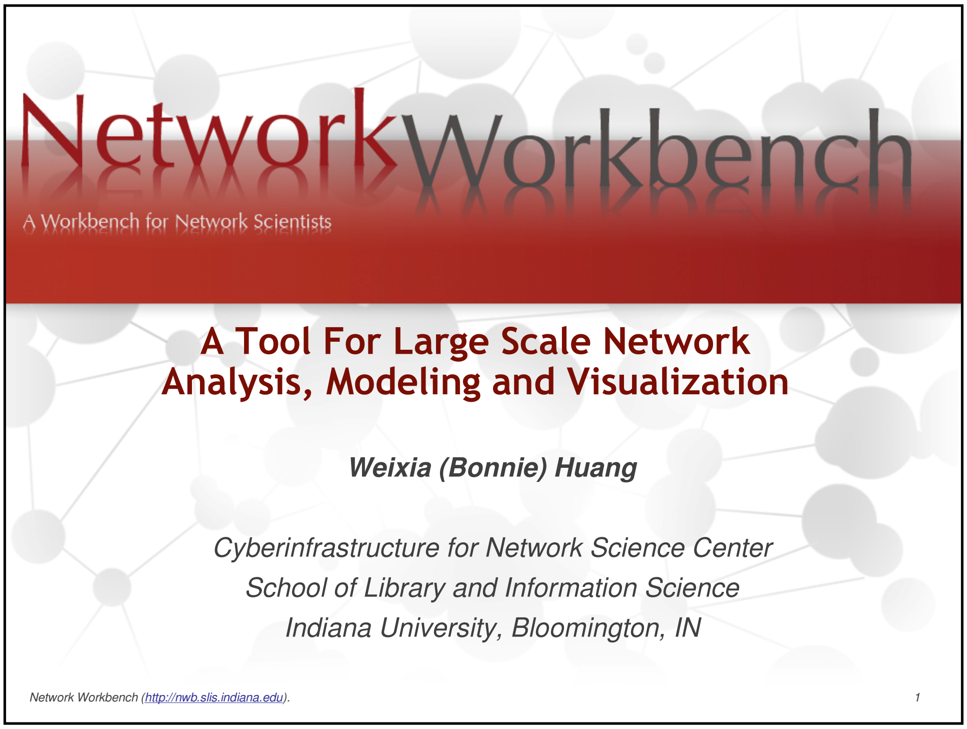


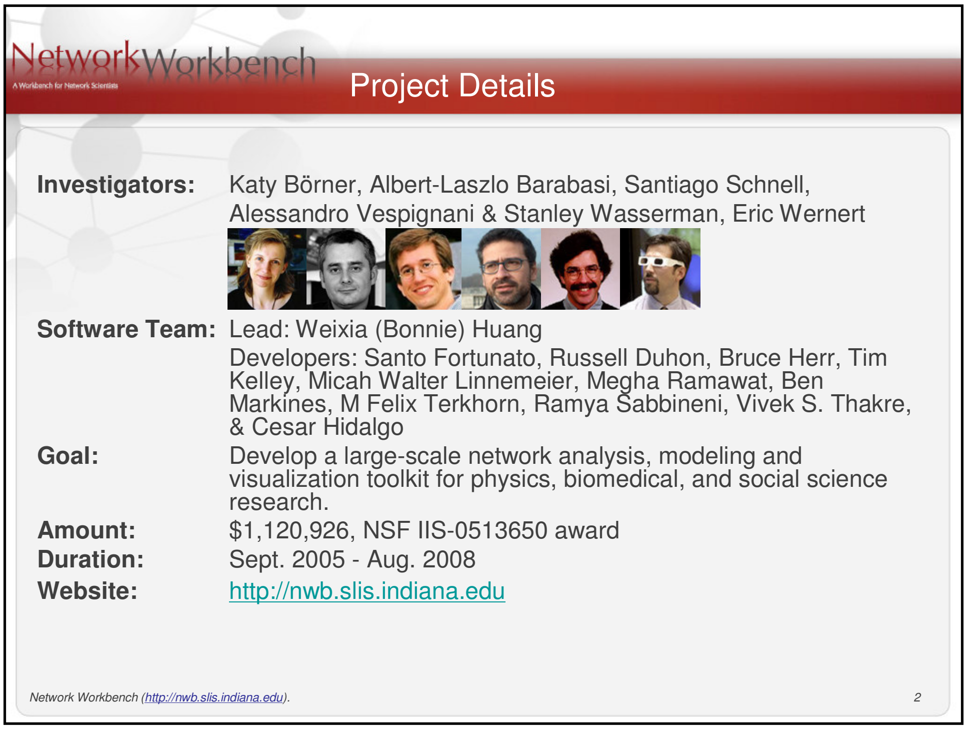
Network Workbench

A Workbench for Network Scientists

A Tool For Large Scale Network Analysis, Modeling and Visualization

Weixia (Bonnie) Huang

Cyberinfrastructure for Network Science Center
School of Library and Information Science
Indiana University, Bloomington, IN

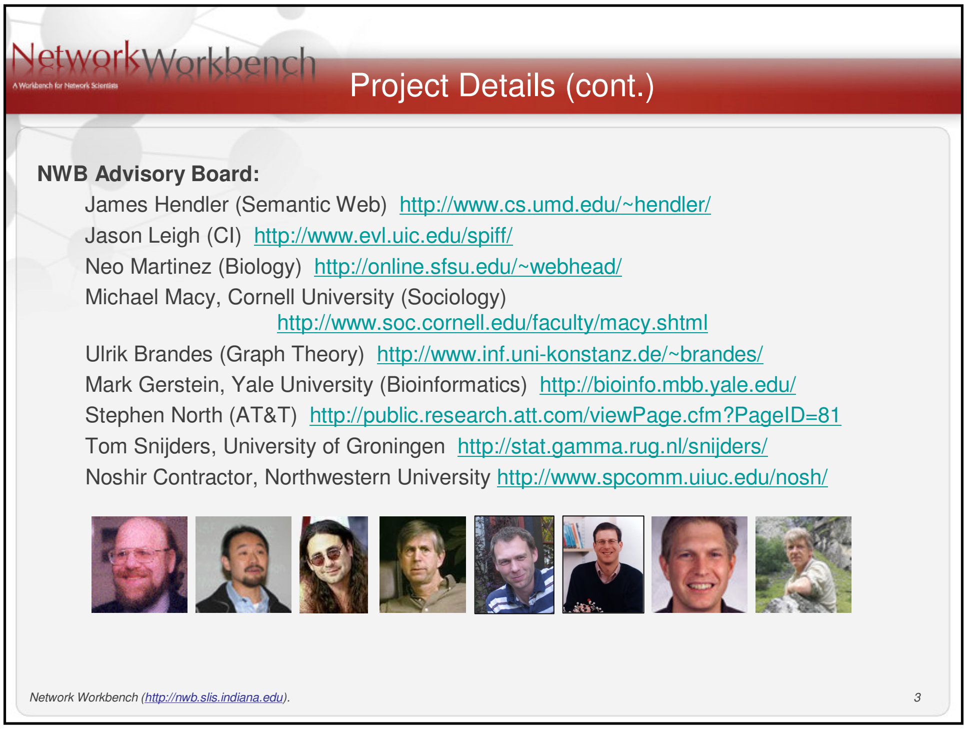


Network Workbench

A Workbench for Network Scientists

Project Details

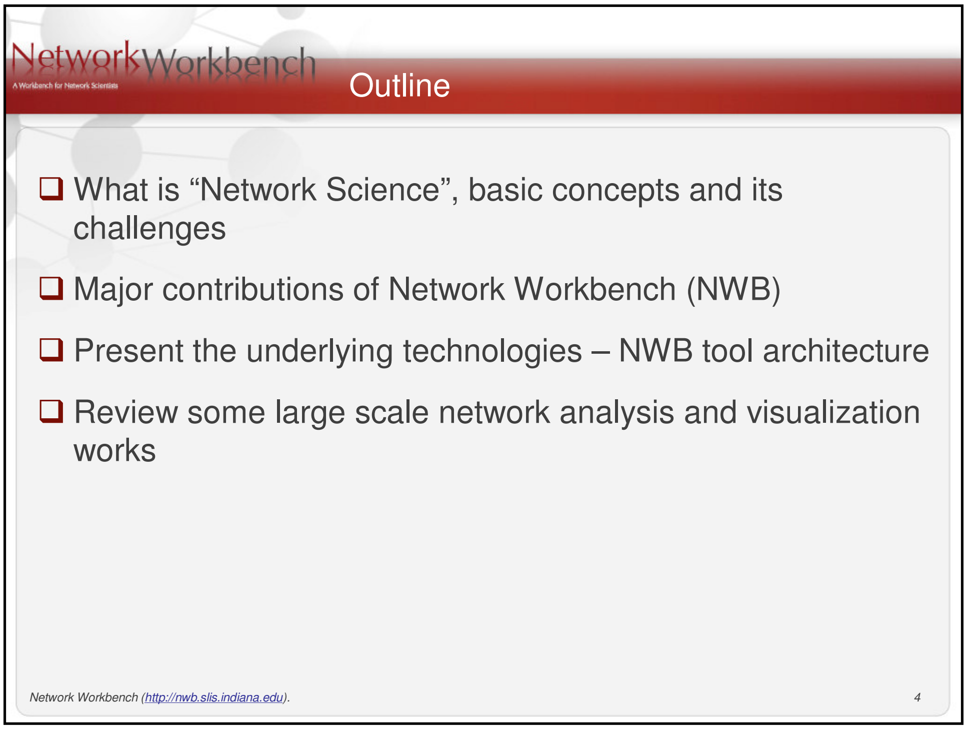
Investigators:	Katy Börner, Albert-Laszlo Barabasi, Santiago Schnell, Alessandro Vespignani & Stanley Wasserman, Eric Wernert
Software Team:	Lead: Weixia (Bonnie) Huang Developers: Santo Fortunato, Russell Duhon, Bruce Herr, Tim Kelley, Micah Walter Linnemeier, Megha Ramawat, Ben Markines, M Felix Terkhorn, 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. 2008
Website:	http://nwb.slis.indiana.edu



Project Details (cont.)

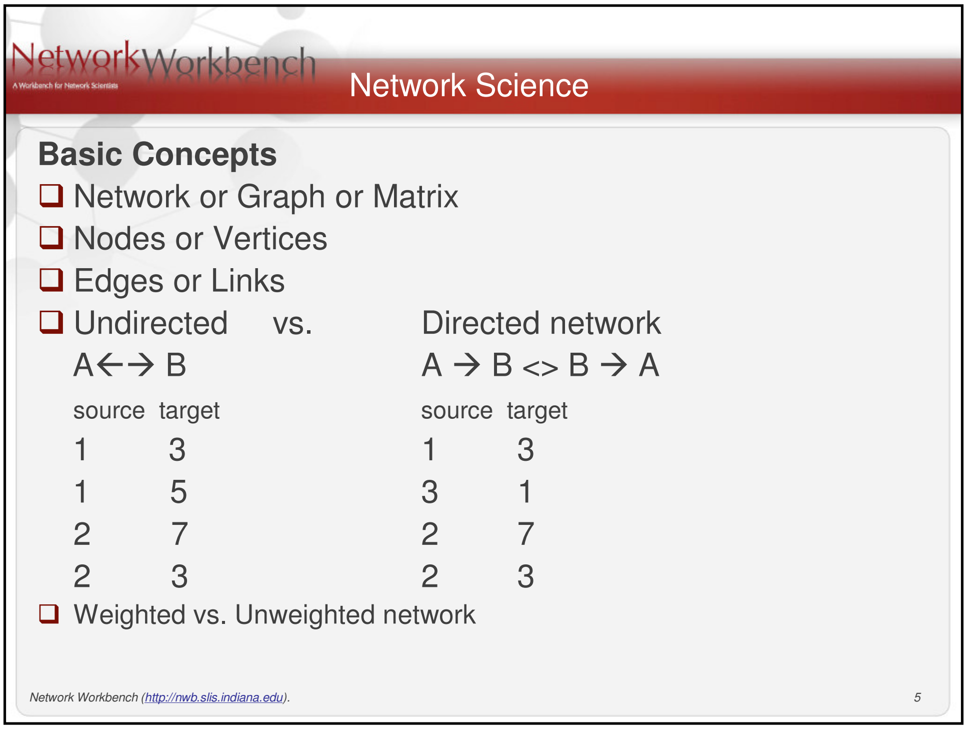
NWB Advisory Board:

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/



Outline

- ❑ What is "Network Science", basic concepts and its challenges
- ❑ Major contributions of Network Workbench (NWB)
- ❑ Present the underlying technologies – NWB tool architecture
- ❑ Review some large scale network analysis and visualization works



Network Science

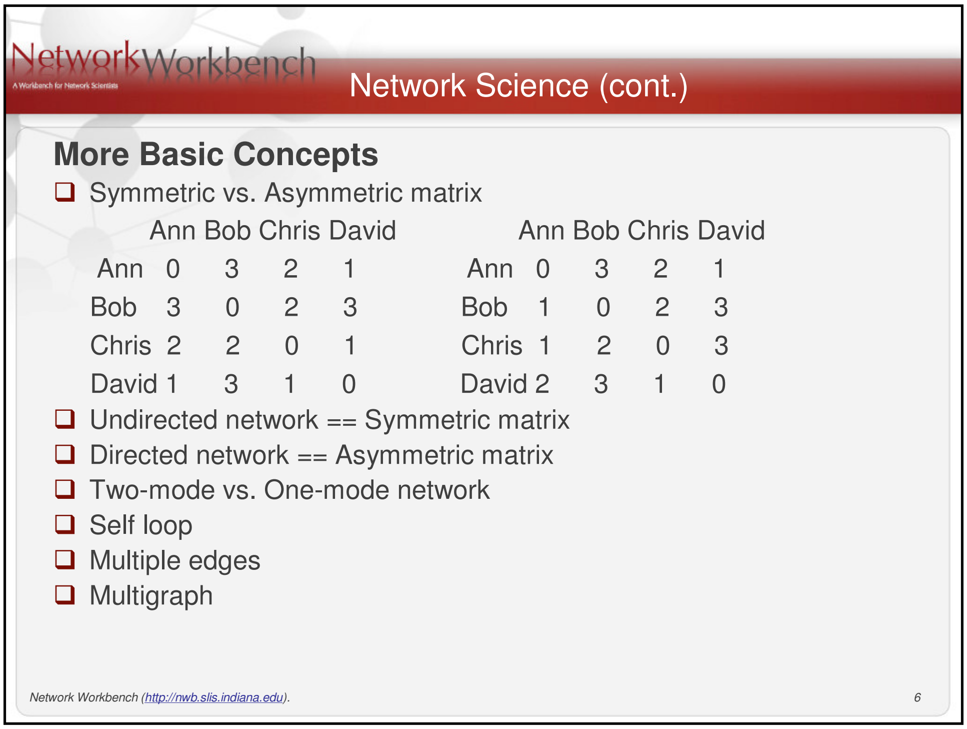
Basic Concepts

- ❑ Network or Graph or Matrix
- ❑ Nodes or Vertices
- ❑ Edges or Links
- ❑ Undirected vs. Directed network

A←→ B   A → B <> B → A

source	target	source	target
1	3	1	3
1	5	3	1
2	7	2	7
2	3	2	3

- ❑ Weighted vs. Unweighted network



Network Science (cont.)

More Basic Concepts

- ❑ Symmetric vs. Asymmetric matrix

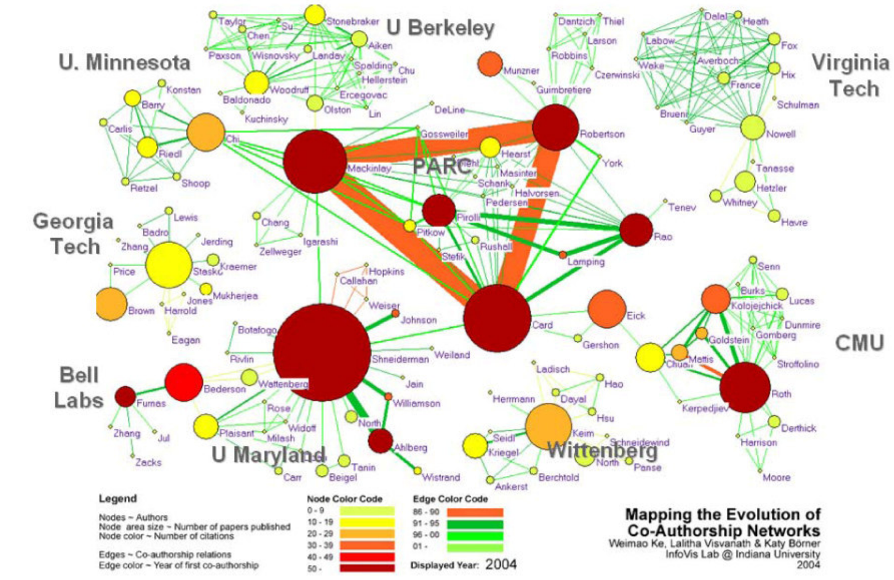
	Ann	Bob	Chris	David
Ann	0	3	2	1
Bob	3	0	2	3
Chris	2	2	0	1
David	1	3	1	0

	Ann	Bob	Chris	David
Ann	0	3	2	1
Bob	1	0	2	3
Chris	1	2	0	3
David	2	3	1	0

- ❑ Undirected network == Symmetric matrix
- ❑ Directed network == Asymmetric matrix
- ❑ Two-mode vs. One-mode network
- ❑ Self loop
- ❑ Multiple edges
- ❑ Multigraph

Mapping the Evolution of Co-Authorship Networks

Ke, Visvanath & Börner, (2004) Won 1st price at the IEEE InfoVis Contest.



