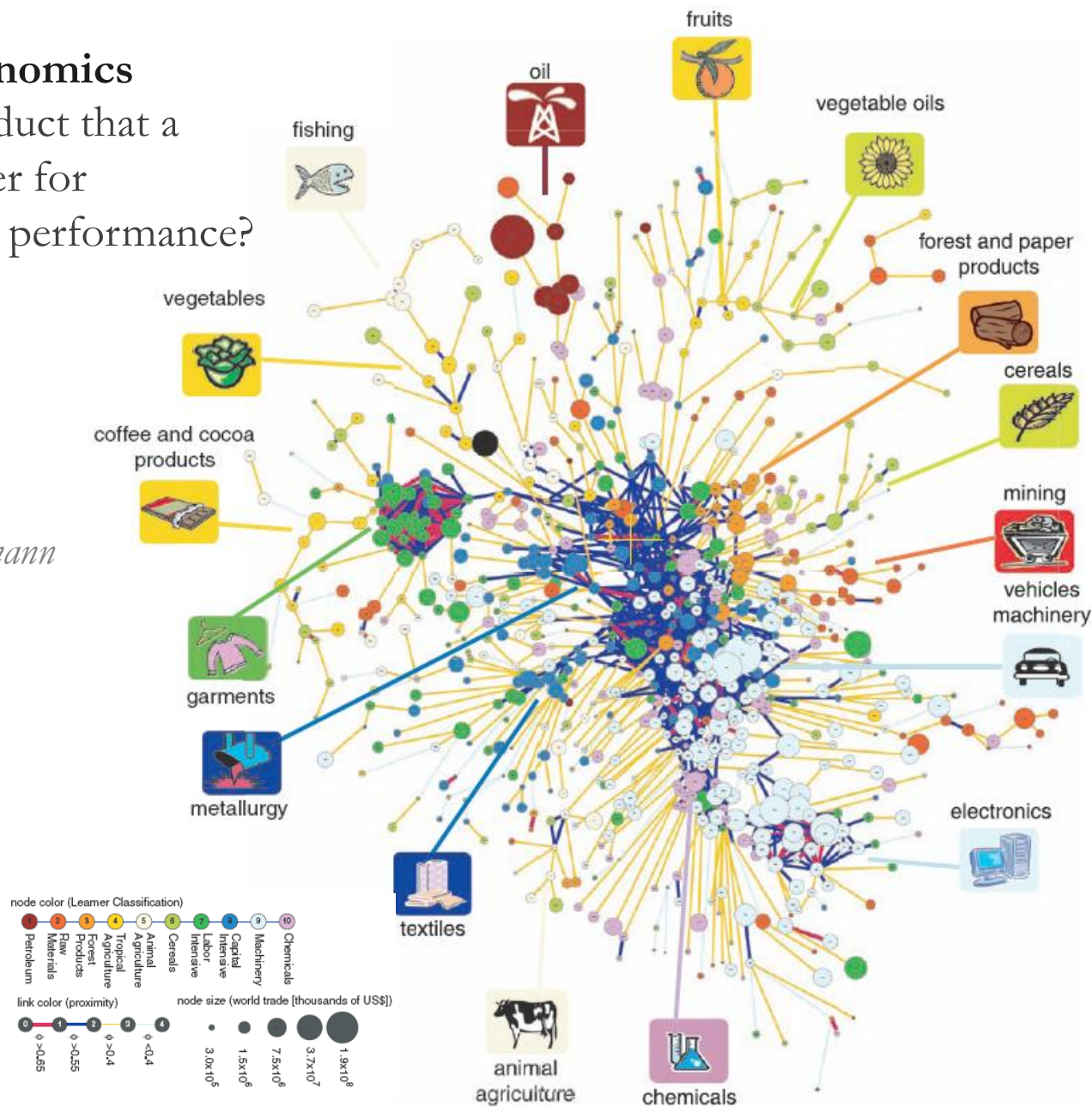


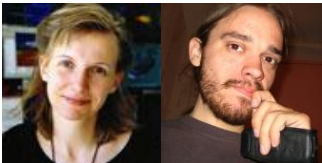
Fig. 1. The product space. (A) Hierarchically clustered proximity (ϕ) matrix representing the 775 SITC-4 product classes exported in the 1998–2000 period. (B) Network representation of the product space. Links are color coded

with their proximity value. The sizes of the nodes are proportional to world trade, and their colors are chosen according to the classification introduced by Leamer.

Computational Social Science

Studying large scale social networks such as Wikipedia

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

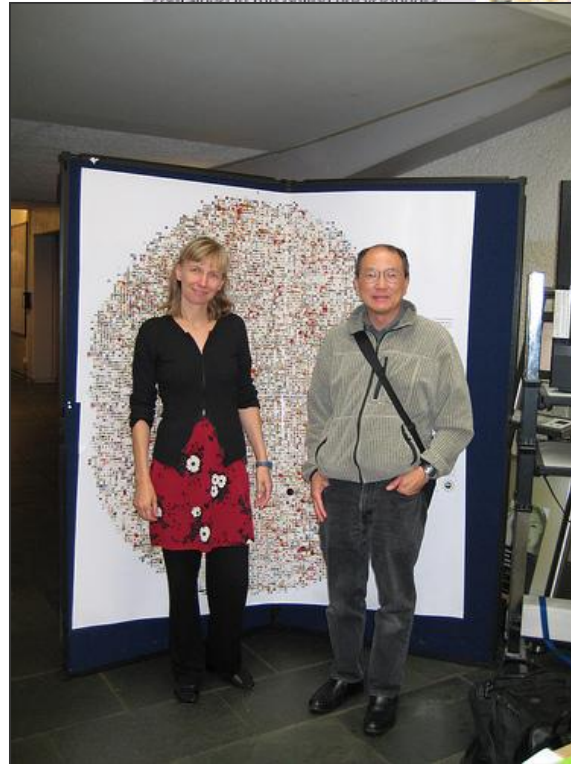


Second sight

Image: Bruce W. Herr and Todd M. Holloway

Power struggle

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



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



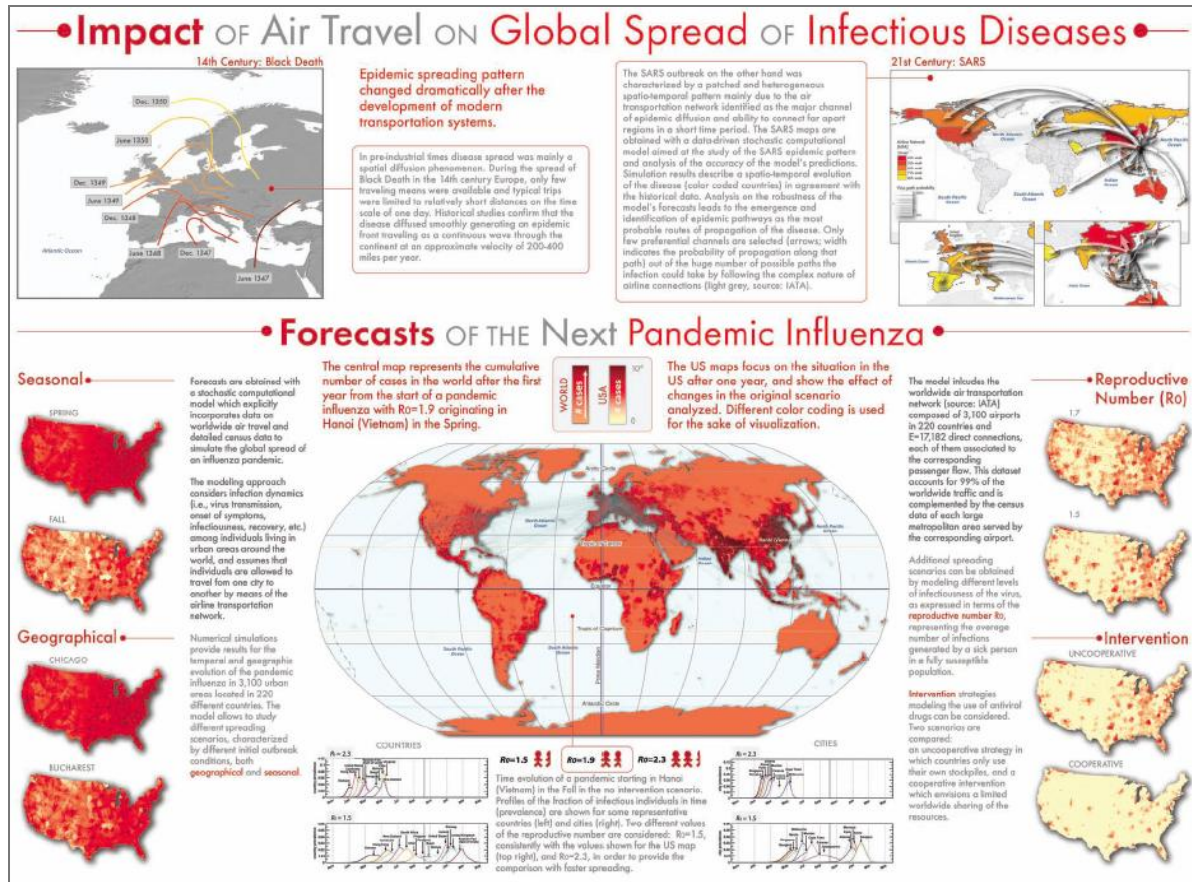
Computational Epidemics

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

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

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

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



NWB Tool 1.0.0

Can be freely downloaded for all major operating systems from <http://nwb.slis.indiana.edu>

Select your operating system from the pull down menu and download.

Unpack into a /nwb directory.

Run /nwb/nwb.exe

Session log files are stored in *'*yournwbdirectory*/logs'* directory.

Cite as

NWB Team. (2006). Network Workbench Tool. Indiana University, Northeastern University, and University of Michigan, <http://nwb.slis.indiana.edu>.

The screenshot shows the Network Workbench website in a browser window. The address bar displays <http://nwb.slis.indiana.edu/download.html>. The page has a red header with the Network Workbench logo and a navigation menu with 'Download' selected. The main content area is titled 'Download' and includes the following information:

- NWB Tool 1.0.0 Official Release** (September 15th, 2009)
- Release Notes** link
- A 'Select Your Operating System' dropdown menu with 'Windows (XP & Vista)' selected and a red 'DOWNLOAD' button.
- NWB-Demo** (Oct 13th, 2009)
- Release Notes** link
- Another 'Select Your Operating System' dropdown menu with 'Windows (XP & Vista)' selected and a red 'DOWNLOAD' button.

On the right side of the page, there is a line graph titled 'Downloads for NWB Tool Releases'. The y-axis is 'Number of Downloads per Month' (0 to 35,000) and the x-axis is 'Time' (Dec-06 to Jul-09). A 'Cumulative Total' line shows a steady increase in downloads over time, with a significant spike in late 2008. The graph also shows individual release points for versions v0.5, v0.6, v0.7, v0.8, v0.9, and v1.0.0.

NWB Tool Interface Components

The screenshot shows the Network Workbench Tool interface with the following components and callouts:

- Console:** Displays data operations (save, load, view, etc.) and algorithm input parameters, selection, & acknowledgements as well as error reporting. It contains a welcome message and a list of primary investigators: Dr. Katy Börner, Dr. Albert-László Barabási, Dr. Santiago Schnell, Dr. Alessandro Vespignani, Dr. Stanley Wasserman, and Dr. Eric A. Wernert. It also lists the developers: Weixia Huang, Russell Duhon, Micah Linnemeier, Timothy Kelley, Duygu Balcan, Mariano Beiró, Bruce Herr, Santo Fortunato, Ben Markines, Felix Terkhorn, Heng Zhang, Megha Ramawat, César Hidalgo, Ramya Sabbineni, Vivek Thakres, Soma Sanyal, Ann McCranie, Alessandro Vespignani, and Katy Börner. The URL <http://nwb.slis.indiana.edu> is provided for citation.
- Scheduler:** Lists what algorithms you've used and displays algorithm progress. It includes buttons for "Remove From List", "Remove all completed", and a play button. Below is a table with columns: Algorithm Name, Date, Time, and % Complete.
- Data Manager:** Keeps track of all datasets that are available for algorithmic visualization or manipulation. It includes a list of visualization options: Table, Matrix, Plot, Text, GUESS, Tree, and Network-.

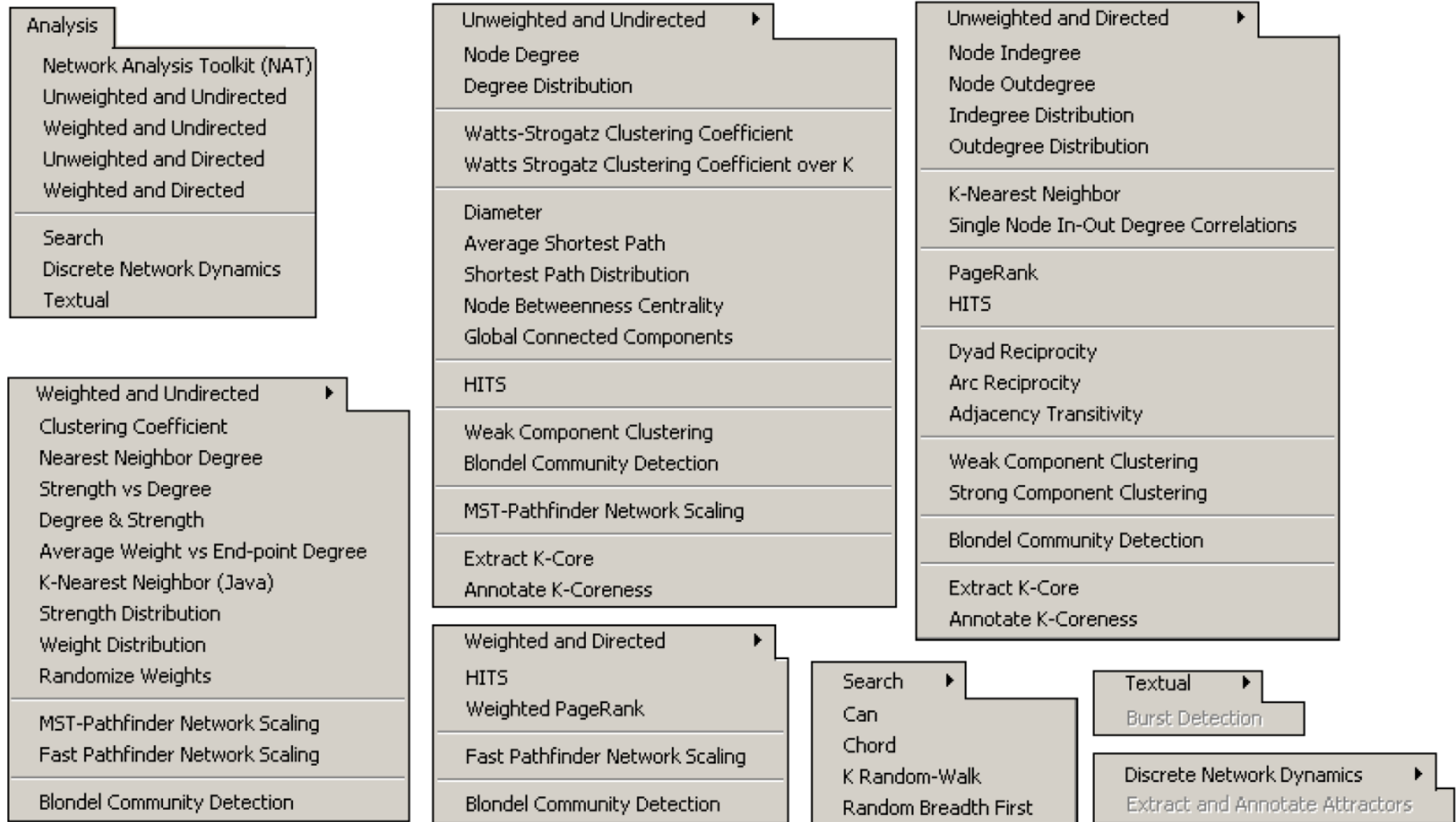
Console shows references to seminal works.
Workflows are recorded into a log file, and soon can be re-run for easy replication.
All algorithms are documented online; workflows are given in tutorials.

File	Preprocessing	Modeling	Visualization
Load...	Extract Top Nodes	Random Graph	GUESS
Load and Clean ISI File	Extract Nodes Above or Below Value	Watts-Strogatz Small World	GnuPlot
Read Directory Hierarchy Datasets	Remove Node Attributes	Barabási-Albert Scale-Free	DrL (VxOrd)
Save...	Delete High Degree Nodes	Can	Specified (prefuse beta)
View...	Delete Random Nodes	Chord	Circular (JUNG)
View with...	Delete Isolates	Hypergrid	Radial Tree/Graph (prefuse alpha)
Merge Node and Edge Files	Extract Top Edges	PRU	Radial Tree/Graph with Annotation (prefuse beta)
Split Graph to Node and Edge Files	Extract Edges Above or Below Value	TARL	Tree Map (prefuse beta)
Tests	Remove Edge Attributes	Discrete Network Dynamics (DND)	Tree View (prefuse beta)
Preferences	Remove Self Loops	Evolving Network (Weighted)	Balloon Graph (prefuse alpha)
Exit	Trim by Degree		Force Directed with Annotation (prefuse beta)
	Snowball Sampling (n nodes)		Kamada-Kawai (JUNG)
	Node Sampling		Fruchterman-Reingold (JUNG)
	Edge Sampling		Fruchterman-Reingold with Annotation (prefuse beta)
	Symmetrize		Spring (JUNG)
	Dichotomize		Small World (prefuse alpha)
	Multipartite Joining		Parallel Coordinates (demo)
	Normalize Text		LaNet
	Slice Table by Time		Circular Hierarchy

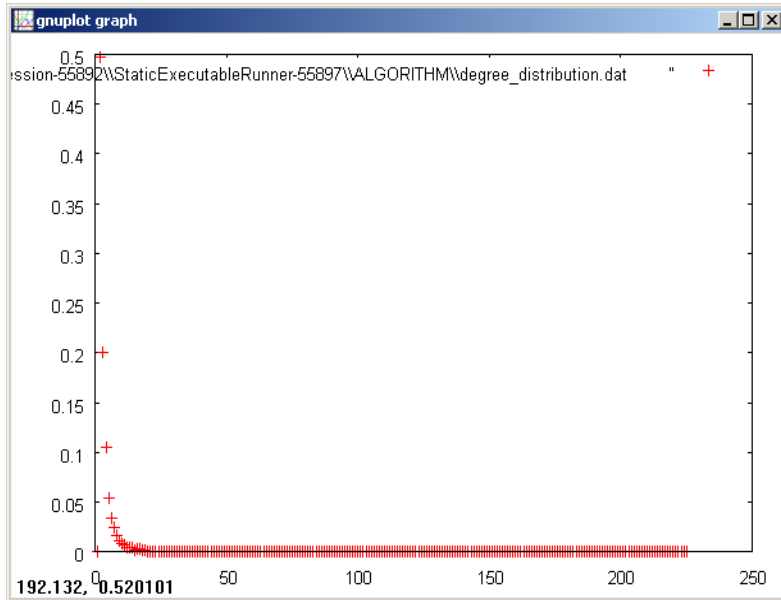
Börner, Katy, Sanyal, Soma and Vespignani, Alessandro (2007). **Network Science**. In Blaise Cronin (Ed.), *ARIST*, Information Today, Inc./American Society for Information Science and Technology, Medford, NJ, Volume 41, Chapter 12, pp. 537-607.

