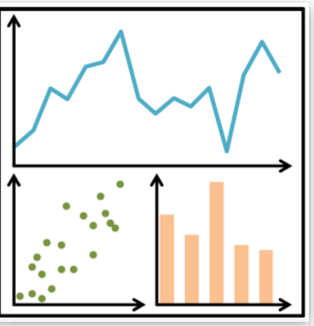
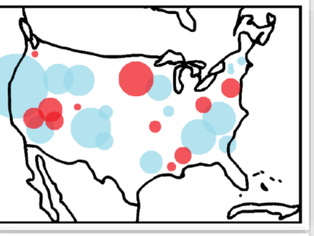
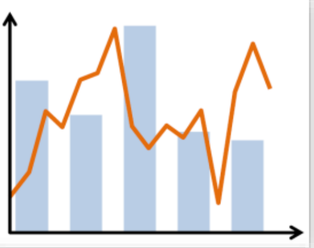
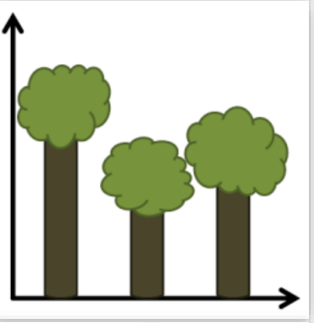
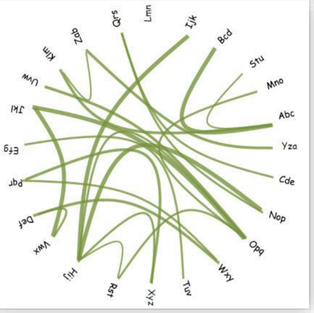






Lecture 14: Trees and Networks

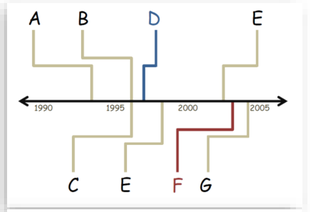
CS 7250

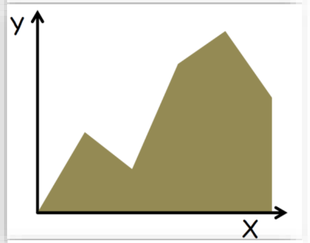
SPRING 2021

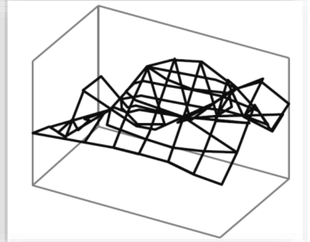
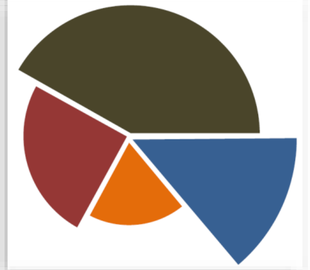
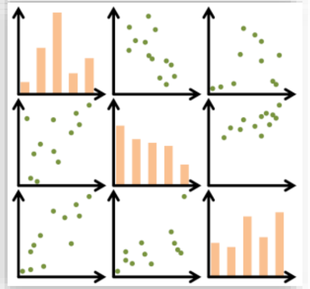
Prof. Cody Dunne

NORTHEASTERN UNIVERSITY

Slides and inspiration from Michelle Borkin, Krzysztof Gajos, Hanspeter Pfister, Miriah Meyer, Jonathan Schwabish, and David Sprague





Checking In

Reading Quiz

Quiz — Trees and Networks

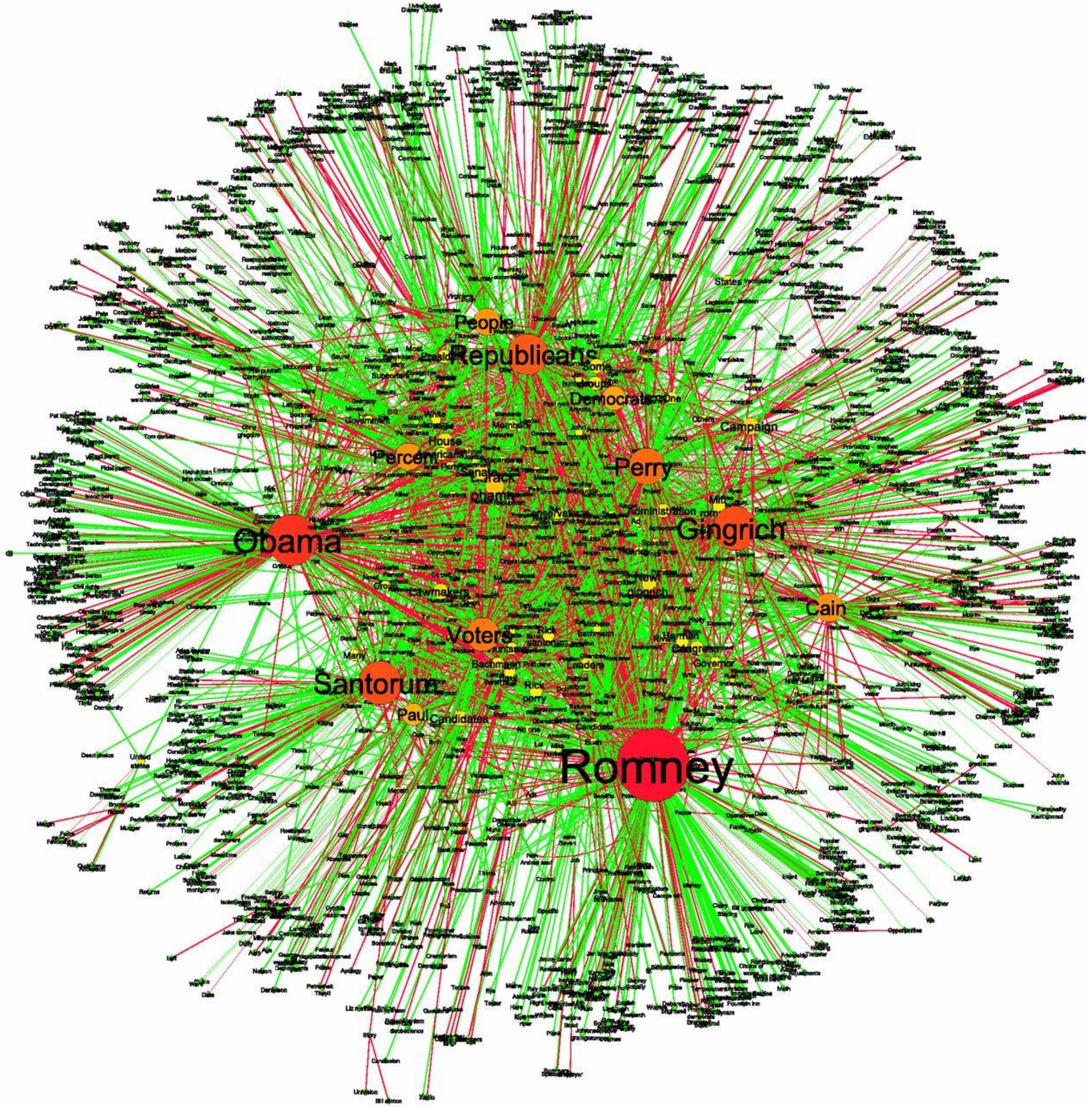
Password: XXXXXXXX

Trees & (Mainly) Networks

Goals for Today

- Learn the definition of a network (including node, edge)
- Learn the definition of a tree
- Learn common visual encoding techniques for network data (i.e., node-link diagram, adjacency matrix), and the advantages of each one.

Hall of Fame or Hall of Shame



US presidential election network for 2012 primaries.

- Nodes: key entities from noun phrases. Sized by degree.

- Edges: relationships from verbs. Colored by positive (green) and negative (red) weights.

