

NetworkWorkbench

A Workbench for Network Scientists

Towards an All-in-One Tool for Network Scientists

Interested in Large Scale Network Analysis, Modeling, and Visualization

Two-Hour Workshop

NWB Team @ IUB

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

Indiana University, Bloomington, IN

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

1

NetworkWorkbench

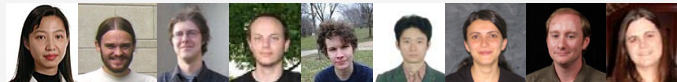
A Workbench for Network Scientists

Project Details

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



Software Team: Lead: Weixia (Bonnie) Huang
Members: Bruce Herr, Russell Duhon, Tim Kelley, Micah Linnemeier, Heng Zhang, Duygu Balcan, Bryan Hook & Ann McCranie
Previous Developers: Ben Markines, Santo Fortunato, Felix Terkhorn, Megha Ramawat, 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. 2008

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

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

2

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/>



- NWB Research Results – *Katy Börner*
- NWB Tool Overview and Demo – *Weixia (Bonnie) Huang*
- NWB Tool in Bioinformatics Research – *Tim Kelley & Santiago Schnell*
- NWB Tool for Scientometrics Research – *Katy Börner & Russell Duhon*
- Discussion of CShell and Future Work – *Bruce Herr*

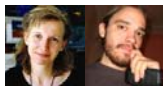
- Computational Social Science
- Computational Scientometrics
- Computational Economics
- Computational Proteomics
- Computational Epidemics

Computational Social Science

Studying large scale social networks such as Wikipedia

Vizzards 2007 Entry

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



locked until the mood cools (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.

Illuminated Diagram Display

W. Bradford Paley,
Kevin W. Boyack,
Richard Klavans,
and Katy Börner
(2007) Mapping,
Illuminating, and
Interacting with
Science.
SIGGRAPH 2007,
San Diego, CA.

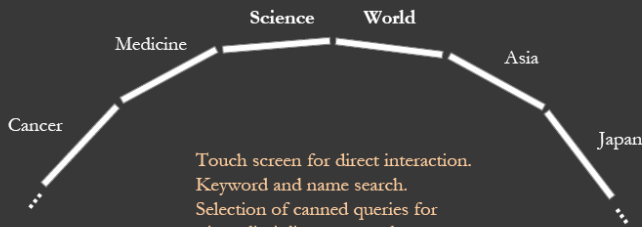


Network Workbench (<http://nwb.slis.indiana.edu>)



Re-implementation of Illuminated Diagram Software (in progress)
by *Advanced Visualization Lab, Indiana University*

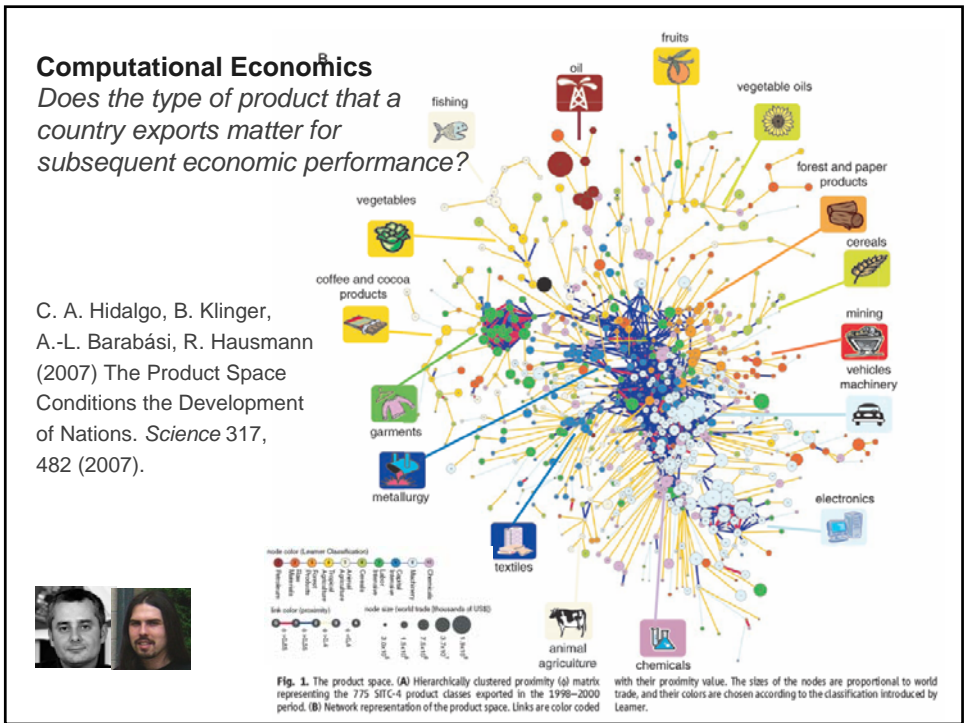
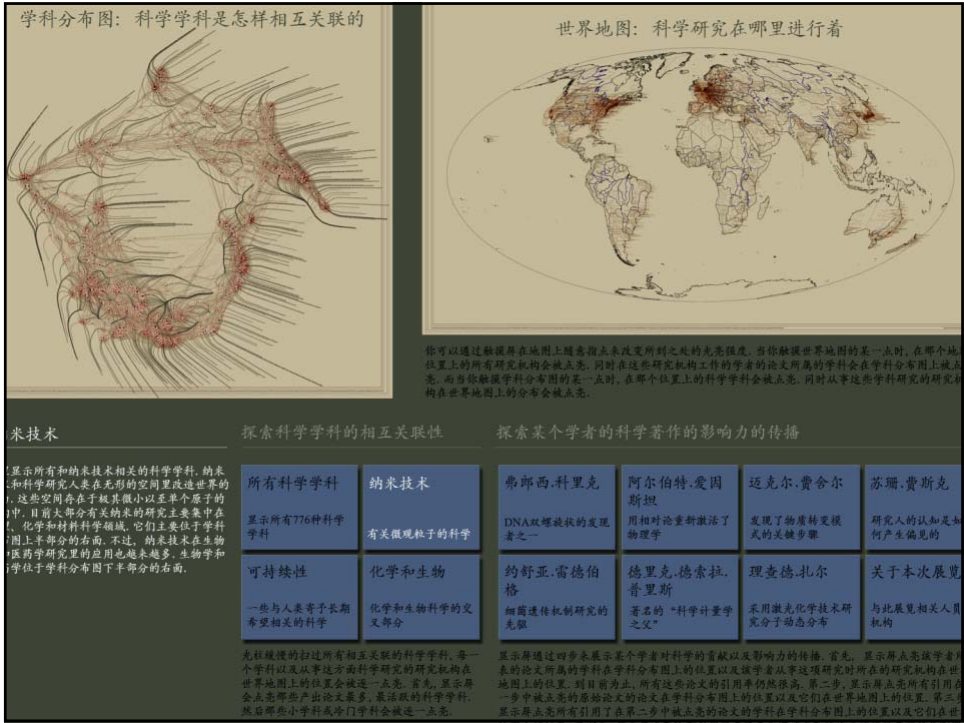
Drives unlimited number of ID screens.



