

Network Analysis using the Network Workbench (NWB) Tool and the Science of Science (Sci2) Tool

Ted Polley and Dr. Katy Börner

CNS & IVL, SLIS, Indiana University, Bloomington, Indiana, USA

Royal Netherlands Academy of Arts and Sciences (KNAW), Amsterdam, The Netherlands

<http://cns.iu.edu>

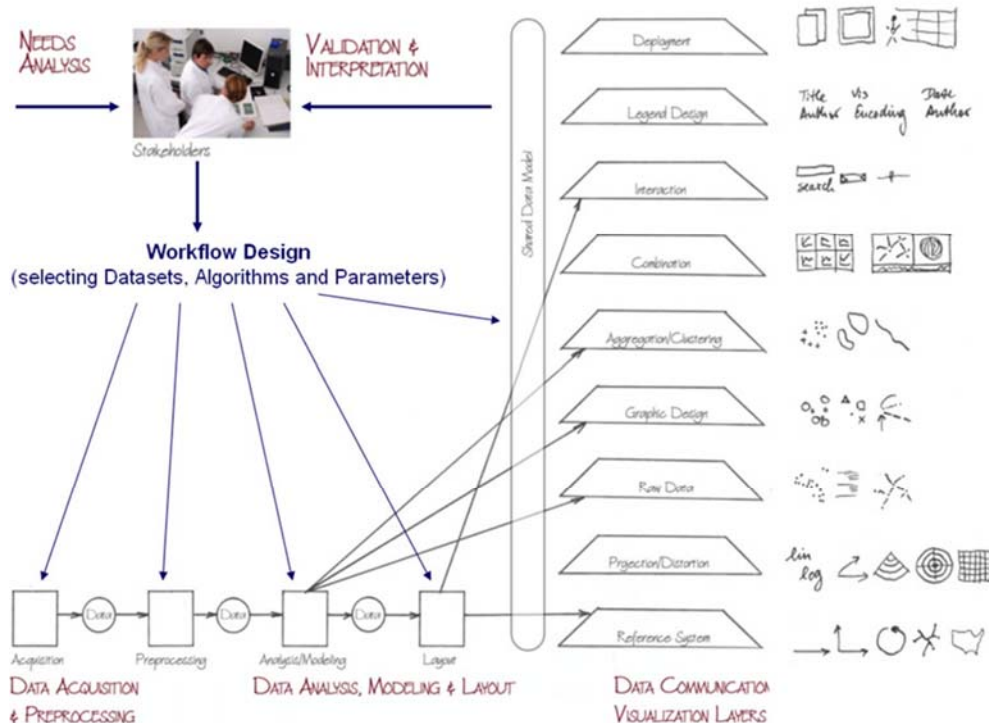
With special thanks to Kevin W. Boyack, Chin Hua Kong, Micah Linnemeier, Russell J. Duhon, Patrick Phillips, Chintan Tank, Thomas Smith, Nianli Ma, Joseph R. Biberstine, David Coe, Scott Weingart, Hanning Guo, Mark A. Price, Angela M. Zoss, Ted Polley, and Sean Lind.

Please download and install each tool prior to the start of the workshop:
 Network Workbench (NWB) Tool from <http://nwb.cns.iu.edu>
 Science of Science (Sci2) Tool from <http://sci2.cns.iu.edu>
 Open the Sci2 Manual <http://wiki.cns.iu.edu/display/SCI2TUTORIAL>

Cyberinfrastructure for Network Science Center
 School of Library and Information Science
 Indiana University Bloomington
 L1001 Wells Library
 Monday September 17, 2012 – 6:00pm-7:00pm



Overview – Workflow Design



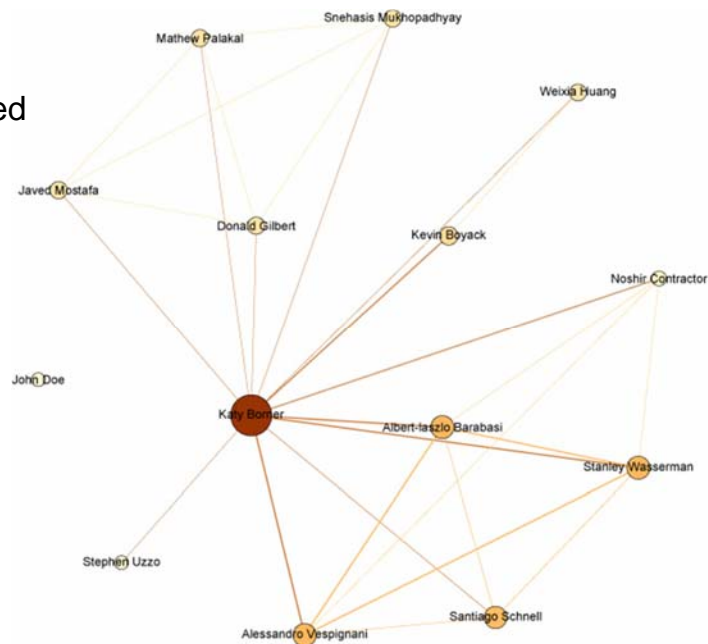
Overview of this workshop

- Brief introduction to Networks
 - Basic Graph Metrics
 - Questions
- Network Workbench (NWB) Tool
 - Introduction to NWB
 - Visualize the Florentine Dataset
 - Questions
- Science of Science (Sci2) Tool
 - Introduction to Sci2
 - Extract Co-Author Network from ISI data
 - Extract Paper-Paper (Citation) Network from ISI data
 - Questions
- Adjourn

3

What is a Network?

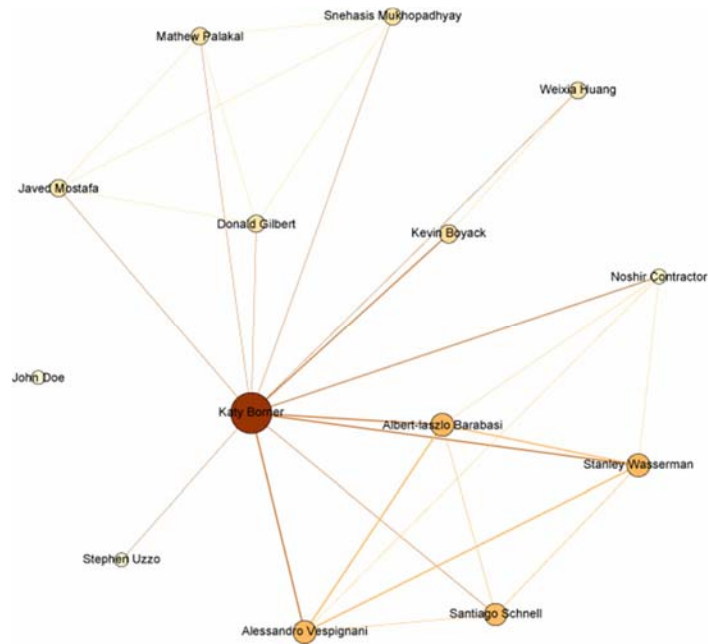
- Graph – network visualized
- Nodes (vertices)
- Edges



4

Graph Metrics - Nodes

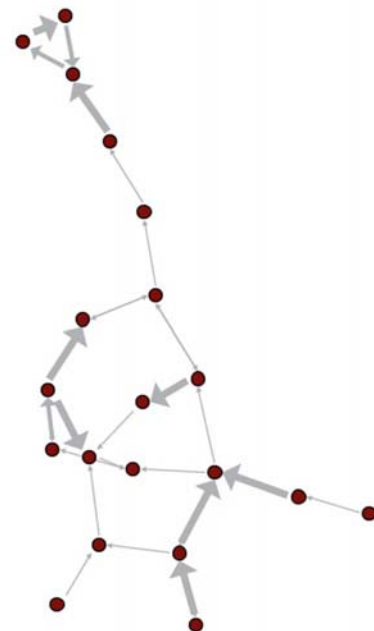
- Betweenness Centrality – number of shortest paths a node sits between
- Degree
- Isolates



5

Graph Metrics - Edges

- Shortest paths – shortest distance between two nodes
- Weight – strength of tie
- Directionality – is the connection one-way or two-way (in-degree vs. out-degree)?
- Bridge – deleting would change structure



6

Q & A

Do you have questions so far?