Sudhahar et al., 2015

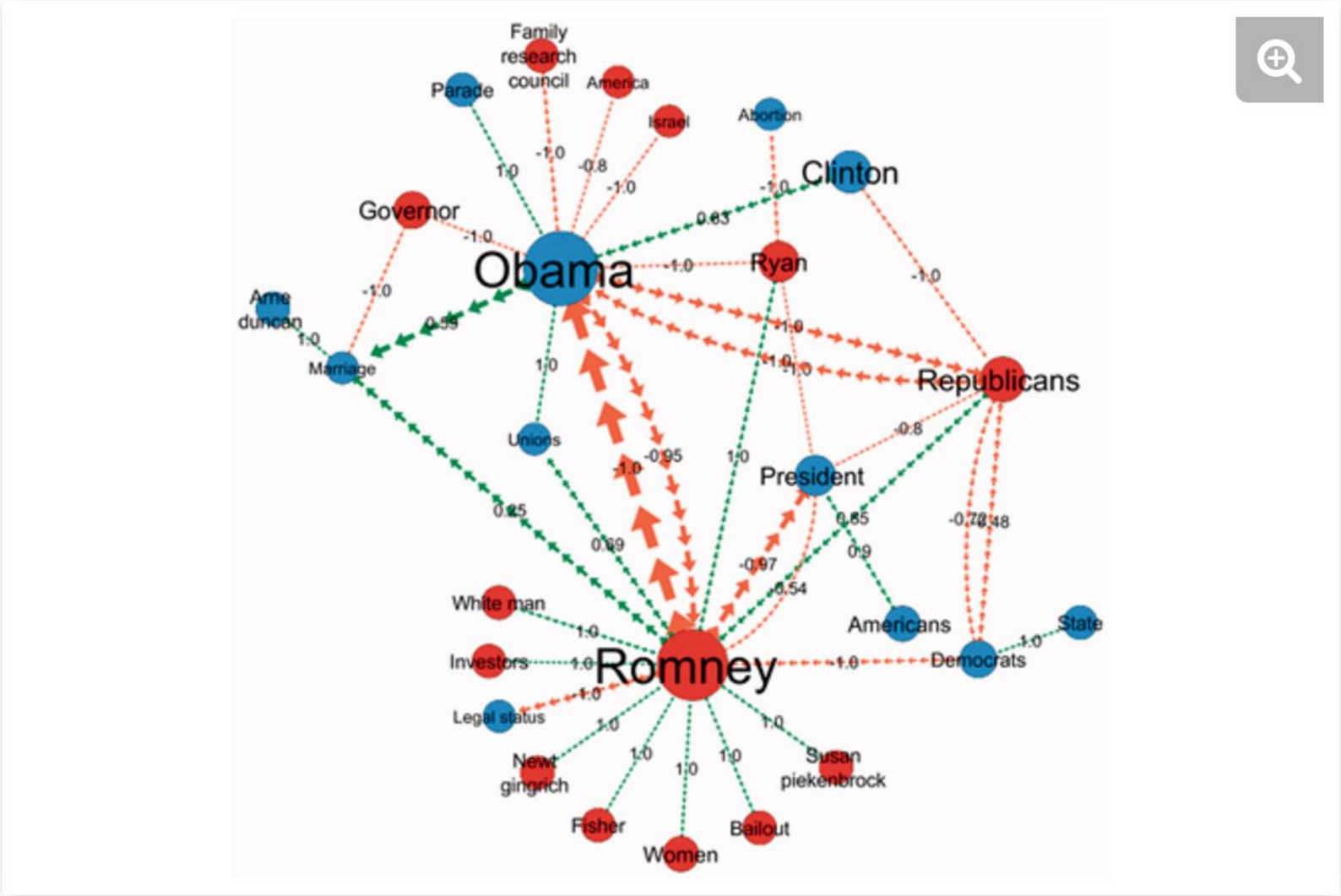


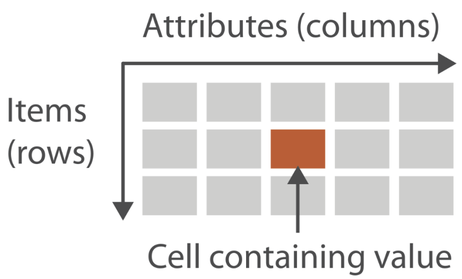
Figure 4.

Download figure | **Open in new tab** | **Download powerpoint**

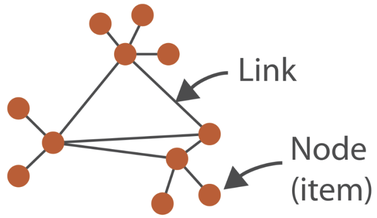
A subset of the election network, coloured by partitioning it via the first eigenvalue of the symmetrised adjacency matrix (see Appendix A8). Note that the split captures well the expected distinction between the Republican (red) and Democratic (blue) camps. The orange and green links show negative and positive relations between entities.

Dataset Types

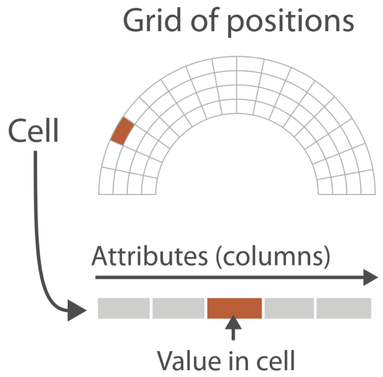
Tables



Networks

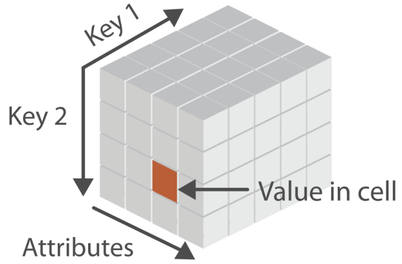


Fields (Continuous)

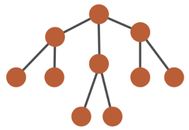


Geometry (Spatial)

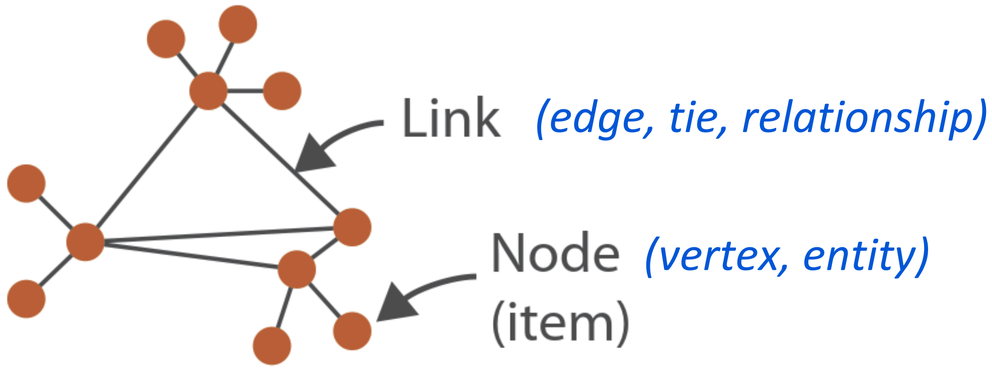
Multidimensional Table



Trees

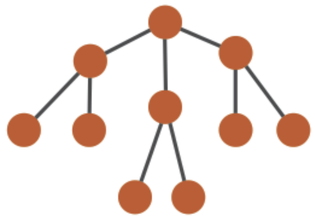


➔ Networks *(graphs)*



Network = entities and relationships between them

➔ *Trees*

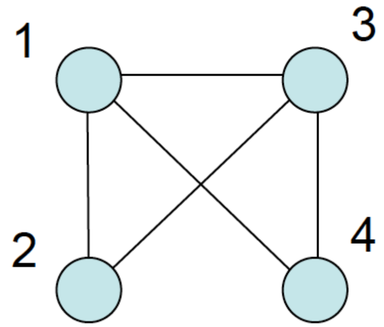


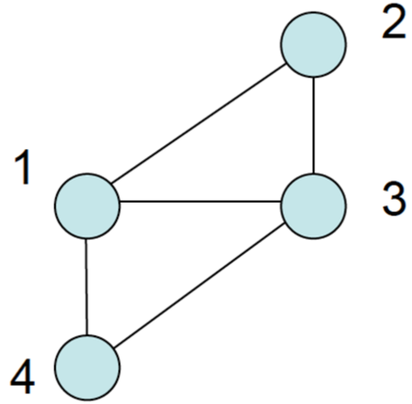
Tree = *undirected*, *connected*, *acyclic* network