Touch screen for direct interaction.
Keyword and name search.
Selection of canned queries for
- interdisciplinary research areas
- famous people
- activity patterns, e.g., bursts, trends, etc.



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



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.



Network Workbench (<http://nwb.sls.indiar>)



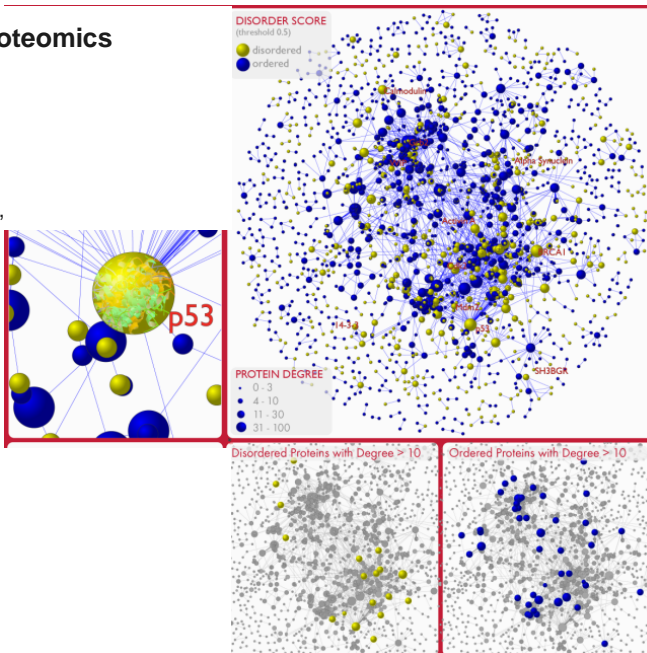
Figure 2 Drug-target network (DT network). The DT network is generated by using the known associations between FDA-approved drugs and their target proteins. Circles and rectangles correspond to drugs and target proteins, respectively. A link is placed between a drug node and a target node if the protein 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 Classification, and the target proteins (rectangular boxes) are colored according to their cellular component obtained from the Gene Ontology database.

Computational Proteomics

S. Schnell, S. Fortunato, and S. Roy (2007). Is the intrinsic disorder of proteins the cause of the scale-free architecture of protein-protein interaction networks? *Proteomics* 7, 961-964.



Network Workbench (<http://nwb.sls.indiana.edu>)



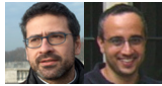
Computational Epidemics

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

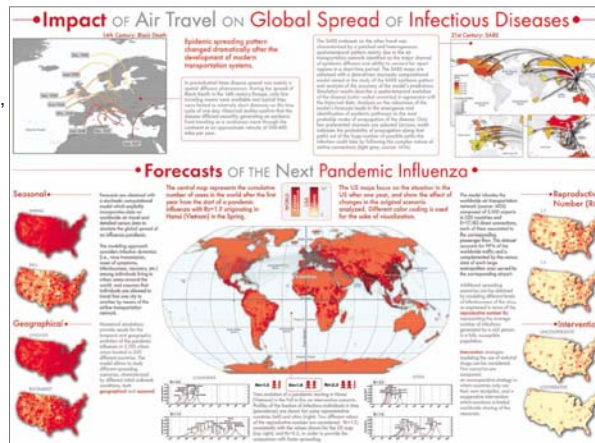
Epidemic Modeling in Complex realities, V. Colizza, A. Barrat, M. Barthelemy, A. Vespignani, *Comptes Rendus Biologie*, 330, 364-374 (2007).

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).



Network Workbench (<http://nwb.sls.indiana.edu>).



NetworkWorkbench The NWB Tool

A Workbench for Network Scientists

Network Workbench (<http://nwb.sls.indiana.edu>)

16

❑ Data

- Different data formats
- Different data models

❑ Algorithms

- Different research purposes (preprocessing, modeling, analysis, visualization, clustering)
- Different implementations of the same algorithm
- Different programming languages

❑ Match between Data and Algorithms

❑ Different communities and practices

❑ Different tools (Pajek, UCInet, Guess, Cytoscape, R, NWB tool)

Network Workbench (NWB) Tool

- A network analysis, modeling, and visualization toolkit for physics, biomedical, and social science research.
- Install and run on multiple Operating Systems.
- Uses Cyberinfrastructure Shell Framework underneath.

