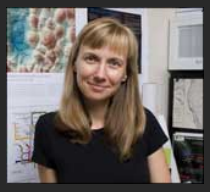
Network Workbench:
A CI-Marketplace for Network Scientists



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

"Flagship Projects in E-Social Science" Panel
Third International Conference on e-Social Science, Ann Arbor, MI
October 9th, 2007

Network Workbench (NWB)

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

Software Team: Lead: Weixia (Bonnie) Huang

Developers: Bruce Herr, Ben Markines, Santo Fortunato, Cesar Hidalgo, Ramya Sabbineni, Vivek S. Thakre, & Russell Duhon

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

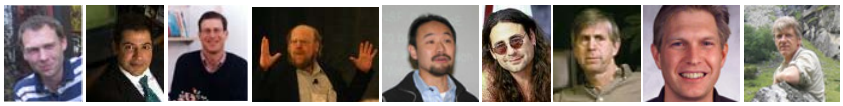
Amount: $1,120,926 NSF IIS-0513650 award.

Duration: Sept. 2005 - Aug. 2008

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

NWB Advisory Board

- Ulrik Brandes, University of Konstanz, Germany (Graph Theory)
- Noshier Contractor, Northwestern University (Communication Theory)
- Mark Gerstein, Yale University (Bioinformatics)
- James Hendler, Rensselaer Polytechnic Institute (Semantic Web)
- Jason Leigh, Electronic Visualization Laboratory, University of Illinois at Chicago (Visualization & CI)
- Neo Martinez, Pacific Ecoinformatics and Computational Ecology Lab (Biology)
- Michael Macy, Cornell University (Sociology)
- Stephen North, AT&T (Graph Visualization)
- Tom Snijders, University of Groningen (Social Network Analysis)



NWB CI Deliverables

Cyberglue:

- CIShell — Core programmer team lead by Bonnie Huang

Tools & Services:

- NWB Tool — Lead by Alex Vespignani with input from other PIs
- SciMaps Service — Lead by Katy Borner
- Bio Tool — Lead by Laszlo Barabasi & Santiago Schnell

All three are prototypical instantiations of CIShell serving as reference implementations.

Documentation/Registry/Market Place:

- NWB Community Wiki — Lead by Katy Borner

Embracing the Diversity of Network Science

See also

Börner, Katy, Sanyal, Soma & Vespignani, Alessandro. (2007). Network Science. In Cronin, Blaise (Eds.), Annual Review of Information Science & Technology (Vol. 41, pp. 537-607), chapter 12, Medford, NJ: Information Today, Inc./American Society for Information Science and Technology.

Computational Social Science
Studying large scale social networks such as Wikipedia

Vizzards 2007 Entr

Second Sight: An En
Mosaic of Wikipedia
The NewScientist, M

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, and those most hotly contested.

task. About
have made
and the English-
contains
ry second a new
day 2000 new

ll, Bruce Herr
ndiana
, created
icles that touch
as a religion or a
a cluster they
e most popular
in a circular grid.
oured dots
how recently
dited. The larger,
intense activity.
s reveals the
Wikipedians:
ndo, Hurricane
d Albert Einstein.
in real time:
s administrators
ts are taking
nival contributors
g each other's
page could be
ools (locked
ting include
entries on Sheffield Wednesday football club, Mikhail Gorbachev and pigs).

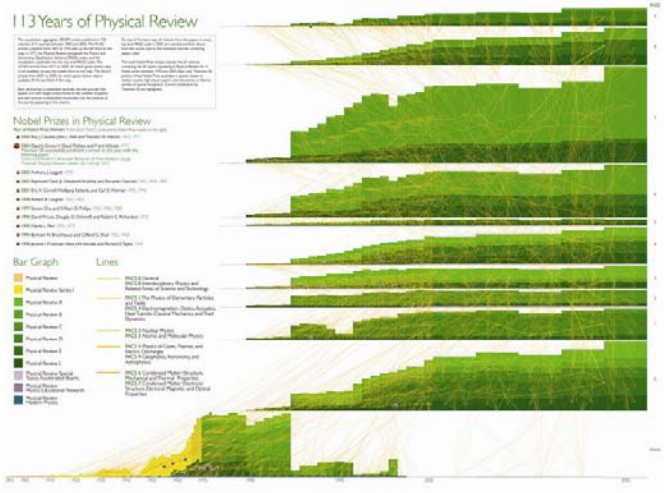
The mosaic has been commended in a competition for images that visualise network dynamics, coinciding with this week's International Workshop and Conference on Network Science in Bloomington.