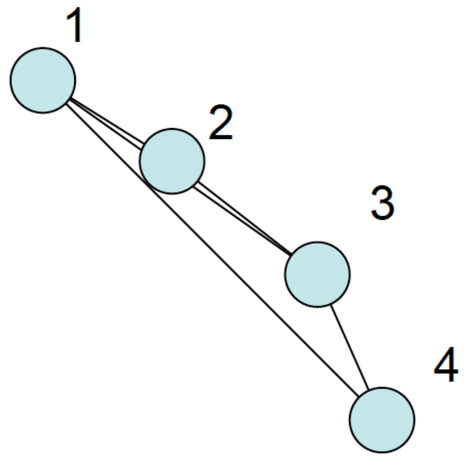
Networks

- A network *G* consists of a set of nodes *N* and a set of edges *E*
- An edge $e_{n1,n2} \in E$ connects two nodes *n1*, *n2* ∈ *N*
- E.g., *G* = {1,2,3,4}, *E* = {(*1*,*2*),(*1*,*3*), (*2*,*3*),(*3*,*4*),(*4*,*1*)}

Note all the same network, just different layouts!

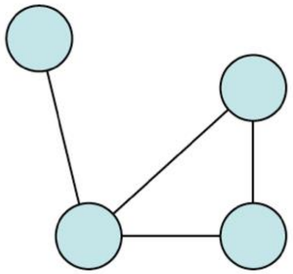




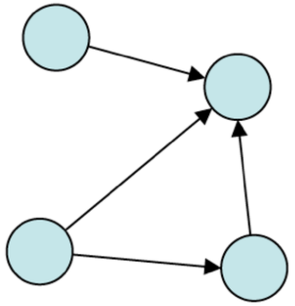


Modified from slide by Frank van Ham

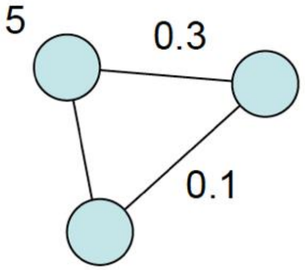
A bunch of definitions



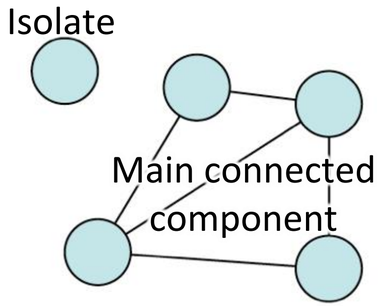
An undirected graph



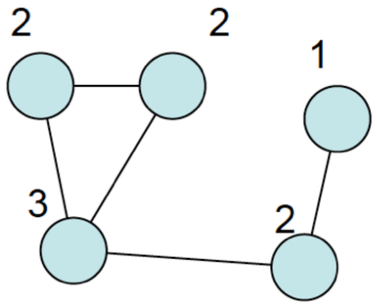
A directed graph



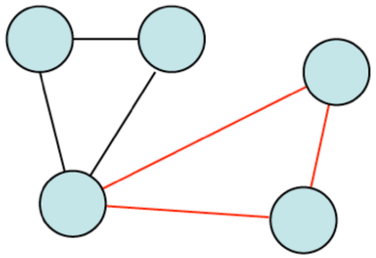
Weighted



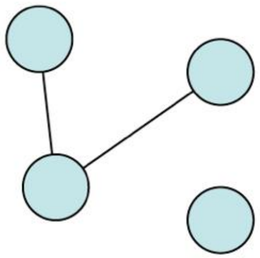
Unconnected



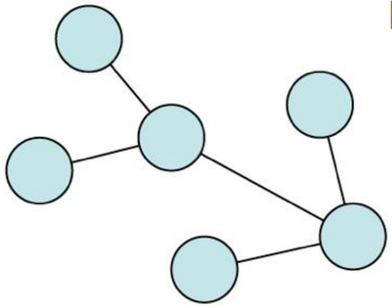
Node degrees



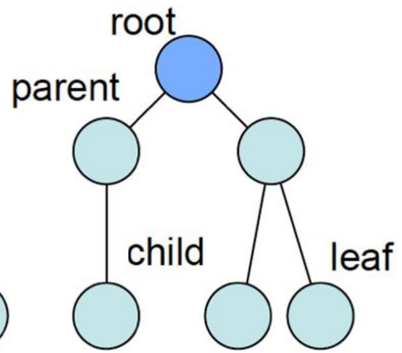
A cycle



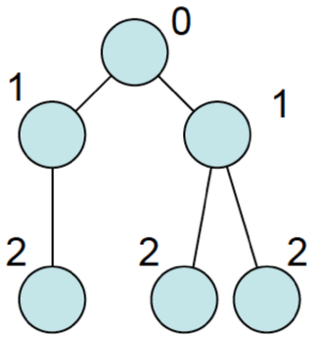
An acyclic graph



A connected acyclic graph, a.k.a. a **tree**



A rooted tree or hierarchy



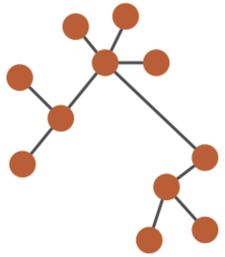
Node depths

Arrange Networks and Trees

Node–Link Diagrams

Connection Marks

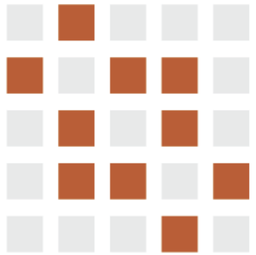
✔ NETWORKS ✔ TREES



Adjacency Matrix

Derived Table

✔ NETWORKS ✔ TREES



Enclosure

Containment Marks

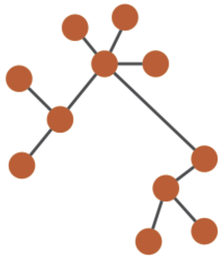
✘ NETWORKS ✔ TREES

"Treemap"