Legend

Nodes ~ Authors
Node area size ~ Number of papers published
Node color ~ Number of citations

Edges ~ Co-authorship relations
Edge color ~ Year of first co-authorship

Node Color Code	
0 - 9	🟩
10 - 19	🟨
20 - 29	🟧
30 - 39	🟧
40 - 49	🟥
50 -	🟥

Edge Color Code	
86 - 90	🟧
91 - 95	🟩
96 - 00	🟩
01 -	🟩

Displayed Year: 2004

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

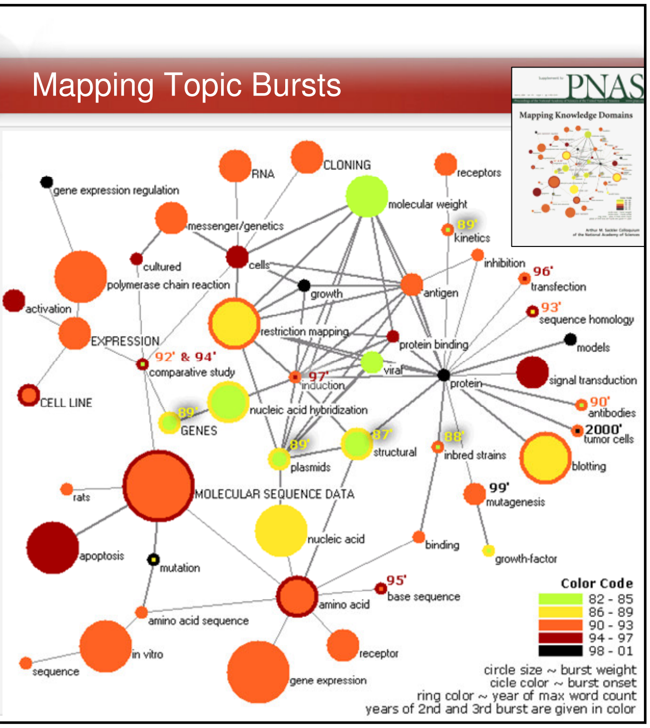
31

Mapping Topic Bursts

Network Workbench
A Workbench for Network Scientists

Co-word space of the top 50 highly frequent and bursty words used in the top 10% most highly cited PNAS publications in 1982-2001.

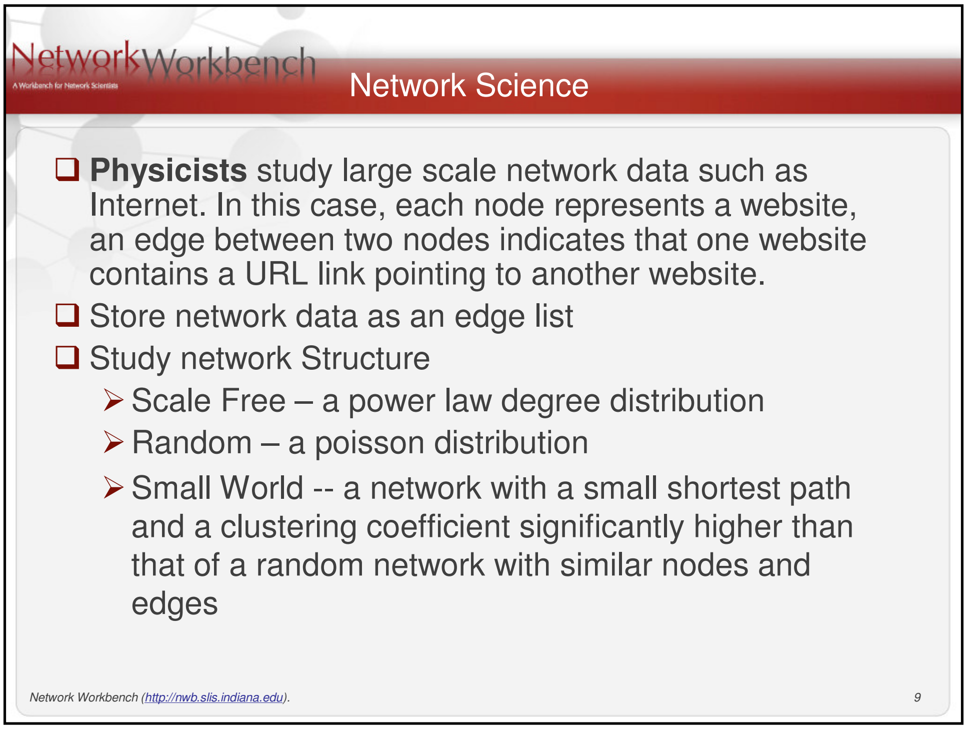
Mane & Börner. (2004) PNAS, 101(Suppl. 1): 5287-5290.



Color Code	
🟩	82 - 85
🟨	86 - 89
🟧	90 - 93
🟥	94 - 97
⬛	98 - 01

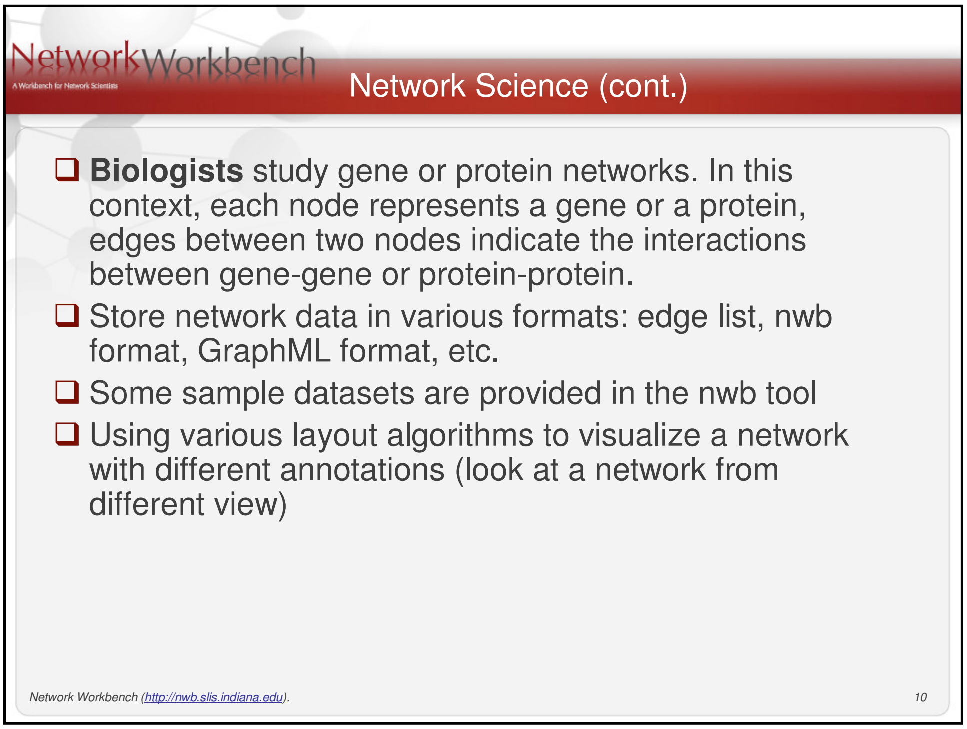
circle size ~ burst weight
cicle color ~ burst onset
ring color ~ year of max word count
years of 2nd and 3rd burst are given in color

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



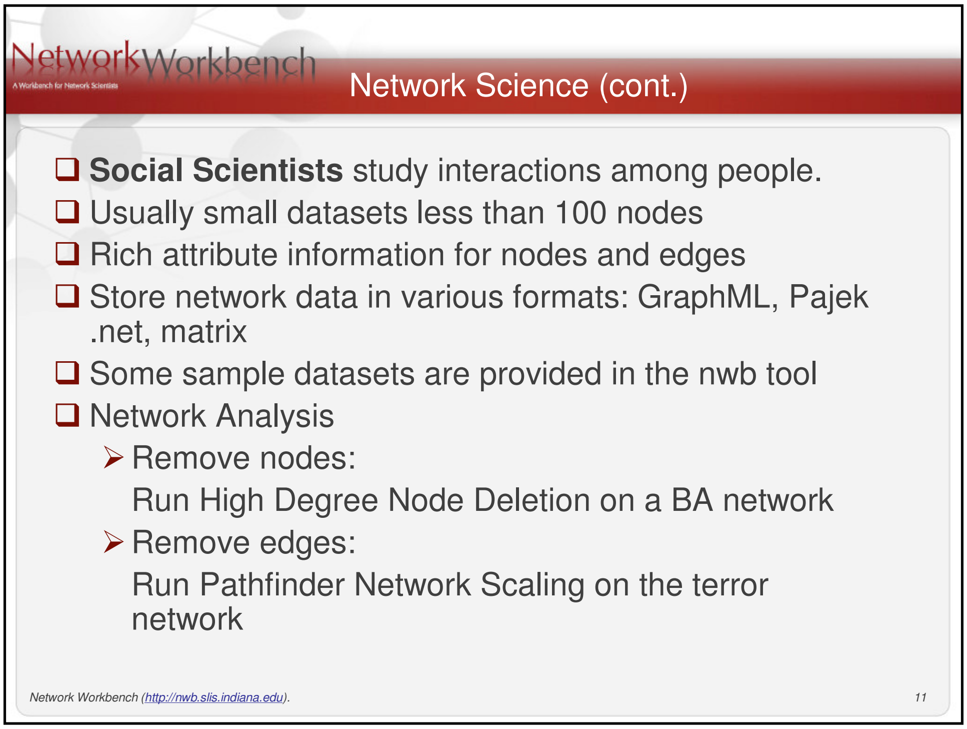
Network Science

- ❑ **Physicists** study large scale network data such as Internet. In this case, each node represents a website, an edge between two nodes indicates that one website contains a URL link pointing to another website.
- ❑ Store network data as an edge list
- ❑ Study network Structure
  - ➢ Scale Free – a power law degree distribution
  - ➢ Random – a poisson distribution
  - ➢ Small World -- a network with a small shortest path and a clustering coefficient significantly higher than that of a random network with similar nodes and edges



Network Science (cont.)

- ❑ **Biologists** study gene or protein networks. In this context, each node represents a gene or a protein, edges between two nodes indicate the interactions between gene-gene or protein-protein.
- ❑ Store network data in various formats: edge list, nwb format, GraphML format, etc.
- ❑ Some sample datasets are provided in the nwb tool
- ❑ Using various layout algorithms to visualize a network with different annotations (look at a network from different view)

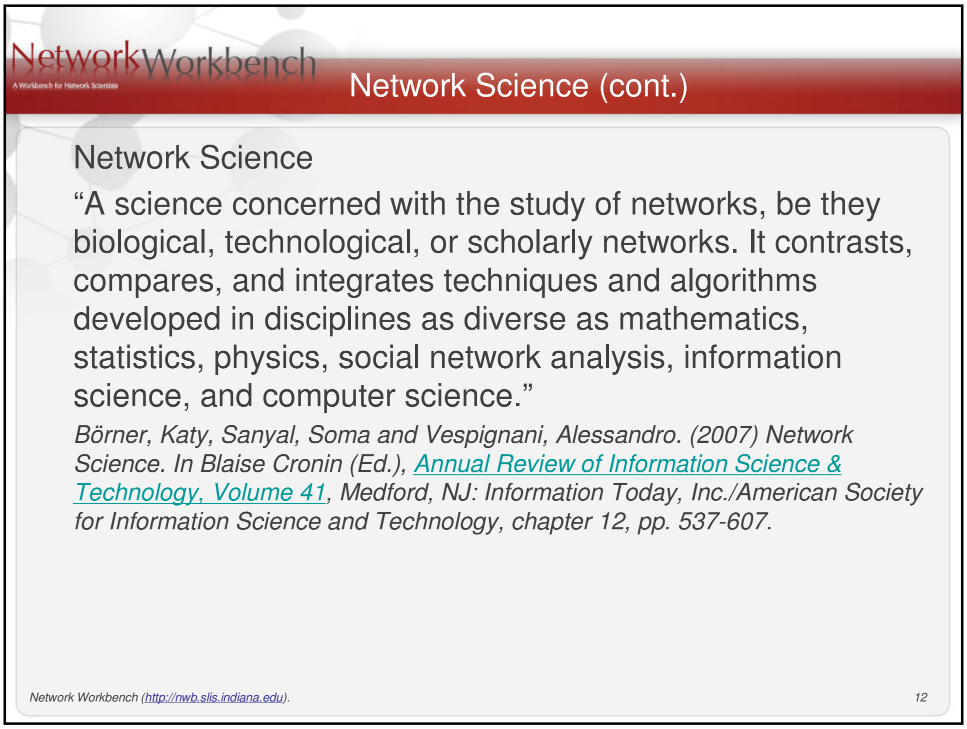


Network Science (cont.)

- ❑ **Social Scientists** study interactions among people.
- ❑ Usually small datasets less than 100 nodes
- ❑ Rich attribute information for nodes and edges
- ❑ Store network data in various formats: GraphML, Pajek .net, matrix
- ❑ Some sample datasets are provided in the nwb tool
- ❑ Network Analysis
  - ➢ Remove nodes:

    Run High Degree Node Deletion on a BA network
  - ➢ Remove edges:

    Run Pathfinder Network Scaling on the terror network

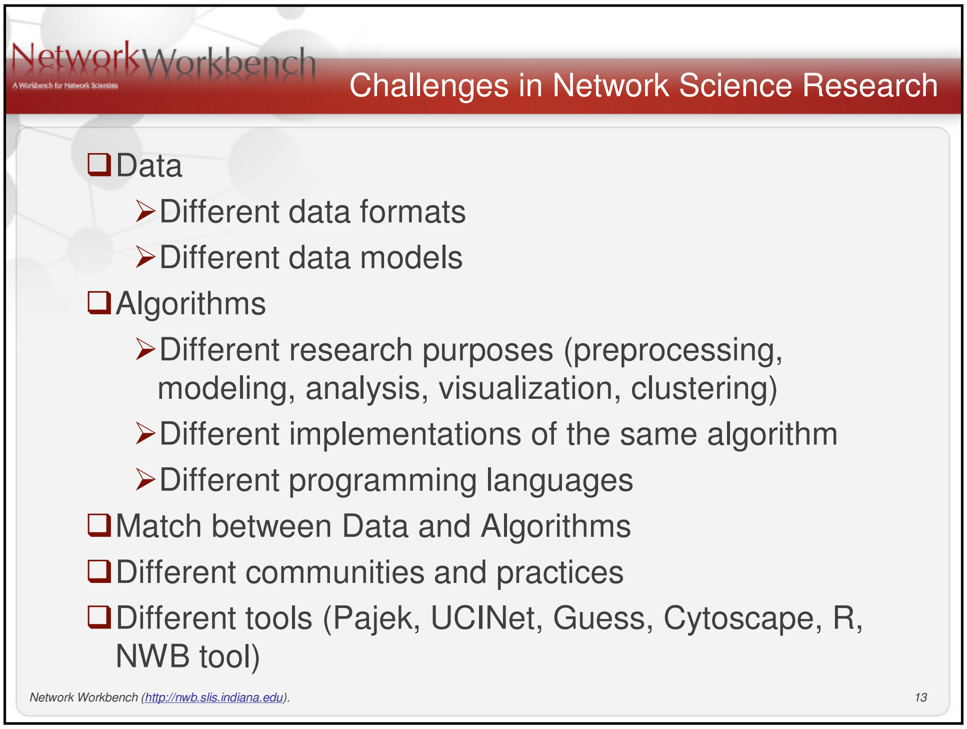


Network Science (cont.)

Network Science

"A science concerned with the study of networks, be they biological, technological, or scholarly networks. It contrasts, compares, and integrates techniques and algorithms developed in disciplines as diverse as mathematics, statistics, physics, social network analysis, information science, and computer science."

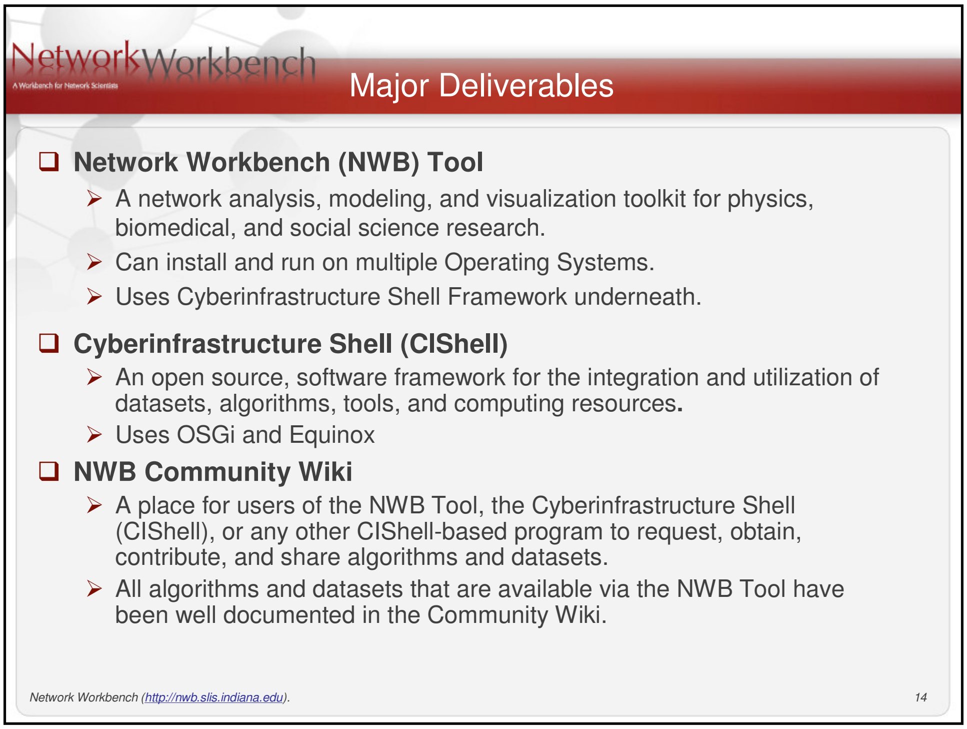
Börner, Katy, Sanyal, Soma and Vespignani, Alessandro. (2007) Network Science. In Blaise Cronin (Ed.), Annual Review of Information Science & Technology, Volume 41, Medford, NJ: Information Today, Inc./American Society for Information Science and Technology, chapter 12, pp. 537-607.



Challenges in Network Science Research

- ❑ 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 (http://nwb.slis.indiana.edu). 13



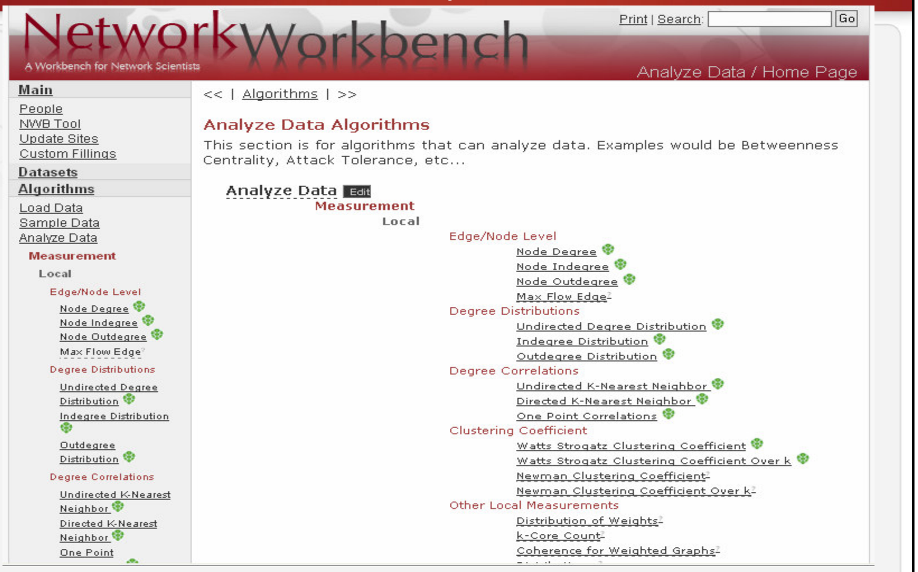
Major Deliverables

- ❑ **Network Workbench (NWB) Tool**
  - ➢ A network analysis, modeling, and visualization toolkit for physics, biomedical, and social science research.
  - ➢ Can 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**.**
  - ➢ Uses OSGi and Equinox
- ❑ **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.