7



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.

Network Workbench Tool
<http://nwb.slis.indiana.edu>

Summary
Network Workbench: A Large-Scale Network Analysis, Modeling and Visualization Toolkit for Biomedical, Social Science and Physics Research. This project will design, evaluate, and operate a unique distributed, shared resources environment for large-scale network analysis, modeling, and visualization, named Network Workbench (NWB). The envisioned data-code-computing resources environment will provide ...
[more](#)
[How to cite this project](#)

News & Updates

- 5.1.09 Kaelble, Steve. 2009. [Mapping the Future of Knowledge](#). *Research & Creative Activity*, 31, 2: 12-15. [Website](#) accessed 5/1/09
- 3.23.09 [1.0.0 beta 5](#) Released
- 1.23.09 Ann Mcranie's [tutorial abstract](#) for Sunbelt 2009
- 11.4.08 Two NWB PIs featured in "[Connected—The Power of Six Degrees](#)." 2008. Anna Maria Talas, Director. Australian Broadcasting Corporation, Ltd. [YouTube](#) [Full Video](#) (300MB)

Download 1.0.0 beta 5 Release
Note: save the download as jar

Select Your Operating System
Windows (XP & Vista) **DOWNLOAD**

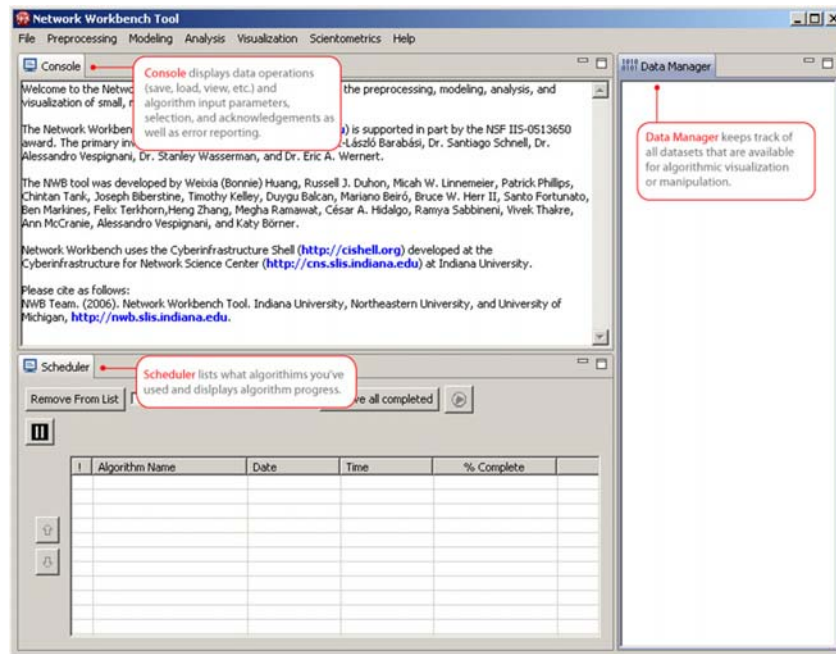
[Getting Started](#)
See more [documentation](#)

Get Involved

Börner, Katy, Huang, Weixia (Bonnie), Linnemeier, Micah, Duhon, Russell Jackson, Phillips, Patrick, Ma, Nianli, Zoss, Angela, Guo, Hanning & Price, Mark. (2010). *Refe-Netzwerk-Red: Analyzing and Visualizing Scholarly Networks Using the Network Workbench Tool*. *Scientometrics*. Vol. 83(3), 863-876.

8

User Interface



9

Visualizing the Florentine dataset

- Data Acquisition & Preprocessing
 - Examine the data
 - Load the data into NWB

- Data Analysis, Modeling, & Layout
 - Since the data is formatted as a network with various attributes added to the nodes and edges – no analysis will be performed in this workflow

- Data Communication & Visualization Layers
 - Visualize the Florentine network with GUESS

10

Florentine families related through business ties (specifically, recorded financial ties such as loans, credits and joint partnerships) and marriage alliances.

Node attributes

- Wealth: Each family's net wealth in 1427 (in thousands of lira).
- Priorates: The number of seats on the civic council held between 1282-1344.
- Totalities: Number of business/marriage ties in complete dataset of 116 families.

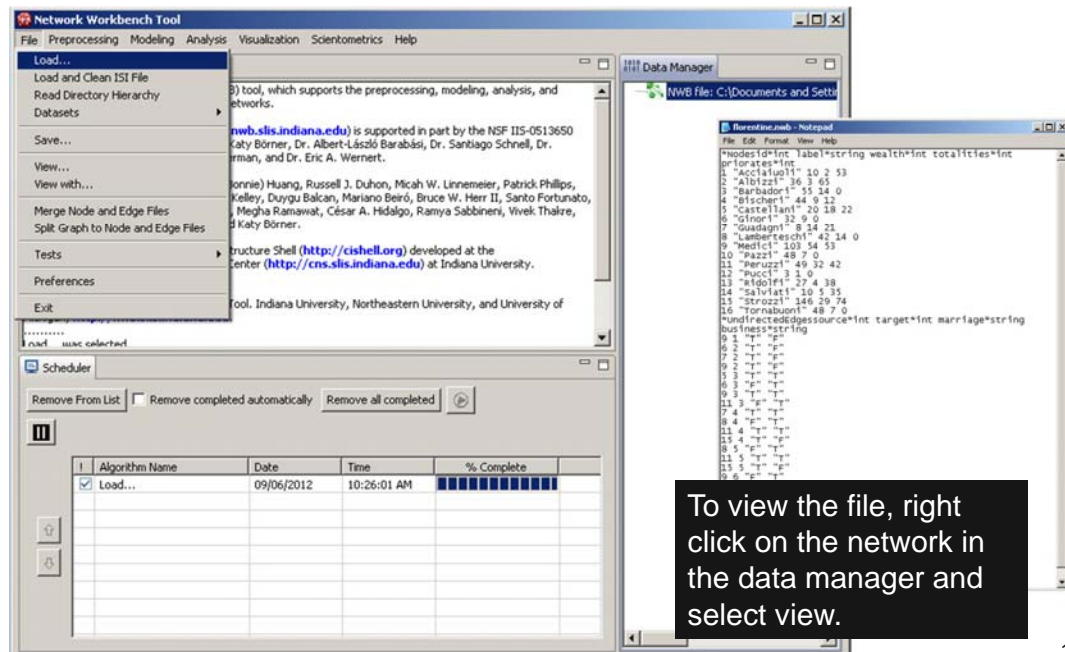
Edge attributes:

- Marriage T/F
- Business T/F

“Substantively, the data include families who were locked in a struggle for political control of the city of Florence around 1430. Two factions were dominant in this struggle: one revolved around the infamous Medicis, the other around the powerful Strozziis.”