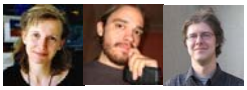
www.newscientist.com

19 May 2007 | NewScientist | 55

113 Years of Physical Review

Bruce W. Herr II and Russell Duhon (Data Mining & Visualization), Elisha F. Hardy (Graphic Design), Shashikant Penumarthy (Data Preparation) and Katy Börner (Concept)

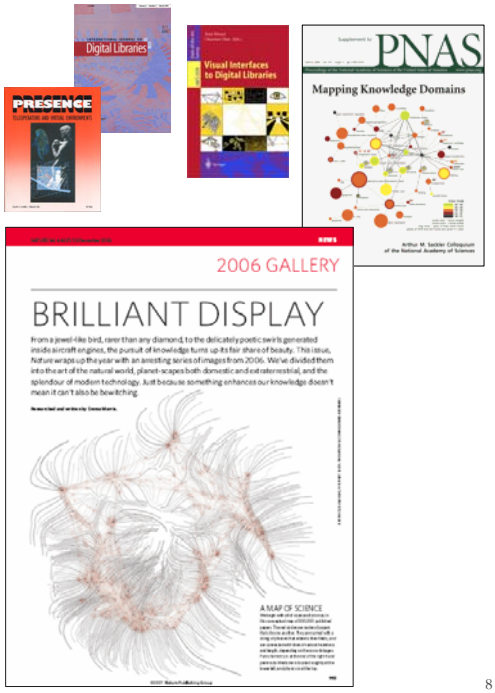




Computational Scientometrics
Studying science by scientific means

Börner, Katy, Chen, Chaomei, and Boyack, Kevin. (2003). **Visualizing Knowledge Domains.** In Blaise Cronin (Ed.), Annual Review of Information Science & Technology, Volume 37, Medford, NJ: Information Today, Inc./American Society for Information Science and Technology, chapter 5, pp. 179-255.
Shiffrin, Richard M. and Börner, Katy (Eds.) (2004). **Mapping Knowledge Domains.** Proceedings of the National Academy of Sciences of the United States of America, 101(Suppl_1).

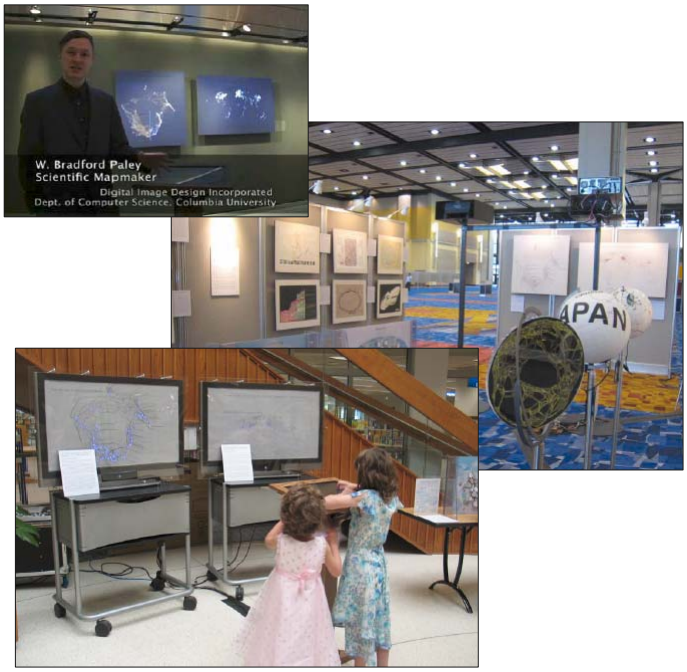
Places & Spaces: Mapping Science exhibit, Currently on display at the American Museum for Science and Energy, Oak Ridge, TN, see also http://scimaps.org.