Network Workbench (http://nwb.slis.indiana.edu). 14

NetworkWorkbench
A Workbench for Network Scientists

Community Wiki



<< | Algorithms | >>

Analyze Data Algorithms

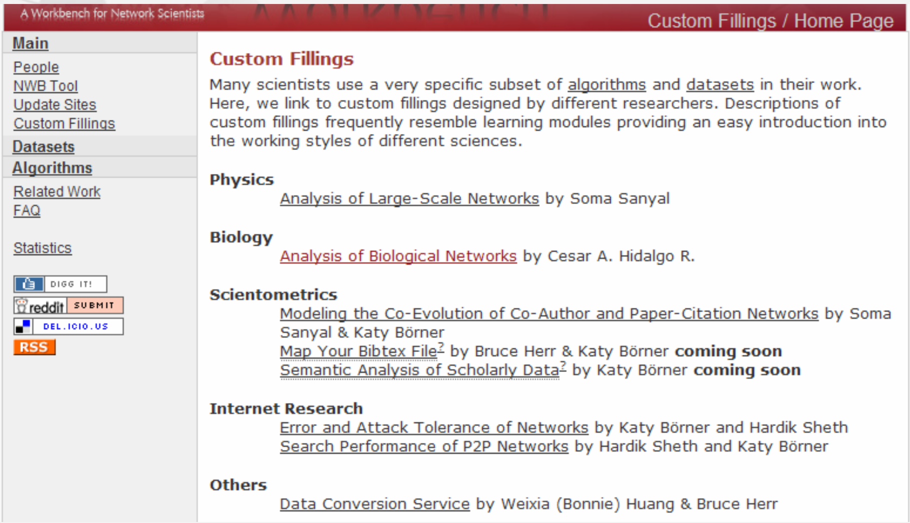
This section is for algorithms that can analyze data. Examples would be Betweenness Centrality, Attack Tolerance, etc...

Analyze Data Edit

- Measurement
  - Local
    - Edge/Node Level
      - Node Degree
      - Node Indegree
      - Node Outdegree
      - Max Flow Edge$^{?}$
    - Degree Distributions
      - Undirected Degree Distribution
      - Indegree Distribution
      - Outdegree Distribution
    - Degree Correlations
      - Undirected K-Nearest Neighbor
      - Directed K-Nearest Neighbor
      - One Point Correlations
    - Clustering Coefficient
      - Watts Strogatz Clustering Coefficient
      - Watts Strogatz Clustering Coefficient Over k
      - Newman Clustering Coefficient$^{?}$
      - Newman Clustering Coefficient Over k$^{?}$
    - Other Local Measurements
      - Distribution of Weights$^{?}$
      - k-Core Count$^{?}$
      - Coherence for Weighted Graphs$^{?}$

NetworkWorkbench
A Workbench for Network Scientists

Community Wiki (cont.)



Custom Fillings

Many scientists use a very specific subset of algorithms and datasets in their work. Here, we link to custom fillings designed by different researchers. Descriptions of custom fillings frequently resemble learning modules providing an easy introduction into the working styles of different sciences.

Physics

Analysis of Large-Scale Networks by Soma Sanyal

Biology

Analysis of Biological Networks by Cesar A. Hidalgo R.

Scientometrics

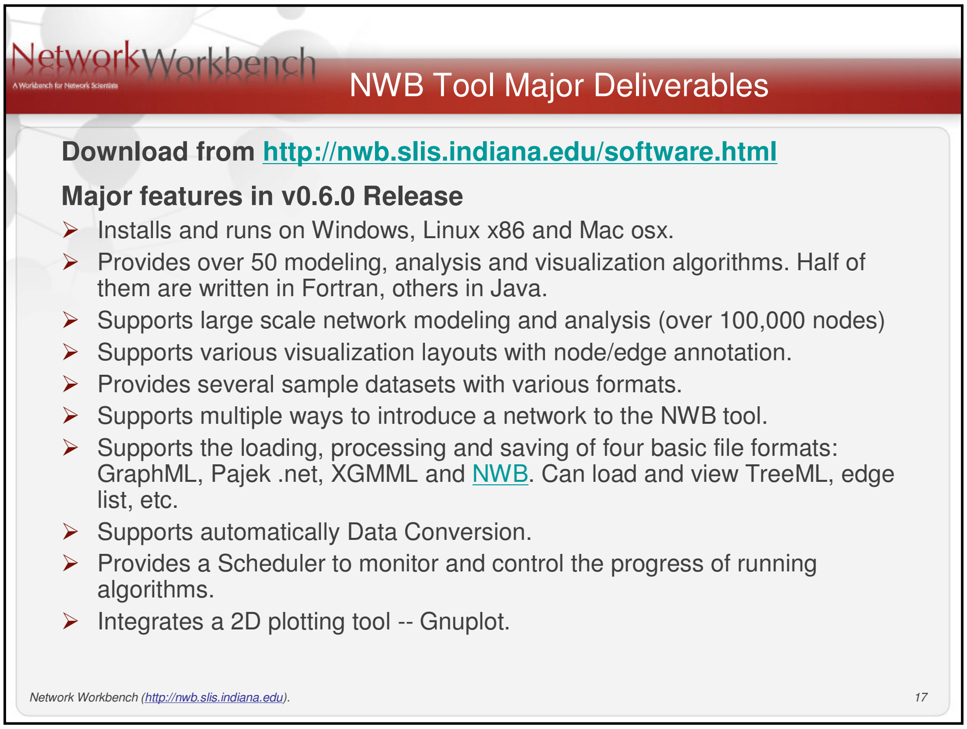
Modeling the Co-Evolution of Co-Author and Paper-Citation Networks by Soma Sanyal & Katy Börner
Map Your Bibtex File$^{?}$ by Bruce Herr & Katy Börner **coming soon**
Semantic Analysis of Scholarly Data$^{?}$ by Katy Börner **coming soon**

Internet Research

Error and Attack Tolerance of Networks by Katy Börner and Hardik Sheth
Search Performance of P2P Networks by Hardik Sheth and Katy Börner

Others

Data Conversion Service by Weixia (Bonnie) Huang & Bruce Herr



NetworkWorkbench

A Workbench for Network Scientists

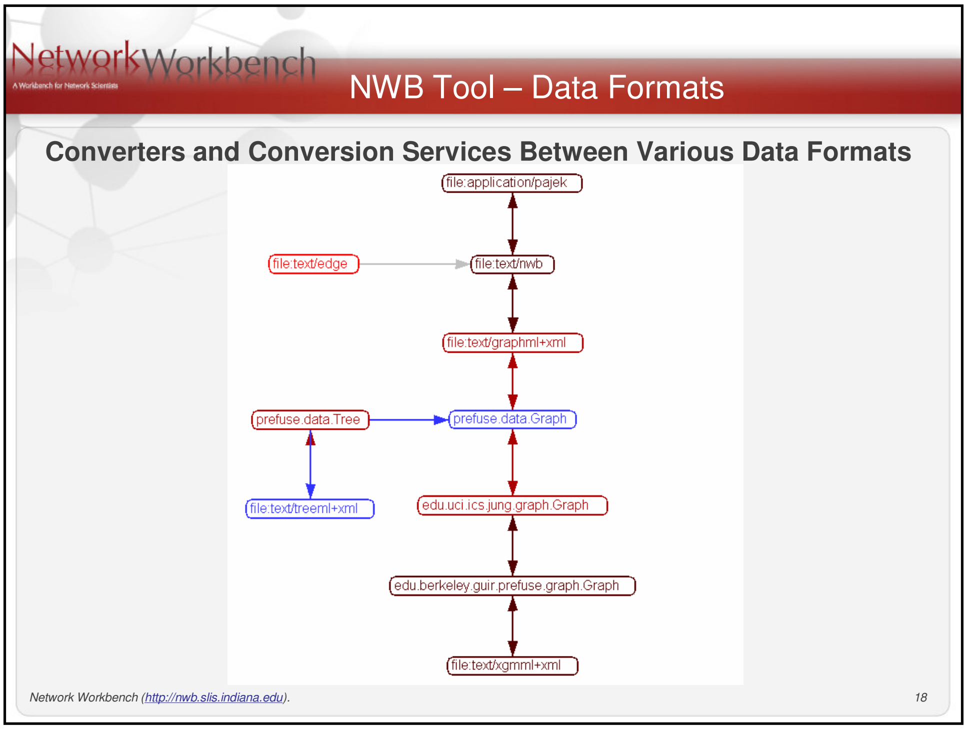
NWB Tool Major Deliverables

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

Major features in v0.6.0 Release

- Installs and runs on Windows, Linux x86 and Mac osx.
- Provides over 50 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 the loading, processing and saving of four basic file formats: GraphML, Pajek .net, XGMML and NWB. Can load and view TreeML, edge list, etc.
- Supports automatically Data Conversion.
- Provides a Scheduler to monitor and control the progress of running algorithms.
- Integrates a 2D plotting tool -- Gnuplot.

Network Workbench (http://nwb.slis.indiana.edu). 17



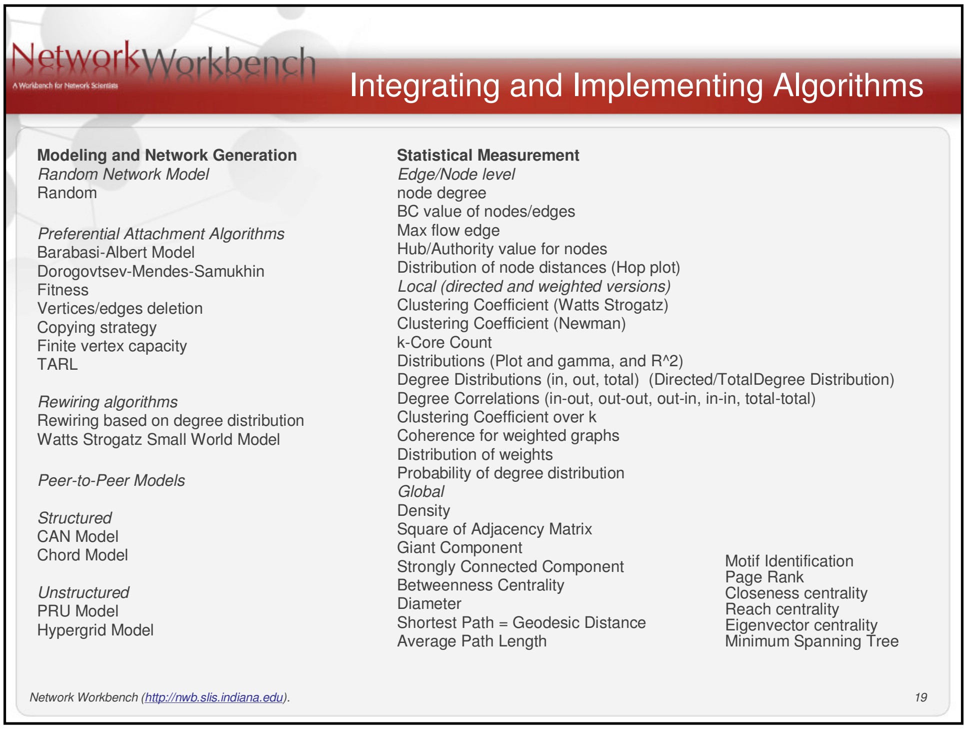
NetworkWorkbench

A Workbench for Network Scientists

NWB Tool – Data Formats

Converters and Conversion Services Between Various Data Formats

Network Workbench (http://nwb.slis.indiana.edu). 18


NetworkWorkbench
A Workbench for Network Scientists

Integrating and Implementing Algorithms

Modeling and Network Generation
Random Network Model
Random

Preferential Attachment Algorithms
Barabasi-Albert Model
Dorogovtsev-Mendes-Samukhin
Fitness
Vertices/edges deletion
Copying strategy
Finite vertex capacity
TARL

Rewiring algorithms
Rewiring based on degree distribution
Watts Strogatz Small World Model

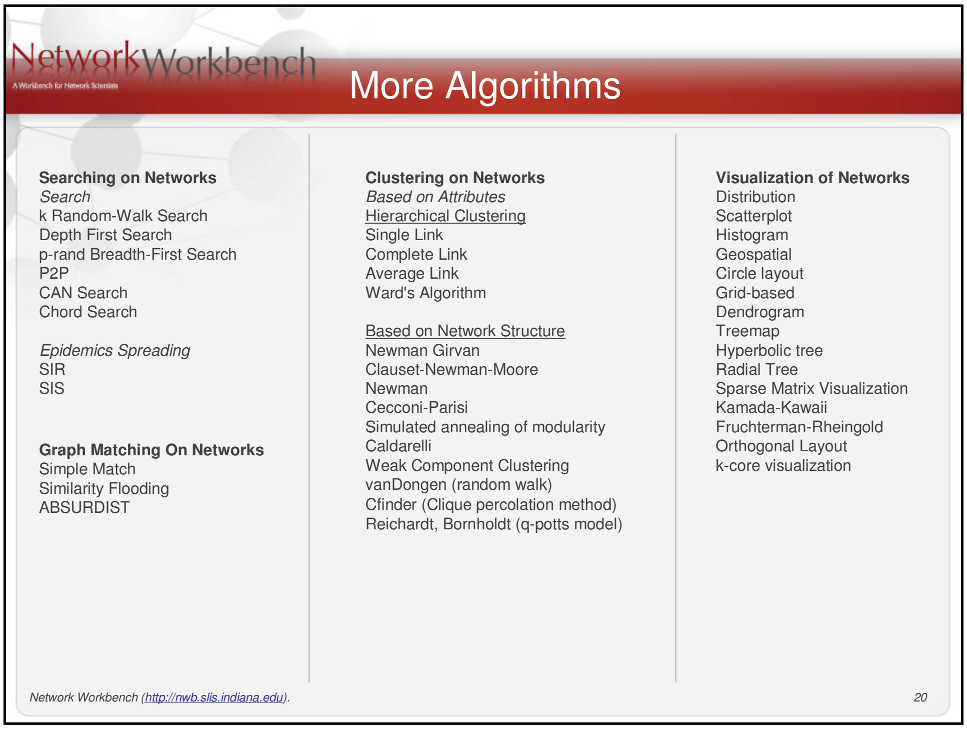
Peer-to-Peer Models

Structured
CAN Model
Chord Model

Unstructured
PRU Model
Hypergrid Model

Statistical Measurement
Edge/Node level
node degree
BC value of nodes/edges
Max flow edge
Hub/Authority value for nodes
Distribution of node distances (Hop plot)
Local (directed and weighted versions)
Clustering Coefficient (Watts Strogatz)
Clustering Coefficient (Newman)
k-Core Count
Distributions (Plot and gamma, and R^2)
Degree Distributions (in, out, total) (Directed/TotalDegree Distribution)
Degree Correlations (in-out, out-out, out-in, in-in, total-total)
Clustering Coefficient over k
Coherence for weighted graphs
Distribution of weights
Probability of degree distribution
Global
Density
Square of Adjacency Matrix
Giant Component
Strongly Connected Component
Betweenness Centrality
Diameter
Shortest Path = Geodesic Distance
Average Path Length

Motif Identification
Page Rank
Closeness centrality
Reach centrality
Eigenvector centrality
Minimum Spanning Tree


NetworkWorkbench
A Workbench for Network Scientists

More Algorithms

Searching on Networks
Search
k Random-Walk Search
Depth First Search
p-rand Breadth-First Search
P2P
CAN Search
Chord Search

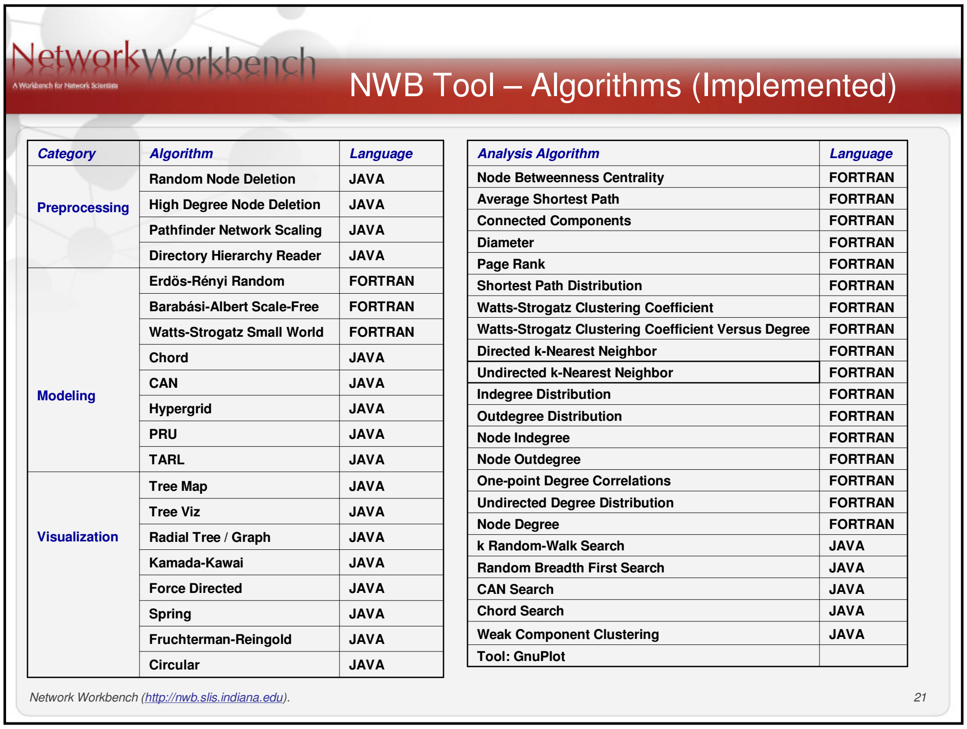
Epidemics Spreading
SIR
SIS

Graph Matching On Networks
Simple Match
Similarity Flooding
ABSURDIST

Clustering on Networks
Based on Attributes
Hierarchical Clustering
Single Link
Complete Link
Average Link
Ward's Algorithm

Based on Network Structure
Newman Girvan
Clauset-Newman-Moore
Newman
Cecconi-Parisi
Simulated annealing of modularity
Caldarelli
Weak Component Clustering
vanDongen (random walk)
Cfinder (Clique percolation method)
Reichardt, Bornholdt (q-potts model)