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

Analysis Menu and Submenus

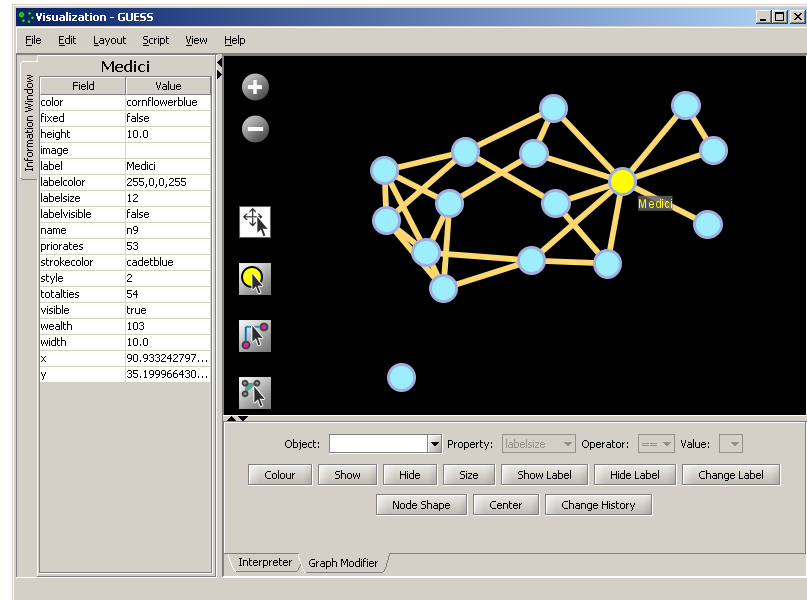


Börner, Katy, Sanyal, Soma and Vespignani, Alessandro (2007). **Network Science**. In Blaise Cronin (Ed.), *ARIST*, Information Today, Inc./American Society for Information Science and Technology, Medford, NJ, Volume 41, Chapter 12, pp. 537-607. <http://ivl.slis.indiana.edu/km/pub/2007-borner-arist.pdf>



Gnuplot

portable command-line driven
interactive data and function plotting
utility <http://www.gnuplot.info/>.



GUESS

exploratory data analysis and visualization tool
for graphs and networks.

<https://nwb.slis.indiana.edu/community/?n=VisualizeData.GUESS>.

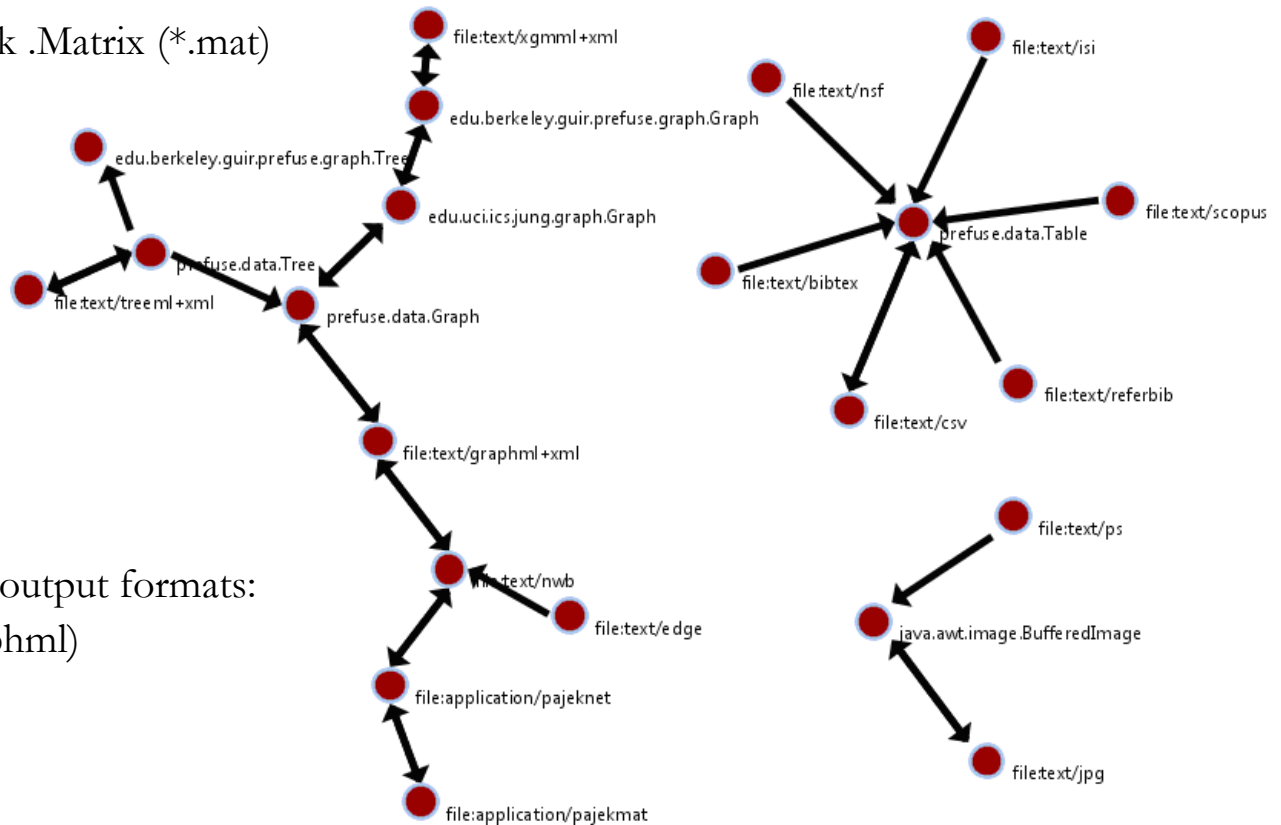
Supported Data Formats

The NWB tool supports loading the following input file formats:

- GraphML (*.xml or *.graphml)
- XGMML (*.xml)
- Pajek .NET (*.net) & Pajek .Matrix (*.mat)
- NWB (*.nwb)
- TreeML (*.xml)
- Edge list (*.edge)
- CSV (*.csv)
- ISI (*.isi)
- Scopus (*.scopus)
- NSF (*.nsf)
- Bibtext (*.bib)
- Endnote (*.enw)

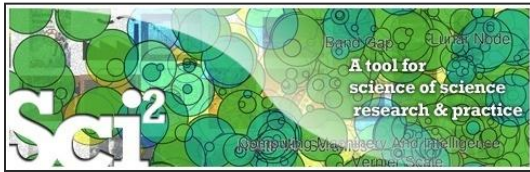
and the following network file output formats:

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



Formats are documented at <https://nwb.slis.indiana.edu/community/?n=DataFormats.HomePage>.

**CIShell Powered Tools:
Science of Science (Sci2) Tool**



Science of Science (Sci2) Tool

<http://sci.slis.indiana.edu>

- Explicitly designed for SoS research and practice, well documented, easy to use.
- Empowers many to run common studies while making it easy for exports to perform novel research.
- Advanced algorithms, effective visualizations, and many (standard) workflows.
- Supports micro-level documentation and replication of studies.
- Is open source—anybody can review and extend the code, or use it for commercial purposes.

nature

OPINION

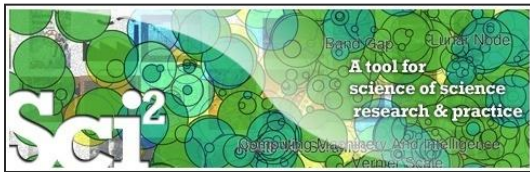
SUMMARY

- Existing metrics have known flaws
- A reliable, open, joined-up data infrastructure is needed
- Data should be collected on the full range of scientists' work
- Social scientists and economists should be involved

Vol 464|25 March 2010

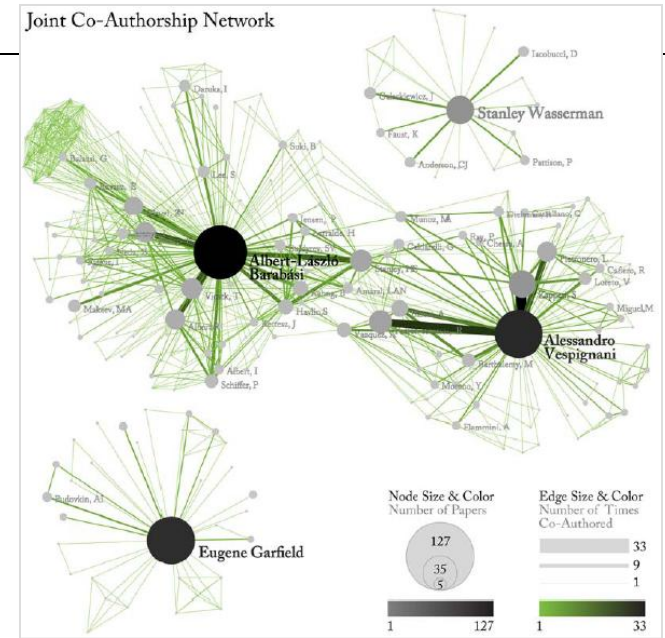
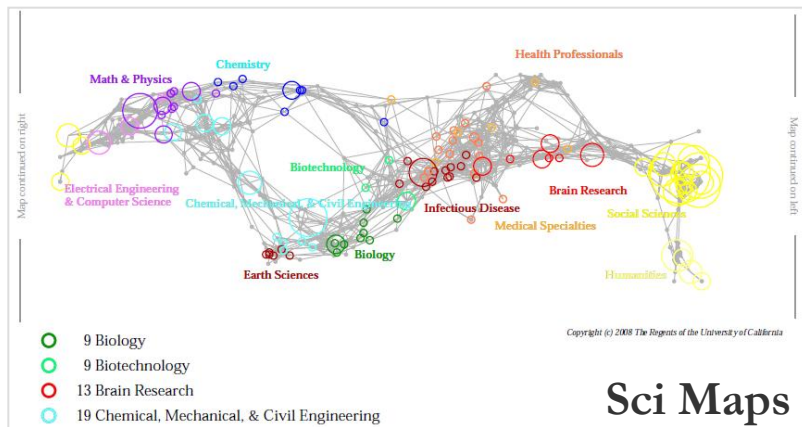
Let's make science metrics more scientific

To capture the essence of good science, stakeholders must combine forces to create an open, sound and consistent system for measuring all the activities that make up academic productivity, says **Julia Lane**.

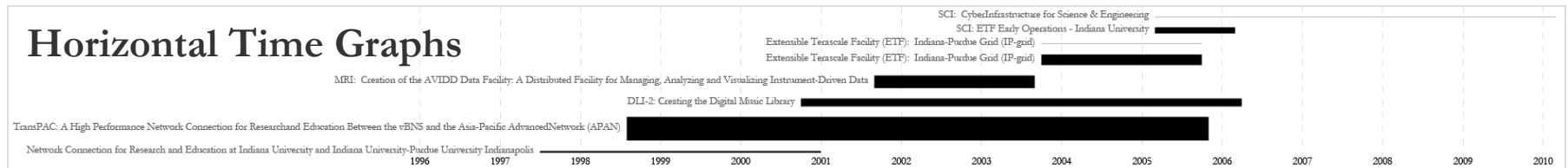


Sci² Tool – “Open Code for S&T Assessment”

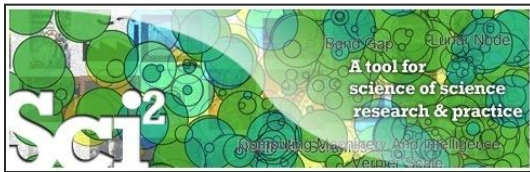
OSGi/CIShell powered tool with NWB plugins and many new scientometrics and visualizations plugins.



Horizontal Time Graphs



Börner, Katy, Huang, Weixia (Bonnie), Linnemeier, Micah, Dubon, Russell Jackson, Phillips, Patrick, Ma, Nianli, Zoss, Angela, Guo, Hanning & Price, Mark. (2009). *Rete-Netzwerk-Red: Analyzing and Visualizing Scholarly Networks Using the Scholarly Database and the Network Workbench Tool*. Proceedings of ISSI 2009: 12th International Conference on Scientometrics and Informetrics, Rio de Janeiro, Brazil, July 14-17. Vol. 2, pp. 619-630.



Sci² Tool

Sci² Tool

File Preprocessing Modeling Analysis Visualization Scientometrics Help

Console

Welcome to the Science of Science Tool (Sci²). The development of this tool is supported in Network Science center and the School of Li Indiana University, the National Science Foundation and IIS-0715303, and the James S. McDonnell Cyberinfrastructure portal (<http://sci.slis.indiana.edu>)

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

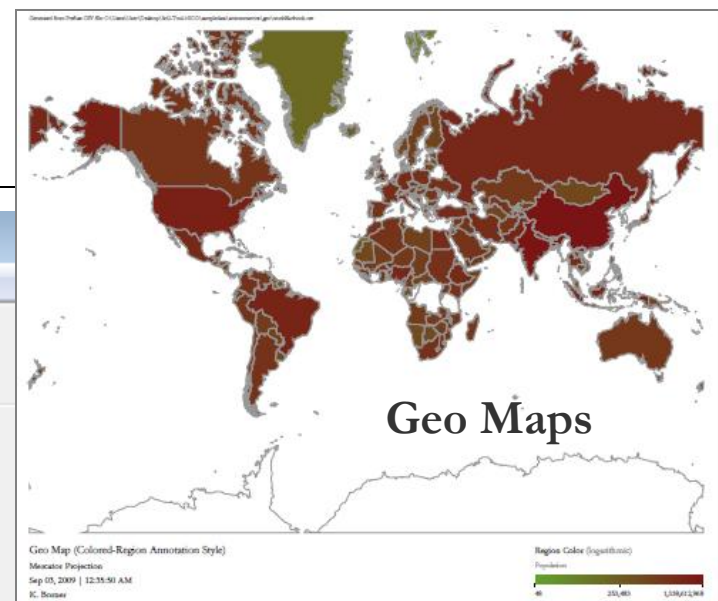
Please cite as follows:
 Sci² Team. (2009). Science of Science Tool. In SciTech Strategies Inc., <http://sci.slis.indiana.edu>.

Scheduler

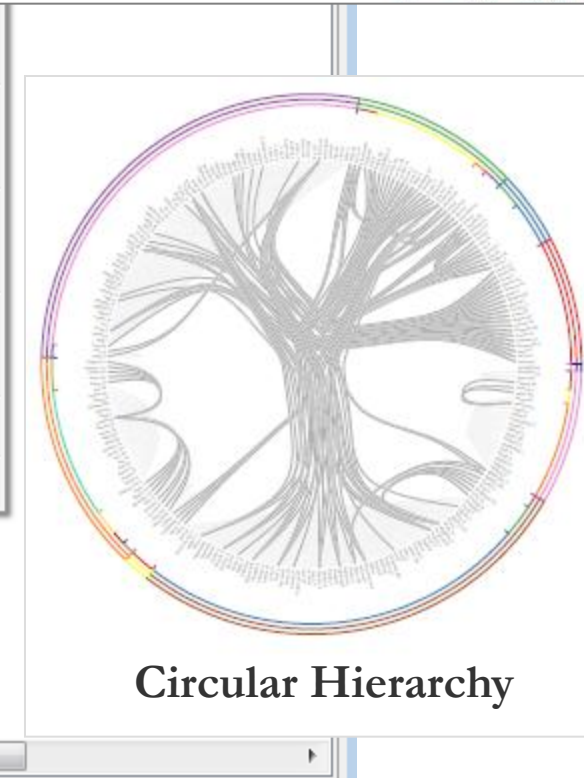
Remove From List Remove completed

!	Algorithm Name	Date	Time	% Comp
<input checked="" type="checkbox"/>	Extract Co-Author Netw...	09/03/2009	00:15:20 AM	<div style="width: 100%; height: 10px; background-color: green;"></div>
<input checked="" type="checkbox"/>	Load and Clean ISI File	09/03/2009	00:15:05 AM	<div style="width: 100%; height: 10px; background-color: green;"></div>

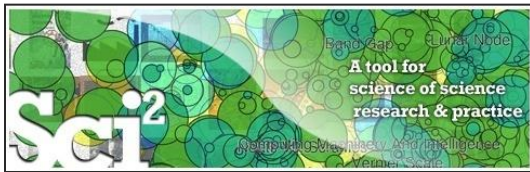
- GUESS
- GnuPlot
- Radial Tree/Graph (prefuse alpha)
- Radial Tree/Graph with Annotation (prefuse beta)
- Tree View (prefuse beta)
- Tree Map (prefuse beta)
- Force Directed with Annotation (prefuse beta)
- Fruchterman-Reingold with Annotation (prefuse beta)
- DrL (VxOrd)
- Specified (prefuse beta)
- Horizontal Line Graph
- Circular Hierarchy
- Geo Map (circle annotations)
- Geo Map (region coloring annotations)
- Image Viewer
- RefMapper



Geo Maps



Circular Hierarchy



Sci² Tool: Download, Install, and Run

Sci² Tool 0.5 Alpha (May 2011)

Can be freely downloaded for all major operating systems from <http://sci.slis.indiana.edu/sci2>

Select your operating system from the pull down menu and download.

Unpack into a /sci2 directory.

Run /sci2/sci2.exe

Session log files are stored in *'*yournwbdirectory*/logs'* directory.

Cite as

Sci² Team. (2009). Science of Science (Sci²) Tool. Indiana University and SciTech Strategies, <http://sci.slis.indiana.edu>

Sci² Registration : Welcome - Windows Internet Explorer
http://sci.slis.indiana.edu/registration/user/

File Edit View Favorites Tools Help

gigapan: no intro Sci² Registration : Welcome

Sci² Tool
A tool for science of science research & practice

Email Address
Password

Forgot your password?
To recover your account password, please visit our [password recovery page](#).

Not registered yet?
[Register now](#)

Tutorials

Scott Weingart, Hanning Guo, Katy Borner, Kevin W. Boyack, Micah W. Linnemeier, Russell J. Duhon, Patrick A. Phillips, Chintan Tank, and Joseph Biberstine (2010) [Science of Science \(Sci²\) Tool User Manual](#). Cyberinfrastructure for Network Science Center, School of Library and Information Science, Indiana University, Bloomington.

Katy Borner and Angela Zoss (2010) [Plug-and-Play Macrosopes Tutorial](#), International Conference on Social Computing, Behavioral Modeling and Prediction, Bethesda, MD.

In the news

The Trustees of Indiana University. [SLIS researcher promotes Innovation Dashboard for policymakers during Capitol Hill visit. \(website accessed 12/22/09\).](#)

Please cite as

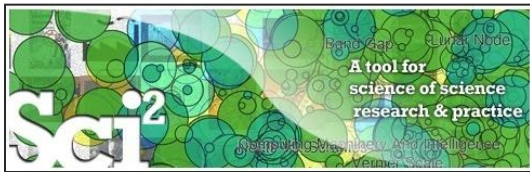
Sci² Team. (2009). Science of Science (Sci²) Tool. Indiana University and SciTech Strategies, <http://sci.slis.indiana.edu>.

Acknowledgements

This work is supported in part by the Cyberinfrastructure for Network Science Center and the School of Library and Information Science at Indiana University, the National Science Foundation under Grant No. SBE-0738111 and IIS-0513650, and the James S. McDonnell Foundation.