Node–Link Diagrams

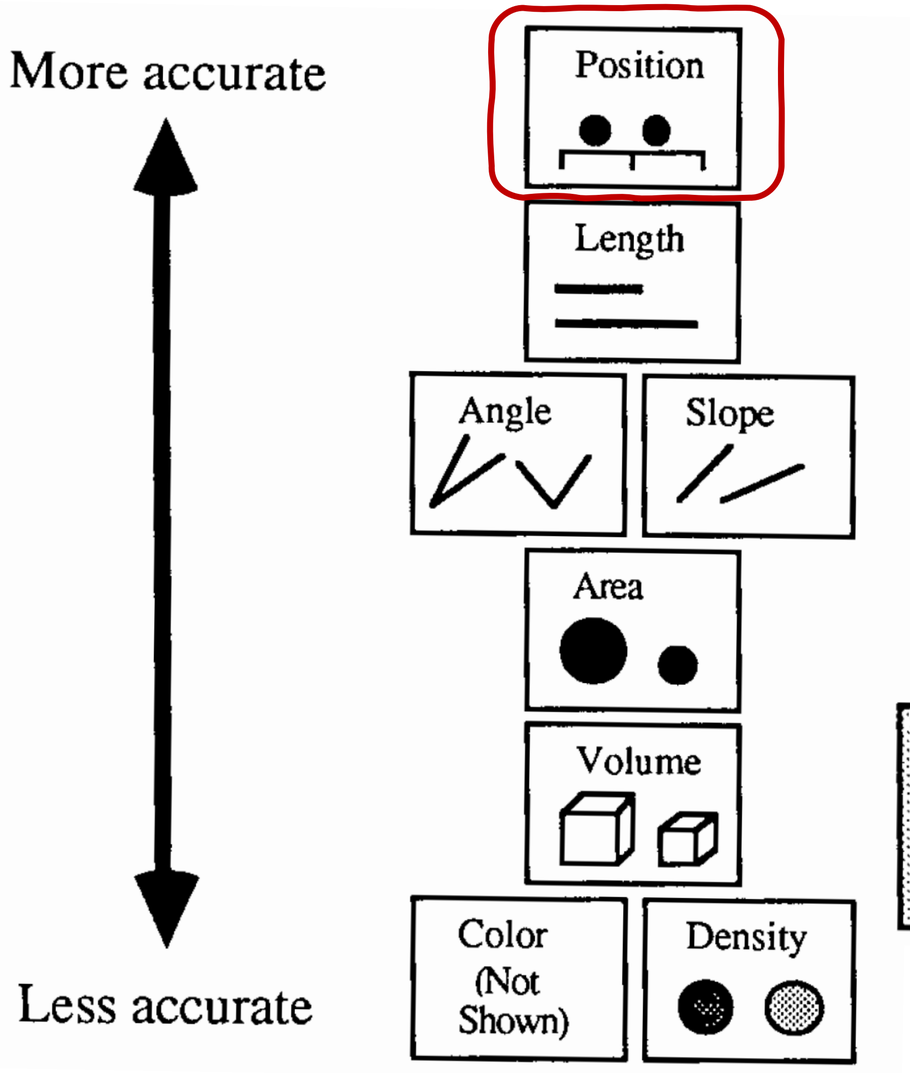
Connection Marks

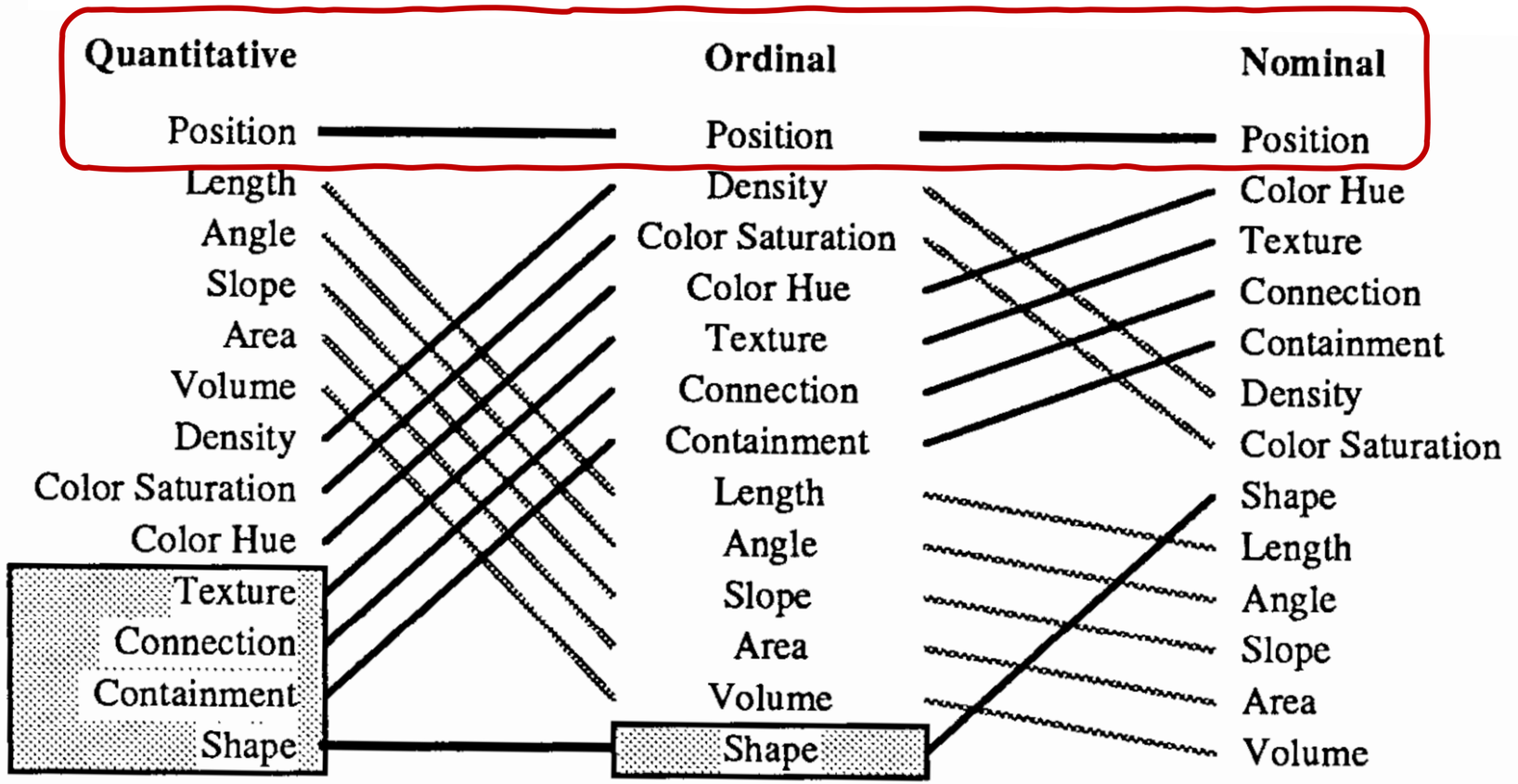
✔ NETWORKS ✔ TREES



- Primary concern is the *spatial layout* of nodes and edges, a.k.a. *graph drawing*
- The goal is often to effectively depict the graph structure for *topology-based tasks*:
  - connectivity, path-following
  - network distance
  - clustering
  - ordering (e.g., hierarchy level)
- But not always topology-based tasks. E.g., understanding attributes, statistics, metrics

Spatial Layout





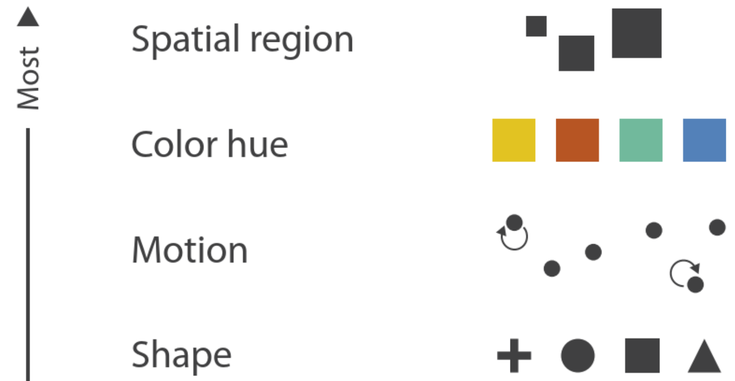
Mackinlay, 1986

Spatial Layout

Channels: Expressiveness Types and Effectiveness Ranks

Magnitude Channels: Ordered Attributes

- Position on common scale
- Position on unaligned scale
- Length (1D size)
- Tilt/angle
- Area (2D size)
- Depth (3D position)
- Color luminance
- Color saturation
- Curvature
- Volume (3D size)