Visualization of Networks
Distribution
Scatterplot
Histogram
Geospatial
Circle layout
Grid-based
Dendrogram
Treemap
Hyperbolic tree
Radial Tree
Sparse Matrix Visualization
Kamada-Kawaii
Fruchterman-Rheingold
Orthogonal Layout
k-core visualization



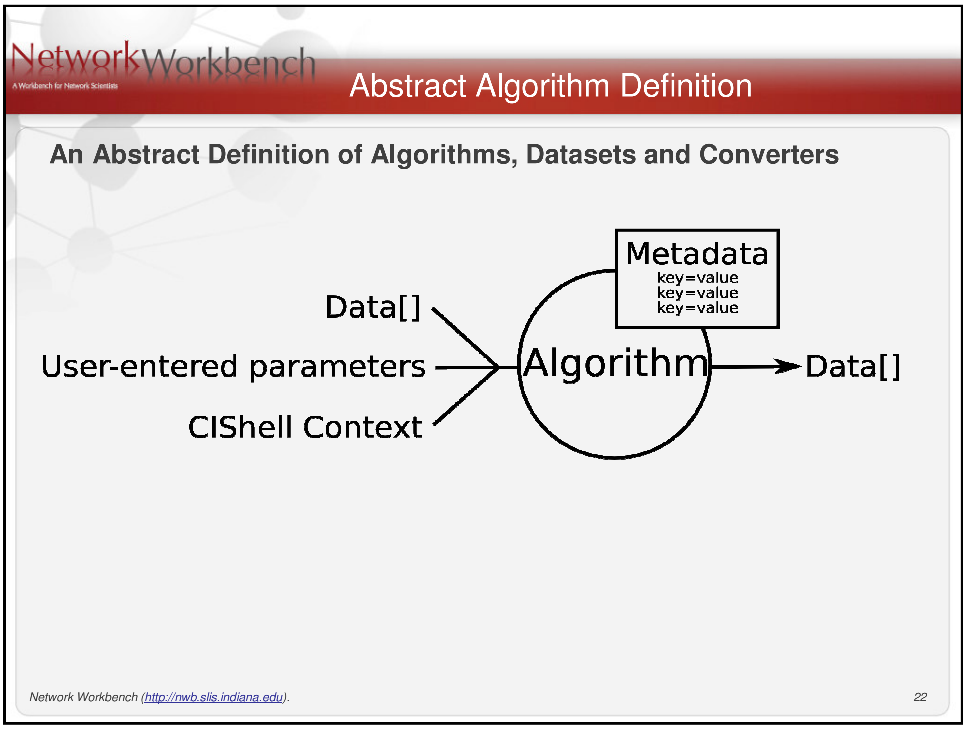
NetworkWorkbench
A Workbench for Network Scientists

NWB Tool – Algorithms (Implemented)

Category	*Algorithm*	*Language*
Preprocessing	Random Node Deletion	JAVA
	High Degree Node Deletion	JAVA
	Pathfinder Network Scaling	JAVA
	Directory Hierarchy Reader	JAVA
Modeling	Erdös-Rényi Random	FORTRAN
	Barabási-Albert Scale-Free	FORTRAN
	Watts-Strogatz Small World	FORTRAN
	Chord	JAVA
	CAN	JAVA
	Hypergrid	JAVA
	PRU	JAVA
	TARL	JAVA
Visualization	Tree Map	JAVA
	Tree Viz	JAVA
	Radial Tree / Graph	JAVA
	Kamada-Kawai	JAVA
	Force Directed	JAVA
	Spring	JAVA
	Fruchterman-Reingold	JAVA
	Circular	JAVA

Analysis Algorithm	*Language*
Node Betweenness Centrality	FORTRAN
Average Shortest Path	FORTRAN
Connected Components	FORTRAN
Diameter	FORTRAN
Page Rank	FORTRAN
Shortest Path Distribution	FORTRAN
Watts-Strogatz Clustering Coefficient	FORTRAN
Watts-Strogatz Clustering Coefficient Versus Degree	FORTRAN
Directed k-Nearest Neighbor	FORTRAN
Undirected k-Nearest Neighbor	FORTRAN
Indegree Distribution	FORTRAN
Outdegree Distribution	FORTRAN
Node Indegree	FORTRAN
Node Outdegree	FORTRAN
One-point Degree Correlations	FORTRAN
Undirected Degree Distribution	FORTRAN
Node Degree	FORTRAN
k Random-Walk Search	JAVA
Random Breadth First Search	JAVA
CAN Search	JAVA
Chord Search	JAVA
Weak Component Clustering	JAVA
Tool: GnuPlot	

Network Workbench (http://nwb.slis.indiana.edu). 21



NetworkWorkbench
A Workbench for Network Scientists

Abstract Algorithm Definition

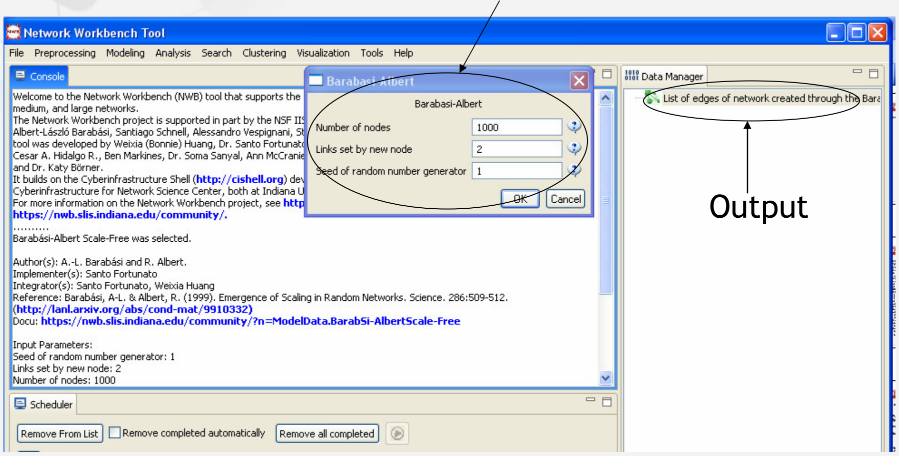
An Abstract Definition of Algorithms, Datasets and Converters

Network Workbench (http://nwb.slis.indiana.edu). 22

Abstract Algorithm Definition (cont.)

Modeling Algorithms

Input



Barabasi-Albert

Number of nodes: 1000

Links set by new node: 2

Seed of random number generator: 1

```
Welcome to the Network Workbench (NWB) tool that supports the
medium, and large networks.
The Network Workbench project is supported in part by the NSF IIS
Albert-László Barabási, Santiago Schnell, Alessandro Vespignani, St
tool was developed by Weixia (Bonnie) Huang, Dr. Santo Fortunato
Cesar A. Hidalgo R., Ben Markines, Dr. Soma Sanyal, Ann McCranie
and Dr. Katy Börner.
It builds on the Cyberinfrastructure Shell (http://cishell.org) dev
Cyberinfrastructure for Network Science Center, both at Indiana U
For more information on the Network Workbench project, see http
https://nwb.slis.indiana.edu/community/.
..........
Barabási-Albert Scale-Free was selected.

Author(s): A.-L. Barabási and R. Albert.
Implementer(s): Santo Fortunato
Integrator(s): Santo Fortunato, Weixia Huang
Reference: Barabási, A-L. & Albert, R. (1999). Emergence of Scaling in Random Networks. Science. 286:509-512.
(http://lanl.arxiv.org/abs/cond-mat/9910332)
Docu: https://nwb.slis.indiana.edu/community/?n=ModelData.BarabSi-AlbertScale-Free

Input Parameters:
Seed of random number generator: 1
Links set by new node: 2
Number of nodes: 1000
```

☐ Remove completed automatically

List of edges of network created through the Bara

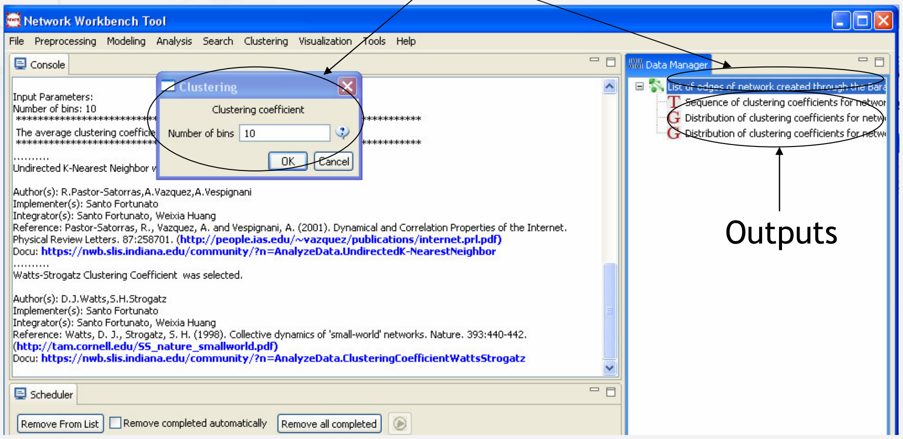
Output

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

Abstract Algorithm Definition (cont.)

Analysis Algorithms

Inputs



Clustering coefficient

Number of bins: 10

```
Input Parameters:
Number of bins: 10
 **************************
 The average clustering coefficie
 **************************
..........
Undirected K-Nearest Neighbor v

Author(s): R.Pastor-Satorras,A.Vazquez,A.Vespignani
Implementer(s): Santo Fortunato
Integrator(s): Santo Fortunato, Weixia Huang
Reference: Pastor-Satorras, R., Vazquez, A. and Vespignani, A. (2001). Dynamical and Correlation Properties of the Internet.
Physical Review Letters. 87:258701. (http://people.ias.edu/~vazquez/publications/internet.prl.pdf)
Docu: https://nwb.slis.indiana.edu/community/?n=AnalyzeData.UndirectedK-NearestNeighbor
..........
Watts-Strogatz Clustering Coefficient  was selected.

Author(s): D.J.Watts,S.H.Strogatz
Implementer(s): Santo Fortunato
Integrator(s): Santo Fortunato, Weixia Huang
Reference: Watts, D. J., Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. Nature. 393:440-442.
(http://tam.cornell.edu/SS_nature_smallworld.pdf)
Docu: https://nwb.slis.indiana.edu/community/?n=AnalyzeData.ClusteringCoefficientWattsStrogatz
```

☐ Remove completed automatically

- List of edges of network created through the Bara
  - T Sequence of clustering coefficients for networ
  - G Distribution of clustering coefficients for netw
  - G Distribution of clustering coefficients for netw

Outputs

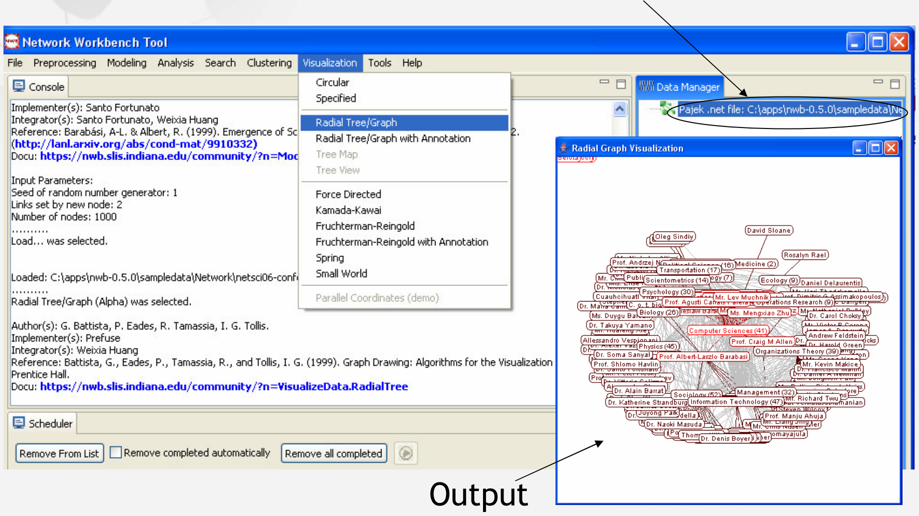
Network Workbench (http://nwb.slis.indiana.edu).

NetworkWorkbench
A Workbench for Network Scientists

Abstract Algorithm Definition (cont.)

Visualization Algorithms

Input



```
Implementer(s): Santo Fortunato
Integrator(s): Santo Fortunato, Weixia Huang
Reference: Barabási, A-L. & Albert, R. (1999). Emergence of Sc                2.
(http://lanl.arxiv.org/abs/cond-mat/9910332)
Docu: https://nwb.slis.indiana.edu/community/?n=Moc

Input Parameters:
Seed of random number generator: 1
Links set by new node: 2
Number of nodes: 1000
..........
Load... was selected.


Loaded: C:\apps\nwb-0.5.0\sampledata\Network\netsci06-conf
..........
Radial Tree/Graph (Alpha) was selected.

Author(s): G. Battista, P. Eades, R. Tamassia, I. G. Tollis.
Implementer(s): Prefuse
Integrator(s): Weixia Huang
Reference: Battista, G., Eades, P., Tamassia, R., and Tollis, I. G. (1999). Graph Drawing: Algorithms for the Visualization
Prentice Hall.
Docu: https://nwb.slis.indiana.edu/community/?n=VisualizeData.RadialTree
```

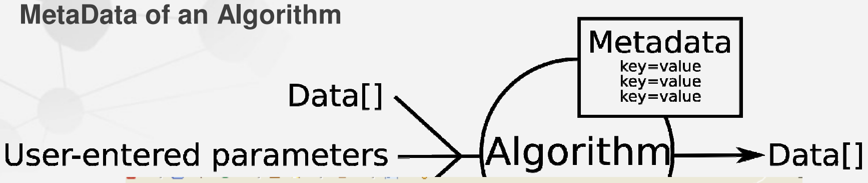
Output

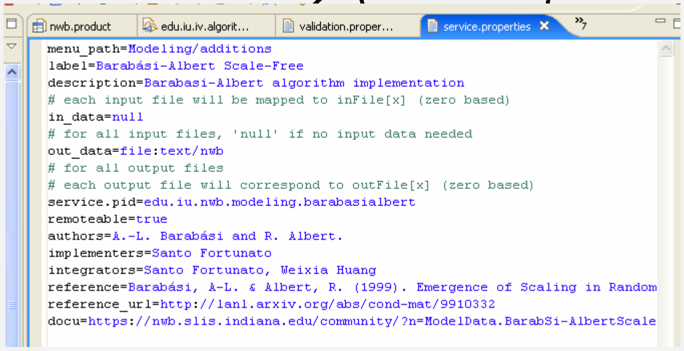
Network Workbench (http://nwb.slis.indiana.edu).

NetworkWorkbench
A Workbench for Network Scientists

For Algorithm Integrators

MetaData of an Algorithm



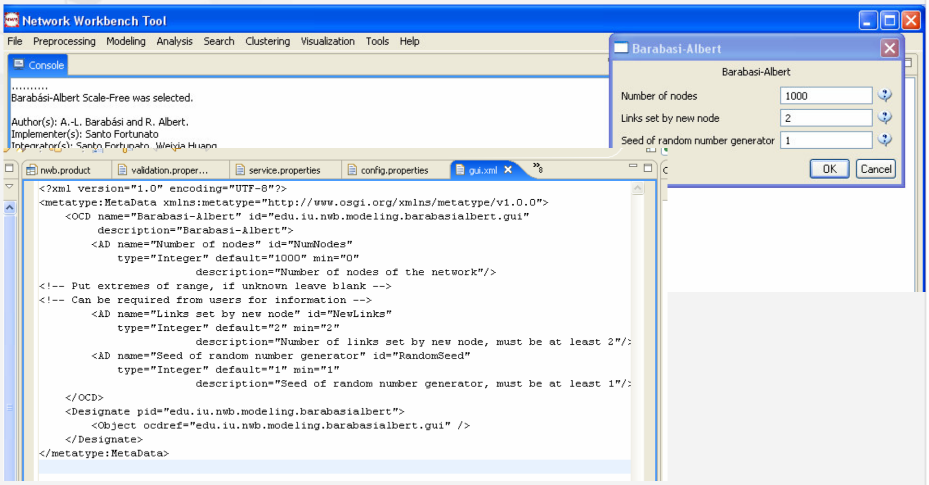


```
menu_path=Modeling/additions
label=Barabási-Albert Scale-Free
description=Barabasi-Albert algorithm implementation
# each input file will be mapped to inFile[x] (zero based)
in_data=null
# for all input files, 'null' if no input data needed
out_data=file:text/nwb
# for all output files
# each output file will correspond to outFile[x] (zero based)
service.pid=edu.iu.nwb.modeling.barabasialbert
remoteable=true
authors=A.-L. Barabási and R. Albert.
implementers=Santo Fortunato
integrators=Santo Fortunato, Weixia Huang
reference=Barabási, A-L. & Albert, R. (1999). Emergence of Scaling in Random
reference_url=http://lanl.arxiv.org/abs/cond-mat/9910332
docu=https://nwb.slis.indiana.edu/community/?n=ModelData.BarabSi-AlbertScale
```

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

For Algorithm Integrators (cont.)

GUI Builder and MetaType Service



....
Barabási-Albert Scale-Free was selected.