More info is at <http://svitsrv25.epfl.ch/R-doc/library/ergm/html/florentine.html>

Load the Florentine dataset: [sampledata/socialscience/florentine.nwb](#)



The screenshot shows the Network Workbench Tool interface. The 'File' menu is open, highlighting the 'Load...' option. The Data Manager window shows the Florentine dataset loaded from 'C:\Documents and Settings\user\My Documents\Florentine.nwb'. A Notepad window displays the dataset's structure and content, including node attributes like 'priorates', 'wealth', and 'totalities', and edge attributes like 'marriage' and 'business'.

```

#nodesId*int label*string wealth*int totalities*int
priorates*int
1 "Acciaiuoli" 10 2 53
2 "Albizi" 36 3 65
3 "Barbadori" 55 14 0
4 "Bischeri" 44 0 11
5 "Castellani" 20 18 22
6 "Ginori" 32 9 0
7 "Guadagni" 9 14 21
8 "Lamberteschi" 42 14 0
9 "Medici" 103 54 53
10 "Pazzi" 48 7 0
11 "Peruzzi" 49 32 42
12 "Pucci" 3 3 0
13 "Ridolfi" 27 4 38
14 "Salviati" 19 5 35
15 "Strozzi" 146 29 74
16 "Tornabuoni" 48 7 0
#undirectededgesource*int target*int marriage*string
business*string
9 3 "m"
9 4 "m"
9 5 "m"
9 6 "m"
9 7 "m"
9 8 "m"
9 9 "m"
11 3 "m"
11 4 "m"
11 5 "m"
11 6 "m"
11 7 "m"
11 8 "m"
11 9 "m"
11 10 "m"
11 11 "m"
11 12 "m"
11 13 "m"
11 14 "m"
11 15 "m"
11 16 "m"
16 3 "m"
16 4 "m"
16 5 "m"
16 6 "m"
16 7 "m"
16 8 "m"
16 9 "m"
16 10 "m"
16 11 "m"
16 12 "m"
16 13 "m"
16 14 "m"
16 15 "m"
16 16 "m"

```

To view the file, right click on the network in the data manager and select view.

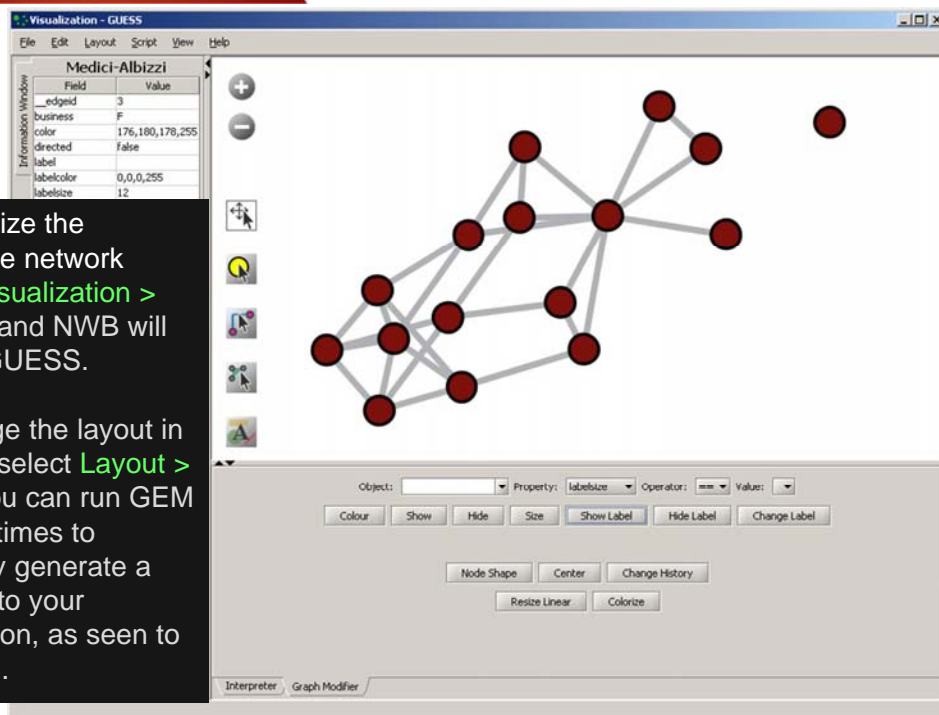
Visualizing the Florentine dataset

```

*Nodes
id*int label*string wealth*int totalities*int priorates*int
1 "Acciaiuoli" 10 2 53
2 "Albizzi" 36 3 65
3 "Barbadori" 55 14 0
4 "Bischeri" 44 9 12
5 "Castellani" 20 18 22
6 "Ginori" 32 9 0
7 "Guadagni" 8 14 21
8 "Lamberteschi" 42 14 0
9 "Medici" 103 54 53
10 "Pazzi" 48 7 0
11 "Peruzzi" 49 32 42
12 "Pucci" 3 1 0
13 "Ridolfi" 27 4 38
14 "Salviati" 10 5 35
15 "Strozzi" 146 29 74
16 "Tornabuoni" 48 7 0
*UndirectedEdges
source*int target*int marriage*string business*string
9 1 "T" "F"
6 2 "T" "F"
7 2 "T" "F"
9 2 "T" "F"
5 3 "T" "T"
  
```

13

Visualizing the Florentine dataset



To visualize the Florentine network select **Visualization > GUESS** and NWB will launch GUESS.

To change the layout in GUESS select **Layout > GEM** (you can run GEM multiple times to randomly generate a network to your satisfaction, as seen to the right).

14

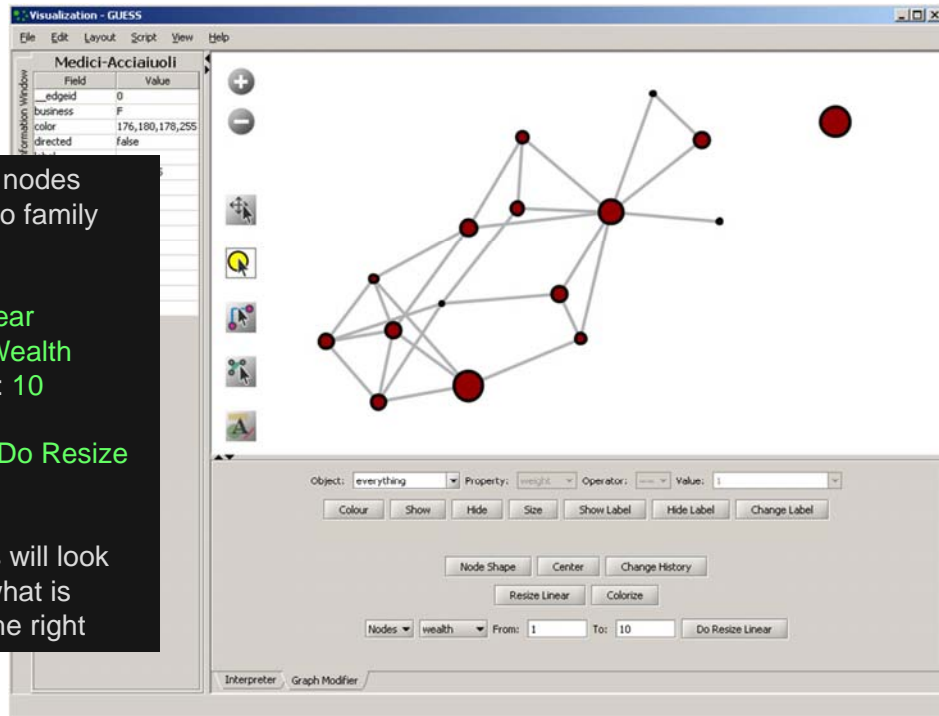
Visualizing the Florentine dataset

Resize the nodes according to family wealth.