8

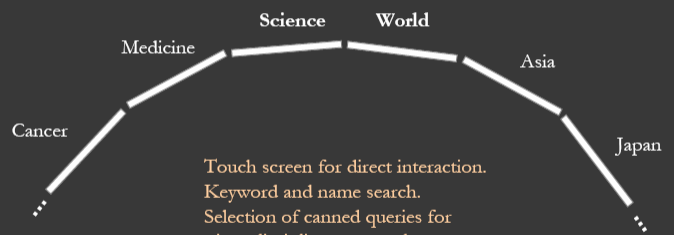
Illuminated Diagram Display

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

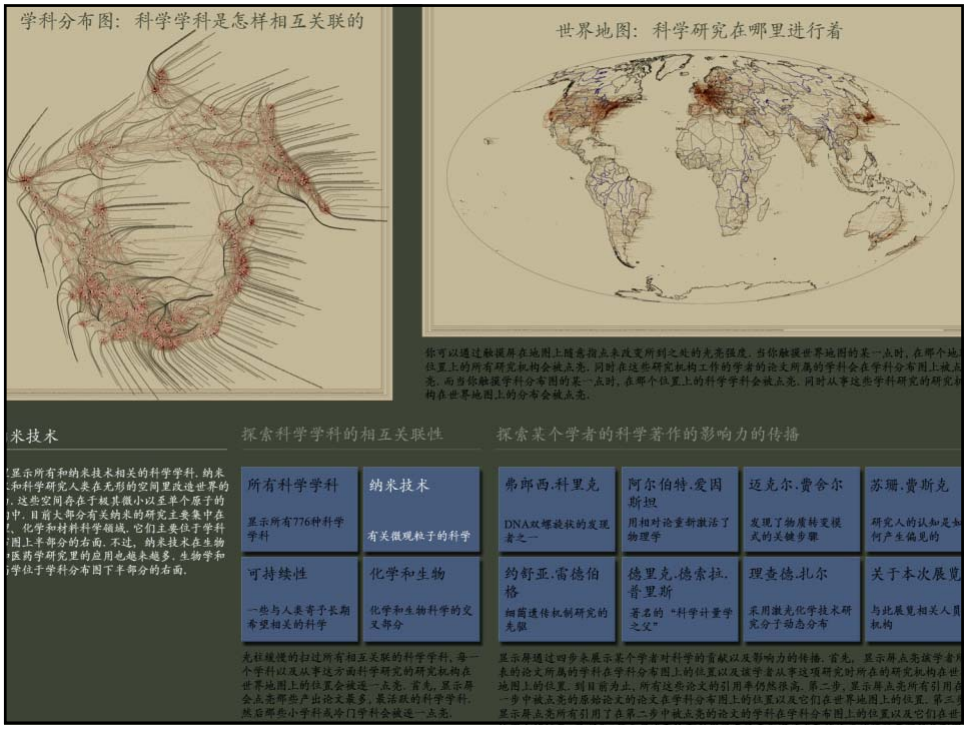


Re-implementation of Illuminated Diagram Software (in progress)
by Advanced Visualization Lab, Indiana University

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.



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).