Author(s): A.-L. Barabási and R. Albert.
Implementer(s): Santo Fortunato
Integrator(s): Santo Fortunato, Weixia Huang

Barabasi-Albert

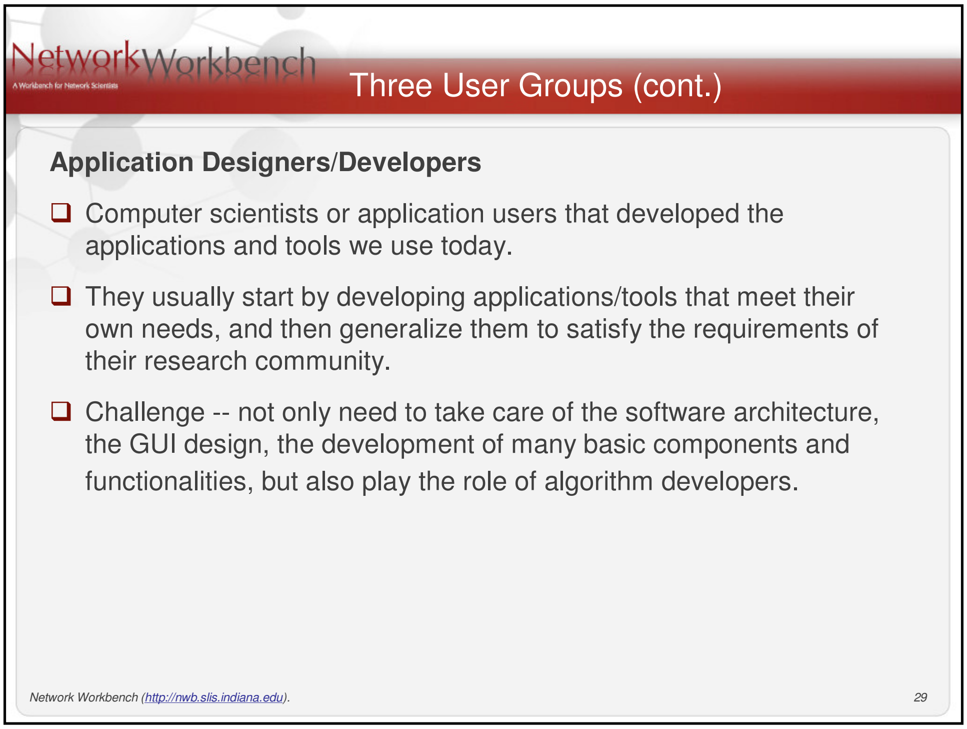
Number of nodes: 1000
Links set by new node: 2
Seed of random number generator: 1

```
<?xml version="1.0" encoding="UTF-8"?>
<metatype:MetaData xmlns:metatype="http://www.osgi.org/xmlns/metatype/v1.0.0">
    <OCD name="Barabasi-Albert" id="edu.iu.nwb.modeling.barabasialbert.gui"
         description="Barabasi-Albert">
        <AD name="Number of nodes" id="NumNodes"
            type="Integer" default="1000" min="0"
                        description="Number of nodes of the network"/>
<!-- Put extremes of range, if unknown leave blank -->
<!-- Can be required from users for information -->
        <AD name="Links set by new node" id="NewLinks"
            type="Integer" default="2" min="2"
                        description="Number of links set by new node, must be at least 2"/
        <AD name="Seed of random number generator" id="RandomSeed"
            type="Integer" default="1" min="1"
                        description="Seed of random number generator, must be at least 1"/
    </OCD>
    <Designate pid="edu.iu.nwb.modeling.barabasialbert">
        <Object ocdref="edu.iu.nwb.modeling.barabasialbert.gui" />
    </Designate>
</metatype:MetaData>
```

Three User Groups

Application Users

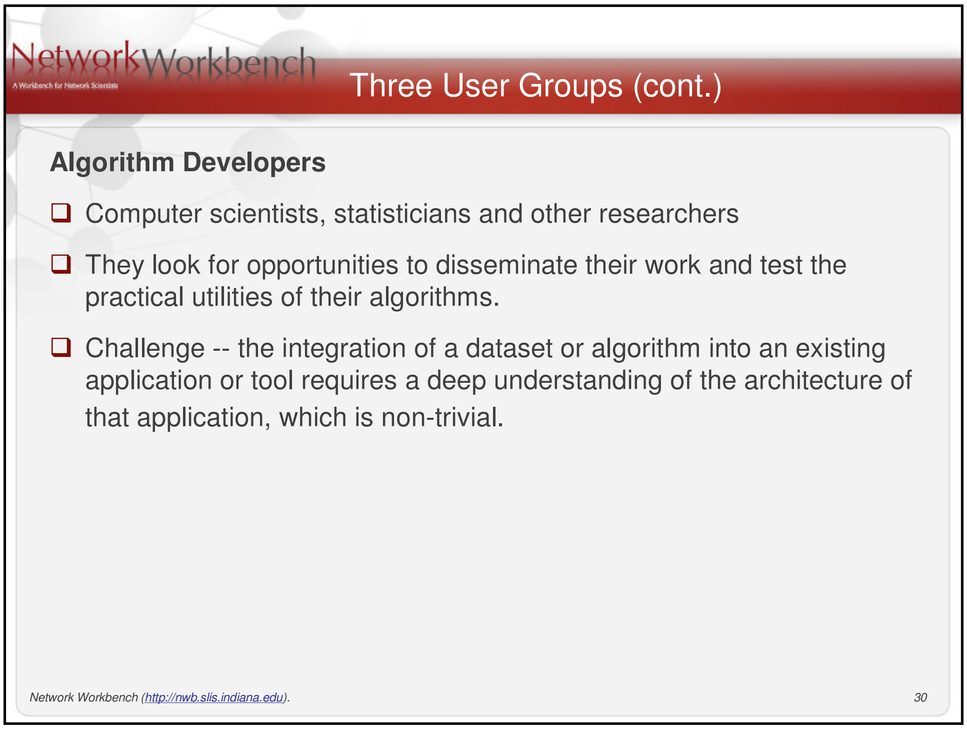
- Scientists in the natural and social sciences (physics, biology, chemistry, psychology, sociology, etc.)
- Their needs -- want to find the best datasets and the most effective algorithms to conduct their research.
- Problem – too many algorithms. Finding a correctly working piece of code is challenging. Frequently, not only one but a sequence of different algorithms needs to be applied to load, parse, clean, mine, analyze, model, visualize, and print data. Today, there is no easy way to extend a tool by adding new algorithms as needed or to customize a tool so that it exactly fits the needs of a specific user (group).



Three User Groups (cont.)

Application Designers/Developers

- Computer scientists or application users that developed the applications and tools we use today.
- They usually start by developing applications/tools that meet their own needs, and then generalize them to satisfy the requirements of their research community.
- Challenge -- not only need to take care of the software architecture, the GUI design, the development of many basic components and functionalities, but also play the role of algorithm developers.



Three User Groups (cont.)

Algorithm Developers

- Computer scientists, statisticians and other researchers
- They look for opportunities to disseminate their work and test the practical utilities of their algorithms.
- Challenge -- the integration of a dataset or algorithm into an existing application or tool requires a deep understanding of the architecture of that application, which is non-trivial.

OSGi – Technical Details

NWB/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.
- Alliance members include IBM (Eclipse), Sun, Intel, Oracle, Motorola, NEC and many others.
- Has successfully been used in the industry from high-end servers to embedded mobile devices for 8 years now.
- Widely adopted in open source realm, especially since Eclipse 3.0 that uses OSGi R4 for its plugin model.

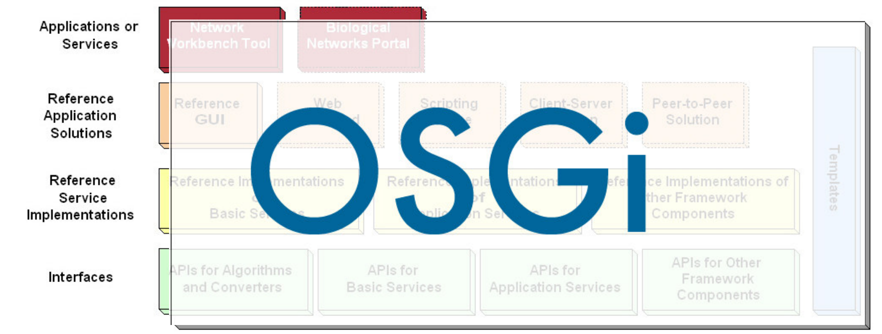
Advantages of Using OSGi

- Directly use many components provided by OSGi framework, such as service registry
- Contribute diverse algorithms to OSGi community -- any CIShell algorithm becomes a service that can be used in any OSGi-based framework.
- Running CIShells/tools can connect to each other via exposed CIShell-defined web services supporting peer-to-peer sharing of data, algorithms, and computing power.

Ideally, CIShell becomes a standard for creating algorithm services in OSGi developed Tools/CI, e.g., IVC&NWB will be using the CIShell reference GUI

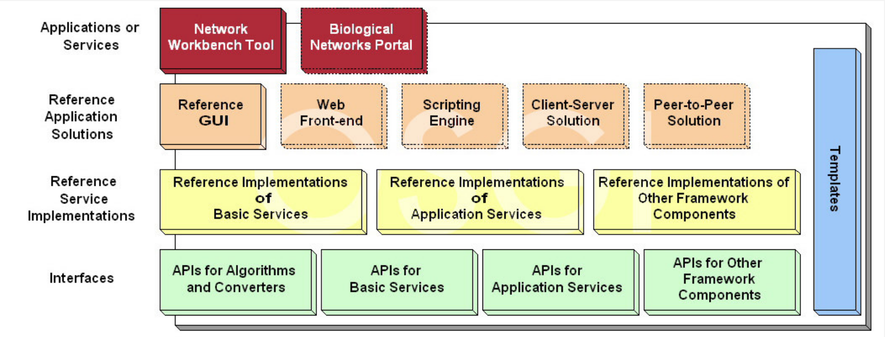
OSGi – Technical Details (cont.)

Service Oriented Component-based Architecture


Applications or Services
Reference Application Solutions
Reference Service Implementations
Interfaces
Network Workbench Tool
Biological Networks Portal
Reference GUI
Web
Scripting
Client-Server
Peer-to-Peer Solution
Reference Implementations of Basic Services
Reference Implementations of Application Services
Reference Implementations of Other Framework Components
APIs for Algorithms and Converters
APIs for Basic Services
APIs for Application Services
APIs for Other Framework Components
Templates
OSGi

NWB/CIShell Architecture

An Overview of NWB/CIShell Architecture



Applications or Services

Network Workbench Tool

Biological Networks Portal

Reference Application Solutions

Reference GUI

Web Front-end

Scripting Engine

Client-Server Solution

Peer-to-Peer Solution

Reference Service Implementations

Reference Implementations of Basic Services

Reference Implementations of Application Services

Reference Implementations of Other Framework Components

Interfaces

APIs for Algorithms and Converters

APIs for Basic Services

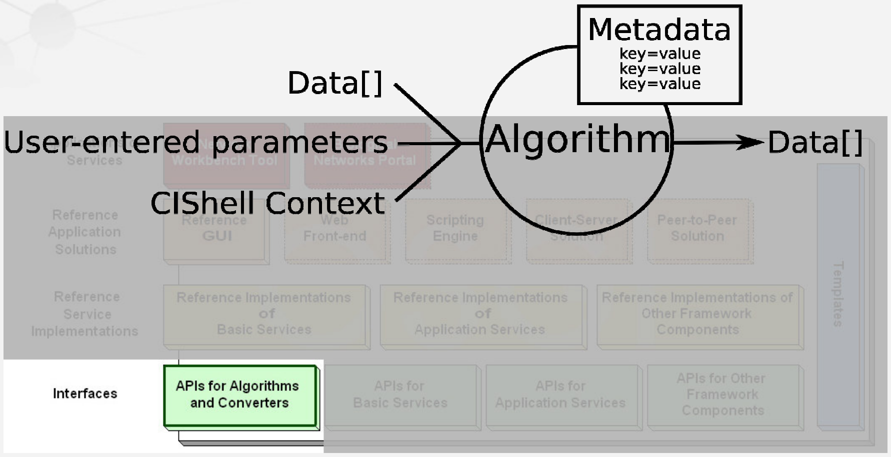
APIs for Application Services

APIs for Other Framework Components

Templates

Interfaces Layer – Algorithm

An Abstract Definition of Algorithms, Datasets and Converters



Metadata
key=value
key=value
key=value

Data[]

User-entered parameters

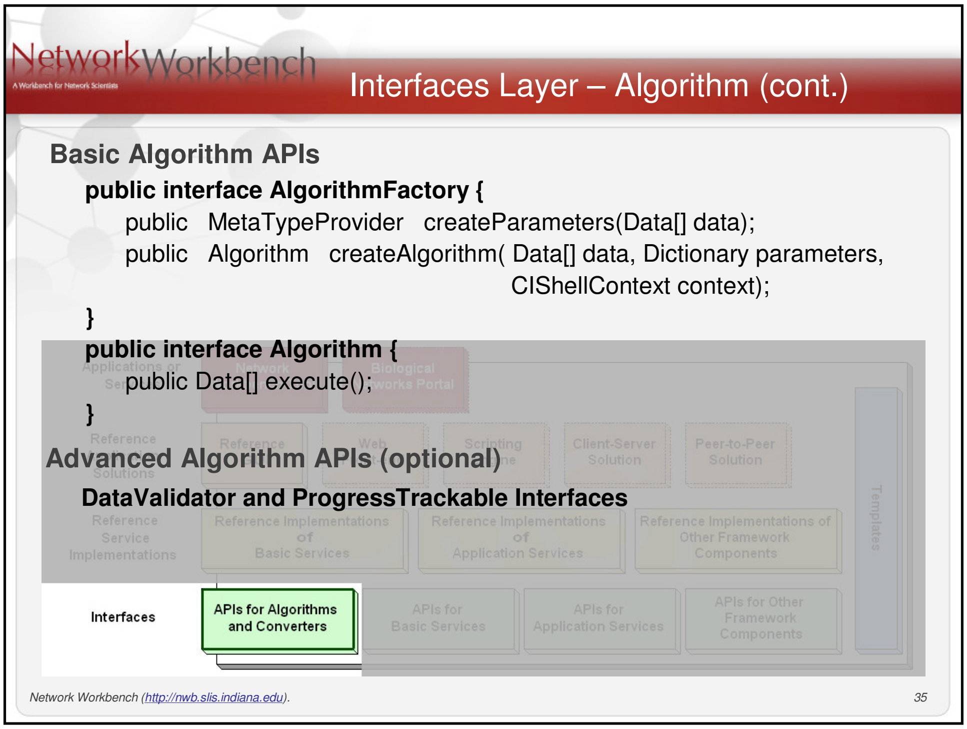
CIShell Context

Algorithm

Data[]

Interfaces

APIs for Algorithms and Converters



Interfaces Layer – Algorithm (cont.)

Basic Algorithm APIs

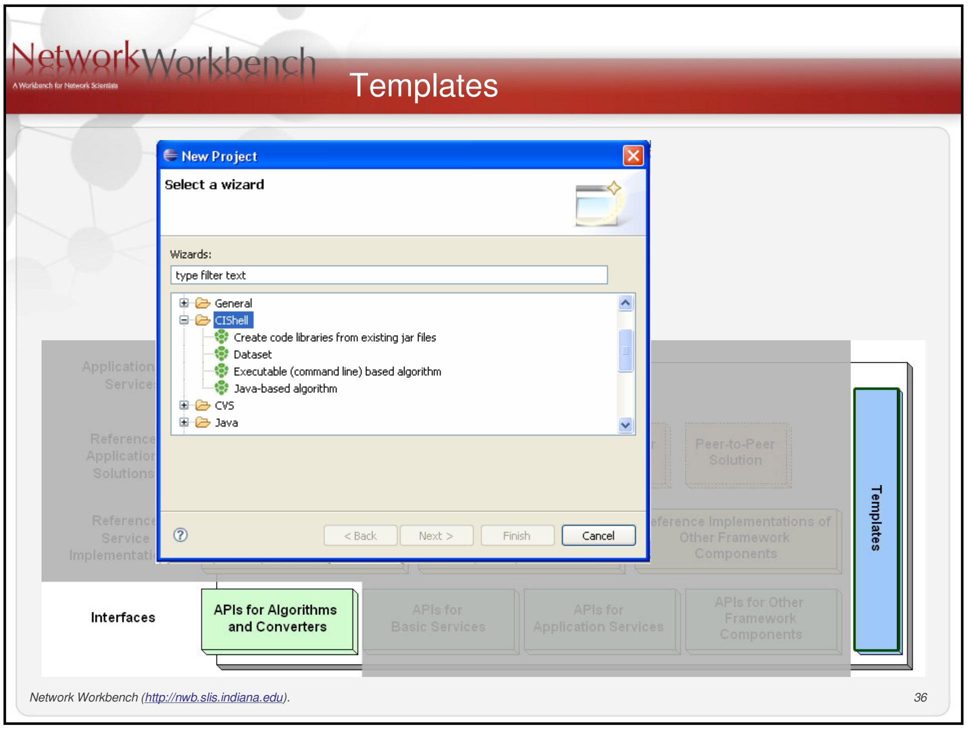
```
public interface AlgorithmFactory {
    public   MetaTypeProvider   createParameters(Data[] data);
    public   Algorithm   createAlgorithm( Data[] data, Dictionary parameters,
                                           CIShellContext context);
}
public interface Algorithm {
    public Data[] execute();
}
```

Advanced Algorithm APIs (optional)

DataValidator and ProgressTrackable Interfaces

Interfaces
APIs for Algorithms and Converters
APIs for Basic Services
APIs for Application Services
APIs for Other Framework Components
Reference Service Implementations
Reference Implementations of Basic Services
Reference Implementations of Application Services
Reference Implementations of Other Framework Components
Reference Solutions
Web
Scripting
Client-Server Solution
Peer-to-Peer Solution
Templates

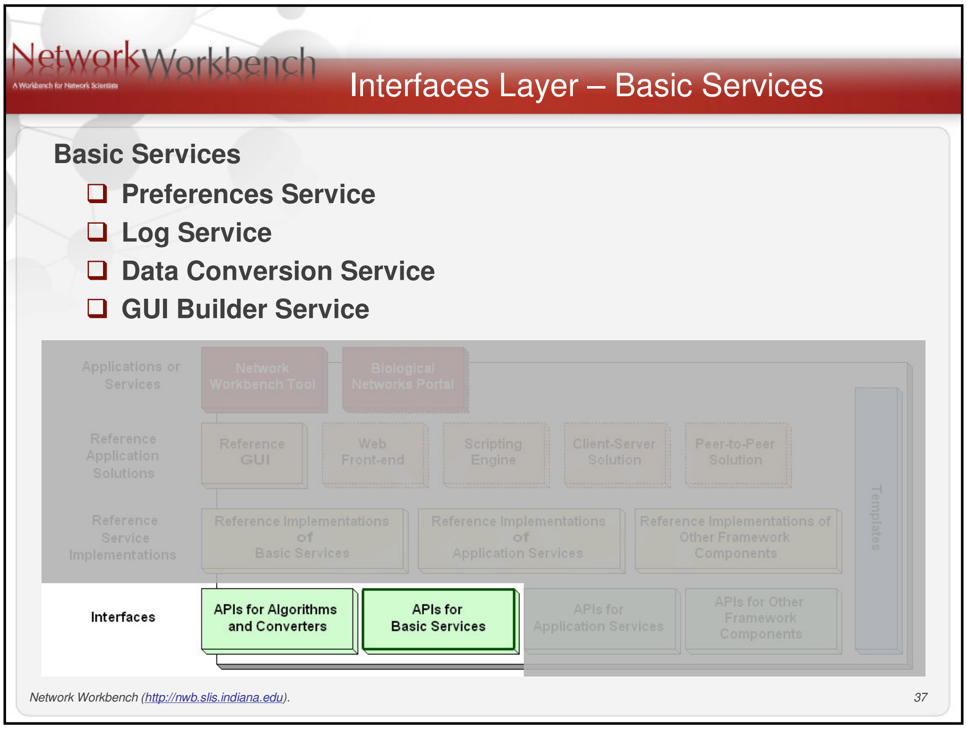
Network Workbench (http://nwb.slis.indiana.edu). 35