Resize Linear
>Nodes> Wealth
From: 1 To: 10

Then click Do Resize Linear

The results will look similar to what is shown to the right



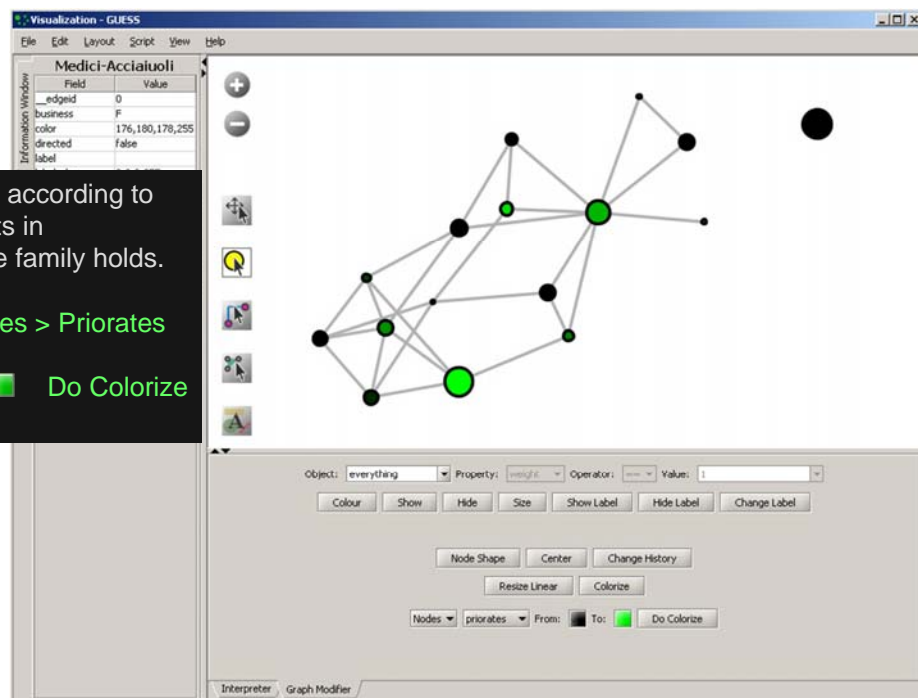
15

Visualizing the Florentine dataset

Colorize nodes according to how many seats in government the family holds.

Colorize > Nodes > Priorates

From : To: Do Colorize



16

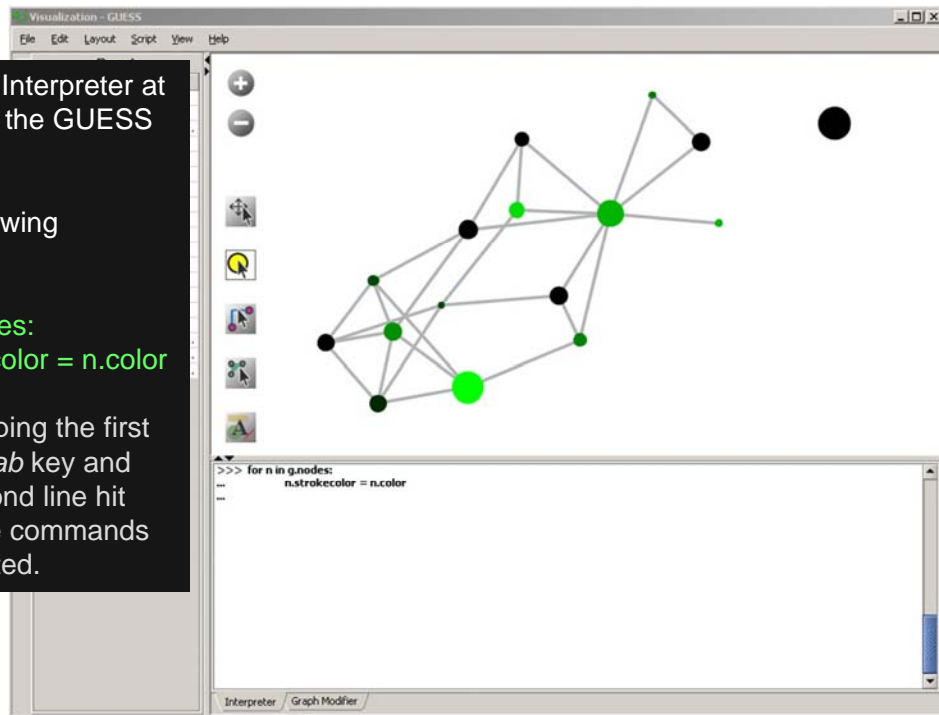
Visualizing the Florentine dataset

Switch to the Interpreter at the bottom of the GUESS window.

Type the following commands:

```
for n in g.nodes:
    n.strokecolor = n.color
```

Note: after typing the first line hit the *Tab* key and after the second line hit *Enter* and the commands will be executed.



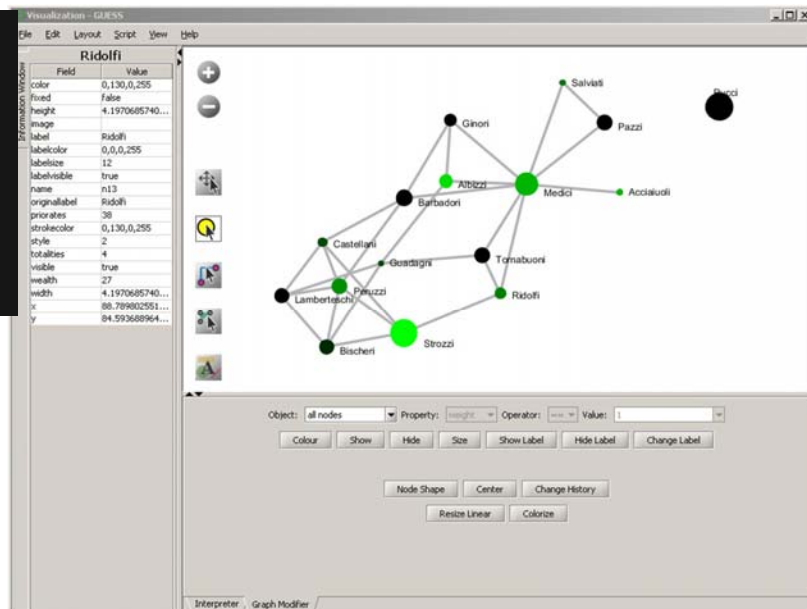
17

Visualizing the Florentine dataset

Add the family labels to the nodes.

Select Object: **All Nodes**
Then Click **Show Label**

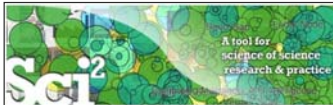
The family names will then appear next to their corresponding nodes.



18

Do you have questions so far?

19



Science of Science (Sci2) Tool

<http://sci2.cns.iu.edu>

- Explicitly designed for science of science 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

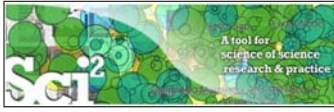
- Existing metrics have known flaws
- A reliable, open, joined-up data infrastructure is needed
- Data should be collected on the full range of scientists' work
- Social scientists and economists should be involved

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**.