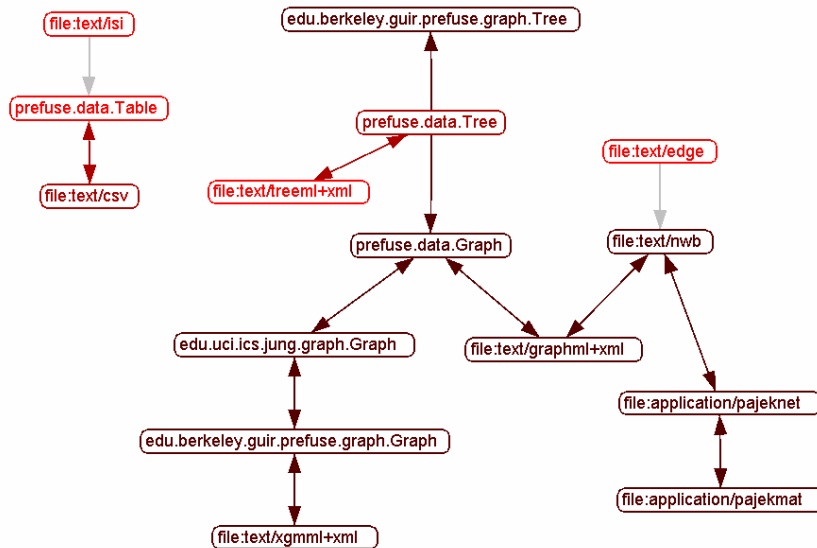
Cyberinfrastructure Shell (CIShell)

- An open source, software framework for the integration and utilization of datasets, algorithms, tools, and computing resources.

NWB Community Wiki

- A place for users of the NWB Tool, the Cyberinfrastructure Shell (CIShell), or any other CIShell-based program to request, obtain, contribute, and share algorithms and datasets.
- All algorithms and datasets that are available via the NWB Tool have been well documented in the Community Wiki.

- ❑ Can load, view, process and save the following file formats:
 - GraphML (.xml or .graphml)
 - XGMML (.xml)
 - Pajek .net (.net)
 - Pajek .mat(.mat)
 - **NWB** (.nwb)
 - TreeML (.xml)
 - Edge list (.edge)
 - CSV (.csv)
 - isi (.isi)
- ❑ Can load two CSV files (node list and edge list) and construct a network.
- ❑ Can load an isi file, extract co-authorship network and update graph by merging nodes if needed.



Download from <http://nwb.slis.indiana.edu/software.html>

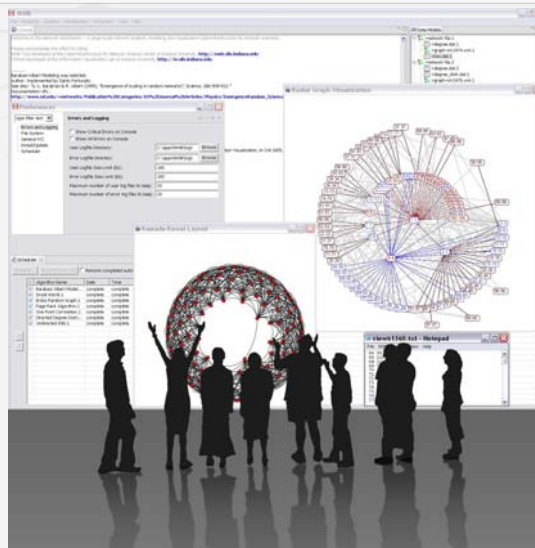
Major features in v0.8.0 Release

- Installs and runs on Windows, Linux x86 and Mac OsX.
- Provides over 60 modeling, analysis and visualization algorithms. Half of them are written in Fortran, others in Java.
- Supports large scale network modeling and analysis (over 100,000 nodes)
- Supports various visualization layouts with node/edge annotation.
- Provides several sample datasets with various formats.
- Supports multiple ways to introduce a network to the NWB tool.
- Supports automatic Data Conversion.
- Provides a Scheduler to monitor and control the progress of running algorithms.
- Integrates a 2D plotting tool – Gnuplot (requires pre-installation on Linux and Mac).
- Integrates GUESS (runs on Linux and Mac. *Windows forthcoming)

Category	Algorithm	Language	Analysis Algorithm	Language
Preprocessing	Random Node Deletion	JAVA	Node Betweenness Centrality	FORTRAN
	High Degree Node Deletion	JAVA	Average Shortest Path	FORTRAN
	Pathfinder Network Scaling	JAVA	Connected Components	FORTRAN
	Directory Hierarchy Reader	JAVA	Diameter	FORTRAN
Modeling	Erdős-Rényi Random	FORTRAN	Page Rank	FORTRAN
	Barabási-Albert Scale-Free	FORTRAN	Shortest Path Distribution	FORTRAN
	Watts-Strogatz Small World	FORTRAN	Watts-Strogatz Clustering Coefficient	FORTRAN
	Chord	JAVA	Watts-Strogatz Clustering Coefficient Versus Degree	FORTRAN
	CAN	JAVA	Directed k-Nearest Neighbor	FORTRAN
	Hypergrid	JAVA	Undirected k-Nearest Neighbor	FORTRAN
	PRU	JAVA	Indegree Distribution	FORTRAN
	TARL	JAVA	Outdegree Distribution	FORTRAN
	Tree Map	JAVA	Node Indegree	FORTRAN
Visualization	Tree Viz	JAVA	Node Outdegree	FORTRAN
	Radial Tree / Graph	JAVA	One-point Degree Correlations	FORTRAN
	Kamada-Kawai	JAVA	Undirected Degree Distribution	FORTRAN
	Force Directed	JAVA	Node Degree	FORTRAN
	Spring	JAVA	k Random-Walk Search	JAVA
	Fruchterman-Reingold	JAVA	Random Breadth First Search	JAVA
	Circular	JAVA	CAN Search	JAVA
			Chord Search	JAVA
		Weak Component Clustering	JAVA	
		Tool: GnuPlot		