Effectiveness: Most → Least

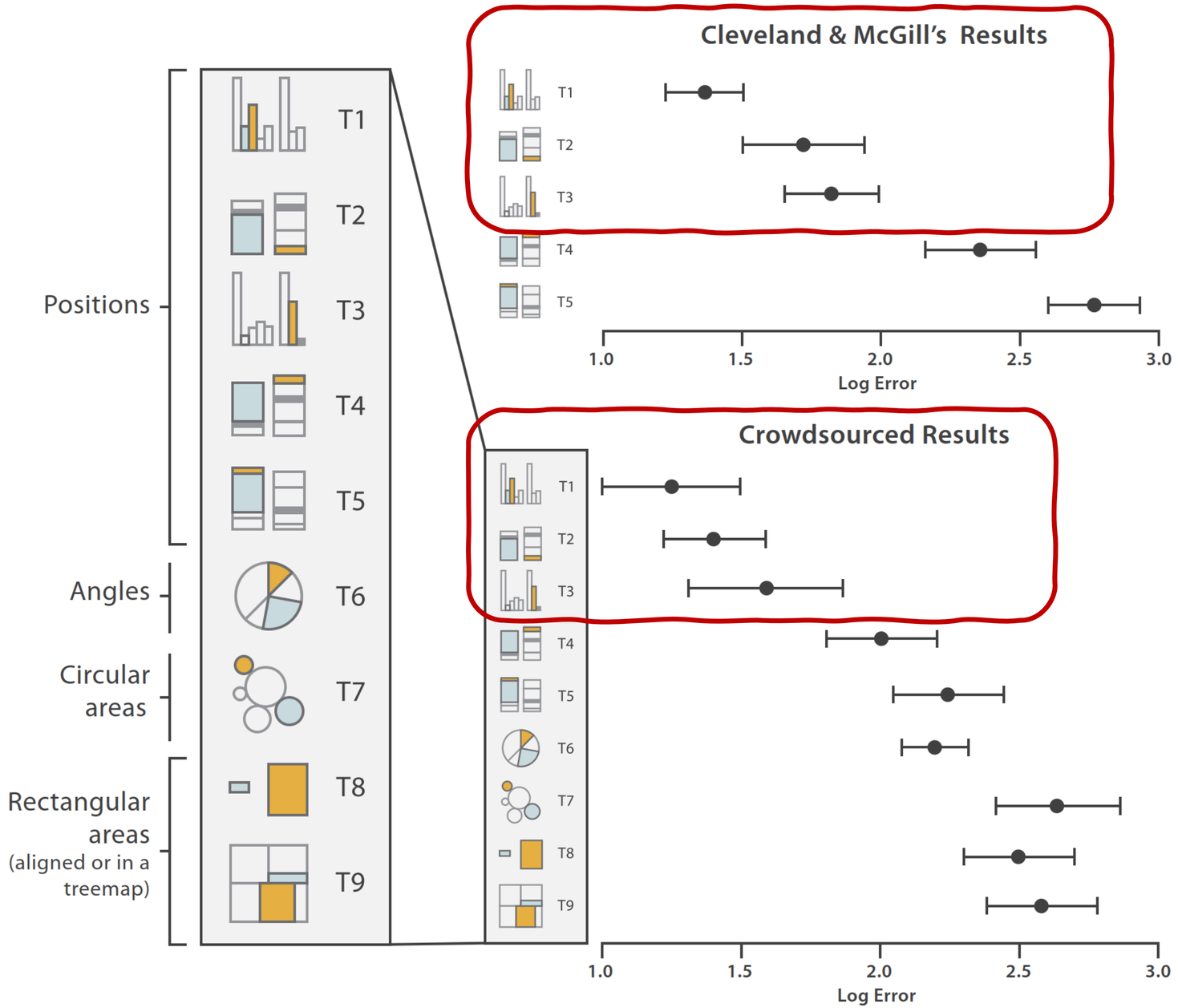
Color luminance, Color saturation: Same

Curvature, Volume (3D size): Same

Identity Channels: Categorical Attributes

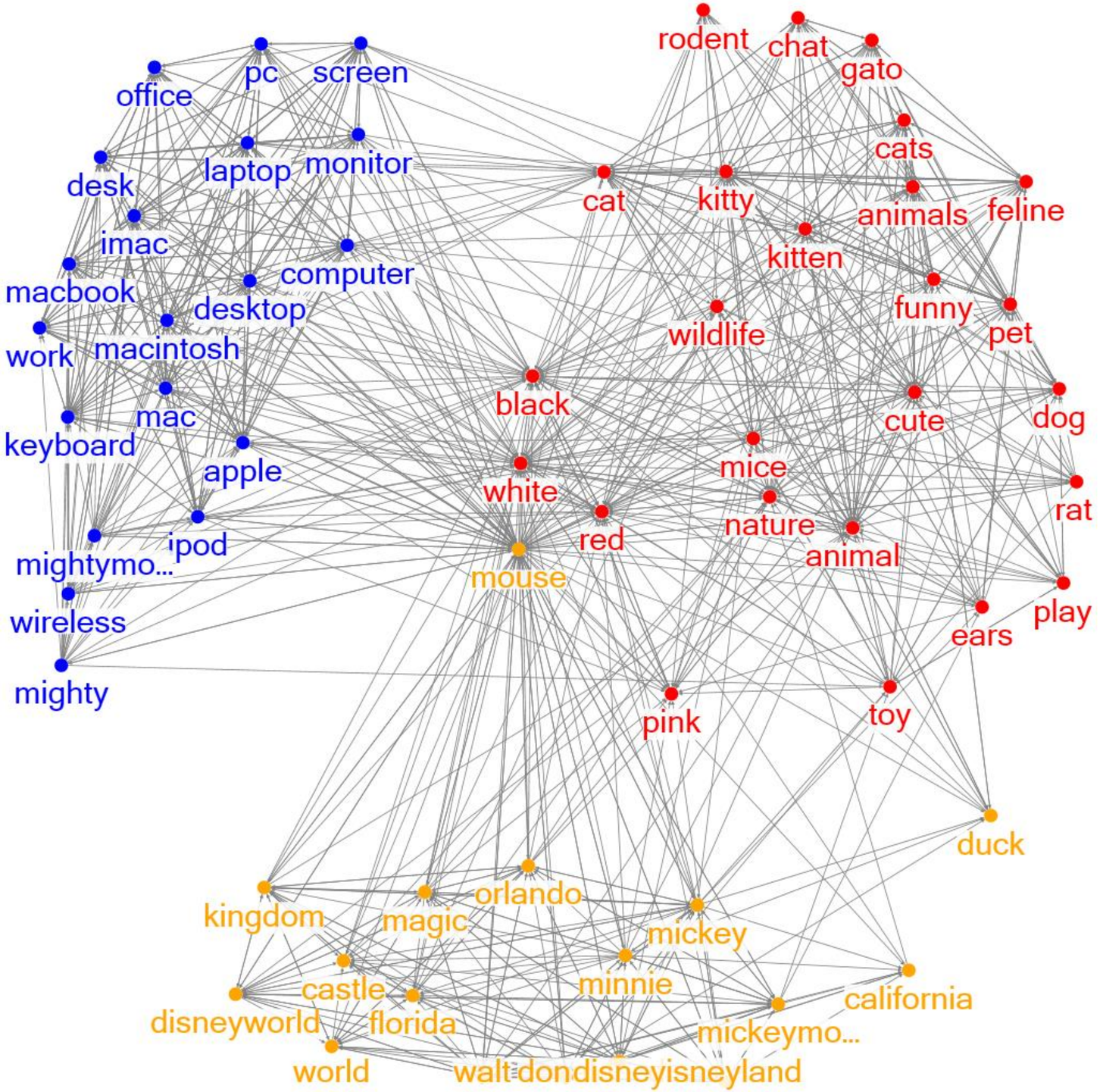
- Spatial region
- Color hue
- Motion
- Shape