20



General Network Extraction: Weighted, Undirected Co-Occurrence Network

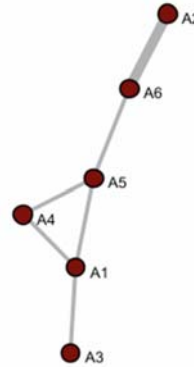
	A	B	C	D
1	Paper	Authors	References	Year
2	P1	A1		1970
3	P2	A2;A6	P1	1980
4	P3	A1;A3	P1;P2	1990
5	P4	A1;A4;A5	P2	1995
6	P5	A5;A6	P1;P2;P3;P4	1995
7	P6	A2;A6	P5	2000

Author co-occurrence network

Extract Network from Table
Extracts a network from a delimited table

Column Name:

Text Delimiter:



*Vertices

6

1 A1

2 A6

3 A2

4 A3

5 A5

6 A4

*Edges 6

2 3 2

1 4 1

1 5 1

5 6 1

1 6 1

2 5 1

21



General Network Extraction: Unweighted, Directed Bipartite Network

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

Paper-author bipartite (2-mode) network

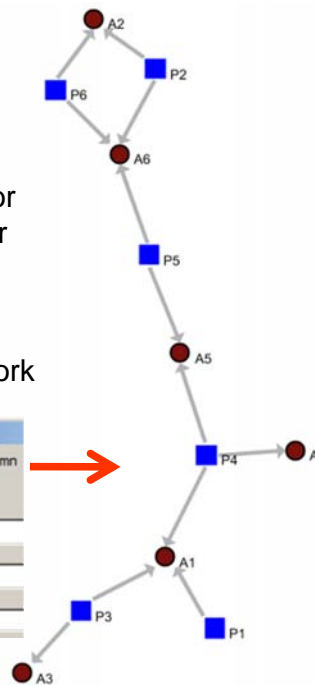
Extract Bipartite Network
Extract a bipartite network from two columns in the table. If the column values may list multiple entries, enter the special text which delimits them.

First column:

Second column:

Text Delimiter:

● Author
■ Paper



*Vertices 12

1 P1 bipartitetype "Paper"

2 A1 bipartitetype "Authors"

3 P2 bipartitetype "Paper"

4 A2 bipartitetype "Authors"

5 A6 bipartitetype "Authors"

6 P3 bipartitetype "Paper"

7 A3 bipartitetype "Authors"

8 P4 bipartitetype "Paper"

9 A4 bipartitetype "Authors"

10 A5 bipartitetype "Authors"

11 P5 bipartitetype "Paper"

12 P6 bipartitetype "Paper"

*Arcs

1 2

3 4

3 5

6 2

6 7

8 2

8 10

8 9

11 5

11 10

12 4

12 5

22



General Network Extraction: Unweighted, Directed Network

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

Extract Directed Network

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

Source Column: Paper

Target Column: Authors

Text Delimiter: ;

Analysis Modeling Visualization R Help

Temporal

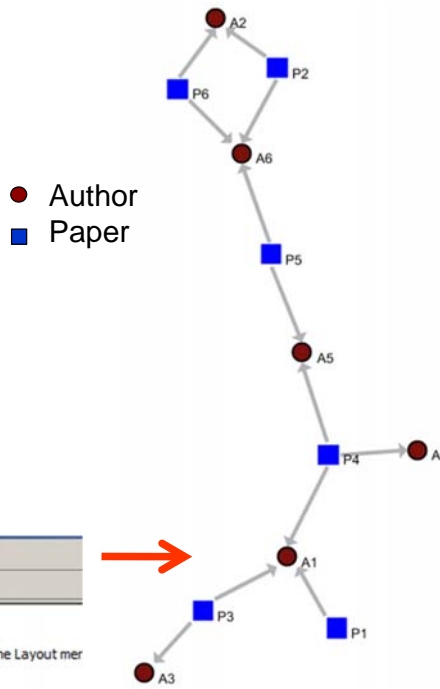
Geospatial

Topical

Networks

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

Node Indegree



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

23



General Network Extraction: Unweighted, Directed Paper-Citation Network

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

Extract Directed Network

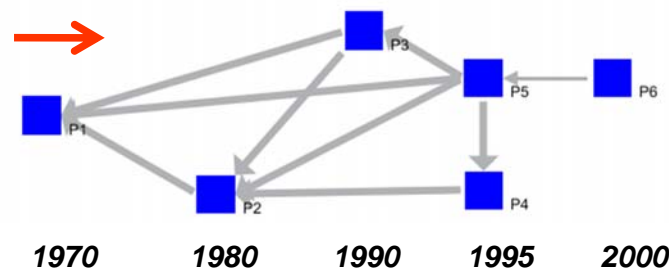
Given a table, this algorithm extracts a directed edge that starts at a column node.

Source Column: Paper

Target Column: References

Text Delimiter: ;

Arcs from papers to references



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

24



General Network Extraction: Unweighted, Directed Bi-Partite Network

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

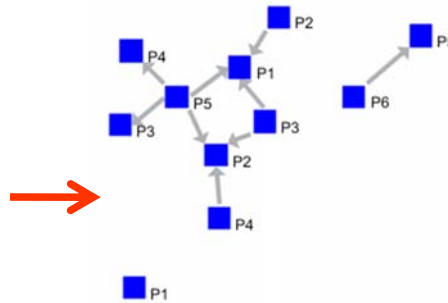
Extract Bipartite Network

Extract a bipartite network values may list multiple ent them.

First column:

Second column:

Text Delimiter:



WRONG!!!

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



ISI Paper-Citation Network Extraction

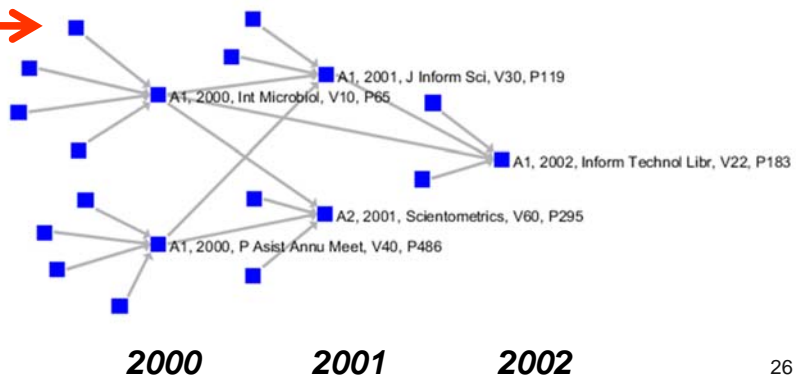
	A	B	C	D	E
1	Authors	Cited References	Publication Year	Title	Cite Me As
2	A1 A2	BENSMAN SJ, 1998, LIBR RESOUR TECH SER, V42, P147 BRO	2000	T1	A1, 2000, INT MICROBIOL, V10, P65
3	A1	BENSMAN SJ, 1999, LIBR RESOUR TECH SER, V42, P147 BRO	2000	T2	A1, 2000, P ASIST ANNU MEET, V40, P486
4	A2 A3	GARFIELD E,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

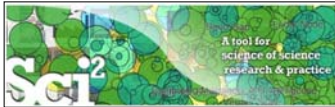
Sci2 Tool

File | Data Preparation | Preprocessing | Analysis

- Remove ISI Duplicate Records
- Remove Rows with Multitudinous Fields
- Extract Directed Network
- Extract Bipartite Network
- Extract Paper Citation Network**
- Extract Author Paper Network

*Arcs from references to papers—
in the direction of information flow*

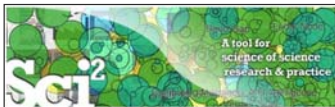




Studying Four Major NetSci Researchers (ISI Data)

- Data Acquisition & Preprocessing
 - Examine the data
 - Load the data into Sci2
 - Extract Co-Author Network
- Data Analysis, Modeling, & Layout
 - Run Network Analysis Toolkit
 - Run Node Degree
- Data Communication & Visualization Layers
 - Visualize the Co-Author network in GUESS

27

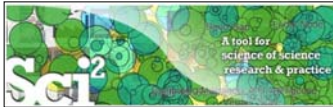


Studying Four Major NetSci Researchers (ISI Data)

Thomson Reuter's Web of Knowledge (WoS) is a leading citation database. 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.