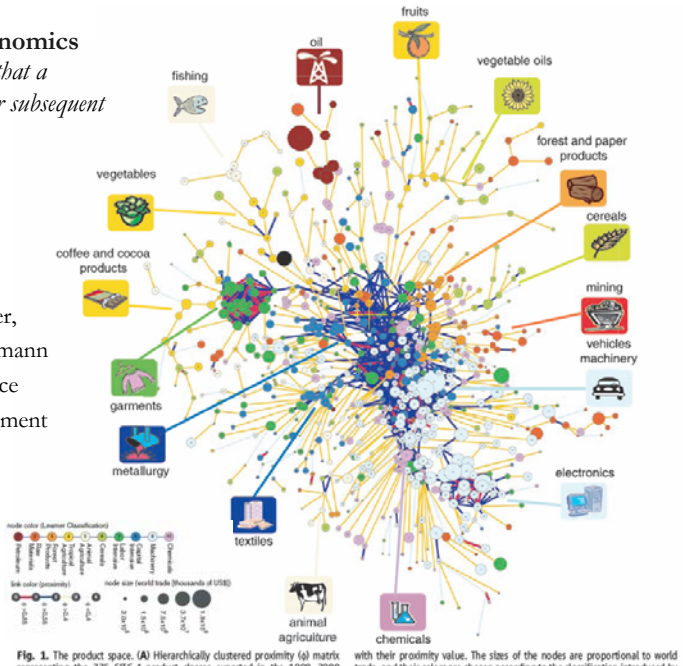
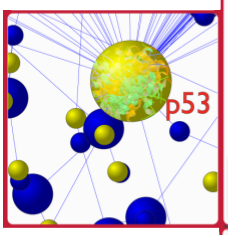
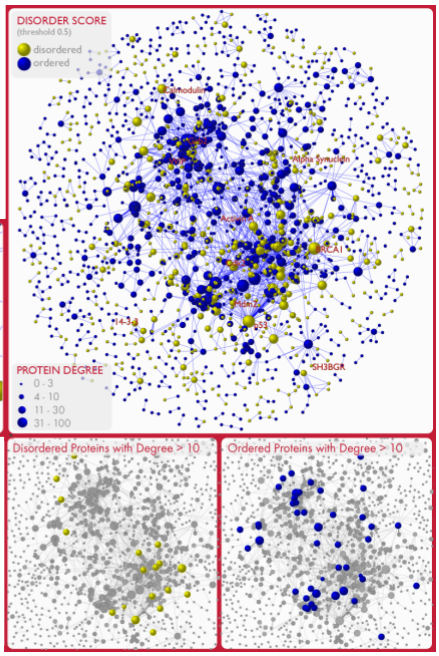


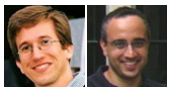
Fig. 1. The product space. (A) Hierarchically clustered proximity (φ) matrix representing the 775 SITC-4 product classes exported in the 1998–2000 period. (B) Network representation of the product space. Links are color coded with their proximity value. The sizes of the nodes are proportional to world trade, and their colors are chosen according to the classification introduced by Leamer.

Computational 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.







13

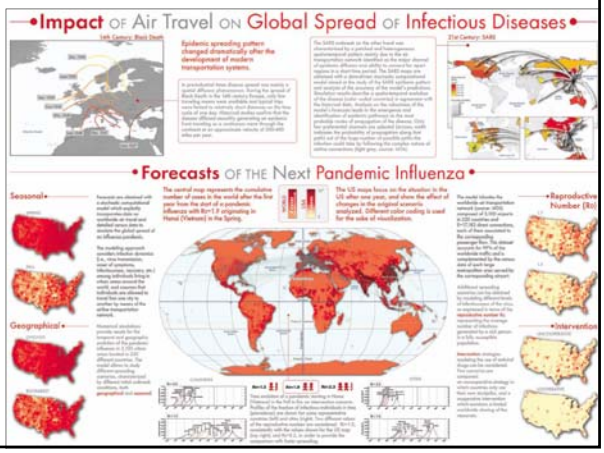
Computational Epidemics

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).