NWB tool and CShell provide

- ❑ A testbed for diverse algorithm implementations
- ❑ A mechanism to quickly integrate an algorithm and disseminate it through the NWB tool and community wiki.
- ❑ A bridge between what application users need and what algorithm developers can provide.



Domain Specific Analysis: Biological Networks

Biological Networks

Types of Networks

- ❑ Protein-Protein Interaction
 - Maps the interaction between proteins.
 - Typically undirected
 - Concerned with co-expression
- ❑ Metabolic
 - Typically directed networks.
 - Map the reactions of proteins and enzymes to their products.
 - Show the chemical pathways for the creation of essential components and the energy required for those reactions

More Networks

- ❑ Cell Signaling Networks
 - Maps the flows of communication proteins between and inside cells
 - Typically directed
- ❑ Gene Regulatory Networks
 - Maps the interactions between genes and proteins to gene expression
 - Typically directed

Critical statistics

- ❑ Degree
 - How many edges to other nodes
- ❑ Degree Distribution
 - Probability a node has k edges.
- ❑ Shortest path and mean path length
 - Smallest number of edges a node A must cross before reaching B .
 - Average of the shortest paths.
 - Gives an idea of how navigable a network is.

Clustering Coefficient

- The number of edges connecting the k neighbors of a node n to one another
- The average $\langle C \rangle$ is taken over all the clustering coefficients
- $C(k)$ is the average clustering coefficient for all nodes with k edges.

- Biological networks demonstrate an amazing ability to survive despite drastic environmental intervention
 - Redundant systems are only a necessary, not a sufficient condition for this robust behavior
 - Homogeneously connected networks are not error-tolerant
- Scale-free networks are error-tolerant, but vulnerable to attacks.
 - Deletion of high-degree nodes leads to rapid increase in diameter and change in topology

Large and Dense data means inferring topology from subgraphs

- ❑ Inferring full graph topology from subgraph samples can lead to false categorization of network topology.
- ❑ Not true in all cases, dependent on coverage of the network
 - Low coverage means low confidence in the inferred topology
- ❑ Limitations in data collection
 - Yeast two-hybrid and Mass Spectrometry methods can lead to false-positives and false negatives
 - These errors in data collection may move the topology more towards scale-free

- ❑ Dynamic Network Analysis
 - Metabolic, Cell Signaling, and Gene regulatory networks are dynamic
 - We want to measure presence or levels of reactants over time.

NetworkWorkbench
A Workbench for Network Scientists

Demo

The screenshot displays the NetworkWorkbench software interface. It features a central workspace with two large network graphs: one is a circular graph with nodes and edges, and the other is a more complex, multi-layered graph. The interface includes several panels: a top-left panel with text and links, a top-right panel with a list of nodes, a middle-left panel with a list of nodes and their properties, and a bottom-right panel with a list of nodes. The background of the interface is a light gray with a faint network pattern. In the foreground, there are silhouettes of several people standing and looking at the software, suggesting a demonstration or collaborative work environment.

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

33

NetworkWorkbench
A Workbench for Network Scientists

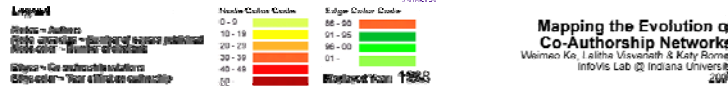
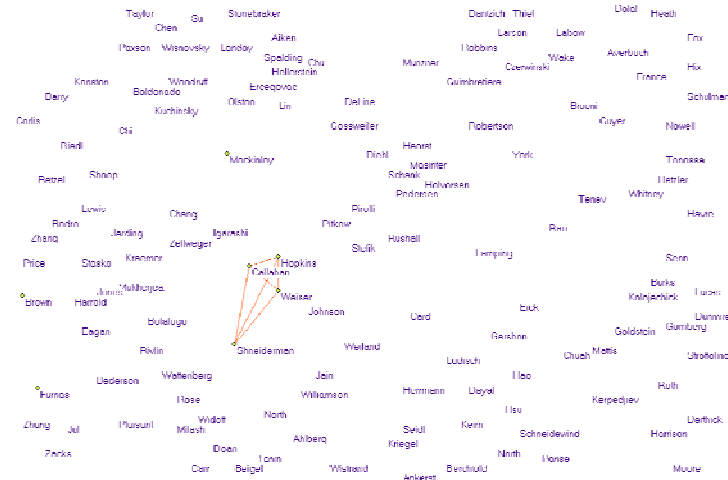
NWB Tool for Scientometrics Research

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

34

Mapping the Evolution of Co-Authorship Networks in Information Visualization, 1988 - 2004

Ke, Viswanath & Börner (2004)



Mapping the Evolution of Co-Authorship Networks
 Weirwan Ke, alpha Viswanath & Katy Börner
 InfoVis Lab @ Indiana University
 2004