[http://sci2.wiki.cns.iu.edu/5.1.4+Studying+Four+Major+NetSci+Researchers+\(ISI+Data\)](http://sci2.wiki.cns.iu.edu/5.1.4+Studying+Four+Major+NetSci+Researchers+(ISI+Data))

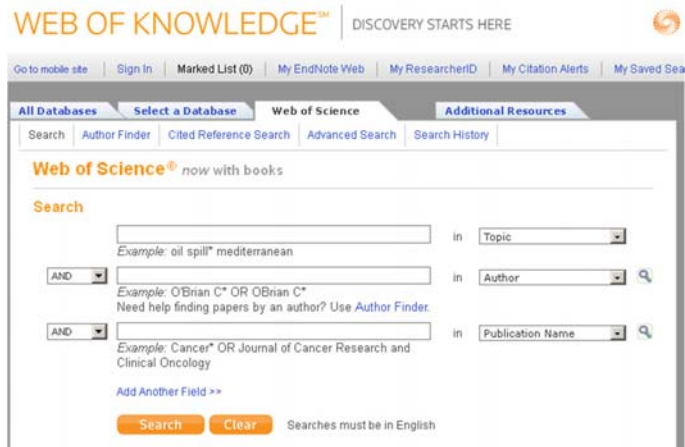
28



Data Acquisition from Web of Science

In Dec 2007, we downloaded 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



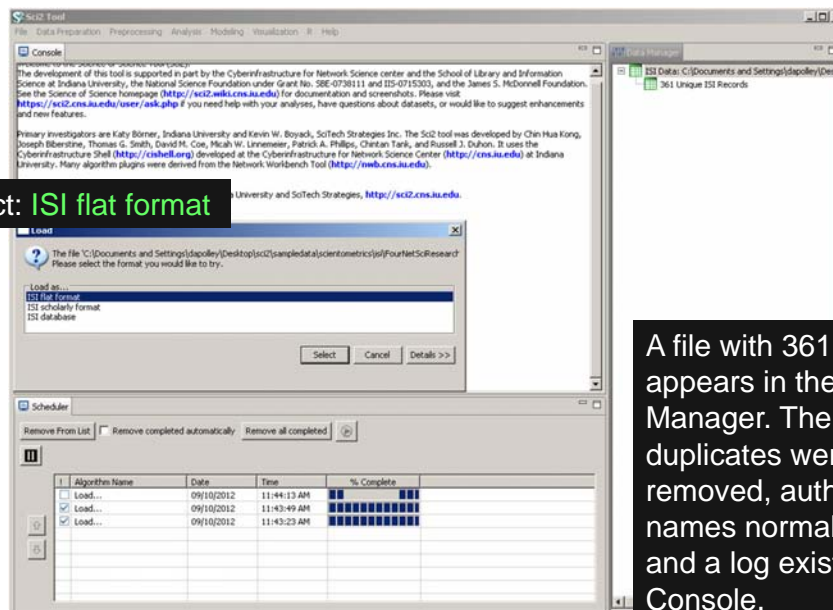
29



Extract Co-Author Network

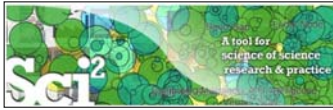
File > Load > *sampledata/scientometrics/isi/FourNetSciResearchers.isi*

Select: ISI flat format



A file with 361 records appears in the Data Manager. The duplicates were removed, author names normalized, and a log exists in the Console.

30



Extract Co-Author Network

To extract a co-author network select run:
Data Preparation > Extract Co-Author Network

Set the input parameter to: **isi**

Click **OK**. The result is an undirected co-author network along with an author information table, which appear in the Data Manager as follows:



The screenshot shows the Sci2 Tool interface. A dialog box titled 'Extract Co-Author Network' is open, with 'File Format' set to 'isi'. The Data Manager window on the right shows a folder named 'ISI Data' containing '361 Unique ISI Records'. Below the dialog, a table shows the progress of the algorithm:

Algorithm Name	Date	Time	% Complete
Extract Co-Author Network	09/10/2012	01:06:39 PM	██████████
Load...	09/10/2012	01:06:21 PM	██████████

31



Extract Co-Author Network

To further analyze the extracted co-author network run:
Analysis > Network > Network Analysis Toolkit (NAT)

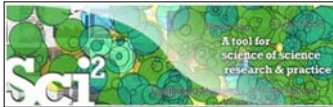
This will calculate the basic properties: 247 nodes and 891 edges.

Then, to calculate the number of neighbors for each node independent of co-authorship weight run:
Analysis > Network > Unweighted and Undirected > Node Degree

The screenshot shows the Sci2 Tool interface with the 'Network Analysis Toolkit (NAT)' menu open. The 'Unweighted & Undirected' sub-menu is selected, and 'Node Degree' is highlighted. The Data Manager window on the right shows the 'Extracted Co-Authorship Network' and 'Author information'.

Algorithm Name	Date	Time	% Complete
Network Analysis Toolkit (...)	09/10/2012	01:37:50 PM	██████████
Co-Author Network	09/10/2012	01:06:39 PM	██████████
Co-Author Network	09/10/2012	01:06:21 PM	██████████

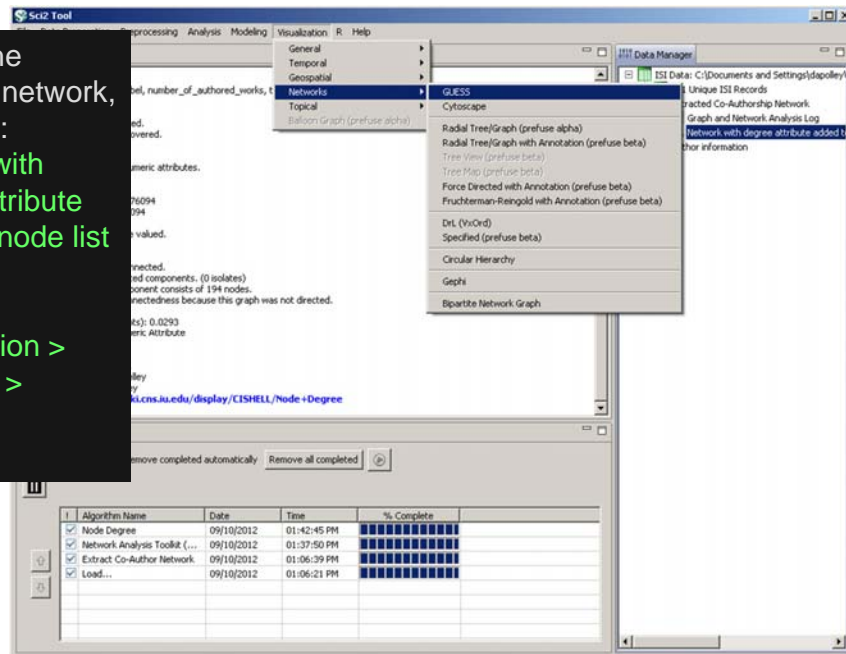
32



Visualizing the Extracted Co-Author Network

To view the complete network, select the:
Network with degree attribute added to node list