Designing Cyberglue

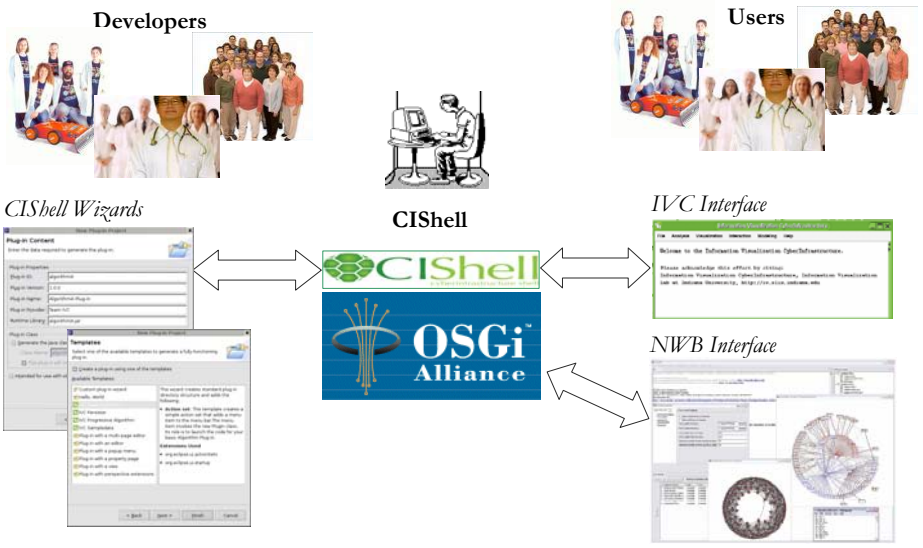
See also

Herr, Bruce W., Huang, Weixia, Penumarthy, Shashikant, Börner, Katy . (2007) Designing Highly Flexible and Usable Cyberinfrastructures for Convergence. In William S. Bainbridge and Mihail C. Roco (Eds.) Progress in Convergence – Technologies for Human Wellbeing. Annals of the New York Academy of Sciences, Boston, MA, volume 1093, pp. 161-179.

CIShell Building Market Places not Cathedrals

- 'Software glue' has to interlink datasets and algorithms written in different languages using different data formats.
- The smaller the glue or 'CI Shell', the more likely it can be maintained.

CIShell – Serving Non-CS Algorithm Developers & Users



CIShell – Build on OSGi Industry Standard

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 7 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.

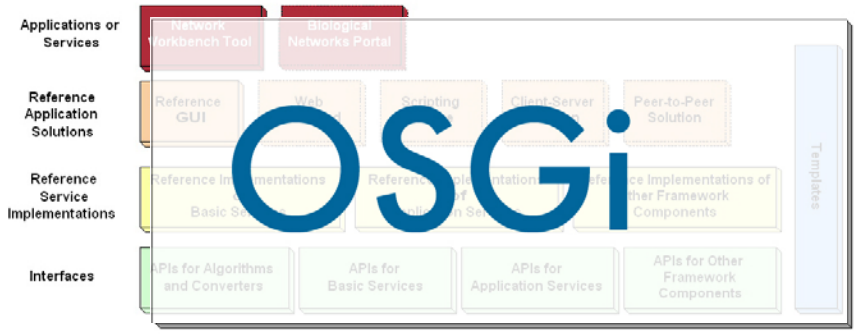
Advantages of Using OSGi

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

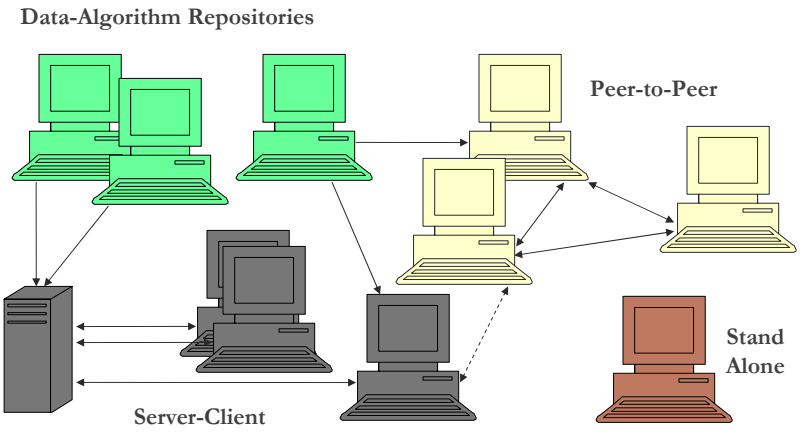
Ideally, CIShell becomes a standard for creating OSGi Services for algorithms.