Templates

New Project
Select a wizard
Wizards:
type filter text
General
CIShell
Create code libraries from existing jar files
Dataset
Executable (command line) based algorithm
Java-based algorithm
CVS
Java
< Back
Next >
Finish
Cancel
Interfaces
APIs for Algorithms and Converters
APIs for Basic Services
APIs for Application Services
APIs for Other Framework Components
Peer-to-Peer Solution
Reference Implementations of Other Framework Components
Templates

Network Workbench (http://nwb.slis.indiana.edu). 36



NetworkWorkbench
A Workbench for Network Scientists

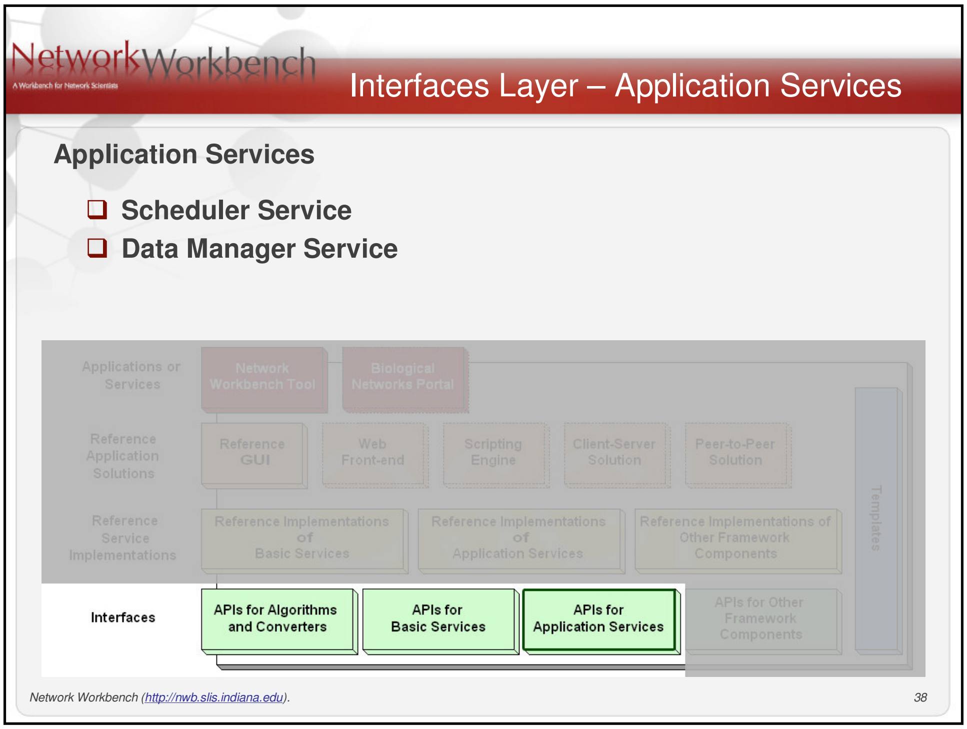
Interfaces Layer – Basic Services

Basic Services

- ❑ Preferences Service
- ❑ Log Service
- ❑ Data Conversion Service
- ❑ GUI Builder Service

Applications or Services
Network Workbench Tool
Biological Networks Portal
Reference Application Solutions
Reference GUI
Web Front-end
Scripting Engine
Client-Server Solution
Peer-to-Peer Solution
Reference Service Implementations
Reference Implementations of Basic Services
Reference Implementations of Application Services
Reference Implementations of Other Framework Components
Interfaces
APIs for Algorithms and Converters
APIs for Basic Services
APIs for Application Services
APIs for Other Framework Components
Templates

Network Workbench (http://nwb.slis.indiana.edu). 37



NetworkWorkbench
A Workbench for Network Scientists

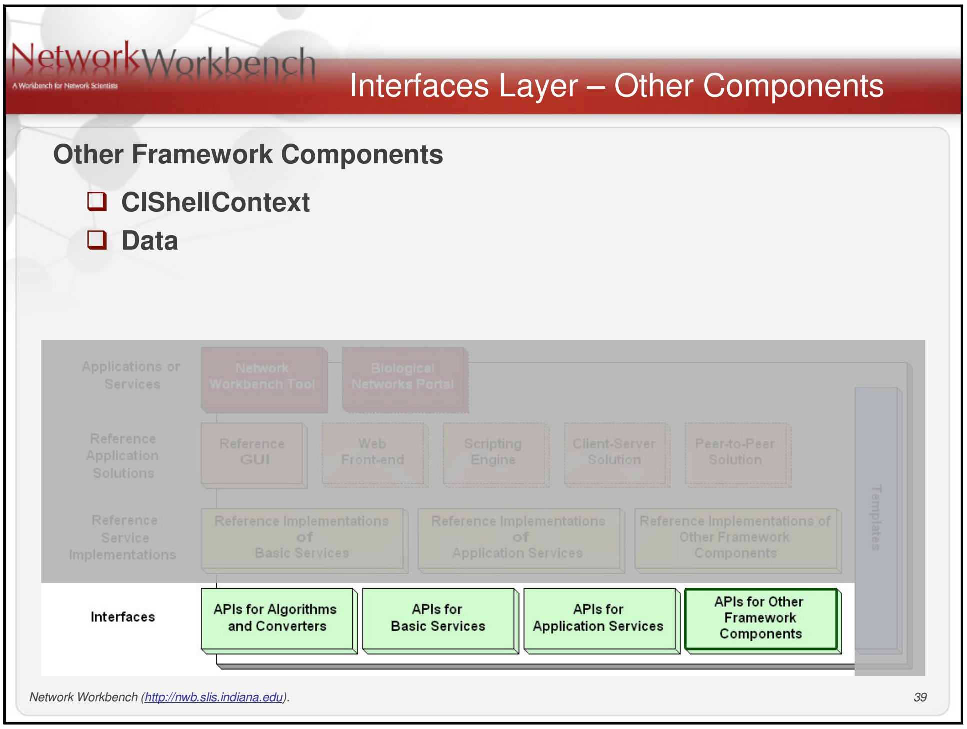
Interfaces Layer – Application Services

Application Services

- ❑ Scheduler Service
- ❑ Data Manager Service

Applications or Services
Network Workbench Tool
Biological Networks Portal
Reference Application Solutions
Reference GUI
Web Front-end
Scripting Engine
Client-Server Solution
Peer-to-Peer Solution
Reference Service Implementations
Reference Implementations of Basic Services
Reference Implementations of Application Services
Reference Implementations of Other Framework Components
Interfaces
APIs for Algorithms and Converters
APIs for Basic Services
APIs for Application Services
APIs for Other Framework Components
Templates

Network Workbench (http://nwb.slis.indiana.edu). 38



NetworkWorkbench
A Workbench for Network Scientists

Interfaces Layer – Other Components

Other Framework Components

- ❑ CIShellContext
- ❑ Data

Applications or Services
Network Workbench Tool
Biological Networks Portal

Reference Application Solutions
Reference GUI
Web Front-end
Scripting Engine
Client-Server Solution
Peer-to-Peer Solution

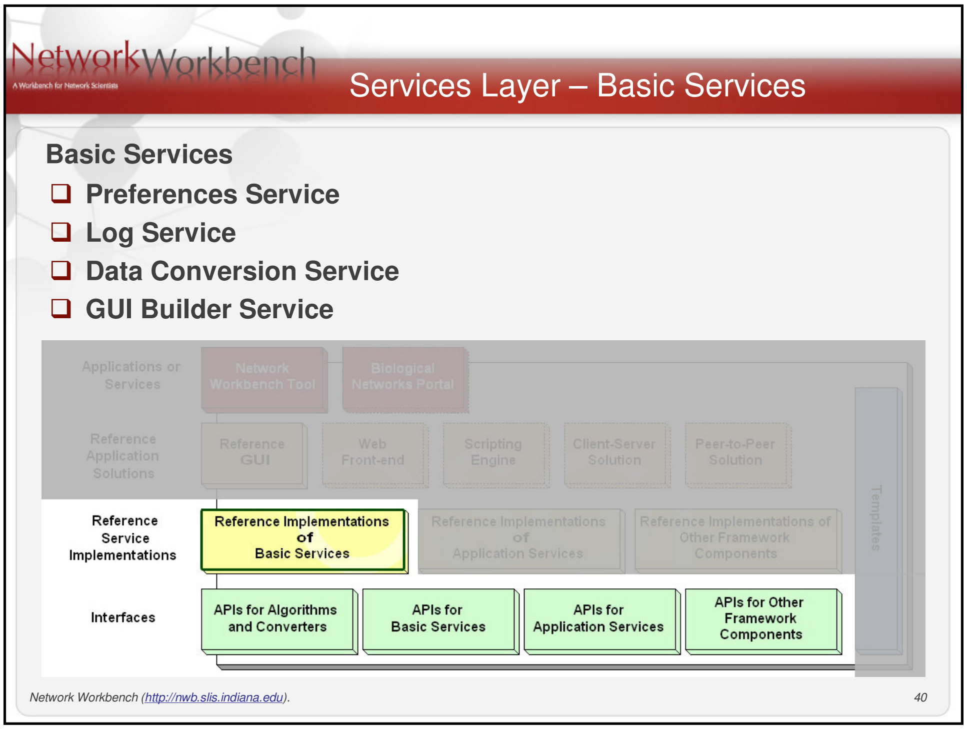
Reference Service Implementations
Reference Implementations of Basic Services
Reference Implementations of Application Services
Reference Implementations of Other Framework Components

Interfaces
APIs for Algorithms and Converters
APIs for Basic Services
APIs for Application Services
APIs for Other Framework Components

Templates

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

39



NetworkWorkbench
A Workbench for Network Scientists

Services Layer – Basic Services

Basic Services

- ❑ Preferences Service
- ❑ Log Service
- ❑ Data Conversion Service
- ❑ GUI Builder Service

Applications or Services
Network Workbench Tool
Biological Networks Portal

Reference Application Solutions
Reference GUI
Web Front-end
Scripting Engine
Client-Server Solution
Peer-to-Peer Solution

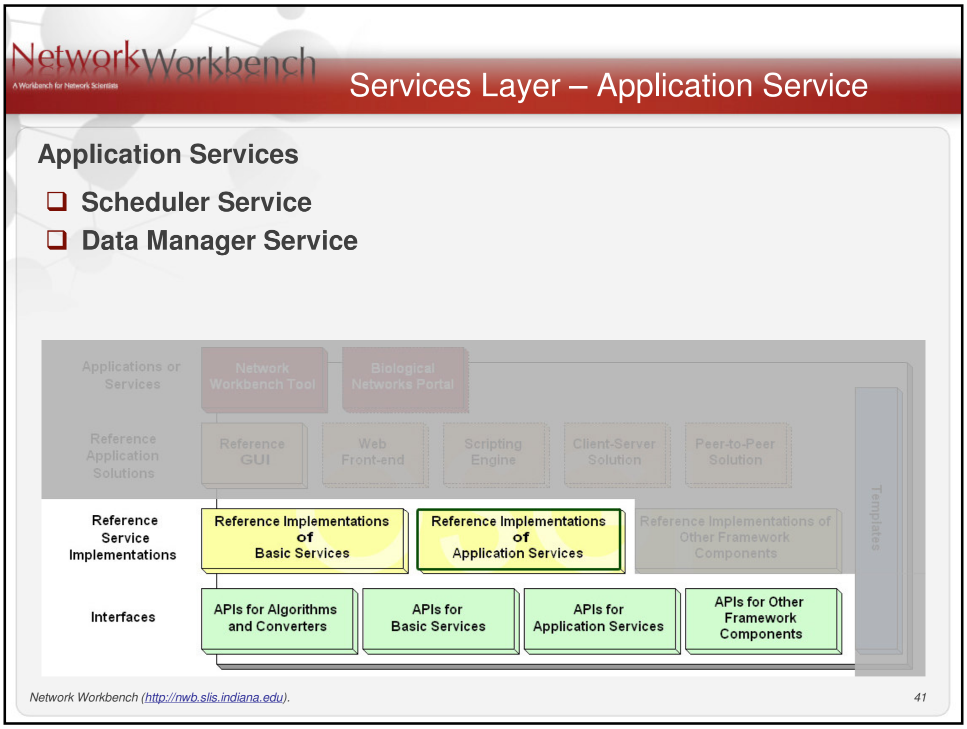
Reference Service Implementations
Reference Implementations of Basic Services
Reference Implementations of Application Services
Reference Implementations of Other Framework Components

Interfaces
APIs for Algorithms and Converters
APIs for Basic Services
APIs for Application Services
APIs for Other Framework Components

Templates

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

40



NetworkWorkbench

A Workbench for Network Scientists

Services Layer – Application Service

Application Services

- ❑ Scheduler Service
- ❑ Data Manager Service

Applications or Services

Network Workbench Tool

Biological Networks Portal

Reference Application Solutions

Reference GUI

Web Front-end

Scripting Engine

Client-Server Solution

Peer-to-Peer Solution

Reference Service Implementations

Reference Implementations of Basic Services

Reference Implementations of Application Services

Reference Implementations of Other Framework Components

Interfaces

APIs for Algorithms and Converters

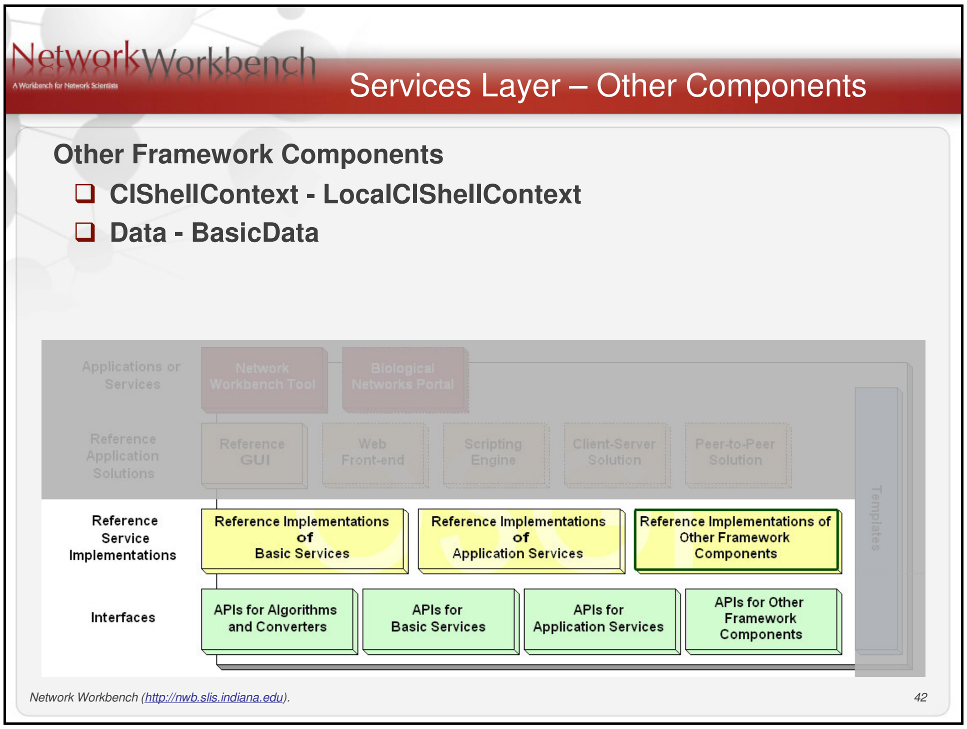
APIs for Basic Services

APIs for Application Services

APIs for Other Framework Components

Templates

Network Workbench (http://nwb.slis.indiana.edu). 41



NetworkWorkbench

A Workbench for Network Scientists

Services Layer – Other Components

Other Framework Components

- ❑ CIShellContext - LocalCIShellContext
- ❑ Data - BasicData

Applications or Services

Network Workbench Tool

Biological Networks Portal

Reference Application Solutions

Reference GUI

Web Front-end

Scripting Engine

Client-Server Solution

Peer-to-Peer Solution

Reference Service Implementations

Reference Implementations of Basic Services

Reference Implementations of Application Services

Reference Implementations of Other Framework Components

Interfaces

APIs for Algorithms and Converters

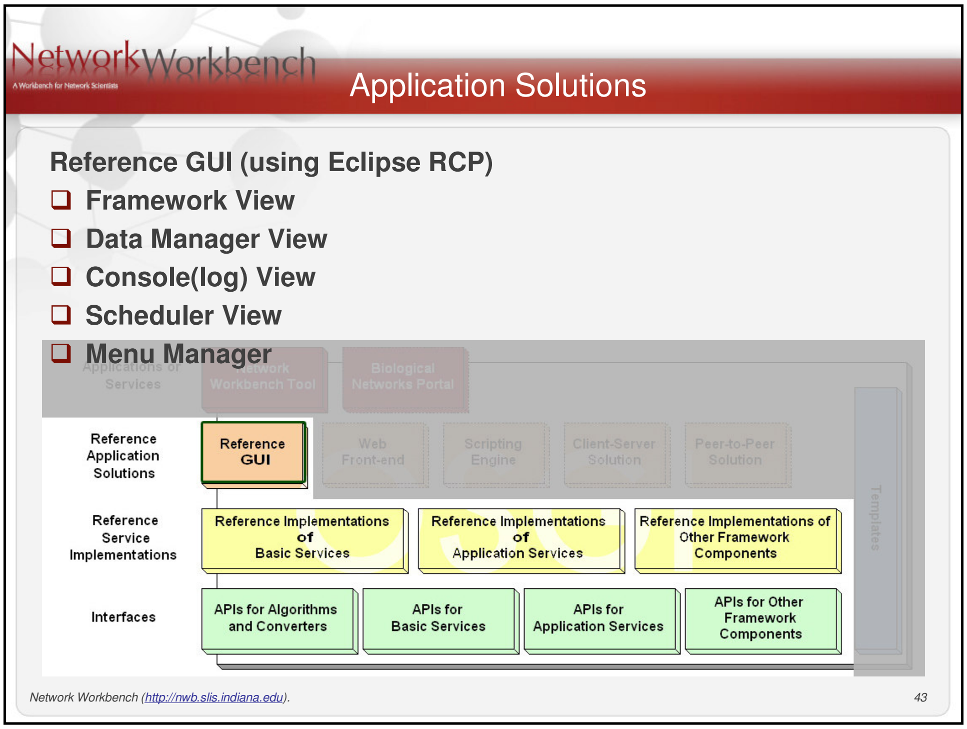
APIs for Basic Services

APIs for Application Services

APIs for Other Framework Components

Templates

Network Workbench (http://nwb.slis.indiana.edu). 42

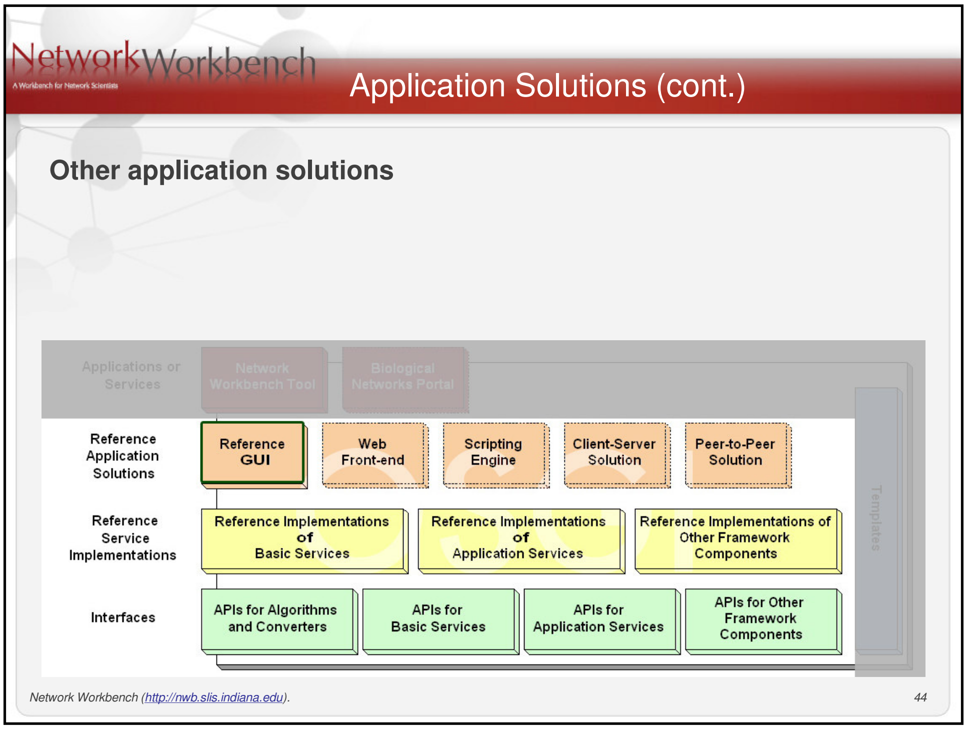


Application Solutions

Reference GUI (using Eclipse RCP)