Spatial Layout

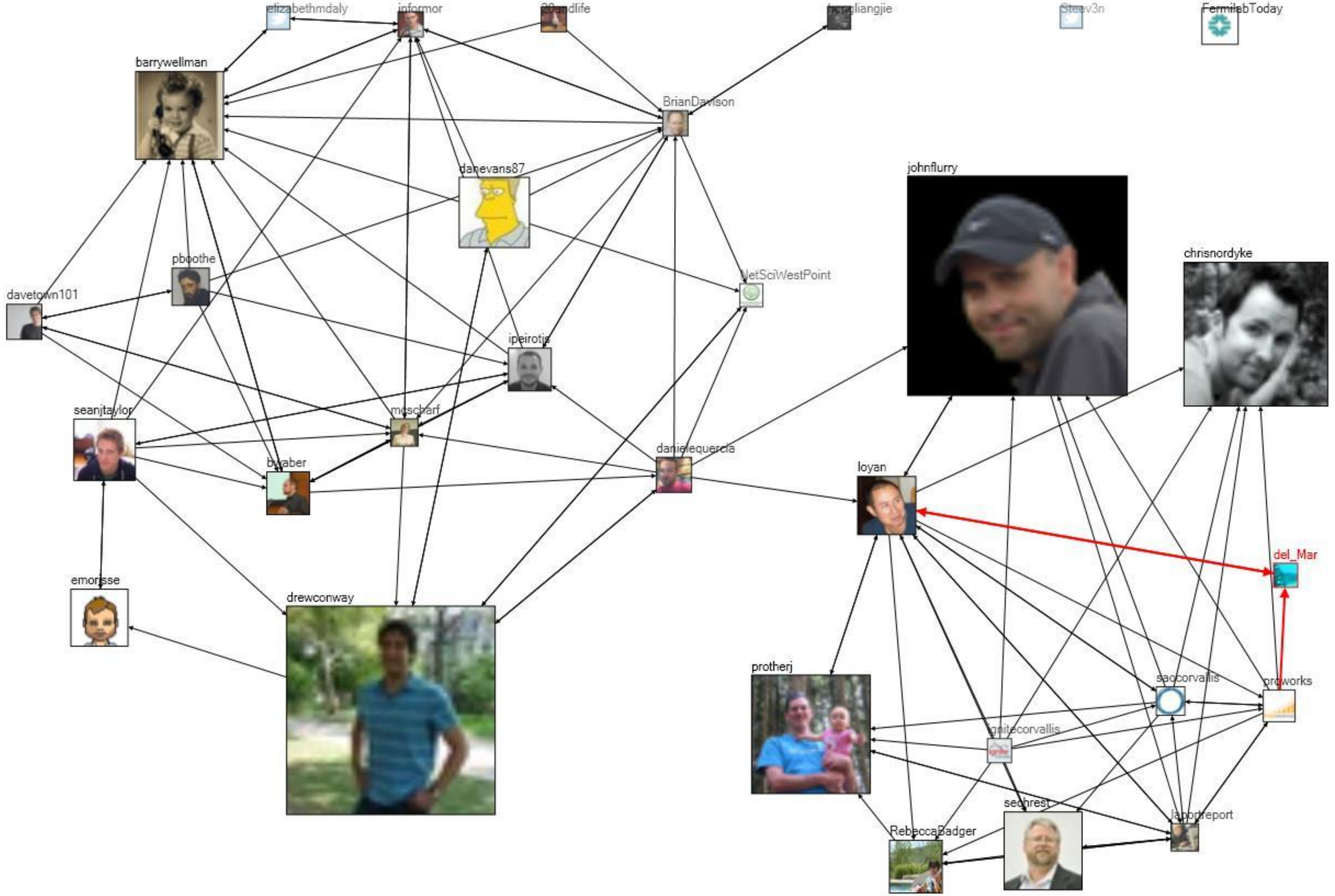


[Cleveland & McGill, 1984](#)
[Heer & Bostock (2010)](#)

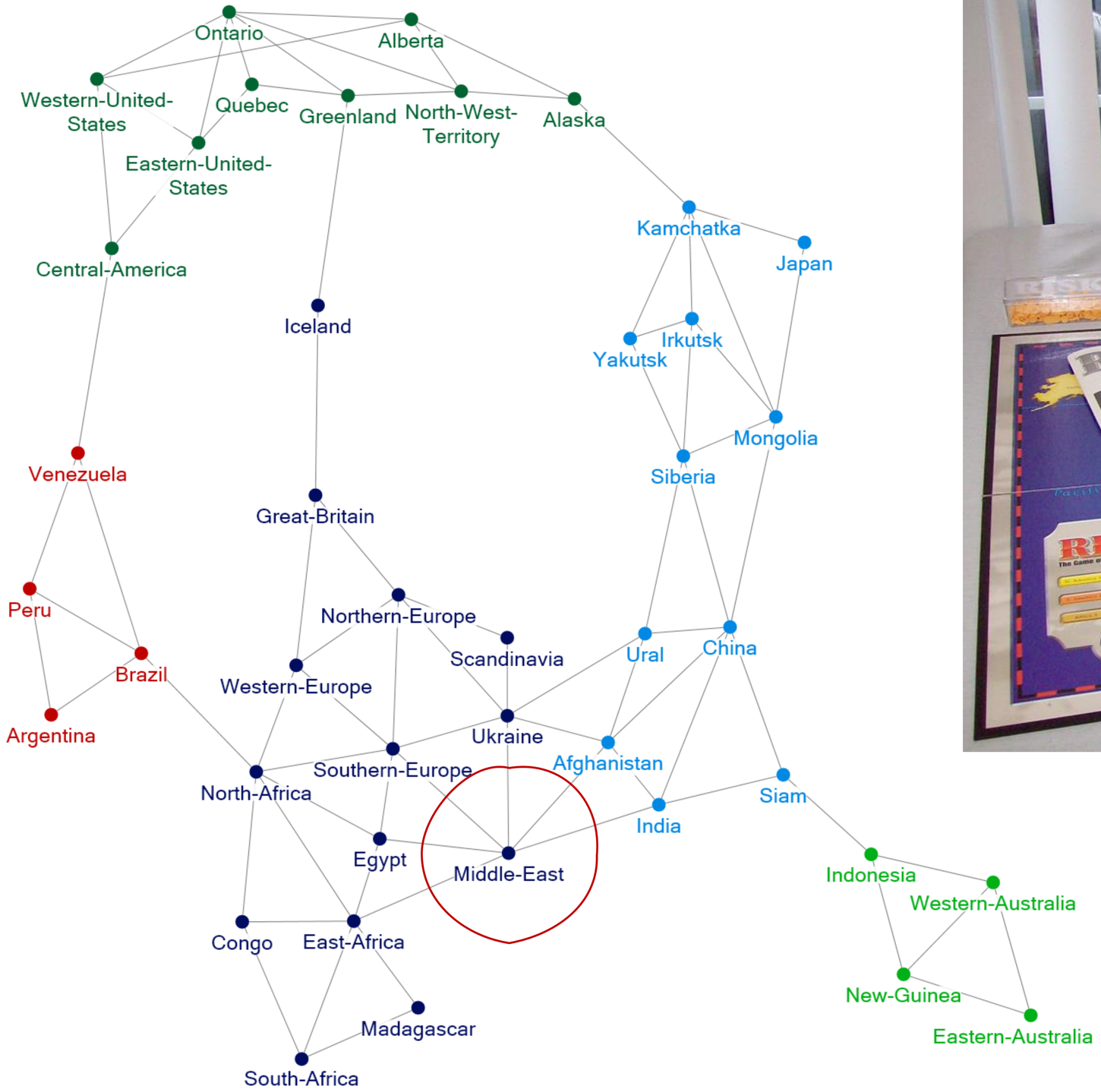
Flickr Query for “Mouse”