INDIANA UNIVERSITY
SCHOOL OF LIBRARY AND INFORMATION SCIENCE

James S. McDonnell Foundation



Sci² Tool: Download, Install, and Run

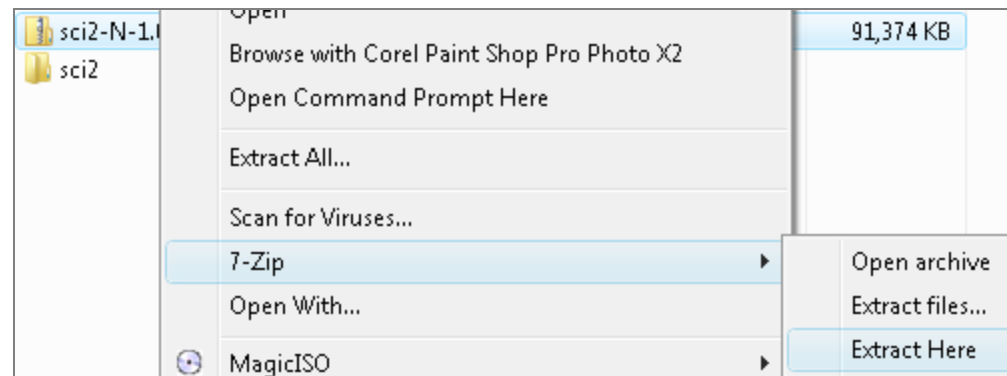
Sci² Tool 0.5 Alpha (May 2011)

Has new features such as

- New Geographic Visualizations
- STAR database (**download separately**)
- Colored Horizontal Bar Graphs
- Supports ASCII UTF-8 characters
- Bug fixes, streamlined workflows

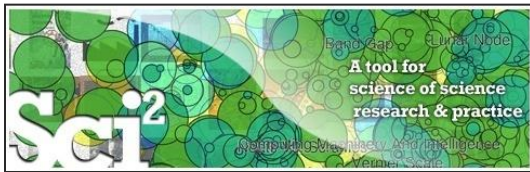
Name	Size
0 - 9 (1)	
2010-03-sci2-manual.pdf	10,947 KB
Q - Z (4)	
sci2-N-1.0.0.201008130505NGT-macosx.carbon.ppc.zip	91,415 KB
sci2-N-1.0.0.201008130505NGT-macosx.carbon.x86.zip	91,414 KB
sci2-N-1.0.0.201008130505NGT-macosx.cocoa.x86_64.zip	89,717 KB
sci2-N-1.0.0.201008130505NGT-win32.win32.x86.zip	91,374 KB

Unzip and run /sci2/sci2.exe



Cite as

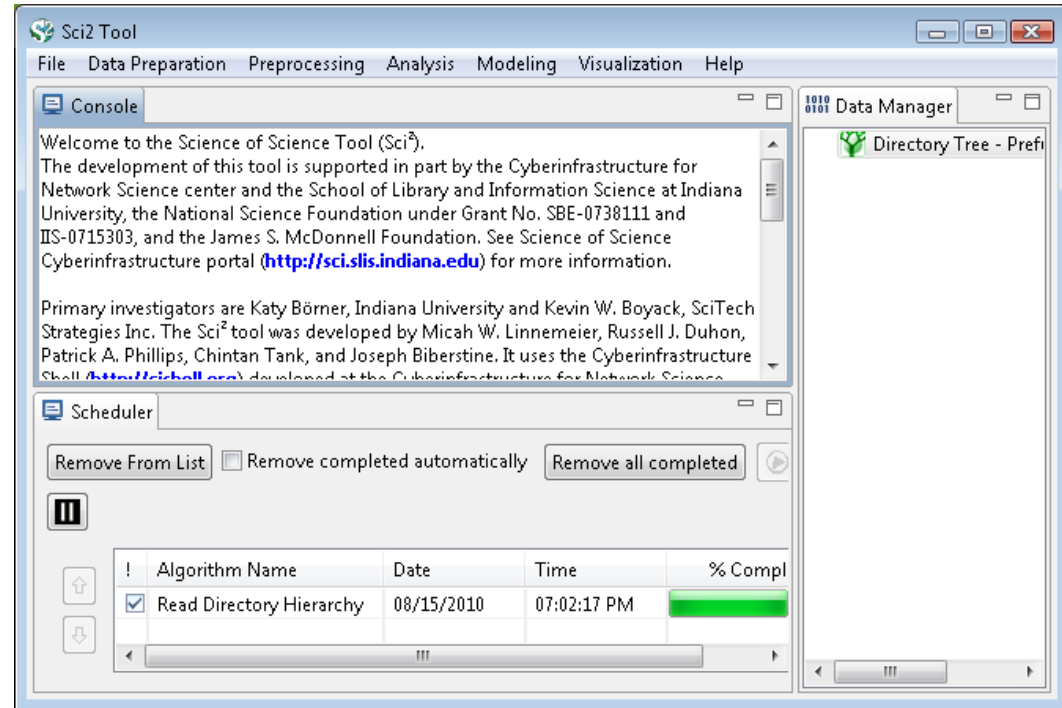
Sci² Team. (2009). Science of Science (Sci²) Tool. Indiana University and SciTech Strategies, <http://sci.slis.indiana.edu>



Sci2 Tool Interface Components

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 soon can be re-run for easy replication. If errors occur, they are saved in a error log to ease bug reporting.

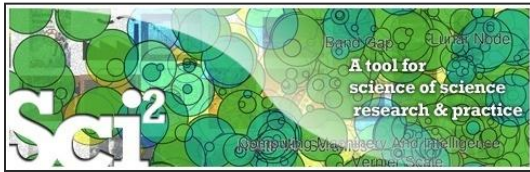
All algorithms are documented online; workflows are given in tutorials, see <http://sci.slis.indiana.edu/sci2> and <http://nwb.slis.indiana.edu> > Community

Contents

Contents.....		
1 Introduction		
2 Getting Started		
2.1 Download, Install, Uninstall.....		
2.2 User Interface		
2.2.1 Menus.....		
2.2.2 Console.....		
2.2.3 Data Manager.....		
2.2.4 Scheduler.....		
2.3 Data Formats		
2.4 Saving Visualizations for Publication		
2.5 Sample Datasets		
3 Algorithm and Tool Plugins		
3.1 Sci ² Tool Plugins.....		
3.2 Load, View, and Save Data.....		
3.3 Memory Allocation		
3.3.1 Windows and Linux.....		
3.3.2 Mac.....		
3.4 Memory Limits.....		
4 Workflow Design		
4.1 Overview.....		
4.2 Data Acquisition and Preparation.....		
4.2.1 Datasets: Publications		
4.2.2 Datasets: Funding.....		
4.2.3 Datasets: Scholarly Database		
4.3 Database Loading and Manipulation.....		
4.4 Summaries and Table Extractions.....		
4.5 Statistical Analysis/Profiling.....		
4.6 Temporal Analysis (When).....		
4.6.1 Burst Detection.....		
4.6.2 Slice Table by Time		
4.7 Geospatial Analysis (Where).....		
4.8 Topical Analysis (What).....		
4.8.1 Word Co-Occurrence Network.....		
4.9 Network Analysis (With Whom?).....	35	
4.9.1 Network Extraction.....	35	
4.9.2 Compute Basic Network Characteristics.....		
4.9.3 Network Analysis.....		
4.9.4 Network Visualization.....		
4.10 Modeling (Why?).....		
4.10.1 Random Graph Model.....		
4.10.2 Watts-Strogatz Small World.....		
4.10.3 Barabási-Albert Scale Free Model.....		
5 Sample Workflows		
5.1 Individual Level Studies - Micro.....		
5.1.1 Mapping Collaboration, Publication and Funding Profiles of One Researcher.....		
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 NetSci Researchers (ISI Data).....		
5.1.5 Studying Four Major NetSci Researchers (ISI Data) using Database		
5.2 Institution Level Studies - Meso		
5.2.1 Funding Profiles of Three Universities (NSF Data).....		
5.2.2 Funding Profiles of Three Universities (NSF Data) Using Database.....		
5.2.3 Mapping CTSA Centers (NIH RePORTER Data).....		
5.2.4 Biomedical Funding Profile of NSF (NSF Data).....		
5.2.5 Mapping Scientometrics (ISI Data)		
5.2.6 Burst Detection in Scientometrics (ISI Data).....		
5.2.7 Mapping the Field of RNAi Research (SDB Data).....		
5.3 Global Level Studies - Macro.....		
5.3.1 Geo USPTO (SDB Data).....		
6 Sample Science Studies & Online Services.....		
6.1 Science Dynamics.....		
6.1.1 Mapping Topics and Topic Bursts in PNAS (2004).....		
6.2 Local Impact-Output / ROI Studies		
6.2.1 Indicator-Assisted Evaluation and Funding of Research: Visualizing the Influence of Grants on the Number and Citation Counts of Research Papers (2003).....	87	
6.2.2 Mapping Transdisciplinary Tobacco Use Research Centers Publications (forthcoming).....	88	
6.3 Local and Global Science Studies.....	89	
6.3.1 Mapping the Evolution of Co-Authorship Networks (2004).....	89	
6.3.2 Studying the Emerging Global Brain: Analyzing and Visualizing the Impact of Co-Authorship Teams (2005).....	90	
6.3.4 Mapping the Diffusion of Information Among Major U.S. Research Institutions (2006).....	92	
6.3.5 Research Collaborations by the Chinese Academy of Sciences (2009)	93	
6.3.6 Mapping the Structure and Evolution of Chemistry Research (2009).....	94	
6.3.7 Science Map Applications: Identifying Core Competency (2007).....	95	
6.4 Modeling Science.....	96	
6.4.1 113 Years of Physical Review: Using Flow Maps to Show Temporal and Topical Citation (2008).....	96	
6.4.2 The Simultaneous Evolution of Author and Paper Networks (2004).....	97	
6.5 Accuracy Studies.....	98	
6.5.1 Mapping the Backbone of Science (2005).....	98	
6.5.2 Toward a Consensus Map of Science (2009).....	99	
6.6 Databases and Tools.....	100	
6.6.1 The Scholarly Database and Its Utility for Scientometrics Research (2009).....	100	
6.6.2 Reference Mapper.....	101	
6.6.3 Rete-Netzwerk-Red: Analyzing and Visualizing Scholarly Networks Using the Scholarly Database and the Network Workbench Tool (2009).....	102	
6.7 Interactive Online Services	103	
6.7.1 The NIH Visual Browser: An Interactive Visualization of Biomedical Research (2009).....	103	
6.7.2 Interactive World and Science Map of S&T Jobs (2010).....	104	
7 Extending the Sci² Tool	105	
7.1 CShell Basics.....	105	
7.2 Read New Data	105	
7.3 Creating and Sharing New Algorithm Plugins.....	105	
7.4 Tools That Use OSGI and/or CShell.....	106	
8 Relevant Datasets and Tools.....	107	
8.1 Datasets.....	107	
8.2 Network Analysis Tools.....	108	
9 References	111	

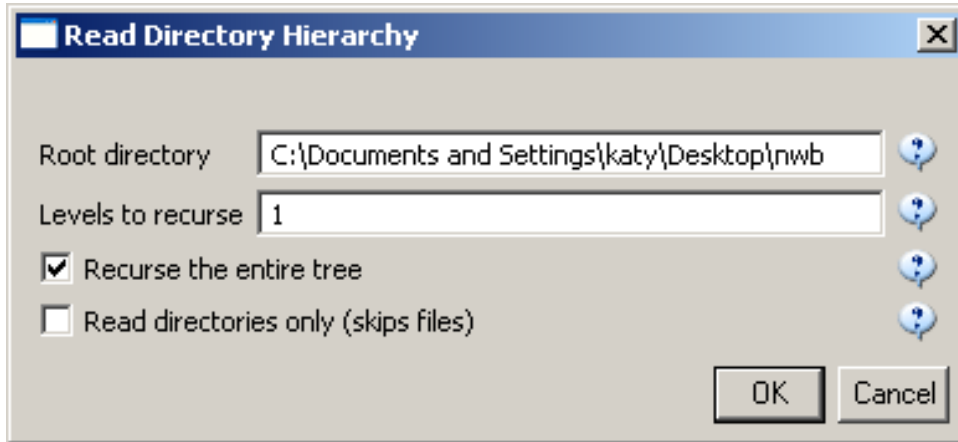
See Sci2 Manual

 2010-03-sci2-manual.pdf



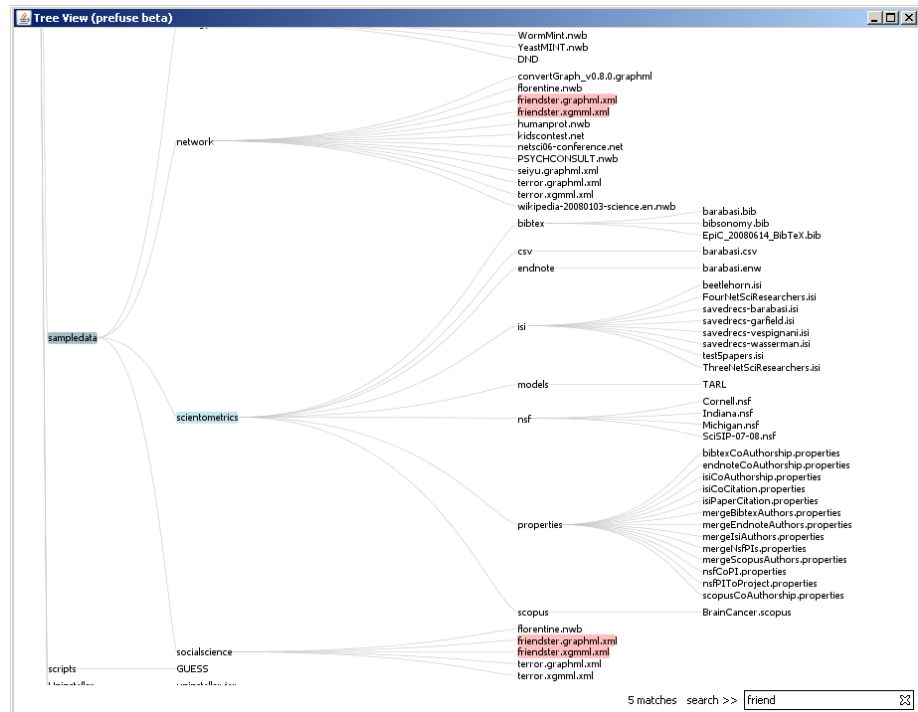
Sci2 Tool – Read+Visualize Sci2 Tool Directory Tree

Use *'File > Read Directory Hierarchy'* with parameters



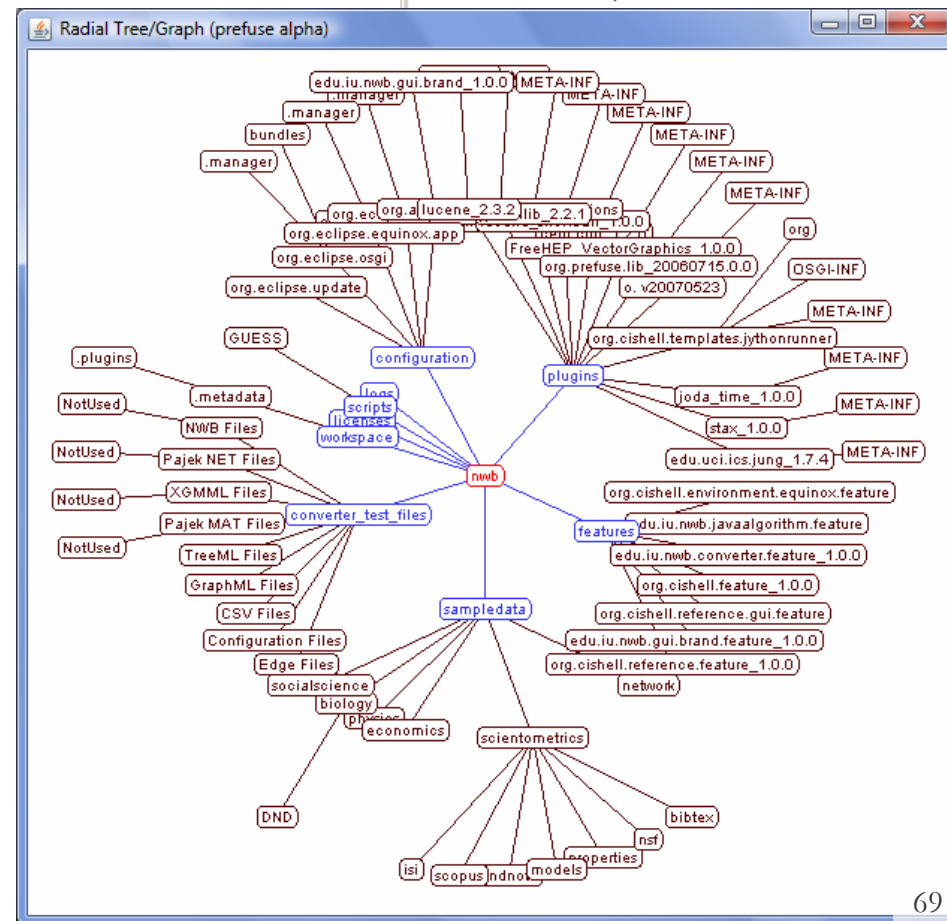
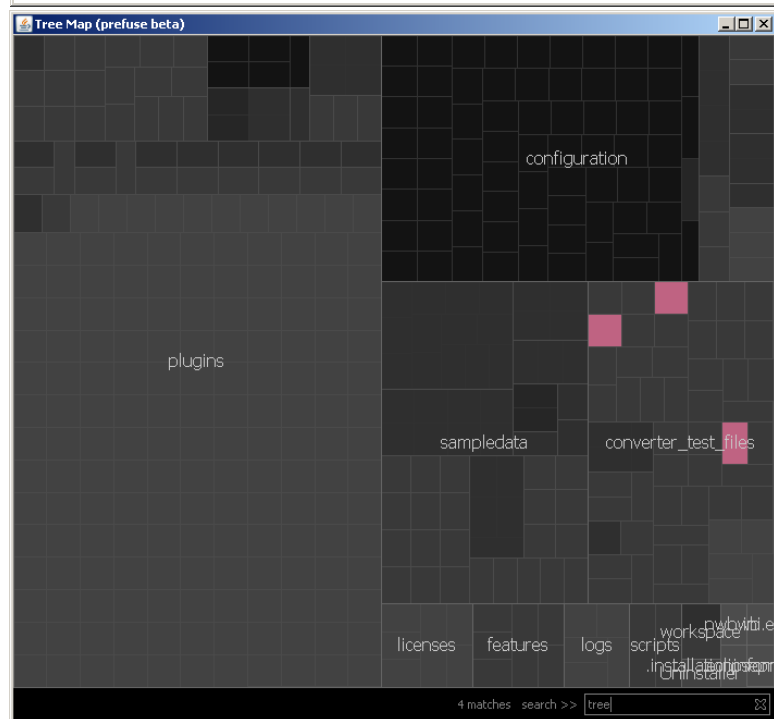
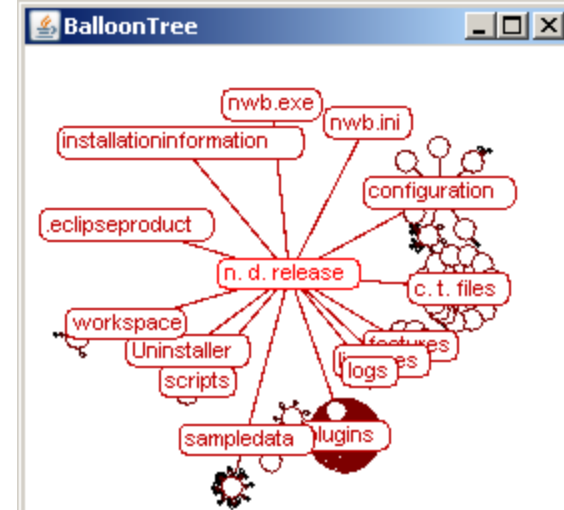
Visualize resulting *'Directory Tree - Prefuse (Beta) Graph'* using

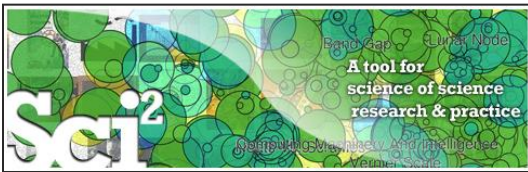
- *'Visualization > Tree View (prefuse beta)'*
- *'Visualization > Tree Map (prefuse beta)'*
- *'Visualization > Balloon Graph (prefuse alpha)'*
- *'Visualization > Radial Tree/Graph (prefuse alpha)'*



Different views of the /nwb directory hierarchy.