NetworkWorkbench

A Workbench for Network Scientists

Data Acquisition from Web of Science

Download all papers by

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

from

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

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

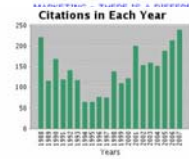
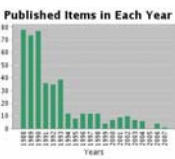
Eugene Garfield

1525 papers

papers/citations for last 20 years

1,525 results found (Set # 14) Go to Page: 1 of 153

- Records 1 -- 10 | Show 10 per page
- Use the checkboxes to select records for output. See the sidebar for options.
- 1. **GARFIELD E**
CITATION ANALYSIS AS A TOOL IN JOURNAL EVALUATION - JOURNALS CAN BE RANKED BY FREQUENCY AND IMPACT OF CITATIONS FOR SCIENCE POLICY STUDIES
SCIENCE 178 (4060): 471-1972
Times Cited: 625
[Full Text](#)
 - 2. **GARFIELD E**
CITATION INDICES FOR SCIENCE - NEW DIMENSION IN DOCUMENTATION THROUGH ASSOCIATION OF IDEAS
SCIENCE 122 (3159): 108-111 1955
Times Cited: 325
[Full Text](#)
 - 3. **Garfield E**
How can impact factors be improved?
BRITISH MEDICAL JOURNAL 313 (7054): 411-413 AUG 17 1996
Times Cited: 126
[Full Text](#) [View Full Text](#)
 - 4. **GARFIELD E**
NEW YEAR, NEW BUILDING
CURRENT CONTENTS (1): 5-8 1980
Times Cited: 256
[Full Text](#)
 - 5. **GARFIELD E**
INTRODUCING CITATION CLASSICS - HUMAN SIDE OF SCIENTIFIC REPORTS
CURRENT CONTENTS (1): 5-7 1977
Times Cited: 258
[Full Text](#)



Results Found: 1,525
Sum of the Times Cited (t): 5,242
Average Citations per Item (i): 3.44
h-index: 31

Network Workbench (<http://nwb.sls.indiana.edu>)

Can download 500 records max.
Exclude Current Contents articles

Refine your results
Subject Categories | Source Titles | Document Types | Authors | Publication Years [more choices](#)

Top Source Titles:

<input checked="" type="checkbox"/> CURRENT CONTENTS (1066)	<input type="checkbox"/> NATURE (12)	<input type="checkbox"/> CURRENT COMMENTS (7)
<input type="checkbox"/> SCIENTIST (143)	<input type="checkbox"/> JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE (11)	<input type="checkbox"/> ANNALS OF INTERNAL MEDICINE (5)
<input type="checkbox"/> CURRENT CONTENTS/LIFE SCIENCES (89)	<input type="checkbox"/> ABSTRACTS OF PAPERS OF THE AMERICAN CHEMICAL SOCIETY (10)	<input type="checkbox"/> BRITISH MEDICAL JOURNAL (5)
<input type="checkbox"/> JOURNAL OF CHEMICAL DOCUMENTATION (12)	<input type="checkbox"/> SCIENCE (9)	<input type="checkbox"/> CURRENT SCIENCE (5)
<input type="checkbox"/> JOURNAL OF INFORMATION SCIENCE (12)	<input type="checkbox"/> SCIENTOMETRICS (9)	more (up to 100)...

For more options, use [Analyze Results](#)

Include only articles. Download 99 articles.

Refine your results
Subject Categories | Source Titles | Document Types | Authors | Publication Years

Top Document Types:

<input type="checkbox"/> EDITORIAL MATERIAL (239)	<input type="checkbox"/> DISCUSSION (4)	<input type="checkbox"/> REVIEW (3)
<input checked="" type="checkbox"/> ARTICLE (99)	<input type="checkbox"/> BIOGRAPHICAL-ITEM (3)	<input type="checkbox"/> CORRECTION (1)
<input type="checkbox"/> LETTER (76)	<input type="checkbox"/> BOOK REVIEW (3)	<input type="checkbox"/> CORRECTION, ADDITION (1)
<input type="checkbox"/> MEETING ABSTRACT (17)	<input type="checkbox"/> NOTE (3)	
<input type="checkbox"/> ITEM ABOUT AN INDIVIDUAL (7)	<input type="checkbox"/> REPRINT (3)	

For more options, use [Analyze Results](#)

Network Workbench (<http://nwb.sls.indiana.edu>)

38

NetworkWorkbench
A Workbench for Network Scientists

Data Acquisition from Web of Science (cont.)

99 results found Go to Page: 1 of 10
 Records 1 -- 10 | Show 10 per page << < | 2 | > >>

Use the checkboxes to select records for output. See the sidebar for options.

- 1. **GARFIELD E**
[CITATION ANALYSIS AS A TOOL IN JOURNAL EVALUATION - JOURNALS CAN BE RANKED BY FREQUENCY AND IMPACT OF CITATIONS FOR SCIENCE POLICY STUDIES](#)
 SCIENCE 178 (4060): 471+ 1972
 Times Cited: **672**
[IU Link](#)
- 2. **GARFIELD E**
[CITATION INDEXES FOR SCIENCE - NEW DIMENSION IN DOCUMENTATION THROUGH ASSOCIATION OF IDEAS](#)
 SCIENCE 122 (3159): 108-111 1955
 Times Cited: 295
[IU Link](#)
- 3. **Garfield E**
[How can impact factors be improved?](#)
 BRITISH MEDICAL JOURNAL 313 (7054): 411-413 AUG 17 1996
 Times Cited: 156
[IU Link](#) [VIEW FULL TEXT](#)
- 4. **GARFIELD E**
[CITATION INDEXING FOR STUDYING SCIENCE](#)
 NATURE 227 (5259): 669& 1970
 Times Cited: 134
[IU Link](#)
- 5. **GARFIELD E**
[SCIENCE CITATION INDEX-NEW DIMENSION IN INDEXING - UNIQUE APPROACH UNDERLIES VERSATILE BIBLIOGRAPHIC SYSTEMS FOR COMMUNICATING + EVALUATING INFORMATION](#)
 SCIENCE 144 (361): 649& 1964
 Times Cited: 109
[IU Link](#)

NetworkWorkbench
A Workbench for Network Scientists

Data Acquisition from Web of Science (cont.)