- ❑ Framework View
- ❑ Data Manager View
- ❑ Console(log) View
- ❑ Scheduler View
- ❑ Menu Manager

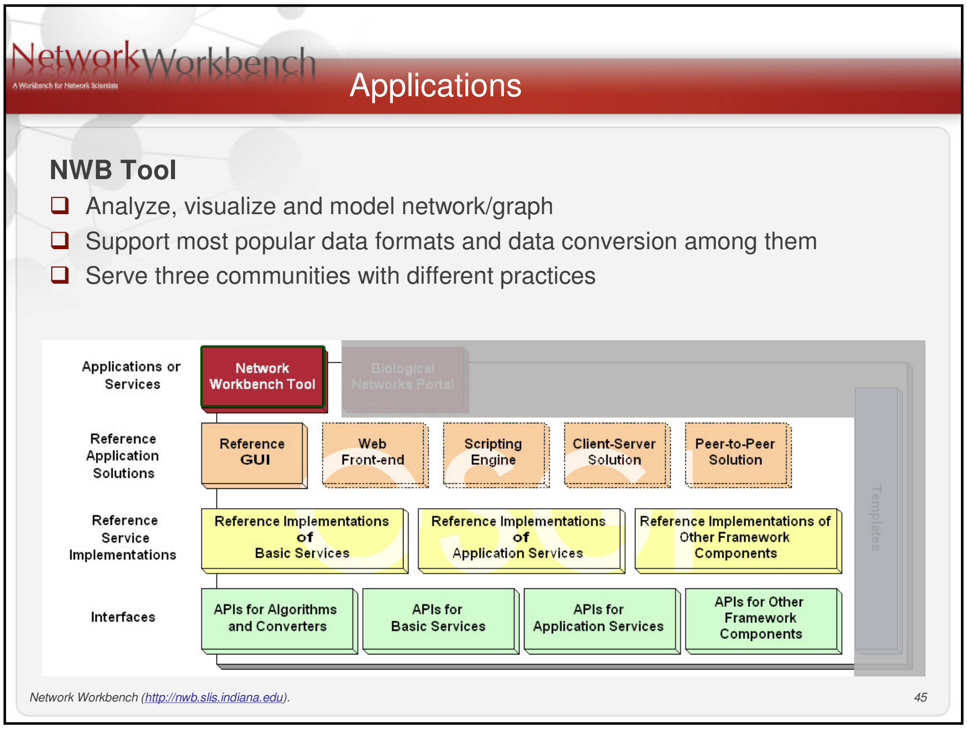
Applications or Services
Network Workbench Tool
Biological Networks Portal
Reference Application Solutions
Reference GUI
Web Front-end
Scripting Engine
Client-Server Solution
Peer-to-Peer Solution
Reference Service Implementations
Reference Implementations of Basic Services
Reference Implementations of Application Services
Reference Implementations of Other Framework Components
Interfaces
APIs for Algorithms and Converters
APIs for Basic Services
APIs for Application Services
APIs for Other Framework Components
Templates



Application Solutions (cont.)

Other application solutions

Applications or Services
Network Workbench Tool
Biological Networks Portal
Reference Application Solutions
Reference GUI
Web Front-end
Scripting Engine
Client-Server Solution
Peer-to-Peer Solution
Reference Service Implementations
Reference Implementations of Basic Services
Reference Implementations of Application Services
Reference Implementations of Other Framework Components
Interfaces
APIs for Algorithms and Converters
APIs for Basic Services
APIs for Application Services
APIs for Other Framework Components
Templates



NetworkWorkbench

A Workbench for Network Scientists

Applications

NWB Tool

- ❑ Analyze, visualize and model network/graph
- ❑ Support most popular data formats and data conversion among them
- ❑ Serve three communities with different practices

Applications or Services: Network Workbench Tool; Biological Networks Portal

Reference Application Solutions: Reference GUI; Web Front-end; Scripting Engine; Client-Server Solution; Peer-to-Peer Solution

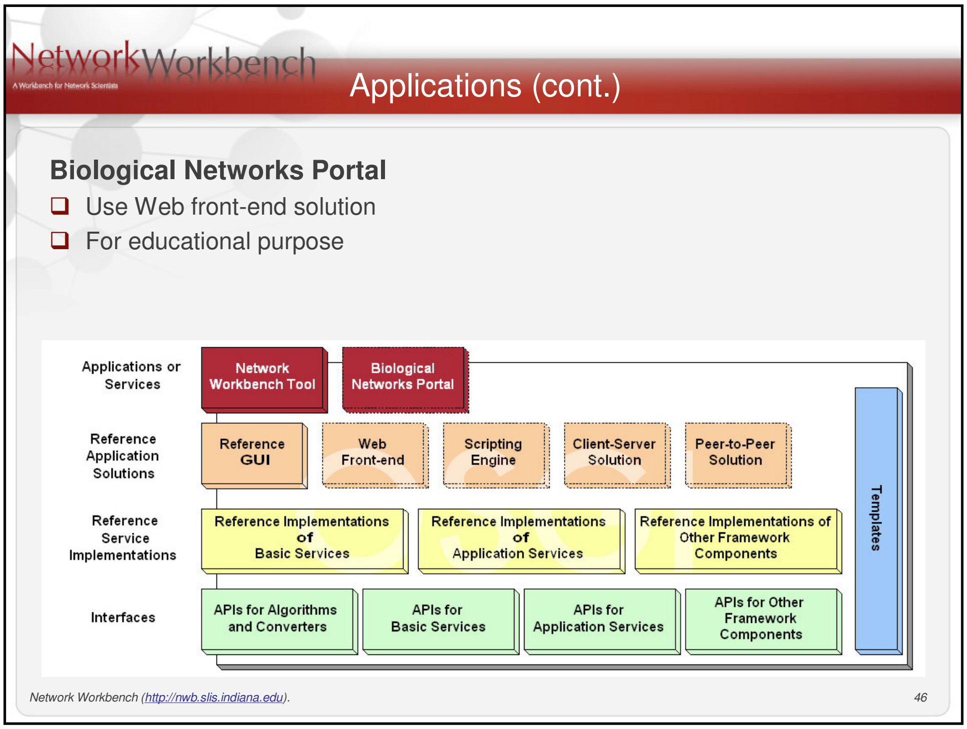
Reference Service Implementations: Reference Implementations of Basic Services; Reference Implementations of Application Services; Reference Implementations of Other Framework Components

Interfaces: APIs for Algorithms and Converters; APIs for Basic Services; APIs for Application Services; APIs for Other Framework Components

Templates

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

45



NetworkWorkbench

A Workbench for Network Scientists

Applications (cont.)

Biological Networks Portal

- ❑ Use Web front-end solution
- ❑ For educational purpose

Applications or Services: Network Workbench Tool; Biological Networks Portal

Reference Application Solutions: Reference GUI; Web Front-end; Scripting Engine; Client-Server Solution; Peer-to-Peer Solution

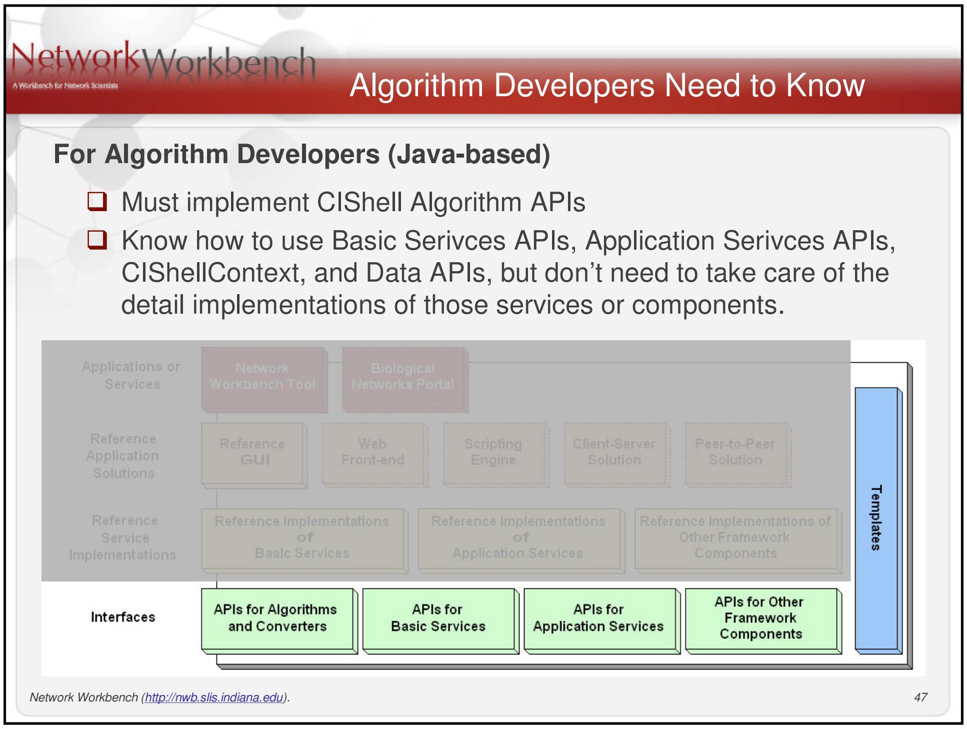
Reference Service Implementations: Reference Implementations of Basic Services; Reference Implementations of Application Services; Reference Implementations of Other Framework Components

Interfaces: APIs for Algorithms and Converters; APIs for Basic Services; APIs for Application Services; APIs for Other Framework Components

Templates

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

46



NetworkWorkbench
A Workbench for Network Scientists

Algorithm Developers Need to Know

For Algorithm Developers (Java-based)

- ❑ Must implement CIShell Algorithm APIs
- ❑ Know how to use Basic Serivces APIs, Application Serivces APIs, CIShellContext, and Data APIs, but don't need to take care of the detail implementations of those services or components.

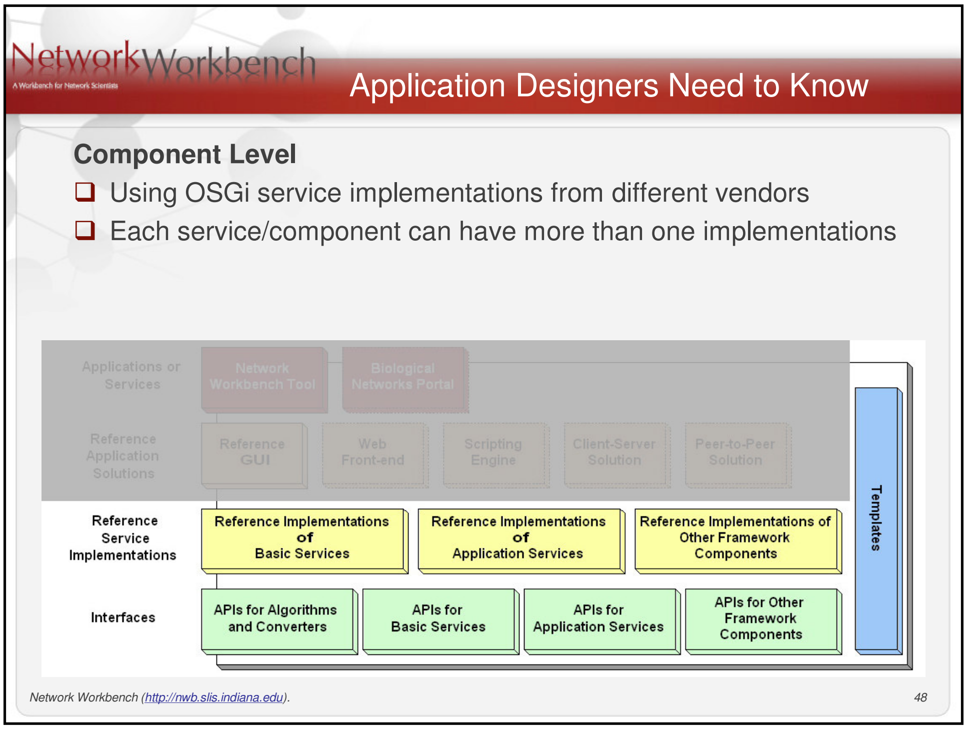
Applications or Services: Network Workbench Tool; Biological Networks Portal

Reference Application Solutions: Reference GUI; Web Front-end; Scripting Engine; Client-Server Solution; Peer-to-Peer Solution

Reference Service Implementations: Reference Implementations of Basic Services; Reference Implementations of Application Services; Reference Implementations of Other Framework Components

Interfaces: APIs for Algorithms and Converters; APIs for Basic Services; APIs for Application Services; APIs for Other Framework Components

Templates



NetworkWorkbench
A Workbench for Network Scientists

Application Designers Need to Know

Component Level

- ❑ Using OSGi service implementations from different vendors
- ❑ Each service/component can have more than one implementations

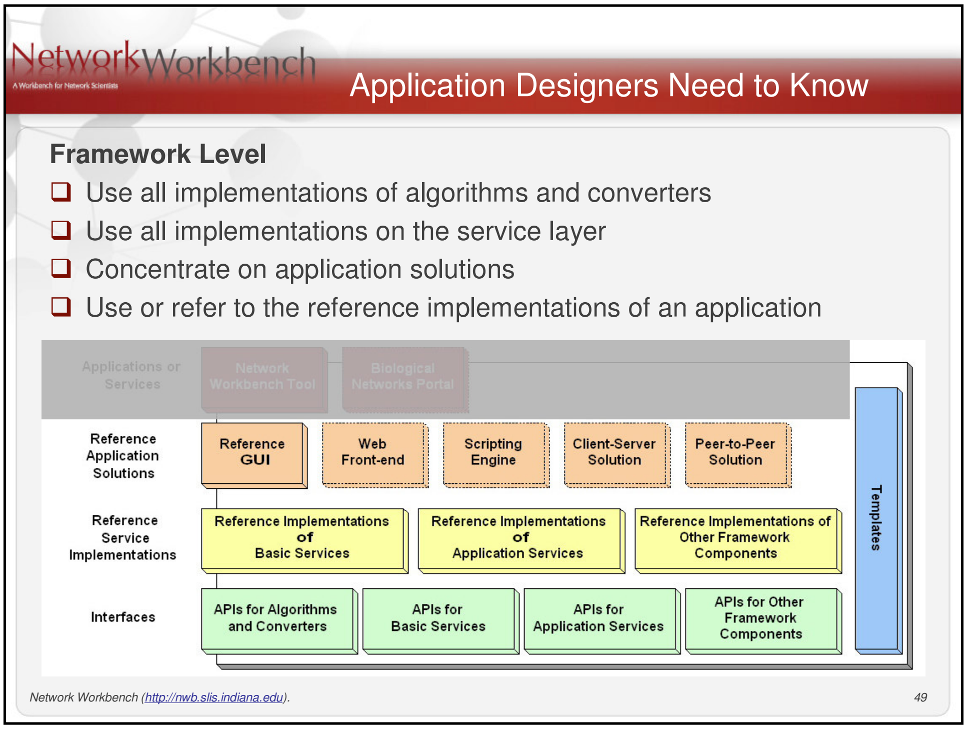
Applications or Services: Network Workbench Tool; Biological Networks Portal

Reference Application Solutions: Reference GUI; Web Front-end; Scripting Engine; Client-Server Solution; Peer-to-Peer Solution

Reference Service Implementations: Reference Implementations of Basic Services; Reference Implementations of Application Services; Reference Implementations of Other Framework Components

Interfaces: APIs for Algorithms and Converters; APIs for Basic Services; APIs for Application Services; APIs for Other Framework Components

Templates



Application Designers Need to Know

Framework Level

- ❑ Use all implementations of algorithms and converters
- ❑ Use all implementations on the service layer
- ❑ Concentrate on application solutions
- ❑ Use or refer to the reference implementations of an application

Applications or Services: Network Workbench Tool, Biological Networks Portal

Reference Application Solutions: Reference GUI, Web Front-end, Scripting Engine, Client-Server Solution, Peer-to-Peer Solution