Note the size of the /plugin directory.



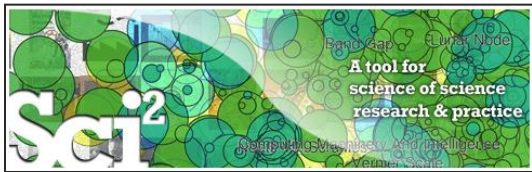


Studying Four Major NetSci Researchers (ISI Data)

FourNetSciResearchers.isi	
Time frame:	1955-2007
Region(s):	Miscellaneous
Topical Area(s):	Network Science
Analysis Type(s):	Paper Citation Network, Co-Author Network, Bibliographic Coupling Network, Document Co-Citation Network, Word Co-Occurrence Network

Thomson Reuter’s Web of Knowledge (WoS) is a leading citation database cataloging over 10,000 journals and over 120,000 conferences. Access it via the “Web of Science” tab at <http://www.isiknowledge.com> (**note:** access to this database requires a paid subscription). Along with Scopus, WoS provides some of the most comprehensive datasets for scientometric analysis.

To find all publications by an author, search for the last name and the first initial followed by an asterisk in the author field.



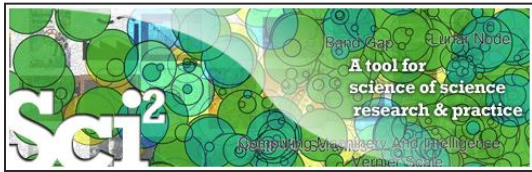
Data Acquisition from Web of Science

Download all papers by

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

from

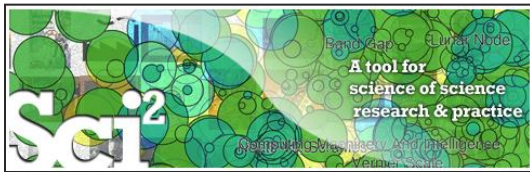
- Science Citation Index Expanded (SCI-EXPANDED)
--1955-present
- Social Sciences Citation Index (SSCI)--1956-present
- Arts & Humanities Citation Index (A&HCI)--1975-present



Comparison of Counts

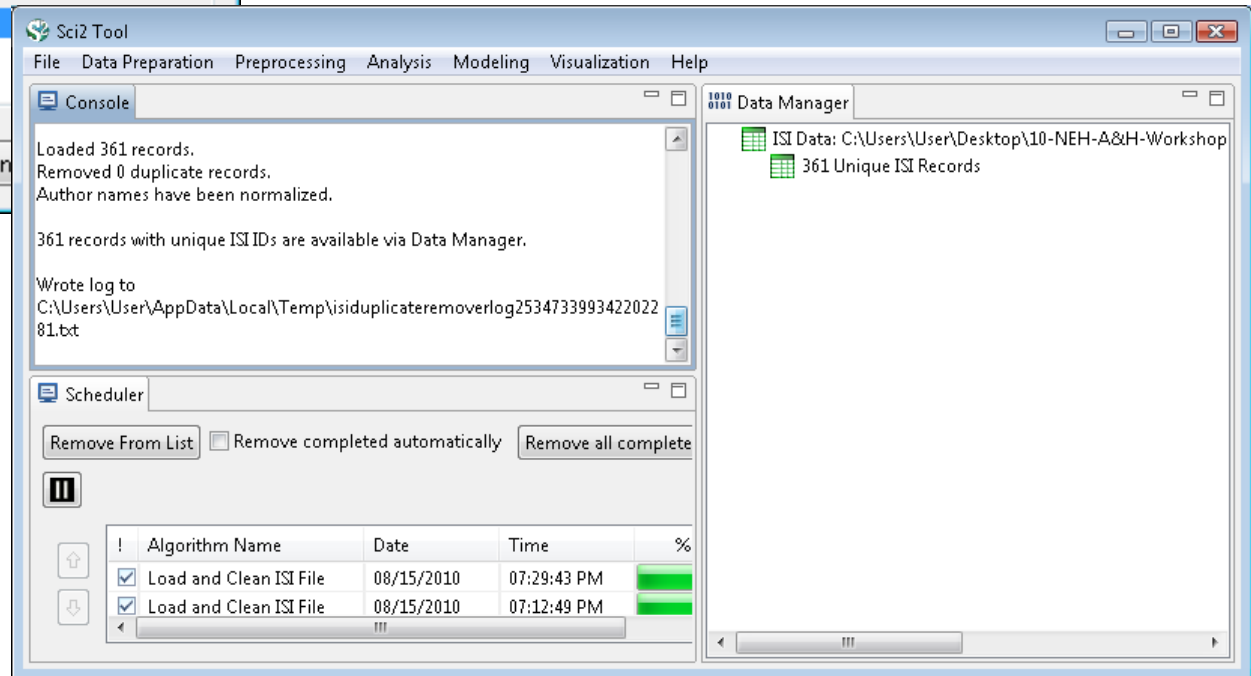
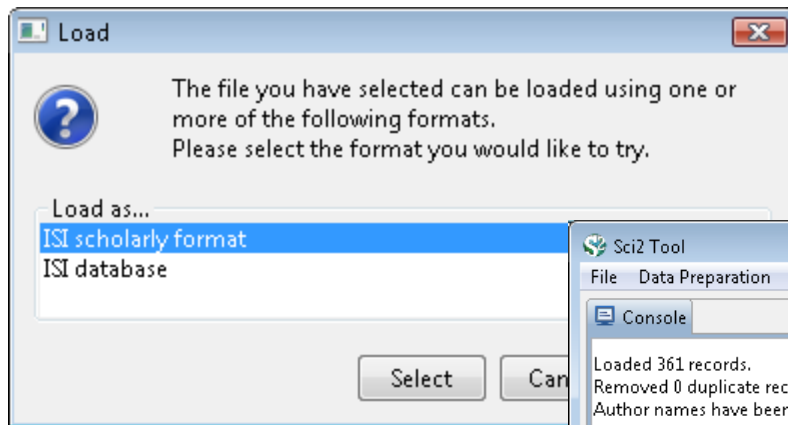
No books and other non-WoS publications are covered.

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

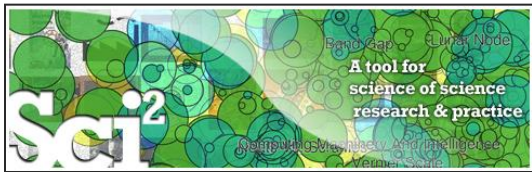


Extract Co-Author Network

Load **yoursci2directory*/sampledata/scientometrics/isi/FourNetSciResearchers.isi*
 using 'File > Load' and parameters



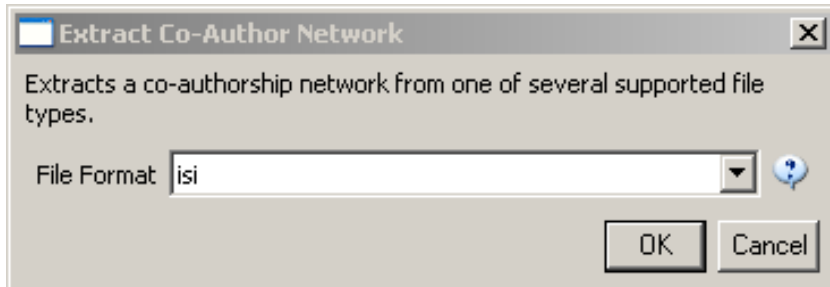
And file with 361 records
 appears in Data Manager.



Extract Co-Author Network

(see section 5.1.4.2 on correcting duplicate/ misspelled author names)

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



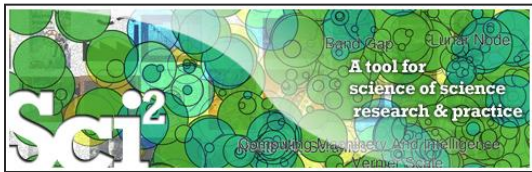
The result is an undirected but weighted network of co-authors in the Data Manager.

Run *'Analysis > Network > Network Analysis Toolkit (NAT)'* to calculate basic properties: the network has 247 nodes and 891 edges.

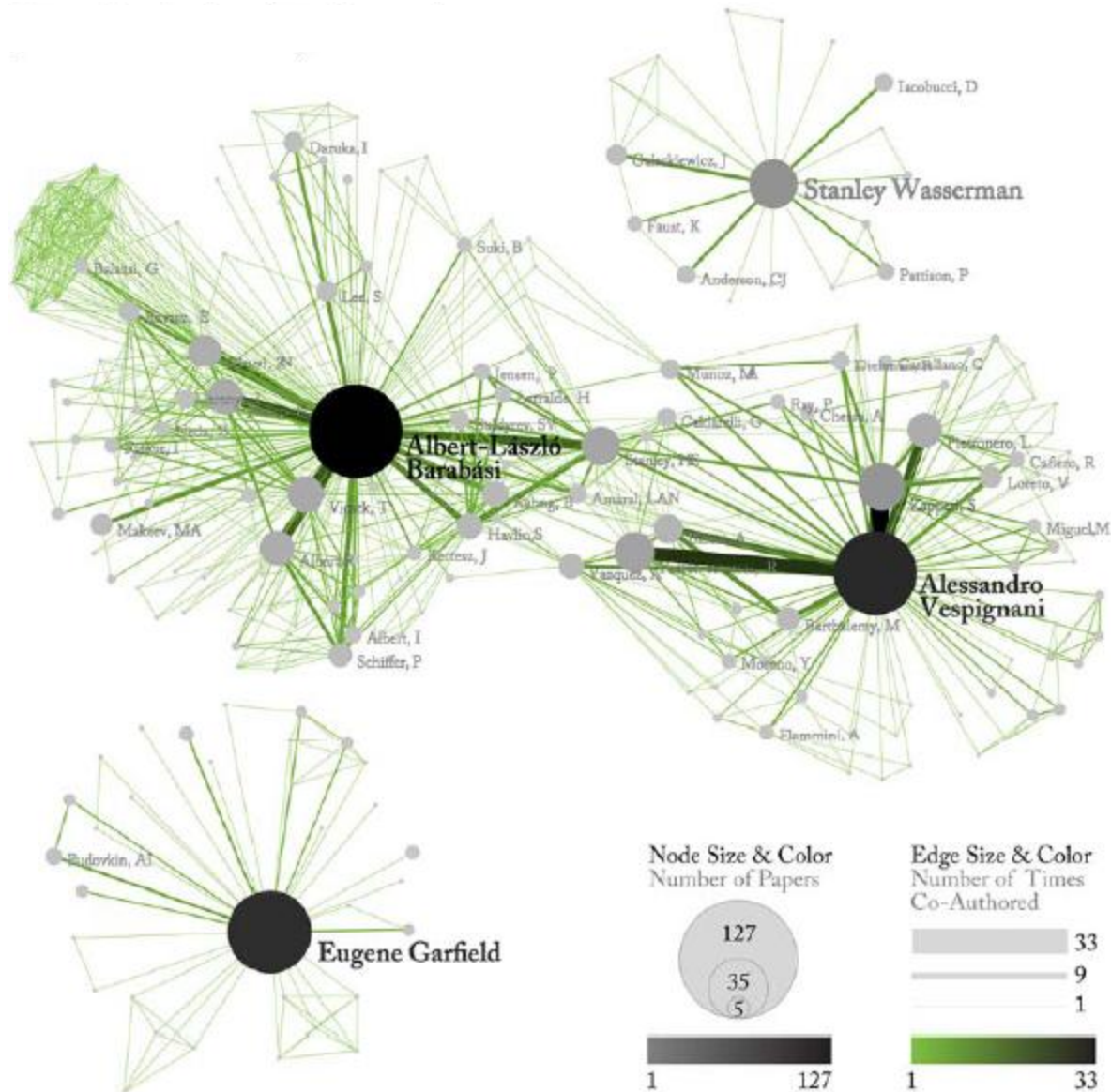
Use *'Analysis > Network > Unweighted and Undirected > Node Degree'* to calculate the number of neighbors for each node.

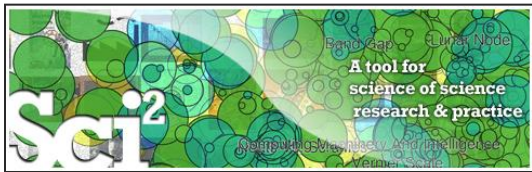
To view the complete network, select the *'Extracted Co-Authorship Network'* and run *'Visualization > Networks > GUESS'*.

Network is loaded with random layout. In GUESS, run *'Layout > GEM'* and *'Layout > Bin Pack'* to improve layout.



Co-Author Network of all Four NetsSci Researchers





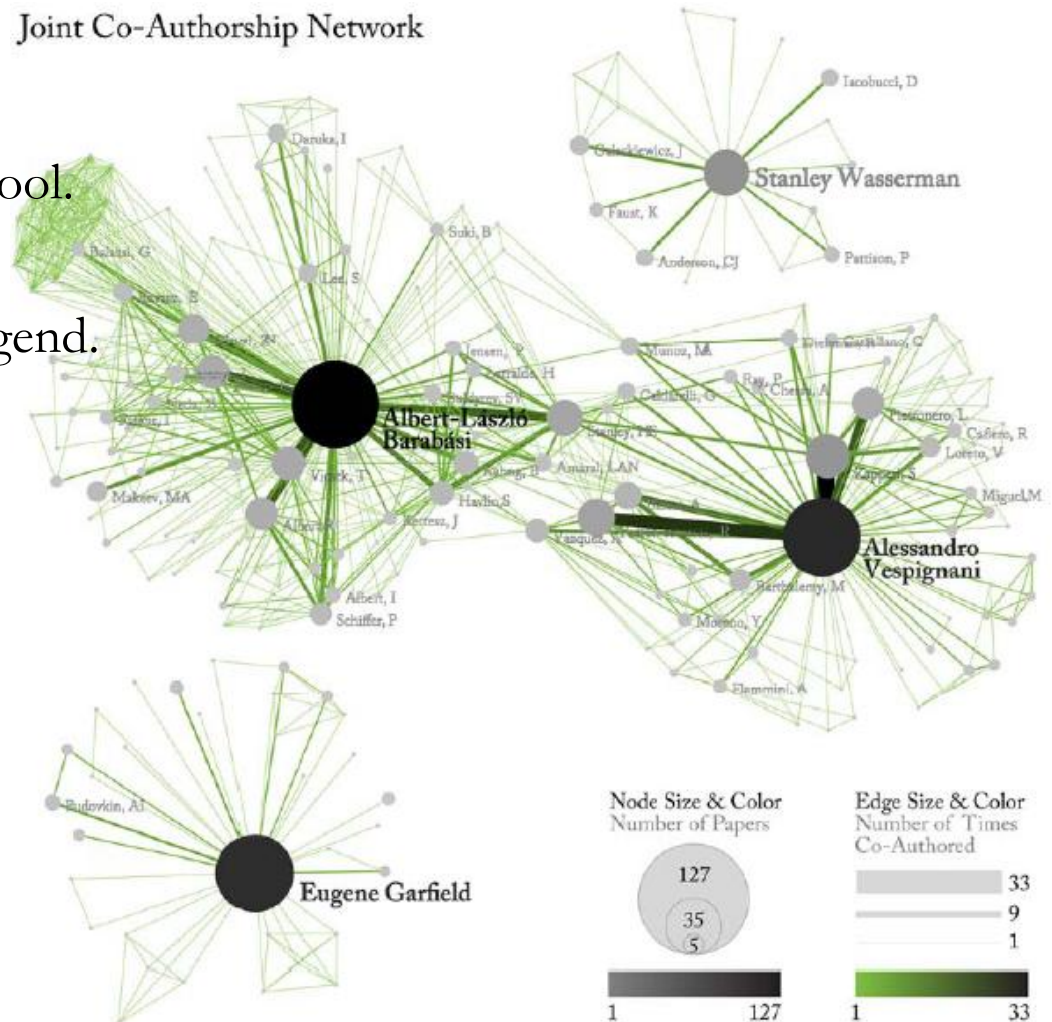
Co-Author Network of all Four NetsSci Researchers

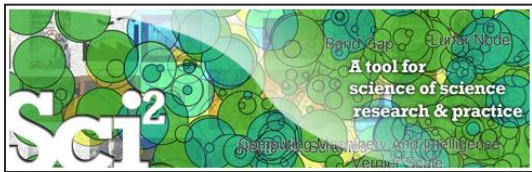
Use the GUESS Graph Modifier to change color and size coding.

Joint Co-Authorship Network

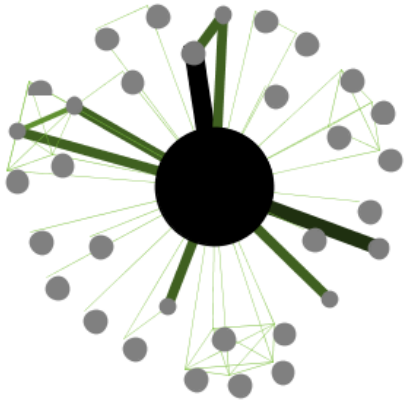
Calculate node degrees in Sci2 Tool.

Use a graphic program to add legend.