Then run:
Visualization > Networks > GUESS

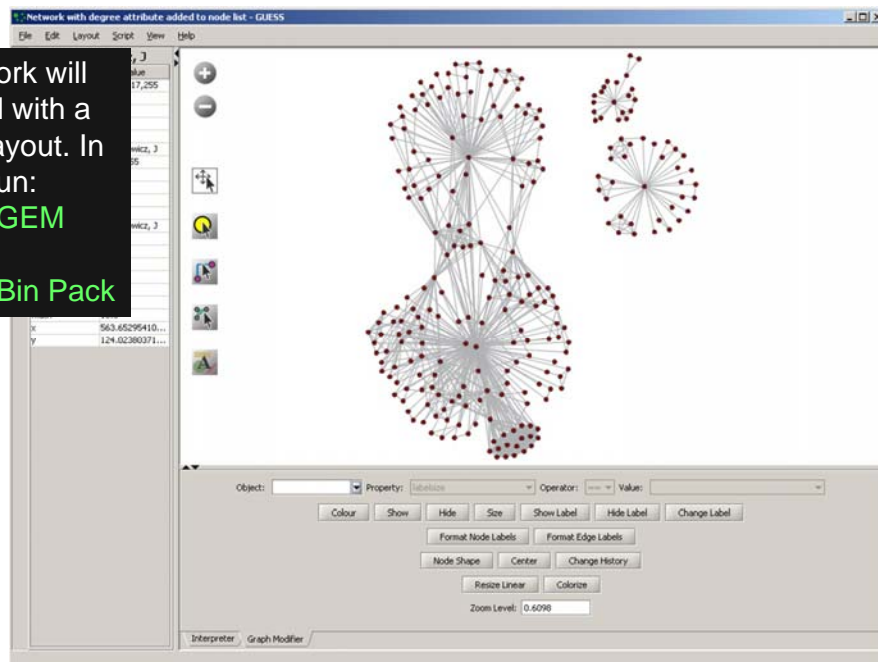


33

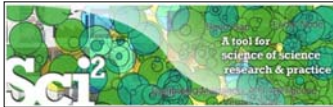


Visualizing the Extracted Co-Author Network

The network will be loaded with a random layout. In GUESS run:
Layout > GEM
 And:
Layout > Bin Pack



34



Visualizing the Extracted Co-Author Network

Follow these steps to enhance the visualization:

Resize Linear > Nodes > totaldegree

From: 5 To: 30

Do Resize Linear

Colorize > Nodes > times_cited

From : To: Do Colorize

Resize Linear > Edges > number_of_coauthored_works

Colorize > Edges > number_of_coauthored_works

From : To: Do Colorize

Object: nodes based on
-> Property: totaldegree
Operator: >= Value: 15
Show Label

In the Interpreter:
>for n in g.nodes:
 n.strokecolor =
 n.color



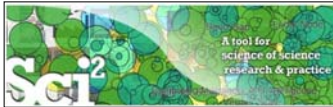
35



Q & A

Do you have questions so far?

36



Extracting a Paper-Paper (Citation) Network

- Data Acquisition & Preprocessing
 - Examine the data
 - Load the data into Sci2
 - Extract Directed Network with a property file -
- Data Analysis, Modeling, & Layout
 - Run Network Analysis Toolkit
- Data Communication & Visualization Layers
 - Visualize the Directed network in GUESS

37



Extracting a Paper-Paper (Citation) Network

For this workflow we are going to use a smaller dataset...

Load the StanleyWasserman.isi file, which can be downloaded from the Sci2 manual section 2.5 Sample Datasets.

Run **Data Preparation > Extract Directed Network** with the following parameters:

Source Column: Cited References

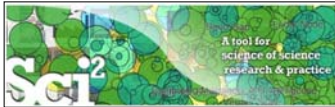
Target Column: Cite Me As

Text Delimiter: |

Aggregate Function File: C:/Documents and Settings/dapolley/Desktop/sci2/sampledata/scientometrics/properties/isiPaperCitation.properties

Note, the **isiPaperCitation.properties** file has been selected for the Aggregate Function File. More on this in the next slide.

38



Extracting a Paper-Paper (Citation) Network: How to use Property Files

The properties files, sometimes referred to as aggregate function files, are plain text files that facilitate analysis by allowing visualization according to certain attributes of nodes and edges. These files can be used where aggregation of data is to be performed based on certain unique values. This ultimately enhances the power of visualization. The properties files are located in Sci2/sampledata/scientometrics/properties

All properties files follow the same pattern:

`{node|edge}.new_attribute = table_column_name. [{target|source}].function`

The first part of file specifies whether action will be performed on a **node** or an **edge**.

The next part, **new_attribute**, will be a name selected by the user, which indicates the name of the attribute.

table_column_name is the name of the attribute we are going to operate on to create a new value for the final node, this can be the name of any of the node attributes in the network

The final part, **function**, will be selected by the user and will determine what function will be performed.

39



Extracting a Paper-Paper (Citation) Network: How to use Property Files

arithmeticmean - finds the average of an independent node attribute

geometricmean - finds the average of a dependent node attribute

count - counts the instances of appearance of a node attribute

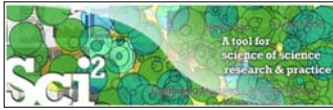
sum - the sum of each node's attribute values. Example use: When you have two author nodes who are really the same author and you want to combine the number of citations they have accumulated under both names.

max - the maximum value of each node's attribute values. Example use: When you have two author nodes who are really the same author, which have two differing author ages, you might want to assume that the younger age was based on an old record, and keep the older age

min - the maximum value of each node's attribute values.

mode - reports the most common value for an attribute

40



Extracting a Paper-Paper (Citation) Network: How to use Property Files

Below is the isiPaperCitation.properties file and a portion of the 361 Unique ISI Records table. Note, some of columns have been moved closer together to facilitate viewing.

	A	B	C	D	E	F	G	H	I	
1	Abstract	Authors	Authors (Full Names)	Beginning Page	Cited References	Book Series Subtitle	Times Cited	Book Series Title	Cited Patent	Cited Re
2	A multirelati	Pattison, P	Wasserman, S	Robins, G Ke	536	ANDERSON CJ, 1992, SOC NETWORKS, V14,	4			
3	The research	Pattison, P	Wasserman, S		169	BESAG J, 1975, STATISTICIAN, V24, P179 BE	30			
4	This paper ge	Robins, G	Pattison, P	Wasserman, S	371	AGRESTI A, 1990, CATEGORICAL DATA ANA	26			
5	A major critici	Anderson, C	Wasserman, S	Crouch, B	37	AGRESTI A, 1996, INTRO CATEGORICAL DA	32			
6	Spanning nea	Wasserman, S	Pattison, P		401	AGRESTI A, 1990, CATEGROICAL DATA ANA	99			
7	The methodo	ANDERSON, C	WASSERMAN, S		96	*SAS INC, 1994, SAS P243 TECHN REP AGRE	4			
8	In this paper	PATTISON, P	WASSERMAN, S		57	AGRESTI A, 1990, CATEGORICAL DATA ANA	4			
9	This paper de	FAUST, K	WASSERMAN, S		177	ANDERSON CJ, 1992, THESIS (UJLLINOIS) BEC				2. SOCIOLOGICAL METHODOLOGY
10	Network anal	GALASKIEW								
11	In recent year	WALKER, M								
12		ANDERSON								
13		ANDERSON								
14	Many method	FAUST, K	V							
15	The literature	ANDERSON, C	WASSERMAN, S	FAUST,	137	*UN, 1984, STAT PAP COMM TRAD D, V34 E	9			
16	A bipartite gr	WASSERMAN, S	IACIOBUCCI, D		13	ANDERSON E, 1987, J MARKETING RES, V22,	3			

```
node.globalCitationCount = Times Cited.target.mode
node.inOriginalDataSet = Authors.target.mode
node.localCitationCount = Cited References.source.count
```