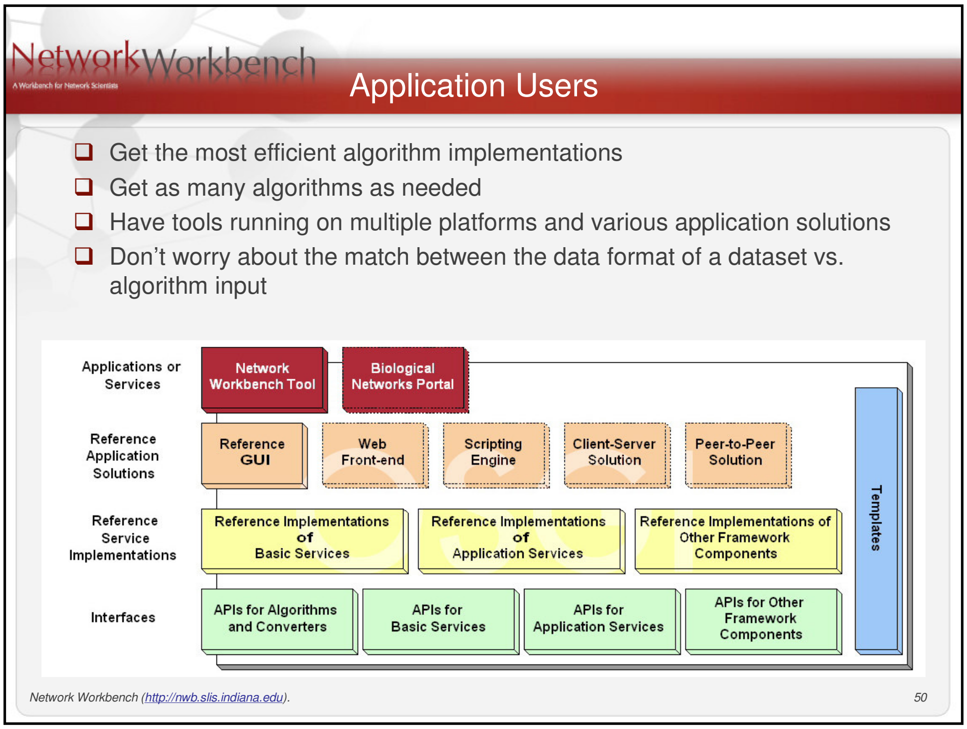
Reference Service Implementations: Reference Implementations of Basic Services, Reference Implementations of Application Services, Reference Implementations of Other Framework Components

Interfaces: APIs for Algorithms and Converters, APIs for Basic Services, APIs for Application Services, APIs for Other Framework Components

Templates

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

49



Application Users

- ❑ Get the most efficient algorithm implementations
- ❑ Get as many algorithms as needed
- ❑ Have tools running on multiple platforms and various application solutions
- ❑ Don't worry about the match between the data format of a dataset vs. algorithm input

Applications or Services: Network Workbench Tool, Biological Networks Portal

Reference Application Solutions: Reference GUI, Web Front-end, Scripting Engine, Client-Server Solution, Peer-to-Peer Solution

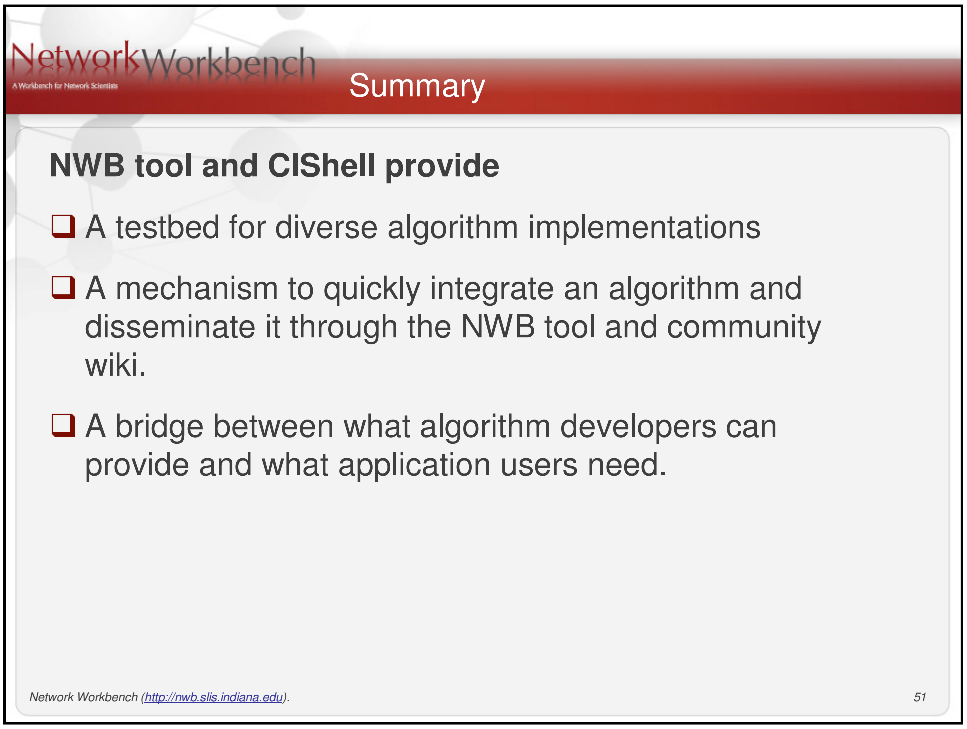
Reference Service Implementations: Reference Implementations of Basic Services, Reference Implementations of Application Services, Reference Implementations of Other Framework Components

Interfaces: APIs for Algorithms and Converters, APIs for Basic Services, APIs for Application Services, APIs for Other Framework Components

Templates

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

50

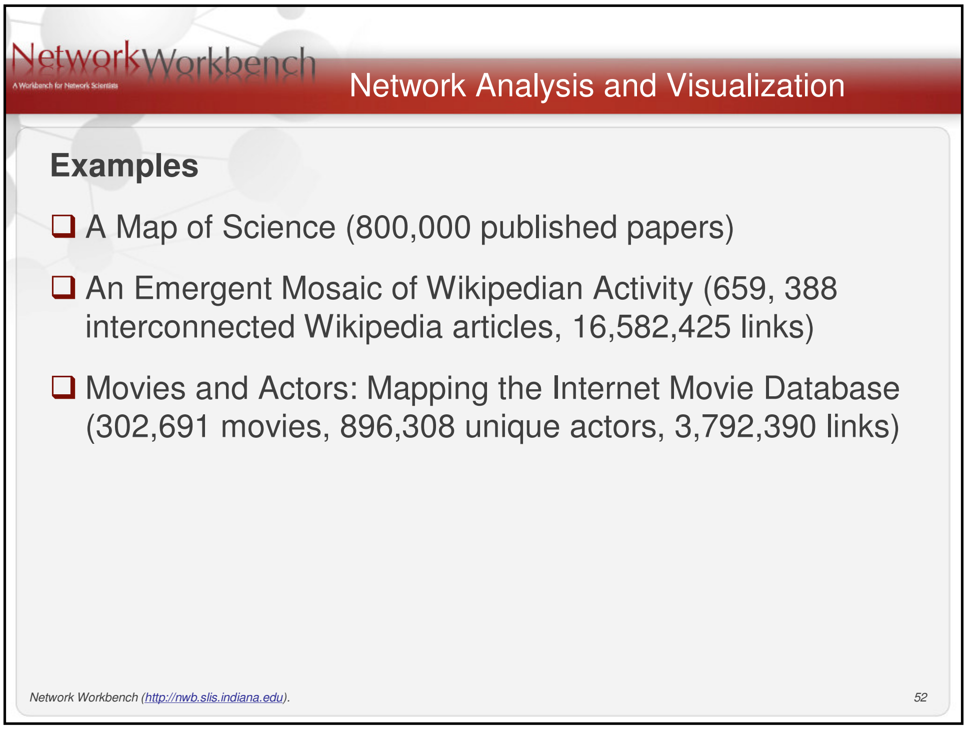


Summary

NWB tool and CIShell 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 algorithm developers can provide and what application users need.

Network Workbench (http://nwb.slis.indiana.edu). 51

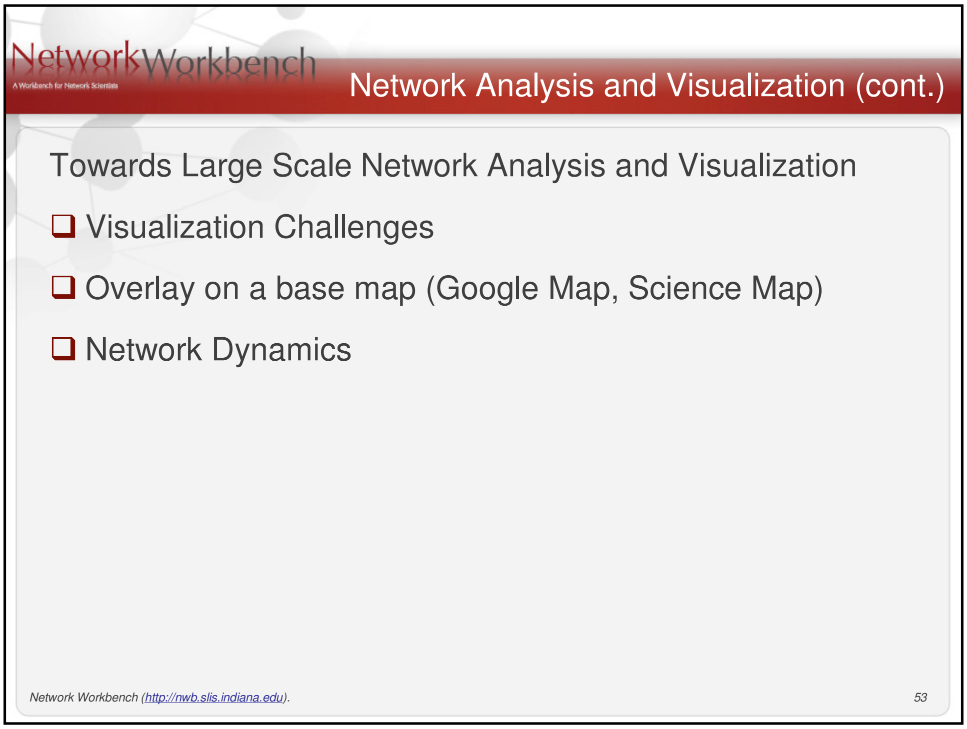


Network Analysis and Visualization

Examples

- ❑ A Map of Science (800,000 published papers)
- ❑ An Emergent Mosaic of Wikipedian Activity (659, 388 interconnected Wikipedia articles, 16,582,425 links)
- ❑ Movies and Actors: Mapping the Internet Movie Database (302,691 movies, 896,308 unique actors, 3,792,390 links)

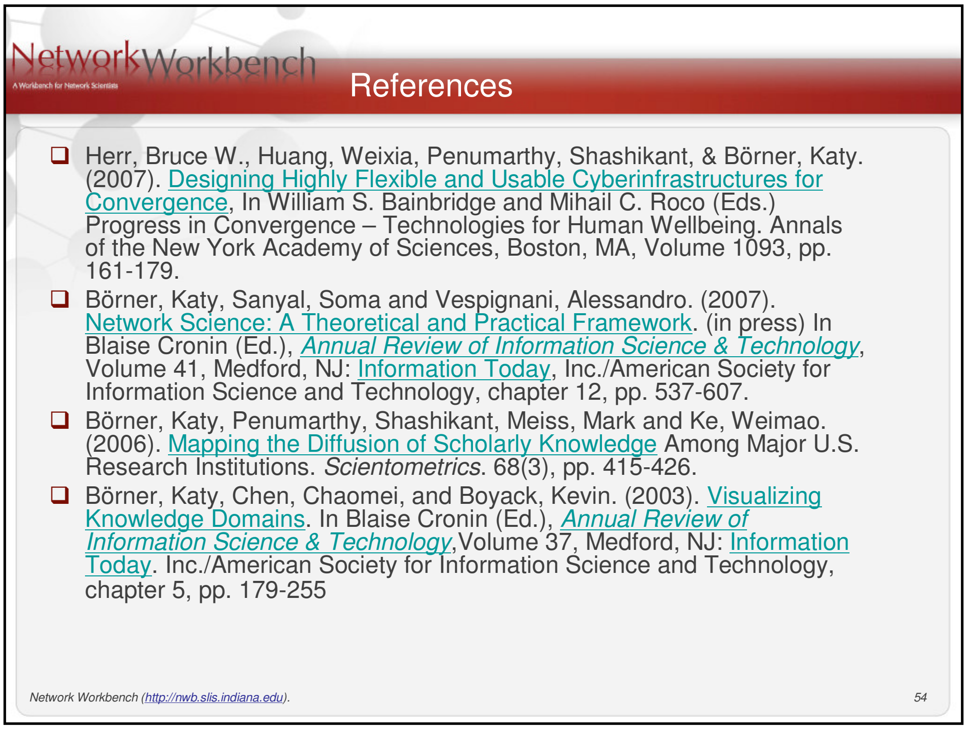
Network Workbench (http://nwb.slis.indiana.edu). 52



Network Analysis and Visualization (cont.)

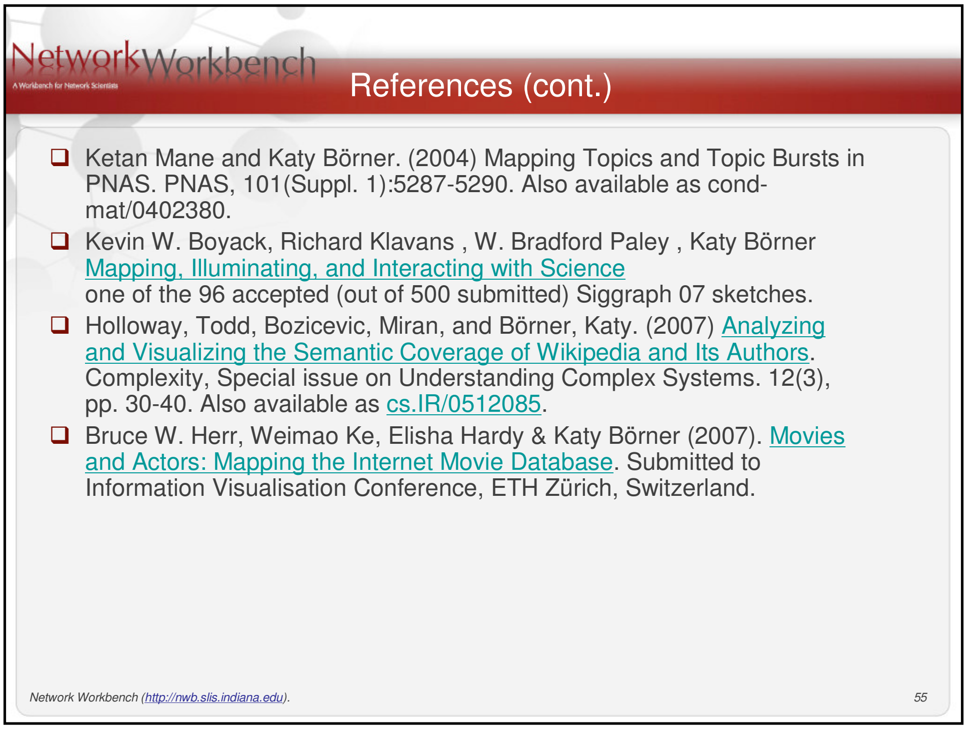
Towards Large Scale Network Analysis and Visualization

- Visualization Challenges
- Overlay on a base map (Google Map, Science Map)
- Network Dynamics



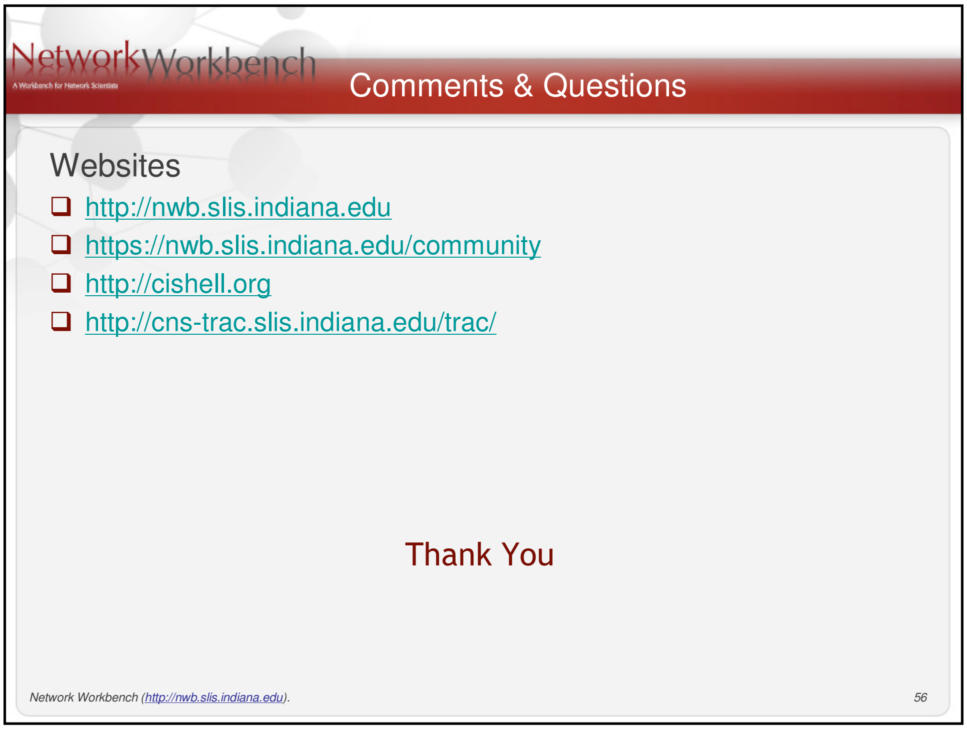
References

- 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.
- Börner, Katy, Sanyal, Soma and Vespignani, Alessandro. (2007). Network Science: A Theoretical and Practical Framework. (in press) In Blaise Cronin (Ed.), *Annual Review of Information Science & Technology*, Volume 41, Medford, NJ: Information Today, Inc./American Society for Information Science and Technology, chapter 12, pp. 537-607.
- Börner, Katy, Penumarthy, Shashikant, Meiss, Mark and Ke, Weimao. (2006). Mapping the Diffusion of Scholarly Knowledge Among Major U.S. Research Institutions. *Scientometrics*. 68(3), pp. 415-426.
- 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



References (cont.)

- ❑ Ketan Mane and Katy Börner. (2004) Mapping Topics and Topic Bursts in PNAS. PNAS, 101(Suppl. 1):5287-5290. Also available as cond-mat/0402380.
- ❑ Kevin W. Boyack, Richard Klavans , W. Bradford Paley , Katy Börner Mapping, Illuminating, and Interacting with Science one of the 96 accepted (out of 500 submitted) Siggraph 07 sketches.
- ❑ Holloway, Todd, Bozicevic, Miran, and Börner, Katy. (2007) Analyzing and Visualizing the Semantic Coverage of Wikipedia and Its Authors. Complexity, Special issue on Understanding Complex Systems. 12(3), pp. 30-40. Also available as cs.IR/0512085.
- ❑ Bruce W. Herr, Weimao Ke, Elisha Hardy & Katy Börner (2007). Movies and Actors: Mapping the Internet Movie Database. Submitted to Information Visualisation Conference, ETH Zürich, Switzerland.



Comments & Questions

Websites

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

Thank You