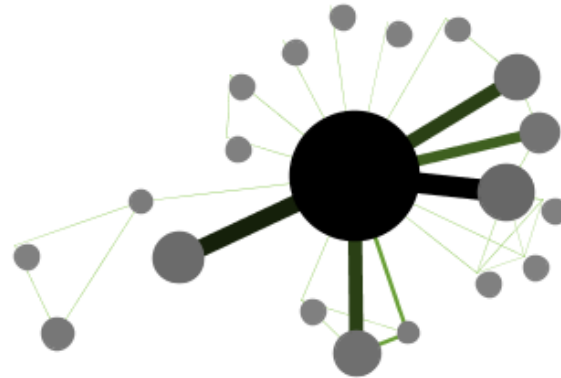




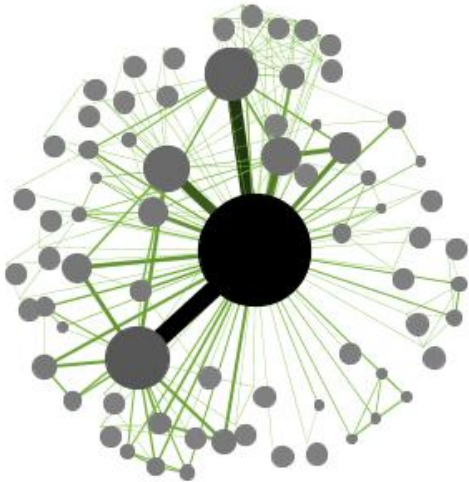
Individual Co-Author Networks (Read/map 4 files separately)



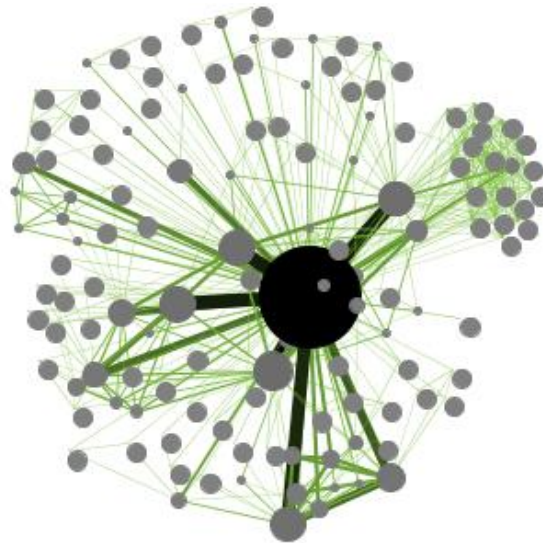
Eugene Garfield



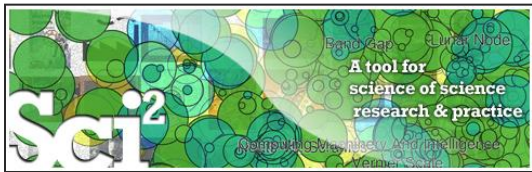
Stanley Wasserman



Alessandro Vespignani



Albert-László Barabási



Network Visualization: Node Layout

Load and Clean ISI File was selected.
Loaded 361 records.
Removed 0 duplicate records.
Author names have been normalized.
361 records with unique ISI IDs are available
via Data Manager.

.....

Extract Co-Author Network was selected.

Input Parameters:

File Format: isi

.....

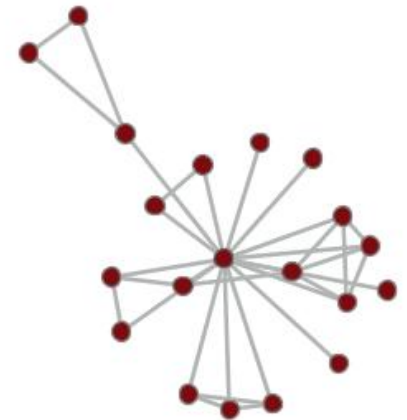
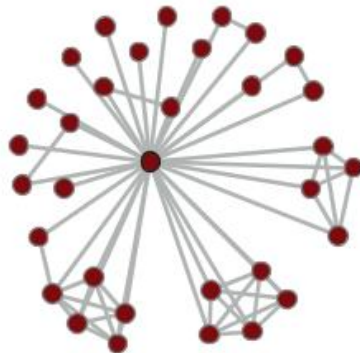
Network Analysis Toolkit (NAT) was selected.

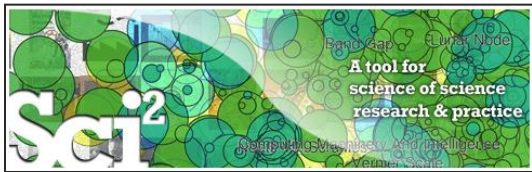
Nodes: 247

Edges: 891

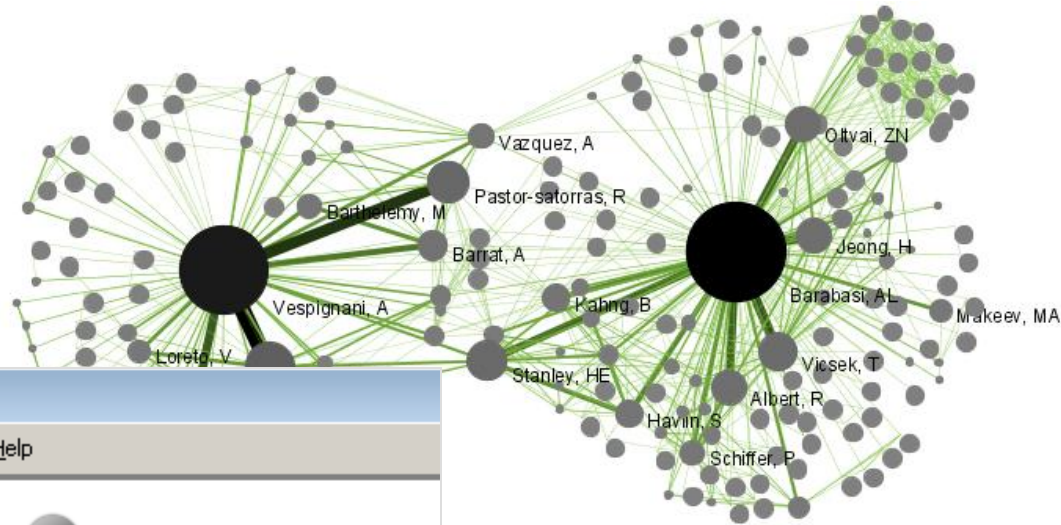
.....

GUESS was selected.





Network Visualization: Color/Size Coding by Data Attribute Values



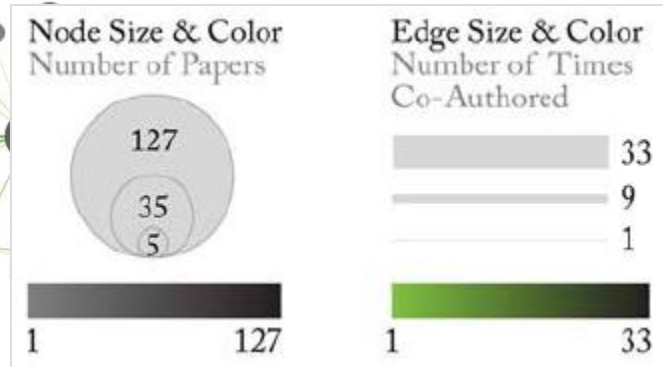
Visualization - GUESS

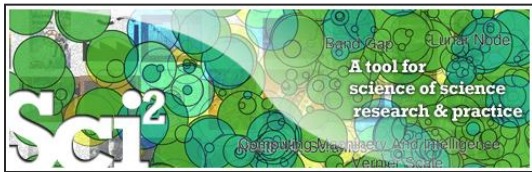
File Edit Layout Script View Help

Vespignani, A

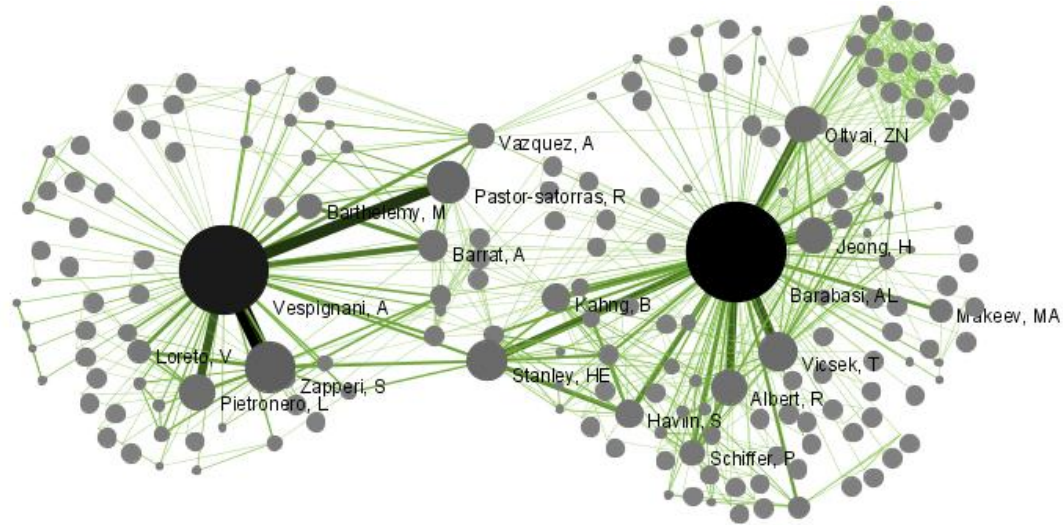
Field	Value
color	125,12,17,255
fixed	false
height	10.0
image	
label	Vespignani, A
labelcolor	0,0,0,255
labelsize	12
labelvisible	false
name	n161
numberofworks	101
originallabel	Vespignani, A
strokecolor	black
style	2
timescited	3811
visible	true
width	10.0
x	586.75
y	107.25

Information Window





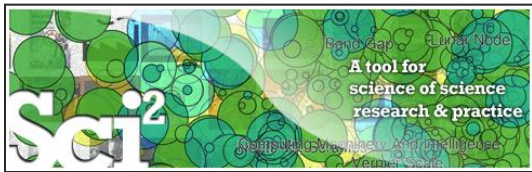
Network Visualization: Giant Component



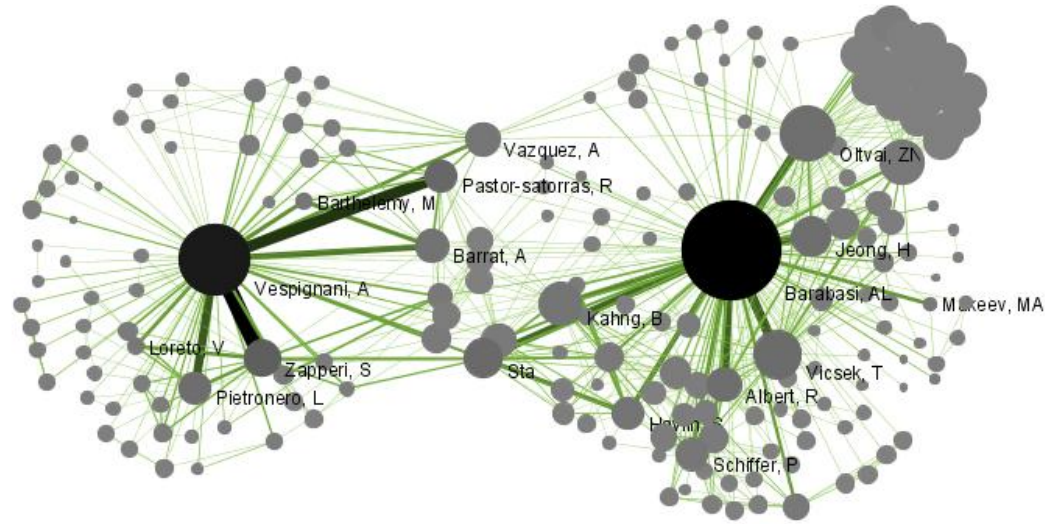
.....
 Weak Component Clustering was selected.
 Implementer(s): Russell Duhon
 Integrator(s): Russell Duhon

Input Parameters:
 Number of top clusters: 10
 3 clusters found, generating graphs for the top 3 clusters.

.....



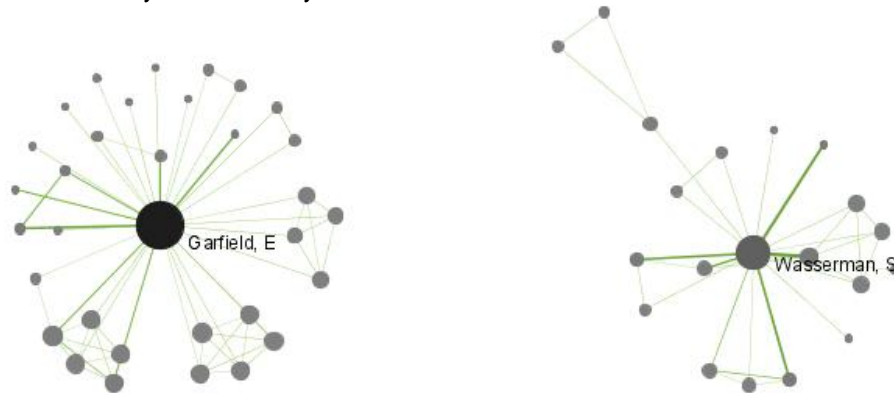
Network Visualization: Color/Size Coding by Degree

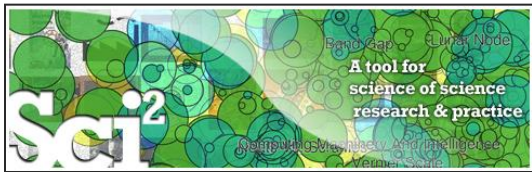


.....
Node Degree was selected.

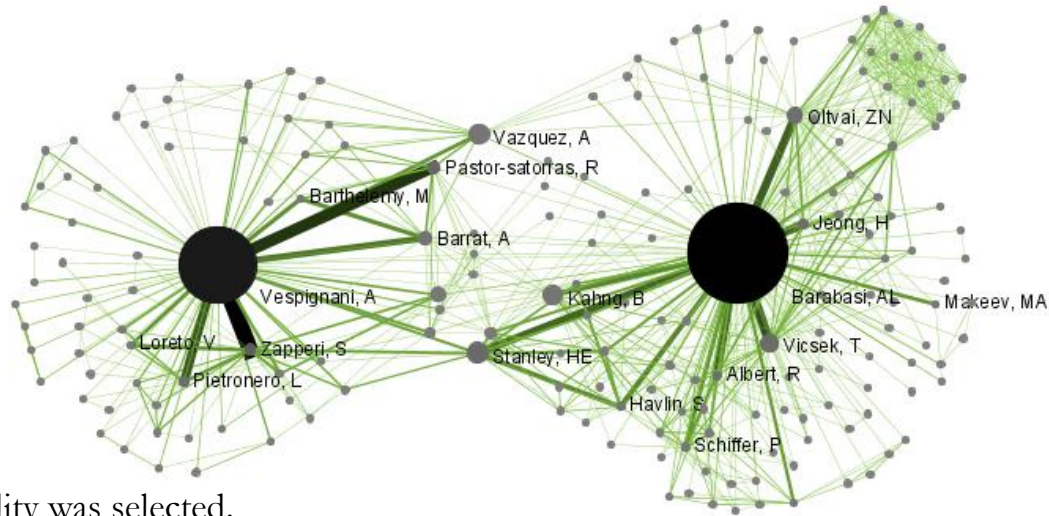
Documentation:
<https://nwb.slis.indiana.edu/community/?n=AnalyzeData.No deDegree>

.....



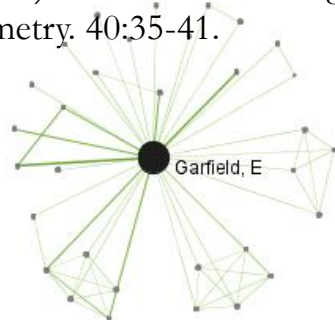


Network Visualization: Color/Size Coding by Betweenness Centrality



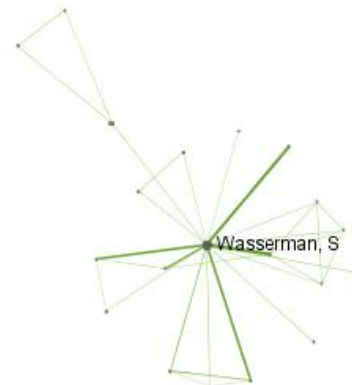
.....
 Node Betweenness Centrality was selected.
 Author(s): L. C. Freeman
 Implementer(s): Santo Fortunato
 Integrator(s): Santo Fortunato, Weixia Huang
 Reference: Freeman, L. C. (1977). A set of measuring centrality based on betweenness. *Sociometry*. 40:35-41.

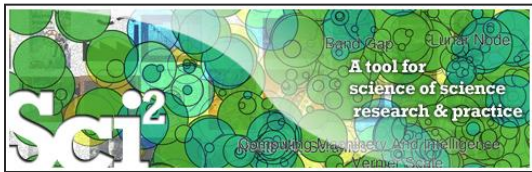
Input Parameters:
 Number of bins: 10



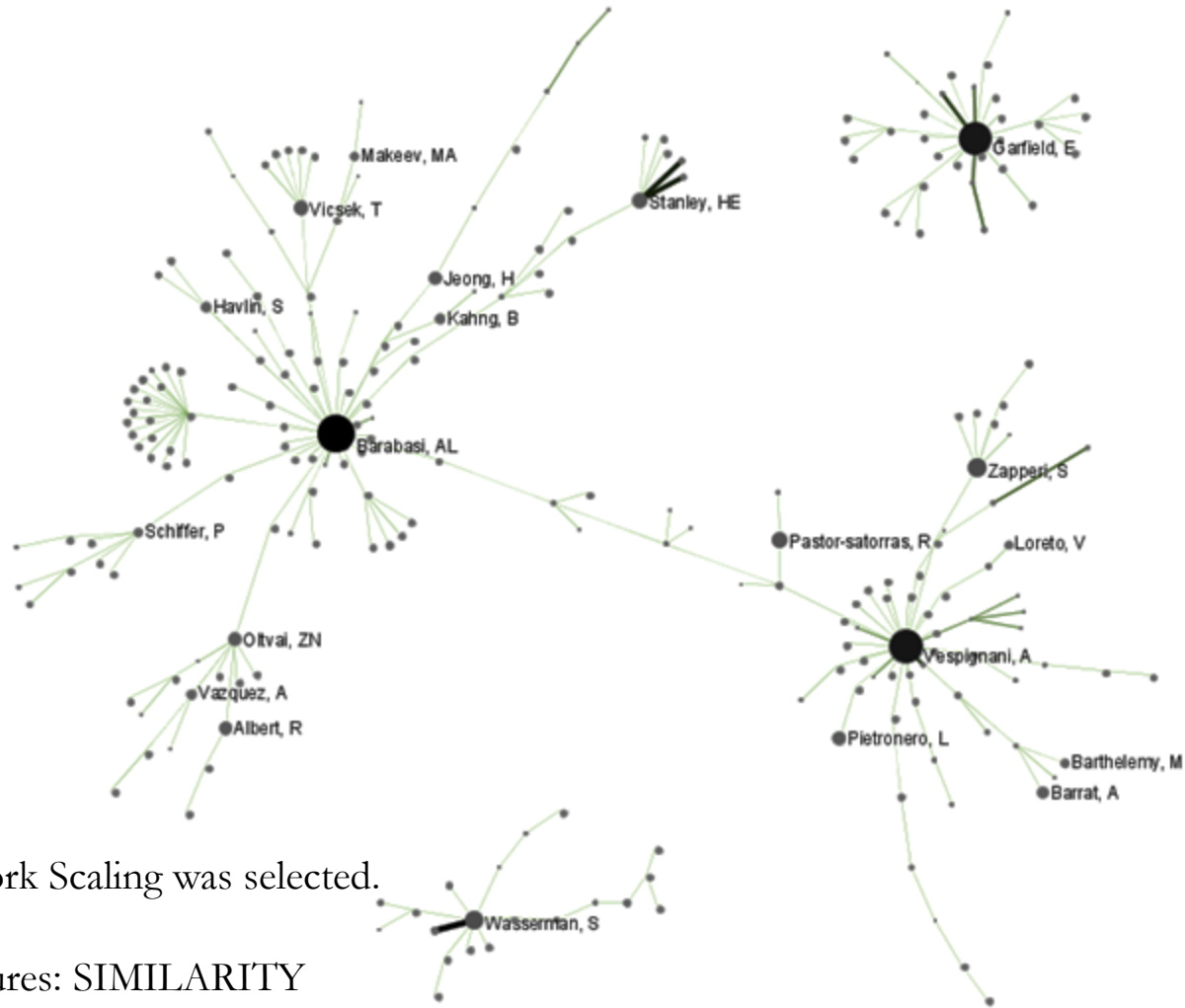
umber of bins: 10

.....