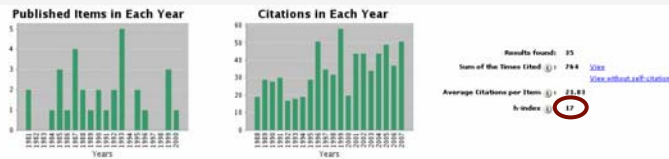
Stanley Wasserman
35 papers

35 results found (Set #7) [Why 35?](#) Go to Page: 1 of 4
 Records 1 -- 10 | Show 10 per page << < | 2 | > >>

Use the checkboxes to select records for output. See the sidebar for options.

- 1. GALASKIEWICZ J, WASSERMAN S
[MIMETIC PROCESSES WITHIN AN INTERORGANIZATIONAL FIELD - AN EMPIRICAL-TEST](#)
 ADMINISTRATIVE SCIENCE QUARTERLY 34 (3): 454-479 SEP 1989
 Times Cited: **122**
[IU Link](#)
- 2. Wasserman S, Pattison P
[Logit models and logistic regressions for social networks .I. An introduction to Markov graphs and p](#)
 PSYCHOMETRIKA 61 (3): 401-425 SEP 1996
 Times Cited: 59
[IU Link](#)
- 3. FIENBERG SE, MEYER MM, WASSERMAN SS
[STATISTICAL-ANALYSIS OF MULTIPLE SOCIOMETRIC RELATIONS](#)
 JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION 80 (389): 51-67 1985
 Times Cited: 47
[IU Link](#)
- 4. WASSERMAN S
[ANALYZING SOCIAL NETWORKS AS STOCHASTIC-PROCESSES](#)
 JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION 75 (370): 280-294 1980
 Times Cited: 38
[IU Link](#)
- 5. JACOBUCCI D, WASSERMAN S
[A GENERAL FRAMEWORK FOR THE STATISTICAL-ANALYSIS OF SEQUENTIAL DYADIC INTERACTION DATA](#)
 PSYCHOLOGICAL BULLETIN 103 (3): 379-390 MAY 1988
 Times Cited: 34
[IU Link](#)

papers/citations for last 20 years

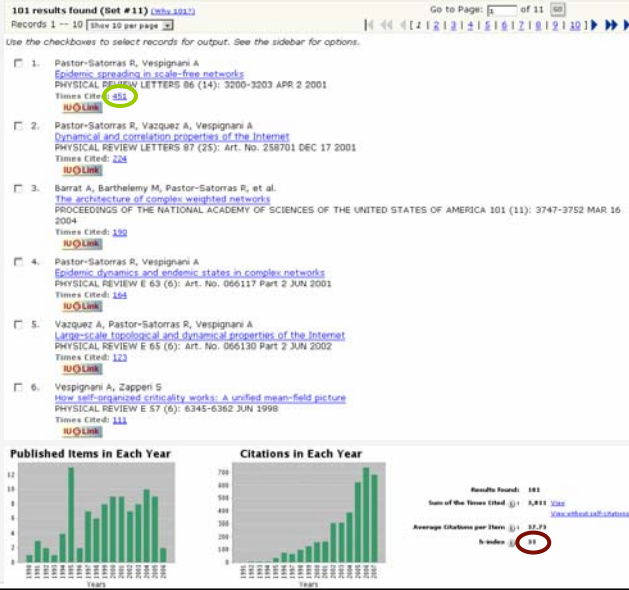


Alessandro Vespignani

101 papers

papers/citations for last 20 years

Network Workbench (<http://nwb.sls.indiana.edu>)



Albert-László Barabási

126 papers

papers/citations for last 20 years

Network Workbench (<http://nwb.sls.indiana.edu>)

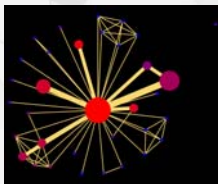


Comparison of Counts

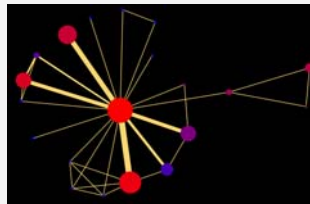
	Age	Highest Cited Paper	H-Index
Eugene Garfield	82	672	31
Stanley Wasserman		122	17
Alessandro Vespignani	42	451	33
Albert-László Barabási	40	2218	47