Tweets of the #Win09 Workshop



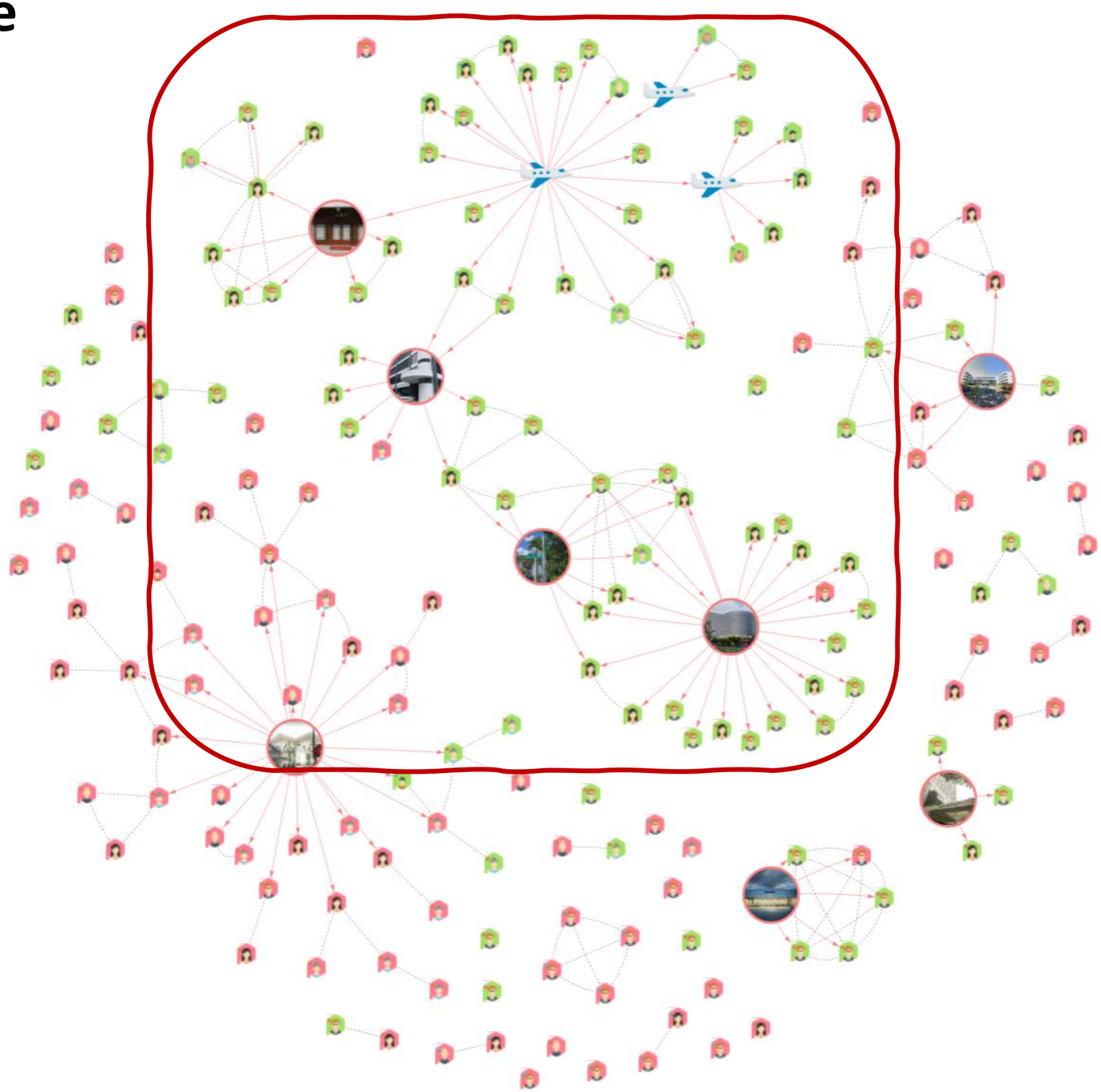

RISK
The Game of Global Domination
RISK
The Game of Global Domination
RISK


Ontario
Alberta
Western-United-States
Quebec
Greenland
North-West-Territory
Alaska
Eastern-United-States
Central-America
Kamchatka
Japan
Iceland
Irkutsk
Yakutsk
Mongolia
Siberia
Venezuela
Great-Britain
Peru
Northern-Europe
Scandinavia
Ural
China
Brazil
Western-Europe
Ukraine
Argentina
Afghanistan
Southern-Europe
North-Africa
Siam
India
Egypt
Middle-East
Indonesia
Western-Australia
Congo
East-Africa
New-Guinea
Madagascar
Eastern-Australia
South-Africa


RISK
The Game of Global Domination

Dashboard of the COVID-19 Virus Outbreak in Singapore

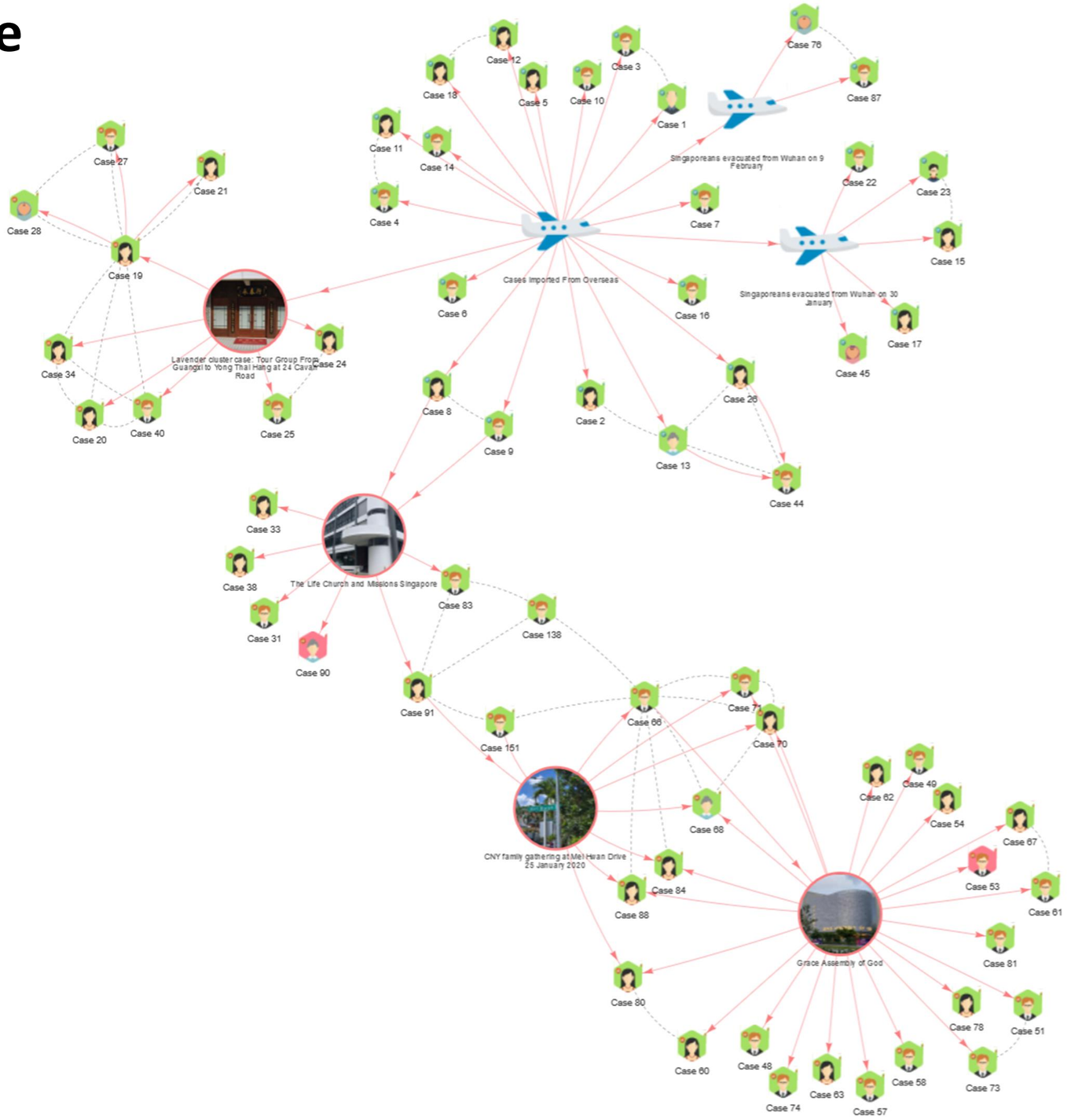
2020-01-21–03-12



[Upcode, 2020](#)

Dashboard of the COVID-19 Virus Outbreak in Singapore

2020-01-21–03-12



IBM **Watson** News Explorer recent news about current events investigate advanced [+]

Details 100 articles loaded

person

Barack Hussein Obama II (US /bəˈrɑːk huːˈseɪn oʊˈbɑːmə/; born August 4, 1961) is an American politician who is the 44th and current President of the United States. He is the first African American to hold the office and the first president born outside the continental United States. Born in Honolulu, Hawaii, Obama is a graduate of Columbia University and Harvard Law School, where he was president of the