Network Visualization: Reduced Network After Pathfinder Network Scaling



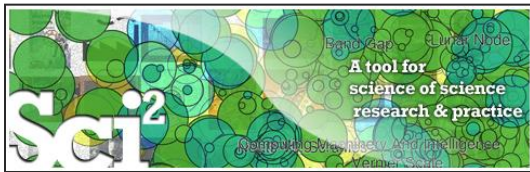
.....
MST-Pathfinder Network Scaling was selected.

Input Parameters:

Weight Attribute measures: SIMILARITY

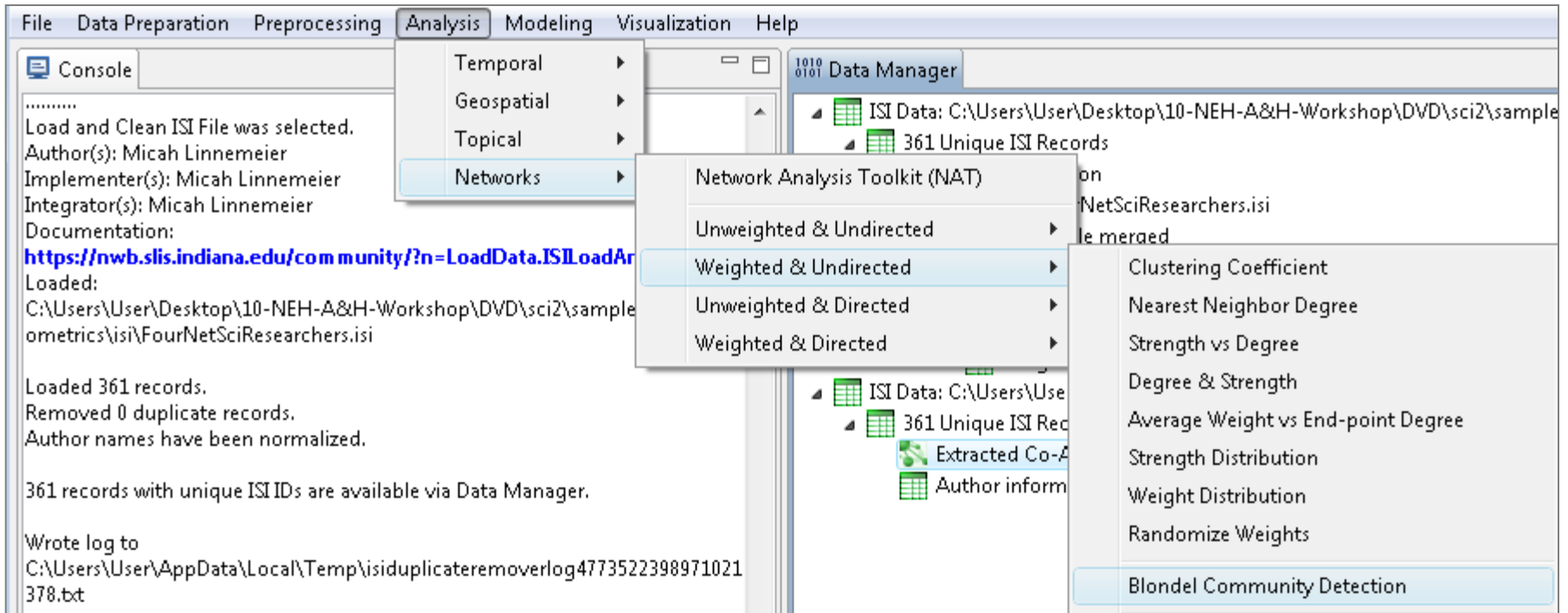
Edge Weight Attribute: weight

.....

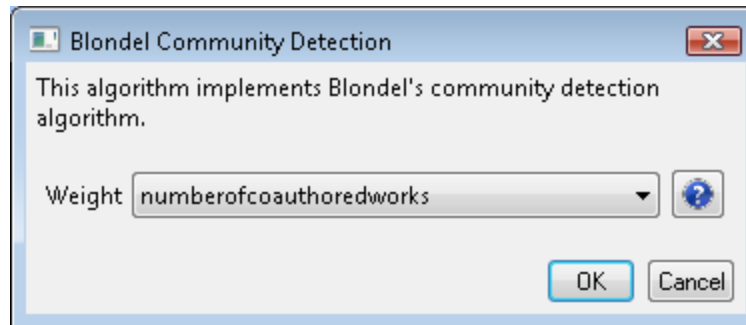


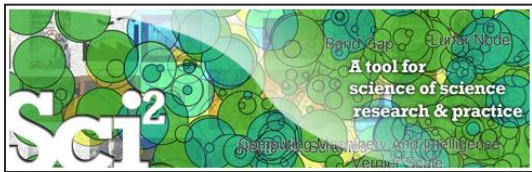
Network Visualization: Circular Hierarchy Visualization

Select Co-Author Network and run Blondel Community detection:



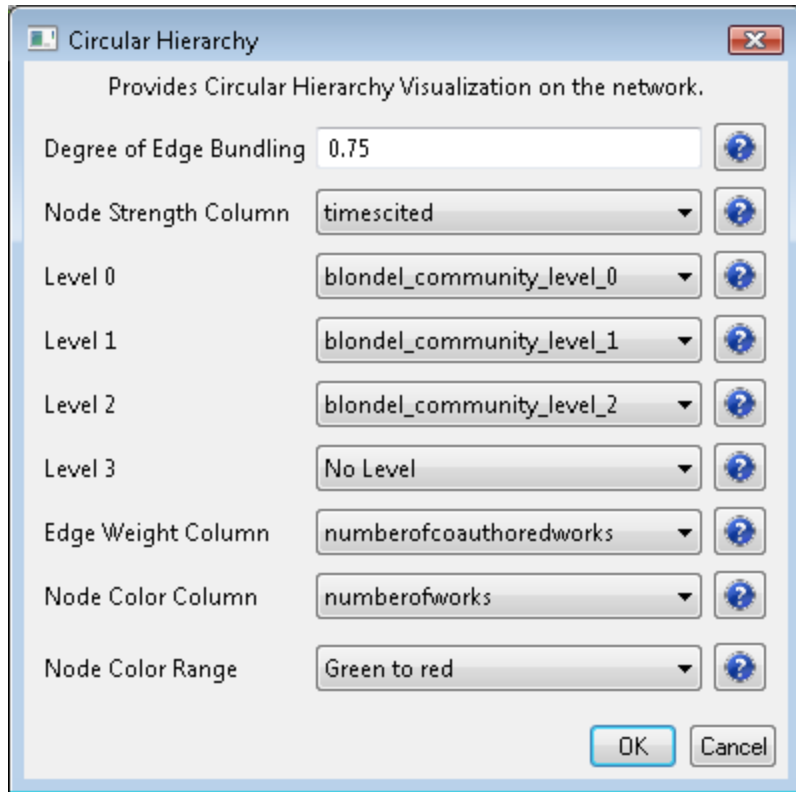
With parameter values

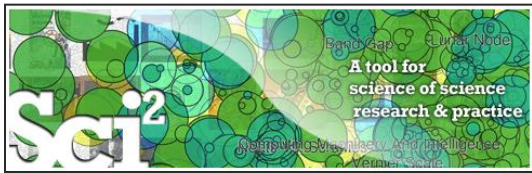




Network Visualization: Circular Hierarchy Visualization

Visualize resulting file using *Visualization > Networks > Circular Hierarchy* with parameter values





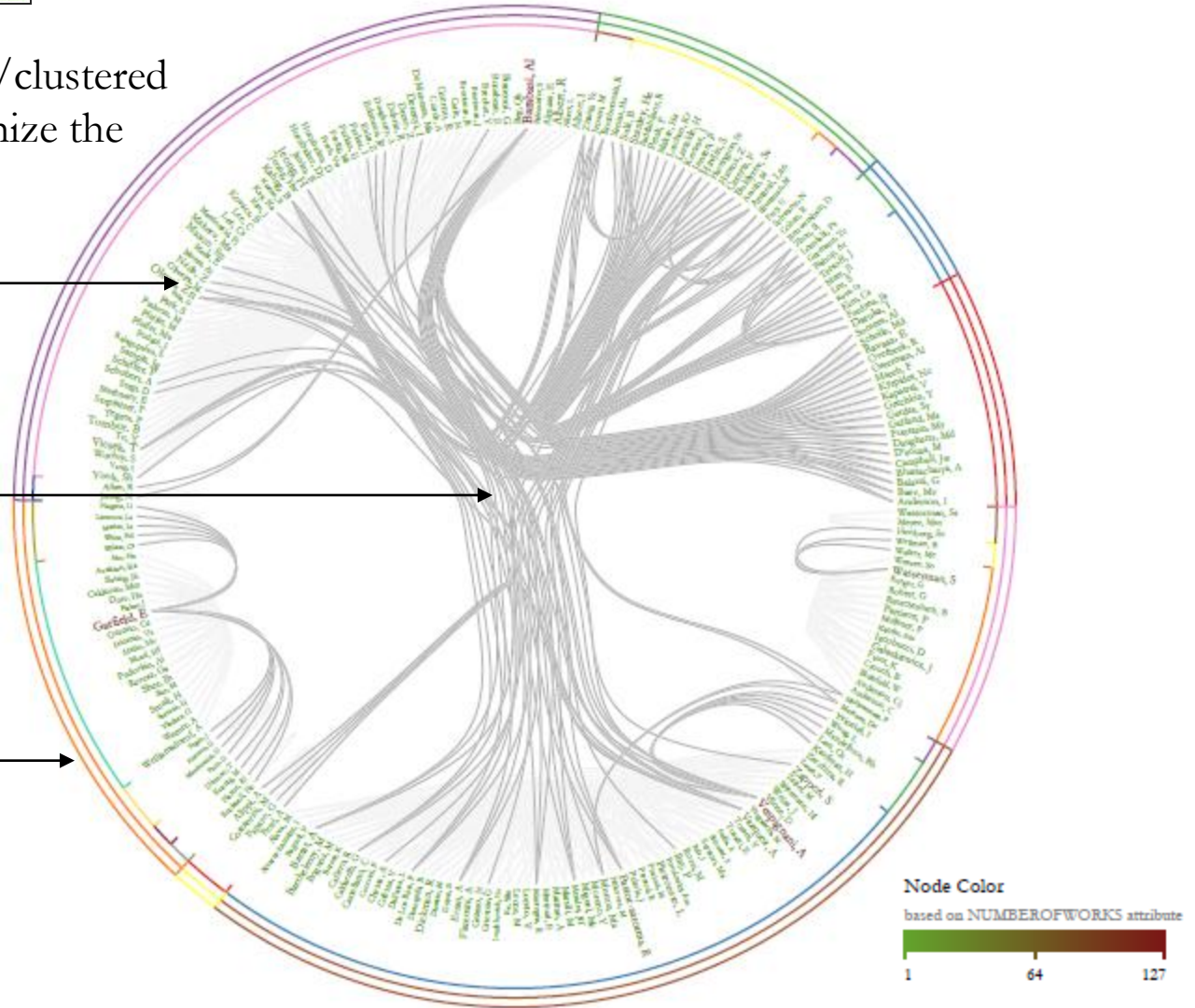
Network Visualization: Circular Hierarchy Visualization

Nodes that are interlinked/clustered are spatially close to minimize the number of edge crossings.

Node labels, e.g., author names.

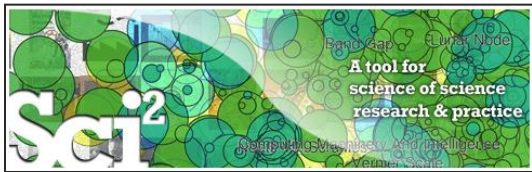
Network structure using edge bundling.

Color coded cluster hierarchy according to Blondel community detection algorithm.



Note:

Header/footer info, legend, and more meaningful color coding are under development.



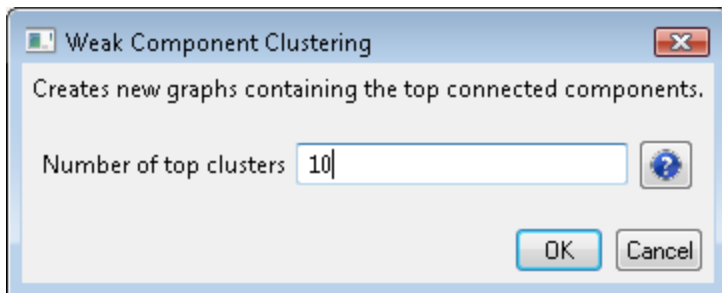
Paper-Citation Network Layout

To extract the paper-citation network, select the *'361 Unique ISI Records'* table and run *'Data Preparation > Text Files > Extract Paper Citation Network.'*

The result is a unweighted, directed network of papers linked by citations, named *Extracted paper-citation network* in the Data Manager.

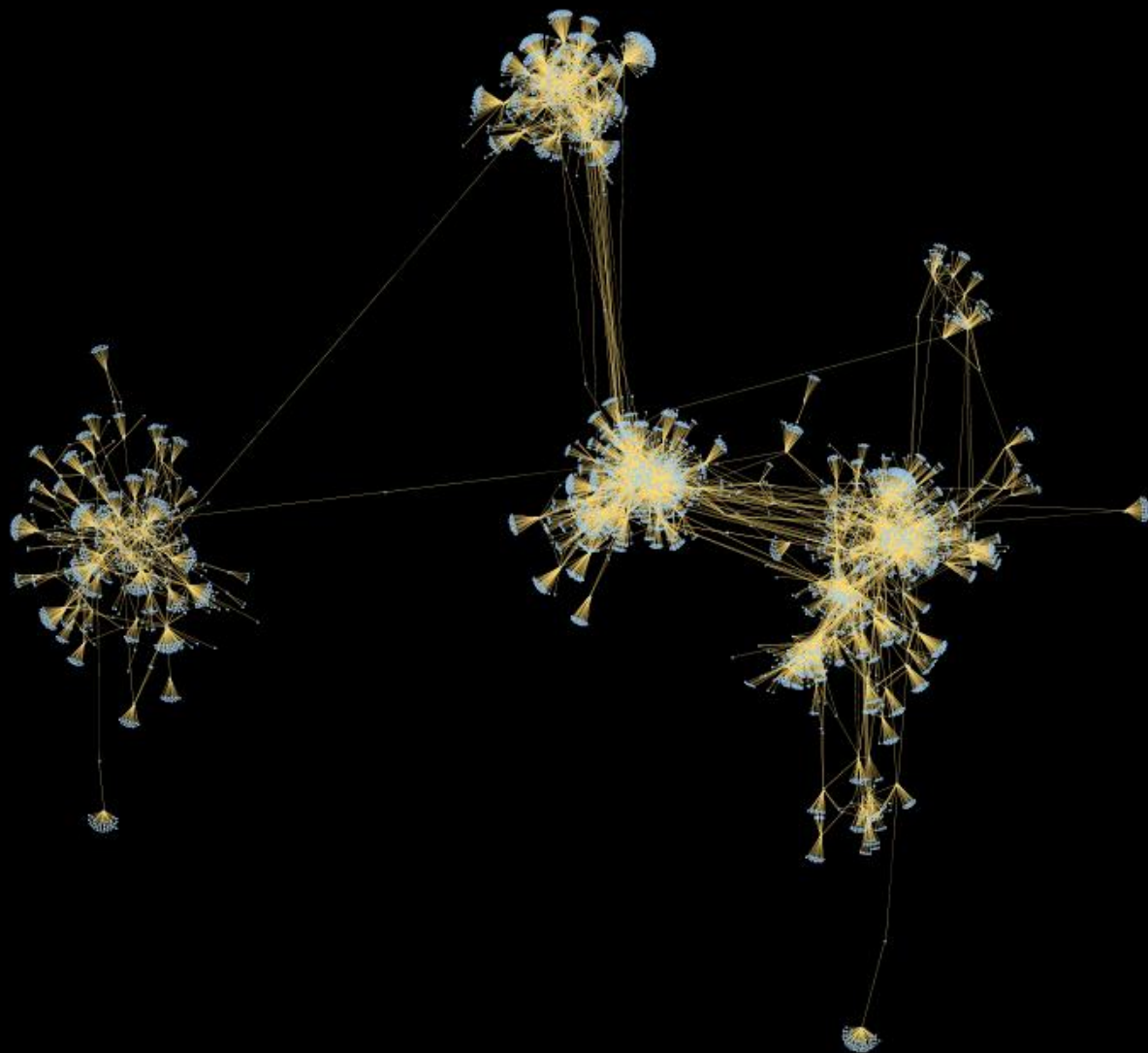
Run *NAT* to calculate that the network has 5,342 nodes and 9,612 edges. There are 15 weakly connected components. (0 isolates)

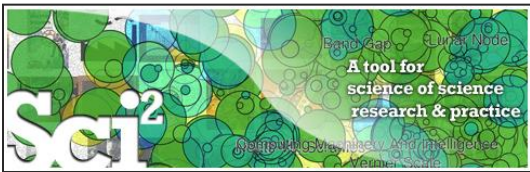
Run *'Analysis > Networks > Unweighted and Directed > Weak Component Clustering'* with parameters



to identify top-10 largest components. The largest (giant) component has 5,151 nodes.

To view the complete network, select the network and run *'Visualization > GUESS'*.





Studying Four Major NetSci Researchers (ISI Data)

Burst Analysis for Abstracts

Run *'Preprocessing > Topical > Lowercase, Tokenize, Stem, and Stopword Text'* with the *'Abstract'* box checked followed by *'Analysis > Topical > Burst Detection'* with parameters on the left and then run *'Visualize > Temporal > Horizontal Line Graph'* with parameters on right.

Perform Burst Detection on time-series textual data.