CIShell – Technical Details

CIShell layer cake.

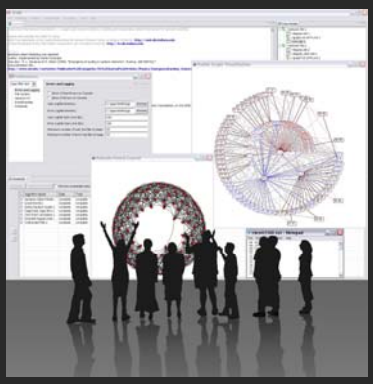


CIShell – Deployment



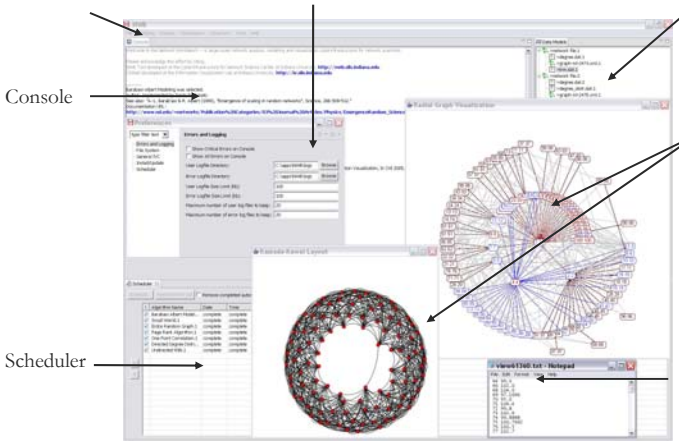
CIShell applications can be deployed as distributed data and algorithm repositories, stand alone applications, peer-to-peer architectures, and server-client architectures.

The NWB Tool



NWB Tool: Interface Elements

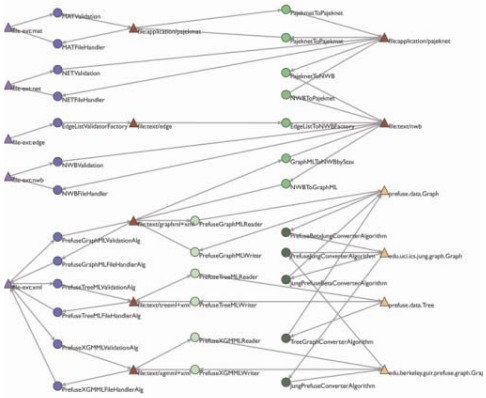
http://nwb.slis.indiana.edu



NWB Ecology of Data Formats and Converters

*Not shown are **15** sample datasets, **45** data preprocessing, analysis, modeling and visualization algorithms, **9** services.*