Comparison of Networks

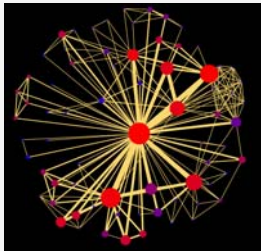
Eugene Garfield



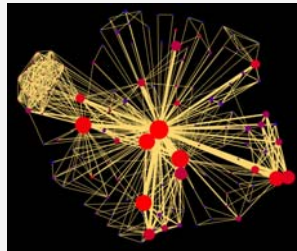
Stanley Wasserman

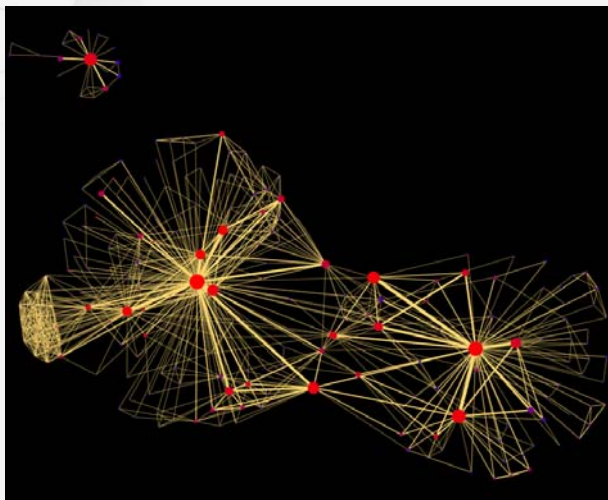


Alessandro Vespignani



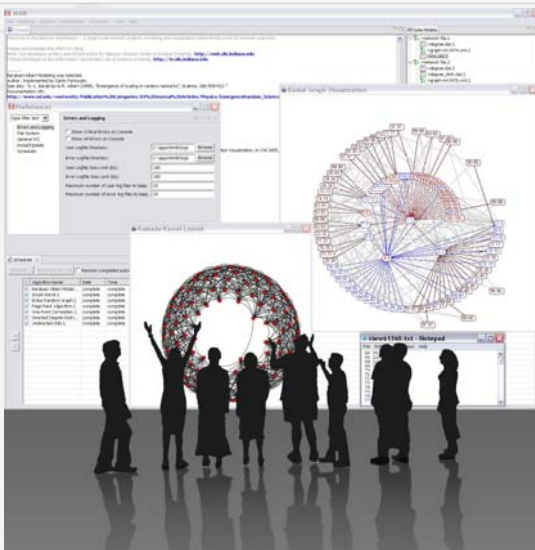
Albert-László Barabási





Network Workbench (<http://nwb.sls.indiana.edu>).

45

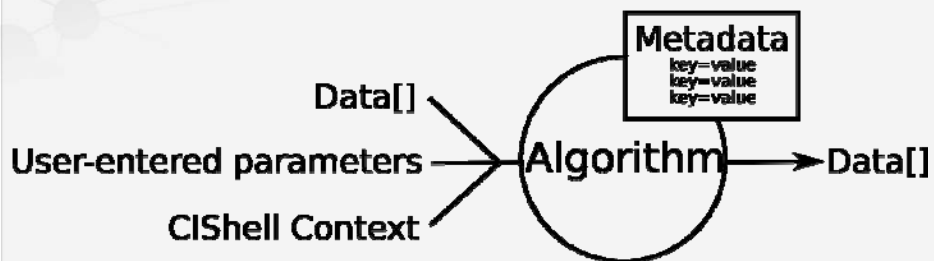


Network Workbench (<http://nwb.sls.indiana.edu>).

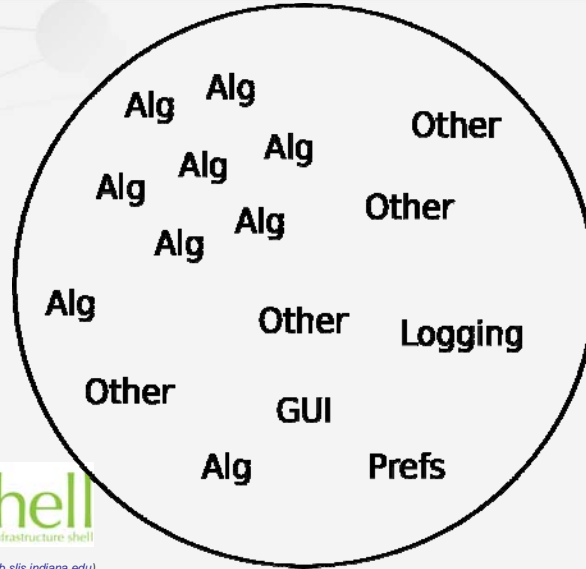
46



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](#).