Gamma	<input type="text" value="1.0"/>	
Density Scaling	<input type="text" value="2.0"/>	
Bursting States	<input type="text" value="1"/>	
Date Column	<input type="text" value="Publication Year"/>	
Date Format	<input type="text" value="yyyy"/>	
Burst Length Unit	<input type="text" value="Years"/>	
Burst Length	<input type="text" value="1"/>	
Text Column	<input type="text" value="Abstract"/>	
Text Separator	<input type="text" value=" "/>	
Document Column	<input type="text" value="Cite Me As"/>	

Ignore Input With Empty Text

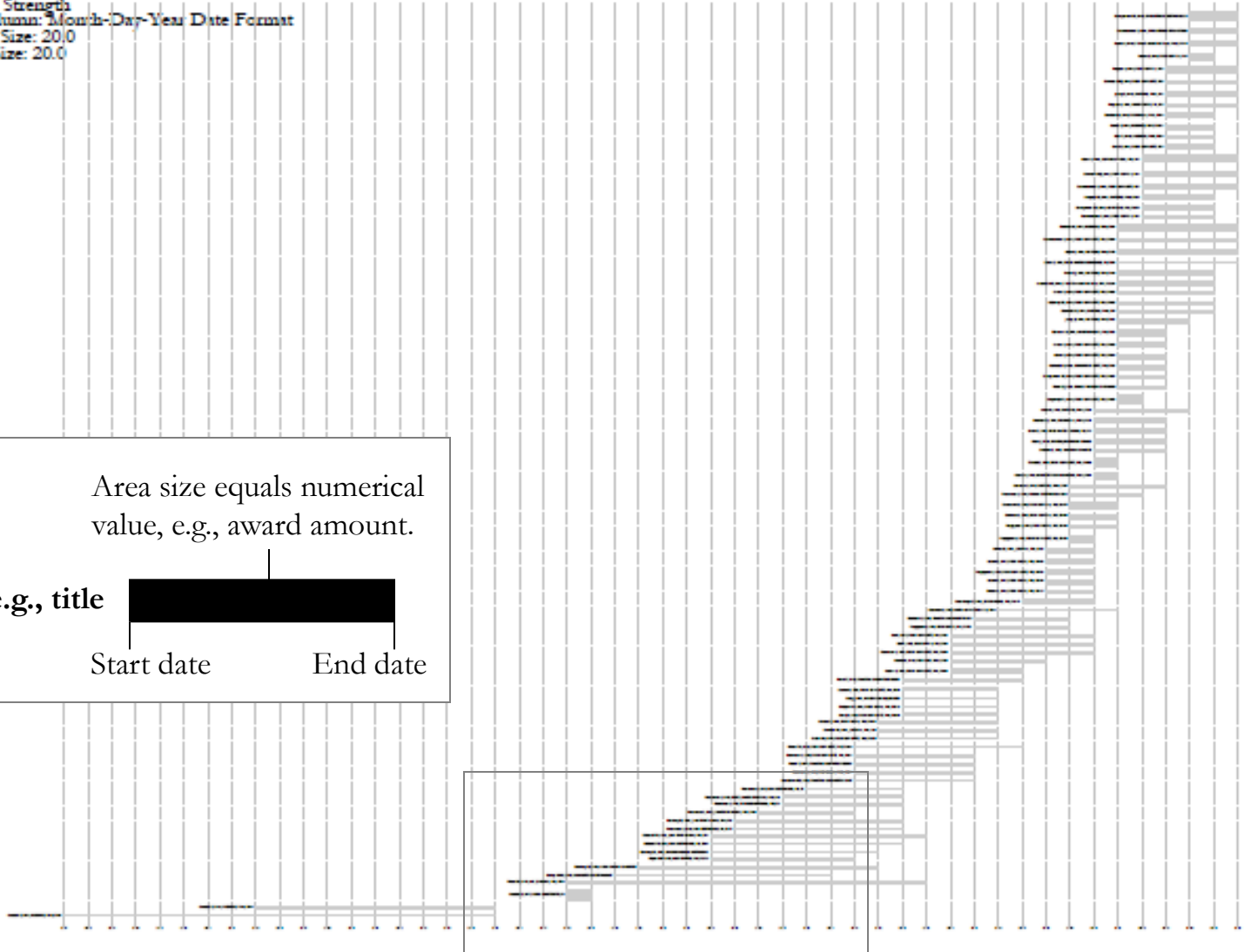
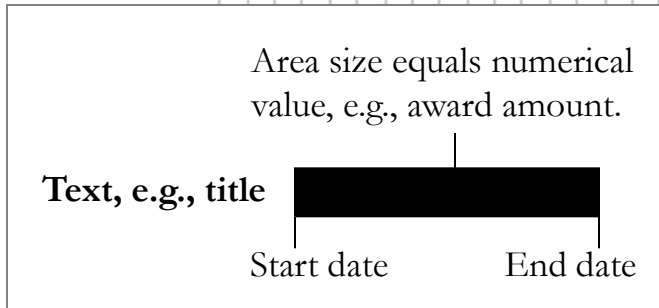
Horizontal Bar Graph

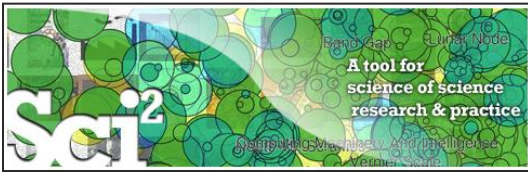
Takes tabular data and generates PostScript for a horizontal bar graph.

Label	<input type="text" value="Word"/>	
Start Date	<input type="text" value="Start"/>	
End Date	<input type="text" value="End"/>	
Size By	<input type="text" value="Strength"/>	
Date Format	<input type="text" value="Month-Day-Year Date Format (U.S., e.g. 10/31/2010)"/>	
Year Label Font Size	<input type="text" value="20.0"/>	
Bar Label Font Size	<input type="text" value="20.0"/>	

Horizontal Bar Graph for maximum burst level 1

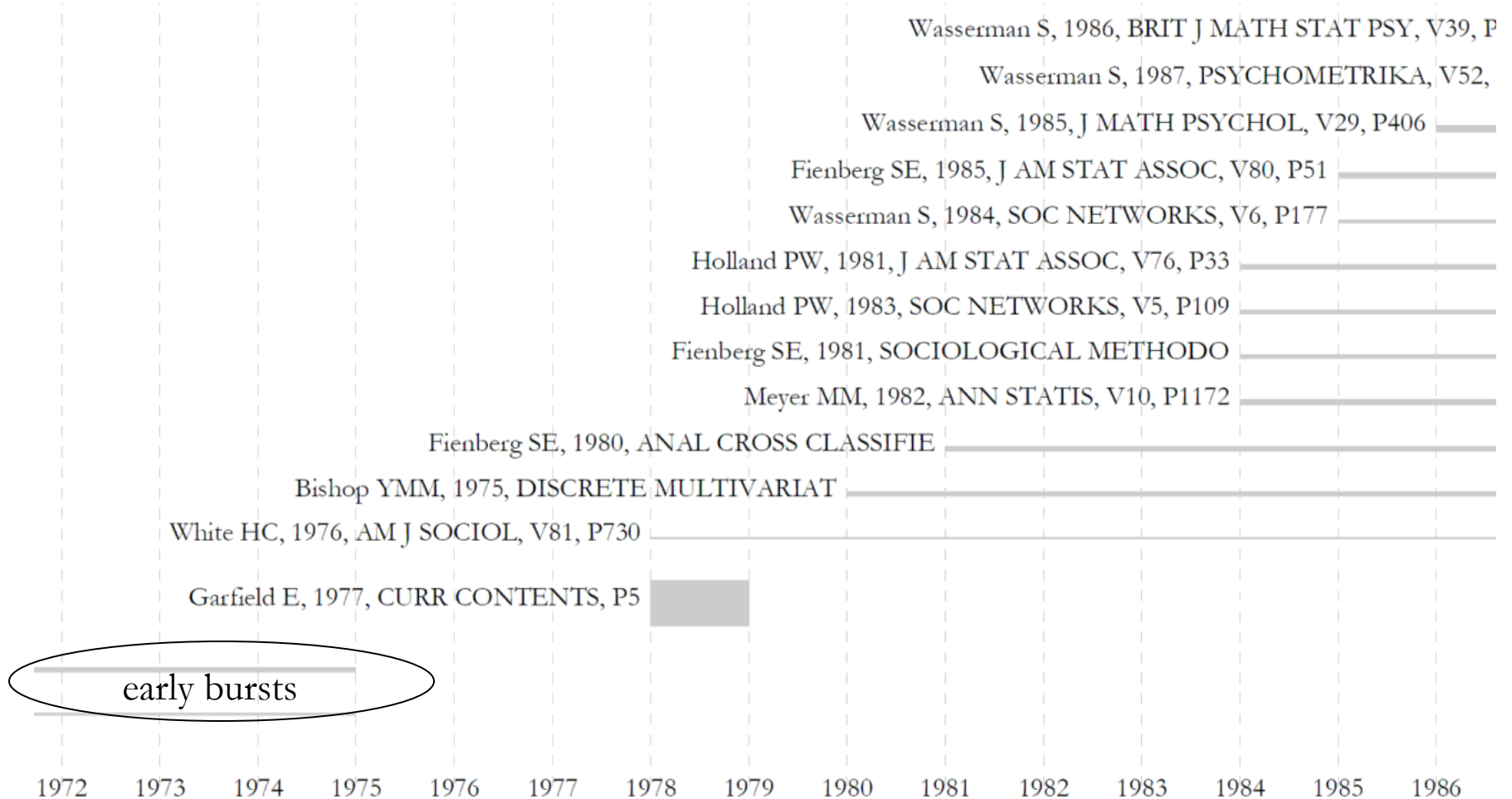
Label Column: Word
Start Date Column: Start
End Date Column: End
Size By Column: Strength
Date Format Column: Month-Day-Year Date Format
Year Label Font Size: 20.0
Bar Label Font Size: 20.0





Studying Four Major NetSci Researchers (ISI Data)

Burst Analysis Result





Workshop Overview

1:00-1:15 Introduction to Network Analysis

1:15-1:45 Network Analysis & Visualization in the Humanities

- Theory, Applications, and Pitfalls.
- Examples In The Wild

1:45-2:15 Collecting, Cleaning & Formatting Data

2:15-2:25 Break

2:25-3:00 Sci2 Tool Basics

- Macroscope Design and Usage.
- Download and run the tool.
- Find basic statistics and run various algorithms over the network.
- Visualize the networks as either a graph or a circular hierarchy.

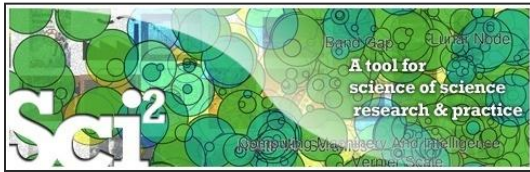
3:00-3:20 Sci2 Workflow Design: Padgett's Florentine Families - Prepare, load, analyze, and visualize family and business networks from 15th century Florence.

3:20-3:35 Break

3:35-4:00 Sci2 Research Demonstration: Mapping the Republic of Letters

4:00-4:30 Q&A and Technical Assistance

**Workflow Design:
Padgett's Florentine Families**



Padgett's Florentine Families - Compute Basic Network Properties & View in GUESS

- Load **yoursci2directory*/sampledata/socialscience/florentine.nwb*
- Run 'Analysis > Network Analysis Toolkit (NAT)' to get basic properties.

This graph claims to be undirected.

Nodes: 16

Isolated nodes: 1

Node attributes present: label, wealth, totalities, priorates

Edges: 27

No self loops were discovered.

No parallel edges were discovered.

Edge attributes:

Nonnumeric attributes:

Example value

marriag...T

busines...F

Average degree: 3.375

There are 2 weakly connected components. (1 isolates)

The largest connected component consists of 15 nodes.

Did not calculate strong connectedness because this graph was not directed.

Density (disregarding weights): 0.225

- **Optional:** Run 'Analysis > Unweighted & Undirected > Node Betweenness Centrality' with default parameters.
- Select network and run 'Visualization > GUESS' to open GUESS with file loaded.
- Apply 'Layout > GEM'.
- Open NWB File

```

.....
GUESS was selected.
Author(s): Eytan Adar
Implementer(s): Eytan Adar (GUESS), Russell Duhon (resizeLinear, colorize fix)
Integrator(s): Russell Duhon
Reference: Adar, Eytan, "GUESS: A Language and Interface for Graph Exploration," CHI 2006 (http://graphexploration.cond.org/)
Documentation: https://nwb.slis.indiana.edu/community/?n=VisualizeData.GUESS
ECHO is off.
Starting GUESS...
ECHO is off.
The initial layout for your visualization is random. For a clearer visualization, please run a layout from the Layout menu. (We recommend GEM.)
ECHO is off.
GUESS log file for this session can be found in
C:\DOCUME~1\katy\LOCALS~1\Temp\CIShell-Session-55892\StaticExecutableRunner-55904\ALGORITHM\guesslog.txt
.....
GUESS was selected.
Author(s): Eytan Adar
Implementer(s): Eytan Adar
Integrator(s): Russell Duhon
Reference: Adar, Eytan, "
Documentation: https://nwb.slis.indiana.edu/community/?n=VisualizeData.GUESS
ECHO is off.
Starting GUESS...
ECHO is off.
The initial layout for your visualization is random. For a clearer visualization, please run a layout from the Layout menu. (We recommend GEM.)
ECHO is off.
GUESS log file for this session can be found in
C:\DOCUME~1\katy\LOCALS~1\Temp\CIShell-Session-55892\StaticExecutableRunner-55904\ALGORITHM\guesslog.txt

```

- NWB file: C:\Documents and Settings\katy\...
- Distribution of degree for network at stu...
- List of edges of network created through th...
- Distribution of degree for network at stu...
- Distribution of degree for network at stu...
- NWB file: C:\Documents and Settings\katy\...
- Graph and Network Analysis Log
- Sequence of site betweennesses for no...
- Distribution of site betweenness for net...
- Distribution of site betweenness for net...
- NWB file with site betweenness attribute

Visualization - GUESS

File Edit Layout Script View Help

Medici-Acciaiuoli

Field	Value
__edgeid	0
business	F
color	dandelion
directed	false
label	
labelcolor	0,0,0,255
labelsize	12
labelvisible	false
marriage	T
node1	n9
node2	n1
visible	true
weight	1.0
width	2.0

Information Window

Object: [] Property: labelsize Operator: == Value: []

Colour Show Hide Size Show Label Hide Label Change Label

Node Shape Center Change History

Interpreter Graph Modifier

- Scheduler
- Remove From List
- GUESS
 - GUESS
 - Node Betw
 - Network Ar
 - Load...
 - GnuPlot
 - Degree Dis
 - Barabási-A
 - GnuPlot
 - Degree Dis
 - GUESS
 - Load...

Visualization - GUESS

File Edit Layout Script View Help

Information Window

Field	Value
color	cornflo...
fixed	false
height	10.0
image	
label	Medici
labelcolor	0,0,0,...
labelsize	12
labelvi...	false
name	n9
original...	Medici
priorates	53
stroke...	cadetb...
style	2
totalities	54
visible	true
wealth	103
width	10.0
x	90.625...
y	44.312...

Object: Property: Operator: Value:

Colour Show Hide Size Show Label Hide Label Change Label

Format Node Labels Format Edge Labels

Node Shape Center Change History

Resize Linear Colorize

Nodes From: To: Do Resize Linear


Interpreter Graph Modifier

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 to modify Color, etc.

Visualization - GUESS

File Edit Layout Script View Help

Pucci

Field	Value
color	255,25...
fixed	false
height	5.0
image	
label	Pucci
labelcolor	0,0,0,...
labelsize	12
labelvi...	true
name	n12
original...	Pucci
priorates	0
stroke...	cadetb...
style	2
totalities	1
visible	true
wealth	3
width	5.0
x	15.0
y	91.100...

Object: Property: Operator: Value:

Colour Show Hide Size Show Label Hide Label Change Label

Format Node Labels Format Edge Labels

Node Shape Center Change History

Resize Linear Colorize

Interpreter Graph Modifier

Graph Modifier:

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

Select ‘Resize Linear > Nodes > totalities’ 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.’

Visualization - GUESS

File Edit Layout Script View Help

Information Window

Field	Value
color	255,24...
fixed	false
height	5.6635...
image	
label	Acciaiu...
labelcolor	0,0,0,...
labelsize	12
labelvi...	true
name	n1
original...	Acciaiuoli
priorates	53
stroke...	cadetb...
style	2
totalities	2
visible	true
wealth	10
width	5.6635...
x	112.01...
y	70.315...

```

>>> resizeLinear(totalities,5,20)
>>> colorize(wealth,white,red)
>>>

```

Interpreter Graph Modifier

Interpreter:

Uses Jython a combination of Java and Python.

Try

`colorize(wealth, white, red)`

`resizeLinear(sitebetweenness, 5, 25)`



Workshop Overview

1:00-1:15 Introduction to Network Analysis

1:15-1:45 Network Analysis & Visualization in the Humanities

- Theory, Applications, and Pitfalls.
- Examples In The Wild

1:45-2:15 Collecting, Cleaning & Formatting Data

2:15-2:25 Break

2:25-3:00 Sci2 Tool Basics

- Macroscope Design and Usage.
- Download and run the tool.
- Find basic statistics and run various algorithms over the network.
- Visualize the networks as either a graph or a circular hierarchy.

3:00-3:20 Sci2 Workflow Design: Padgett's Florentine Families - Prepare, load, analyze, and visualize family and business networks from 15th century Florence.

3:20-3:35 Break

3:35-4:00 Sci2 Research Demonstration: Mapping the Republic of Letters

4:00-4:30 Q&A and Technical Assistance



Workshop Overview

1:00-1:15 Introduction to Network Analysis

1:15-1:45 Network Analysis & Visualization in the Humanities

- Theory, Applications, and Pitfalls.
- Examples In The Wild

1:45-2:15 Collecting, Cleaning & Formatting Data

2:15-2:25 Break

2:25-3:00 Sci2 Tool Basics

- Macroscope Design and Usage.
- Download and run the tool.
- Find basic statistics and run various algorithms over the network.
- Visualize the networks as either a graph or a circular hierarchy.

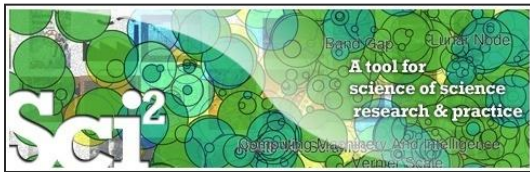
3:00-3:20 Sci2 Workflow Design: Padgett's Florentine Families - Prepare, load, analyze, and visualize family and business networks from 15th century Florence.

3:20-3:35 Break

3:35-4:00 Sci2 Research Demonstration: Mapping the Republic of Letters

4:00-4:30 Q&A and Technical Assistance

Sci2 Research Demonstration: Mapping the Republic of Letters



Mapping the Republic of Letters

- Load *Sample Data/sampleLettersNetwork.nwb*
- Run ‘*Analysis > Network Analysis Toolkit (NAT)*’ to get basic properties.

This graph claims to be directed.

Nodes: 9

Isolated nodes: 0

Node attributes present: label, totaldegree

Edges: 15

No self loops were discovered.

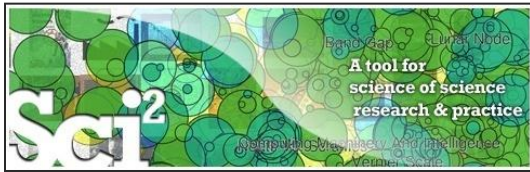
No parallel edges were discovered.

Edge attributes:

Did not detect any nonnumeric attributes.

...

- Select *Analysis > Networks > Unweighted & Directed > Node Betweenness Centrality* with ‘weight’ for Weight Attribute
- Select *Analysis > Networks > Unweighted & Directed > Adjacency Transitivity*
- Select *Analysis > Networks > Unweighted & Directed > Dyad Reciprocity*
- Select network and run ‘*Visualization > GUESS*’ to open GUESS with file loaded.
- Apply ‘*Layout > GEM*’.
- Export / Import Node Positions – notice that full network is needed before doing this



Mapping the Republic of Letters

- Load *Sample Data/CEN1640.nwb*, *Sample Data/CEN1641.nwb*, and *Sample Data/CEN1642.nwb*
- Run '*Analysis > Network Analysis Toolkit (NAT)*' to get basic properties.