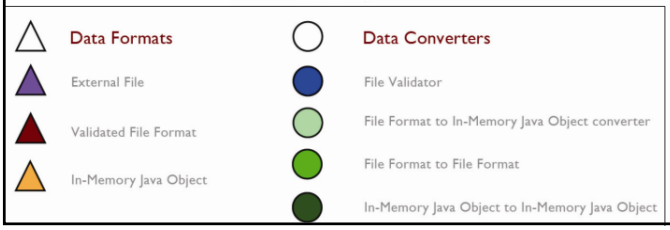
5
Supported data formats



5
Output formats for diverse visualization algorithms

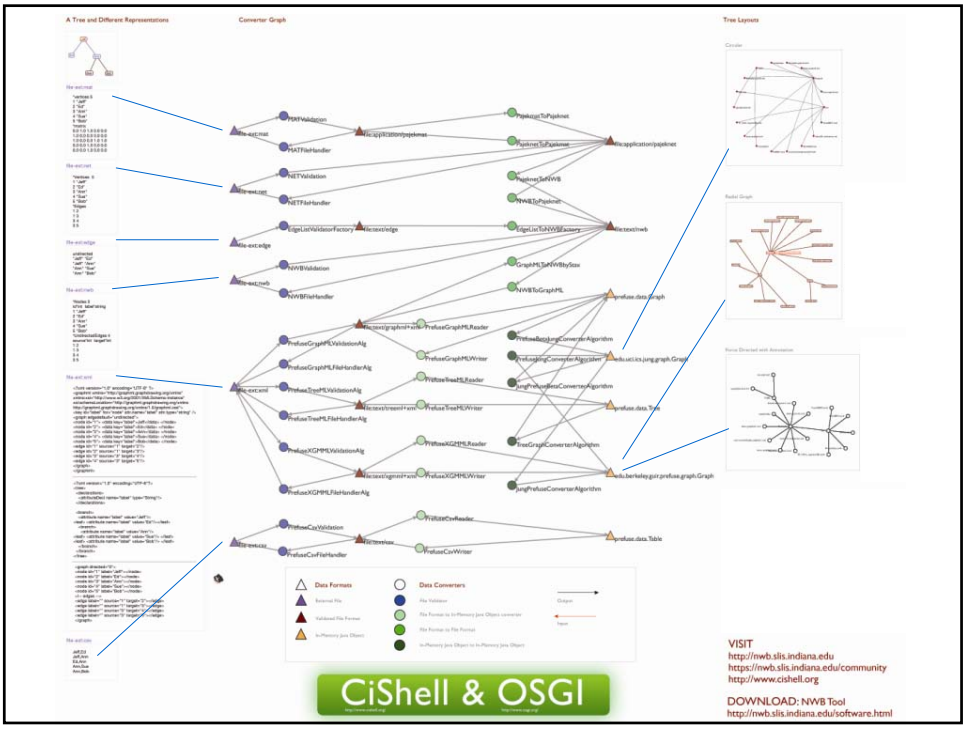
8
Intermediate data formats

Supported by
35
data converters.



VISIT
http://nwb.slis.indiana.edu
https://nwb.slis.indiana.edu/community
http://www.cishell.org

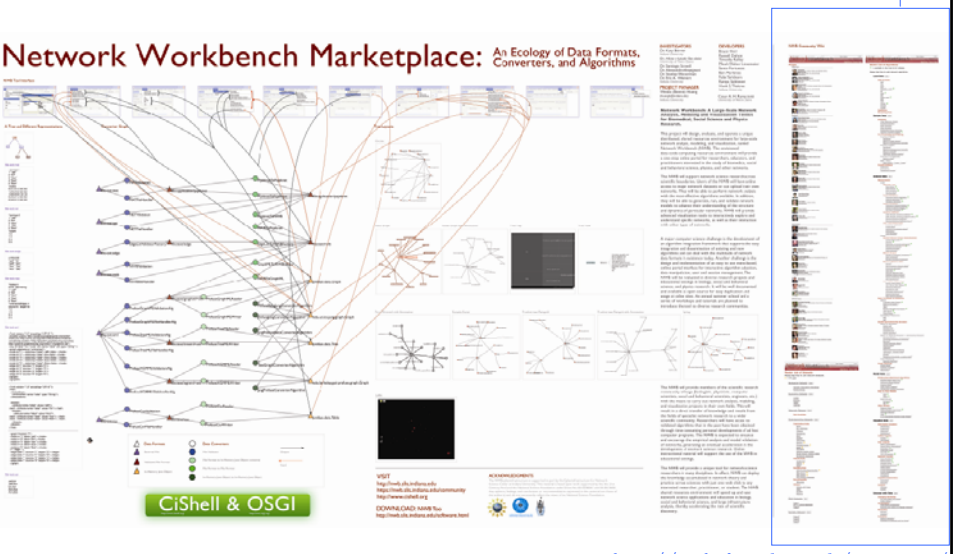
DOWNLOAD: NWB Tool
http://nwb.slis.indiana.edu/software.html



VISIT
http://nwb.slis.indiana.edu
https://nwb.slis.indiana.edu/community
http://www.cishell.org

DOWNLOAD: NWB Tool
http://nwb.slis.indiana.edu/software.html

NWB
Community Wiki



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

Education – Learning Modules, NWB User and Developer Workshops

Huang with students at the Complex System Summer School in Beijing, China

The End.