41



Extracting a Paper-Paper (Citation) Network: How to use Property Files

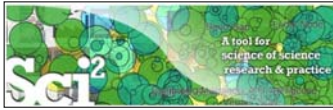
...Now back in Sci2

To view the complete network, select the:
Network with directed edges from Cited References to Cite Me As

Then run:
Visualization > Networks > GUESS

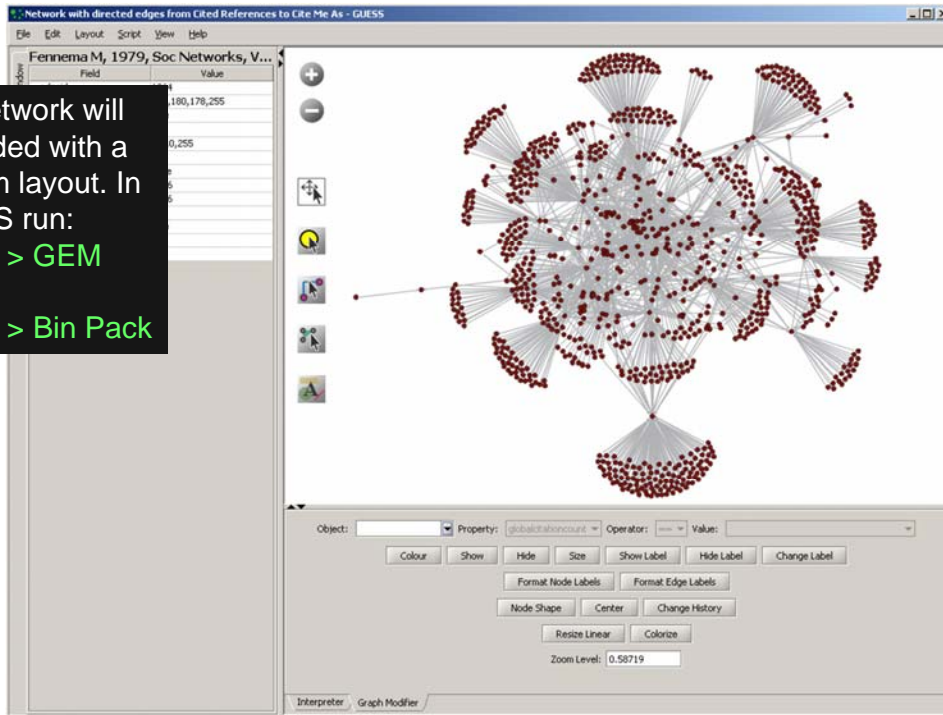
The screenshot shows the Sci2 Tool application window. The 'Networks' menu is open, and 'GUESS' is selected. The main workspace displays a network visualization. The task list at the bottom shows the progress of the 'Extract Directed Network' and 'Load...' tasks.

40



Extracting a Paper-Paper (Citation) Network: How to use Property Files

The network will be loaded with a random layout. In GUESS run:
Layout > GEM
And:
Layout > Bin Pack



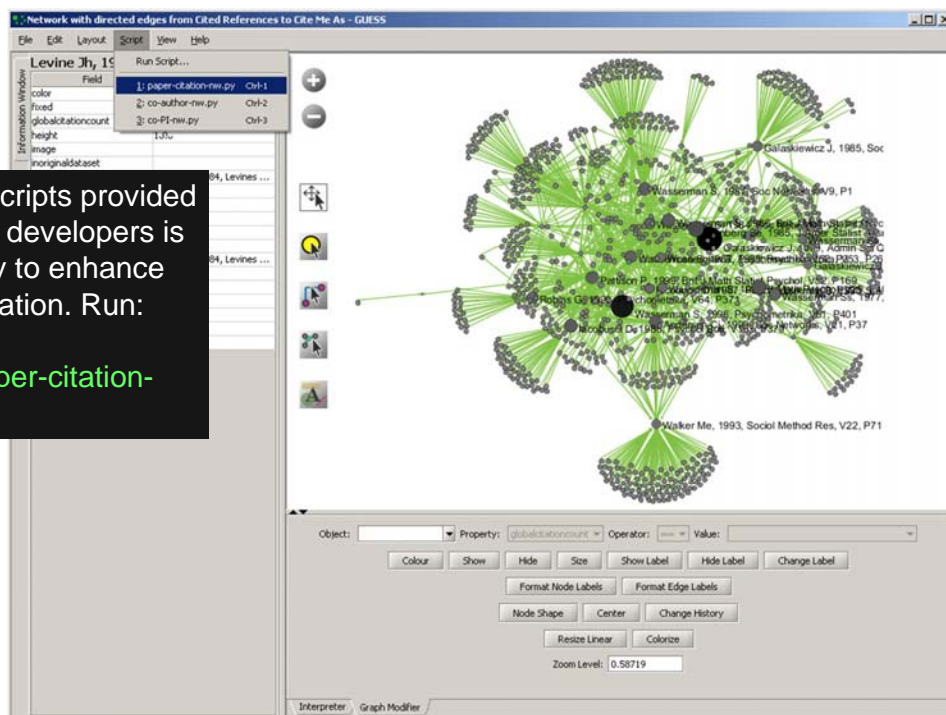
43



Extracting a Paper-Paper (Citation) Network: How to use Property Files

Using the scripts provided by the Sci2 developers is a quick way to enhance the visualization. Run:

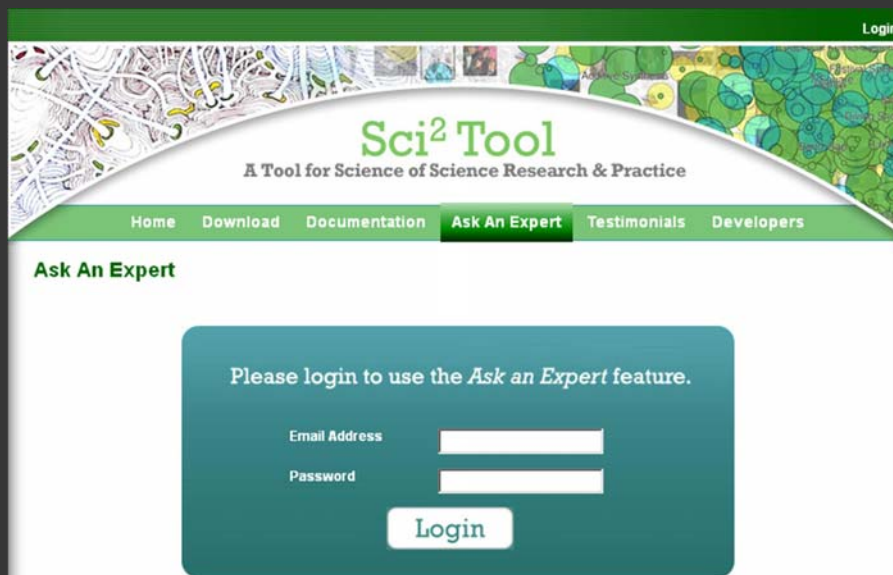
Script > paper-citation-nw.py



44



Q & A



If you have questions feel free to email me [dapolley\[at\]indiana\[dot\]edu](mailto:dapolley[at]indiana[dot]edu)