Nodes: 868

Isolated nodes: 0

Edges: 898

No self loops were discovered.

No parallel edges were discovered.

Average total degree: 2.0691

Average in degree: 1.0346

Average out degree: 1.0346

This graph is not weakly connected.

There are 95 weakly connected components. (0 isolates)

The largest connected component consists of 607 nodes.

This graph is not strongly connected.

...

- Select network and run '*Visualization > GUESS*' to open GUESS with file loaded.
- Apply '*Layout > GEM*'.
- Export / Import Node Positions – notice that full network is needed before doing this



Workshop Overview

1:00-1:15 Introduction to Network Analysis

1:15-1:45 Network Analysis & Visualization in the Humanities

- Theory, Applications, and Pitfalls.
- Examples In The Wild

1:45-2:15 Collecting, Cleaning & Formatting Data

2:15-2:25 Break

2:25-3:00 Sci2 Tool Basics

- Macroscope Design and Usage.
- Download and run the tool.
- Find basic statistics and run various algorithms over the network.
- Visualize the networks as either a graph or a circular hierarchy.

3:00-3:20 Sci2 Workflow Design: Padgett's Florentine Families - Prepare, load, analyze, and visualize family and business networks from 15th century Florence.

3:20-3:35 Break

3:35-4:00 Sci2 Research Demonstration: Mapping the Republic of Letters

4:00-4:30 Q&A and Technical Assistance



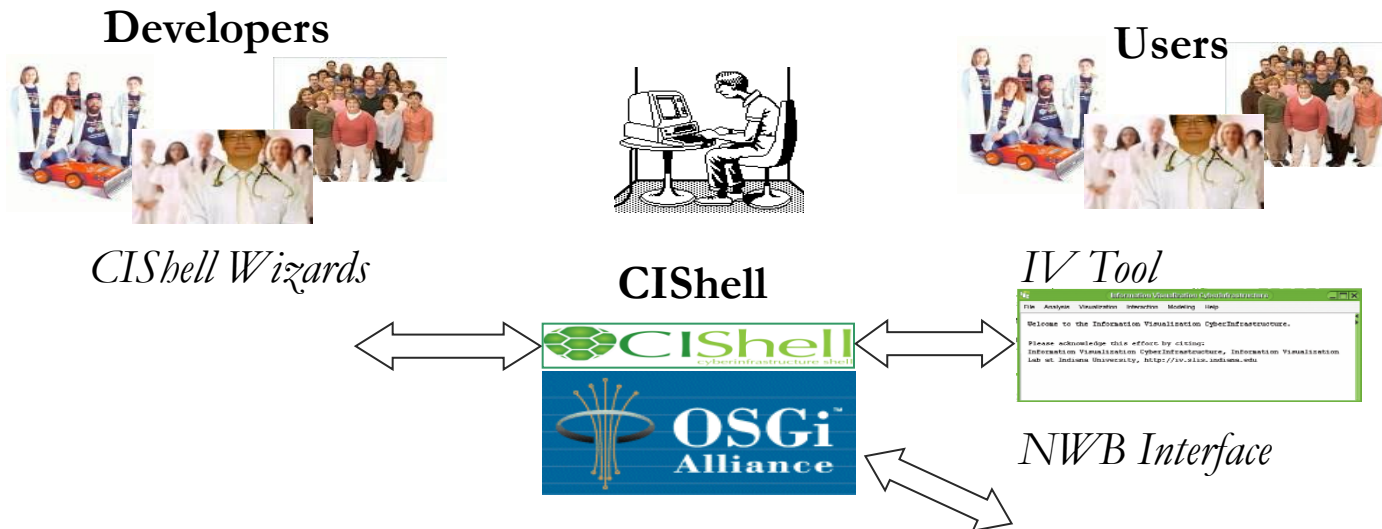
Possible Workflows

- **Geographic Visualizations**
- **Word Co-Occurrence Analysis**
- **Your Data**

Extraneous Slides

- Adding Plugins to CShell Powered Tools
- OSGi/CShell Adoption

- CIShell is an open source software specification for the integration and utilization of datasets, algorithms, and tools.
- It extends the Open Services Gateway Initiative (OSGi) (<http://www.osgi.org>), a standardized, component oriented, computing environment for networked services widely used in industry since 10 years.
- Specifically, CIShell provides “sockets” into which existing and new datasets, algorithms, and tools can be plugged using a wizard-driven process.



CIShell is built upon the Open Services Gateway Initiative (OSGi) Framework.

OSGi (<http://www.osgi.org>) is

- A standardized, component oriented, computing environment for networked services.
- Successfully used in the industry from high-end servers to embedded mobile devices since 8 years.
- Alliance members include IBM (Eclipse), Sun, Intel, Oracle, Motorola, NEC and many others.
- Widely adopted in open source realm, especially since Eclipse 3.0 that uses OSGi R4 for its plugin model.

Advantages of Using OSGi

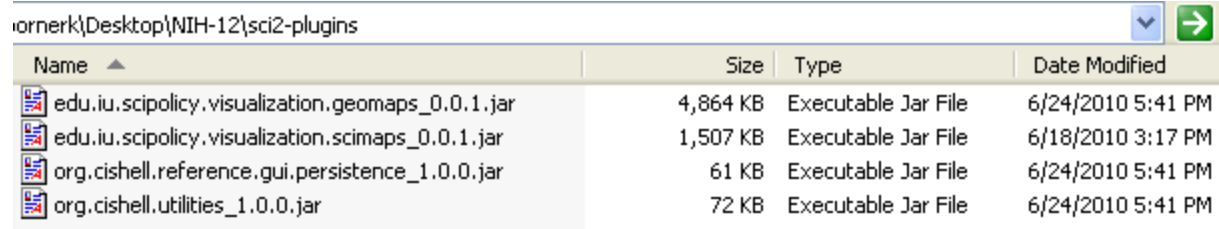
- Any CIShell algorithm is a service that can be used in any OSGi-framework based system.
- Using OSGi, running CIShells/tools can be connected via RPC/RMI supporting peer-to-peer sharing of data, algorithms, and computing power.

Ideally, CIShell becomes a standard for creating OSGi Services for algorithms.

- No central data format.
- Sci² Tool has 26 external and internal data formats and 35 converters.
- Their relationships can be derived by running ‘File > Converter Graph’ and plotted as shown here. Note that some conversions are symmetrical (double arrow) while others are one-directional (arrow).



- Not all code can be shared freely (yet).
- To make the UCSD Science Map and new geomaps available via the Sci² menu, simply add



Name	Size	Type	Date Modified
edu.iu.scipolicy.visualization.geomaps_0.0.1.jar	4,864 KB	Executable Jar File	6/24/2010 5:41 PM
edu.iu.scipolicy.visualization.scimaps_0.0.1.jar	1,507 KB	Executable Jar File	6/18/2010 3:17 PM
org.cishell.reference.gui.persistence_1.0.0.jar	61 KB	Executable Jar File	6/24/2010 5:41 PM
org.cishell.utilities_1.0.0.jar	72 KB	Executable Jar File	6/24/2010 5:41 PM

The files were made available in /sci2-plugins directory on the computers in the tutorial room.

to the ‘*yourdirectory/plugin*’ directory and restart the tool.

The rights to the UCSD map are owned by the Regents of UCSD. Usage does not require a separate, signed agreement or an additional request to our office if consistent with the permission. As a courtesy, please send information on how the map is being used to

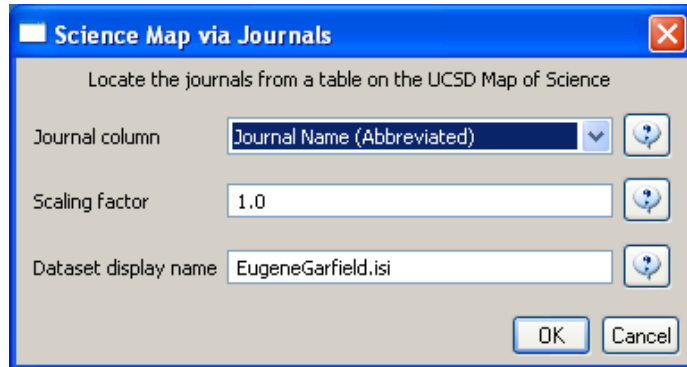
William J. Decker, Ph.D., Associate Director, Technology Transfer Office
 University of California, San Diego, 9500 Gilman Drive Dept. 0910, La Jolla, CA 92093
 phone:858-822-5128, fax: 858-534-7345, e-mail: wjdecker@ucsd.edu

- To delete algorithms that you do not use, simply delete the corresponding *.jar files in the plugin directory.
- Customize your menu structure accordingly—see next slide.

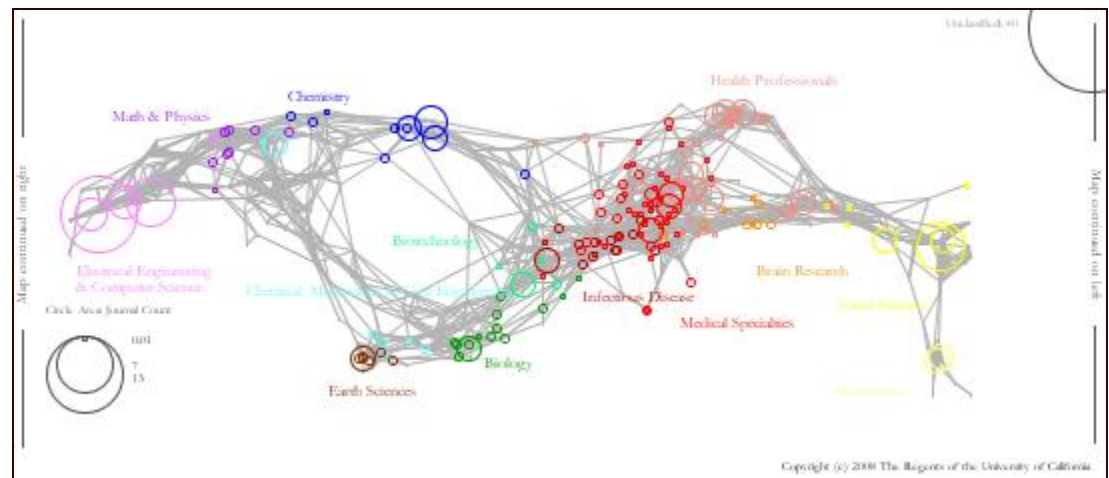
- After you added the new plugins, load an ISI file using ‘File > Load and Clean ISI File > EugeneGarfield.isi.’

The file can be found in the */sampledata/scientometrics/isi* directory.

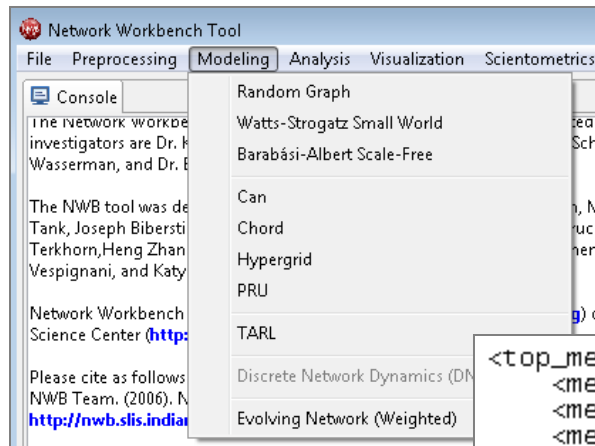
- Select ‘99 Unique ISI Records’ file in Data Manger and run ‘Visualization > Topical > Science Map via Journals’ with parameters:



- The result is a science map overlay of Garfield’s papers and a listing of journals in 13 fields of science below. See details in **Tutorial #6**.



- The file *'yourtooldirectory/configuration/default_menu.xml'* encodes the structure of the menu system.
- In NWB Tool, the Modeling menu (left) is encoded by the following piece of xml code:



```
<top_menu name="Modeling">
  <menu pid="edu.iu.nwb.modeling.erdosrandomgraph"/>
  <menu pid="edu.iu.nwb.modeling.smallworld"/>
  <menu pid="edu.iu.nwb.modeling.barabasiAlbert"/>
  <menu type="break"/>
  <menu pid="edu.iu.iv.modeling.p2p.can.CanAlgorithm"/>
  <menu pid="edu.iu.iv.modeling.p2p.chord.ChordAlgorithm"/>
  <menu pid="edu.id.iv.modeling.p2p.hypergrid.Hypergrid"/>
  <menu pid="edu.iu.iv.modeling.p2p.pru.PruAlgorithm"/>
  <menu type="break"/>
  <menu pid="edu.iu.iv.modeling.tarl.TarlAlgorithm"/>
  <menu type="break"/>
  <menu pid="edu.iu.nwb.modeling.discretenetworkdynamics.DNDAAlgorithm"/>
  <menu type="break"/>
  <menu pid="edu.iu.nwb.modeling.weighted.evolvingnetwork"/>
</top_menu>
```


Algorithm Developer's Guide

Overview

The Cyberinfrastructure Shell (CIShell) is an open source, community-driven platform for the integration and utilization of datasets, algorithms, tools, and computing resources. Algorithm integration support is built in for Java and most other programming languages. Being Java based, it will run on almost all platforms. The software and specification is released under an [Apache 2.0 License](#).

This guide attempts to aid algorithm developers in creating algorithms for CIShell (and applications built on CIShell).

This guide tries to contain all the information a new developer needs, but where necessary, it may cite the [CIShell 1.0 Specification \(API\)](#) or the [OSGi Service Platform Specification, Release 4 \(API\)](#). While the guide tries to make beginning algorithm development easier, the CIShell Specification has the last word on how the CIShell Platform works.

Table of Contents

1. [CIShell Basics](#)
2. Getting Started
 1. [Tutorial 0: Setting Up the Development Environment](#)
 2. [Tutorial 1: Creating a Hello World Java Algorithm](#)
 3. [Tutorial 2: Practical Java Algorithm Development](#)
 4. [Tutorial 3: Integrating a Non-Java Program As An Algorithm](#)
 5. [Mini-Tutorial: Integrating 3rd-party libraries](#)
 6. [Where to Learn More](#)
3. Reference
 1. [How Algorithms Work: A guide to algorithm plugins in CIShell](#)
 2. [Accessing the OSGi Console in CIShell tools](#)

<http://cishell.org/?n=DevGuide.NewGuide>



OSGi/CIShell Adoption

CIShell/OSGi is at the core of different CIs and a total of 169 unique plugins are used in the

- **Information Visualization** (<http://iv.slis.indiana.edu>),
- **Network Science (NWB Tool)** (<http://nwb.slis.indiana.edu>),
- **Scientometrics and Science Policy (Sci² Tool)** (<http://sci.slis.indiana.edu>), and
- **Epidemics** (<http://epic.slis.indiana.edu>) research communities.

Most interestingly, a number of other projects recently adopted OSGi and one adopted CIShell:

Cytoscape (<http://www.cytoscape.org>) lead by Trey Ideker, UCSD is an open source bioinformatics software platform for visualizing molecular interaction networks and integrating these interactions with gene expression profiles and other state data (Shannon et al., 2002).

Taverna Workbench (<http://taverna.sourceforge.net>) lead by Carol Goble, University of Manchester, UK is a free software tool for designing and executing workflows (Hull et al., 2006). Taverna allows users to integrate many different software tools, including over 30,000 web services.

MAEviz (<https://wiki.ncsa.uiuc.edu/display/MAE/Home>) managed by Shawn Hampton, NCSA is an open-source, extensible software platform which supports seismic risk assessment based on the Mid-America Earthquake (MAE) Center research.

TEXTrend (<http://www.textrend.org>) lead by George Kampis, Eötvös University, Hungary develops a framework for the easy and flexible integration, configuration, and extension of plugin-based components in support of natural language processing (NLP), classification/mining, and graph algorithms for the analysis of business and governmental text corpuses with an inherently temporal component.

As the functionality of OSGi-based software frameworks improves and the number and diversity of dataset and algorithm plugins increases, the capabilities of custom tools will expand.



The Changing Scientific Landscape

Star Scientist -> **Research Teams** might have 100 or more members & exist few months only.

Users -> **Contributors** students, faculty, practitioners.

Disciplinary -> **Cross-disciplinary** with different cultures, languages, approaches.

One Specimen -> **Data Streams** updated nightly or even more frequently

High Quality Open Data



Scholarly Database: 23 million scholarly records

<http://sdb.slis.indiana.edu>



VIVO National Researcher Networking

<http://vivoweb.org>

Static Instrument -> **Evolving Cyberinfrastructure (CI)** daily learning and documentation.

Macrosopes can make a major difference if they support:

Division of Labor – proper incentive structures are key.

Ease of Use – learn from YouTube, Flickr, Wikipedia

Modularity – plug-and-play helps reduce costs; increases flexibility, augmentation, customization

Standardization – speeds up ‘translation’ into products/practice.

Open Data and Open Code – use the minds of millions!



Epidemics Marketplace

The screenshot shows the EpiC Marketplace website. At the top left is the EpiC logo, a stylized globe with a network of nodes. To its right are navigation links: Browse, Upload, Request, and About. Further right is a 'My Account' link with a user icon. Below the navigation is a search bar with a 'Search' button. The main header area features a large pink banner with the text 'The EpiC Marketplace' and a sub-header 'A community to browse, request, and share epidemics data. Learn More >>'. Below this banner are three main action buttons: 'Browse' (with a magnifying glass icon), 'Upload' (with an upward arrow icon), and 'Request' (with a question mark icon). Each button has a brief description of its function. To the left of the main content is a 'Categories' sidebar with a list of categories: Demographics, Infectious Diseases, Social Contagion, Transportation, and Uncategorized. Below the categories is a word cloud containing various terms related to epidemiology and health, such as '1855', '1918', '2008', 'Adolescent', 'AIDS', 'airline', 'Behavior', 'bodily', 'Canada', 'cde', 'database', 'Demographic', 'Epidemiologic', 'family', 'flight', 'flu', 'fluids', 'Food and V...', 'geography', 'grad', 'health', 'Health law', 'HIV', 'Housing Vac...', 'income', 'influenza', 'insurance', 'location', 'Obesity', 'Pandemic', 'Physical Ac...', 'Policy/str...', 'regional', 'seasonal', 'Smoking', 'spanish', 'Third', 'Tobacco Use', 'travel', 'U.S.', 'United', 'Stat...', 'USA', 'world', 'yearly'. To the right of the main content is a 'Recent Activity' sidebar. It contains three sections: 'Data Requests' with two entries (one about health insurance coverage data and another about HIV/AIDS data), and 'Data Uploads' with two entries (one about housing vacancies and another about a behavioral risk factor surveillance system). At the bottom of the main content area is a 'Location of datasets' map showing a world map with several red location pins. The map includes navigation controls (directional arrows, zoom in/out) and a legend for 'Map', 'Satellite', and 'Hybrid' views. The map is powered by Google and includes a 'Terms of Use' link.

<http://dev.epic.slis.indiana.edu>



cyberinfrastructure for NETWORK SCIENCE CENTER

School of Library and Information Science | Indiana University Bloomington



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