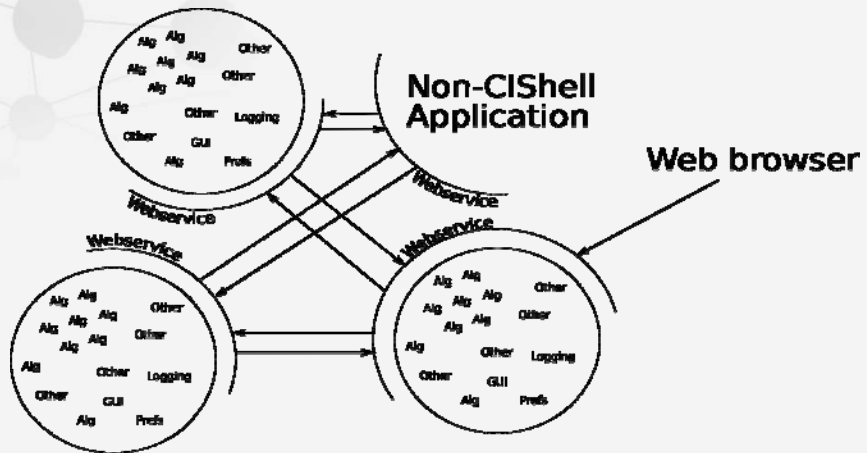


Pooling Algorithms



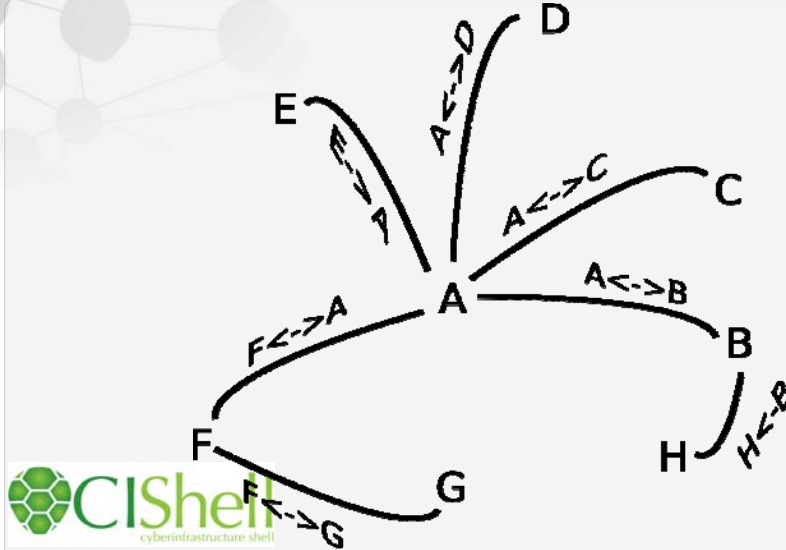
Network Workbench (<http://nwb.sls.indiana.edu>)

Inter-Pool Interaction



Network Workbench (<http://nwb.sls.indiana.edu>)

Data Conversion



Adding New Plugins

- Using update sites
- Using OSGi Console Magick™!
- Dropping plugins into the plugins directory
- Using the NWB Community Wiki

- Wizard-driven templates ease development
- Documentation Forthcoming
 - CIShell Specification
 - CIShell Developer's Guide
- Some preliminary documentation is available at <http://cishell.org>
- A future workshop will address this
- We are available for consulting

- New release (v0.8.0) of the NWB tool and a complete user manual with tutorials (v1.0) will be ready after Christmas.
- An end-user workshop is scheduled in the middle of January at IUB (Alex for physics and internet research, Ann & Stan for social network research)
- Ann McCranie will run another end-user workshop in late January during the Sunbelt Conference
- CIShell specification and CIShell/NWB algorithm developer guide will be available in late January.
- Workshop for algorithm developers will be planned accordingly.

Add features to serve communities including Physics, Biology, Social Science, and Scientometrics.

- Integrate classic datasets
- Support the most popular data formats for biology and social science research.
- Develop the converters to bridge those formats to the current formats supported by NWB tool.
- Design and deliver better visualization algorithms and modularity
- Develop components to connect and query SDB
- R bridge
- Customize Menu – Users can re-organize the algorithms for their needs
- Continue integrating best algorithm implementations

- ❑ Hidalgo, César A. and C. Rodriguez-Sickert. [Persistence, Topology and Sociodemographics of a Mobile Phone Network](#). 2007. (Submitted to Physica A)
- ❑ Hidalgo, C.A., B. Klinger, A. L. Barabási, and R. Hausmann. [The Product Space and its Consequences for Economic Growth](#). *Science*. Vol. 317 (2007, July 27): 482-487.
- ❑ Börner, Katy. [Making Sense of Mankind's Scholarly Knowledge and Expertise: Collecting, Interlinking, and Organizing What We Know and Different Approaches to Mapping \(Network\) Science](#). *Environment and Planning B: Planning and Design*. Vol. 34(5), 808-825, Pion.
- ❑ 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.
- ❑ Vespignani, Alessandro, Soma Sanyal, and Katy Börner. (2007). [Network Science](#). In *Annual Review of Information Science & Technology*, vol. 41, ed. Blaise Cronin, 537-607. Medford, NJ: Information Today, Inc./American Society for Information Science and Technology.
- ❑ Herr II, Bruce W., Weixia (Bonnie) Huang, Shashikant Penumarthy, and Katy Börner. (2007). [Designing Highly Flexible and Usable Cyberinfrastructures for Convergence](#). In *Progress in Convergence – Technologies for Human Wellbeing*, vol. 1093, eds. William S. Bainbridge and Mihail C. Roco, 161-179. Boston: Annals of the New York Academy of Sciences.

- ❑ Colizza, V., A. Barrat, M. Barthelemy, and A. Vespignani. (2007). [Epidemic modeling in complex realities](#). *Comptes Rendus Biologie* 330: 364-374. Elsevier.
- ❑ Colizza, Vittoria, Romualdo Pastor-Satorras, and Alessandro Vespignani. (2007). [Reaction-diffusion processes and metapopulation models in heterogeneous networks](#). *Nature Physics* 3: 276-282. Nature Publishing Group.
- ❑ Vermeirssen, Vanessa, M. Inmaculada Barrasa, César A. Hidalgo, Jenny Aurelle B. Babon, Reynaldo Sequerra, Lynn Doucette-Stamm, Albert-László Barabási, and Albertha J. M. Walhout. (2007). [Transcription factor modularity in a gene-centered C. elegans core neuronal protein-DNA interaction network](#). *Network Genome Research*. Cold Spring Harbor Laboratory Press.
- ❑ Börner, Katy, Elisha F. Hardy, Bruce W. Herr II, Todd Holloway, and W. Bradford Paley. (2007). [Taxonomy Visualization in Support of the Semi-Automatic Validation and Optimization of Organizational Schemas](#). *Journal of Informetrics* 1 (3): 214-225. Elsevier.
- ❑ More papers at <http://nwb.slis.indiana.edu/papers.html>

Websites

- ❑ <http://nwb.slis.indiana.edu>
- ❑ <https://nwb.slis.indiana.edu/community>
- ❑ <http://cishell.org>
- ❑ <http://cns-trac.slis.indiana.edu/trac/nwb/>

NSF IIS-0513650 award

Thank You