California endures more wildfires, 1 sparked by a hot car

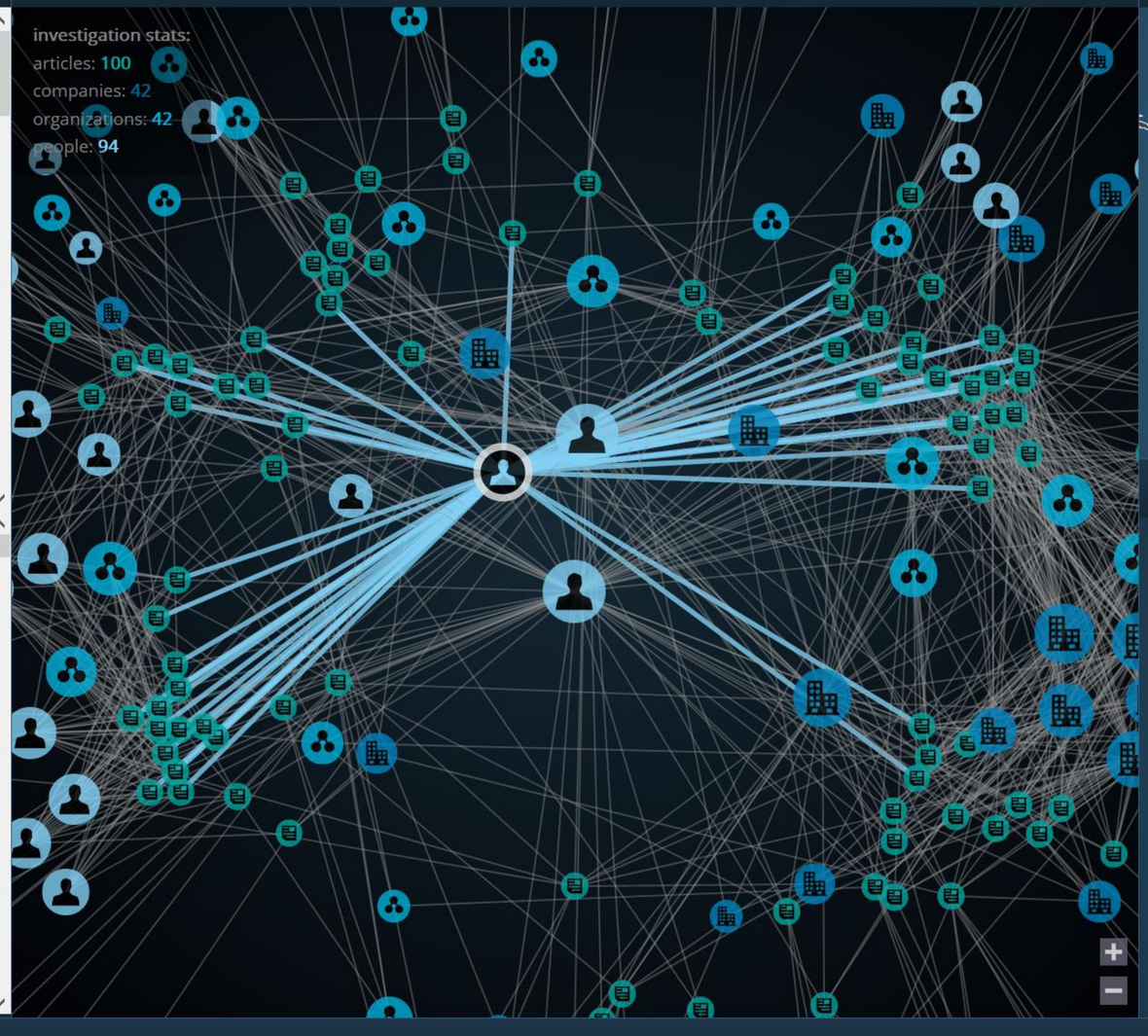
Mass shooting at Halloween party leaves at least 4 dead in California

US role in Syria grows more complex with Trump claim to oil

What is Dia de los Muertos and when is it celebrated?

Chicago girl, 7, shot while trick-or-treating, in critical condition

News Network show/hide: companies, organizations, people

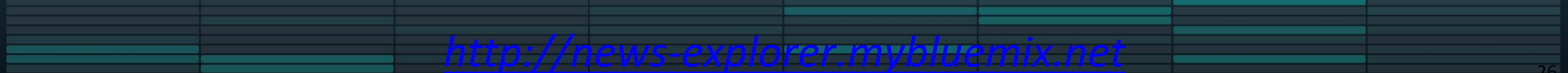


Locations 41 found, view in map list

Topics People Companies Organizations

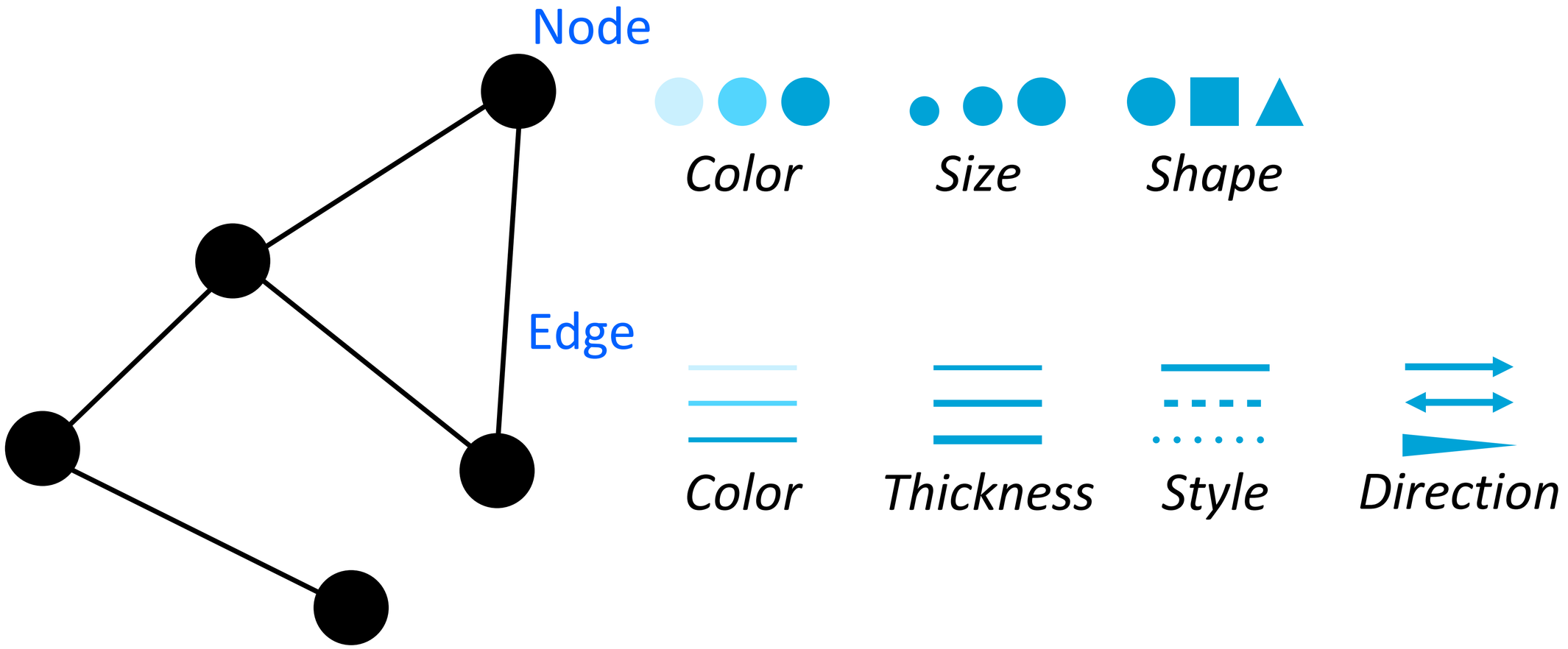
Timeline news articles across 7 days, 11 hrs, 2 min, 0 sec up to the current date: 11/1/2019

Powered By IBM Watson™



http://news-explorer.mybluemix.net

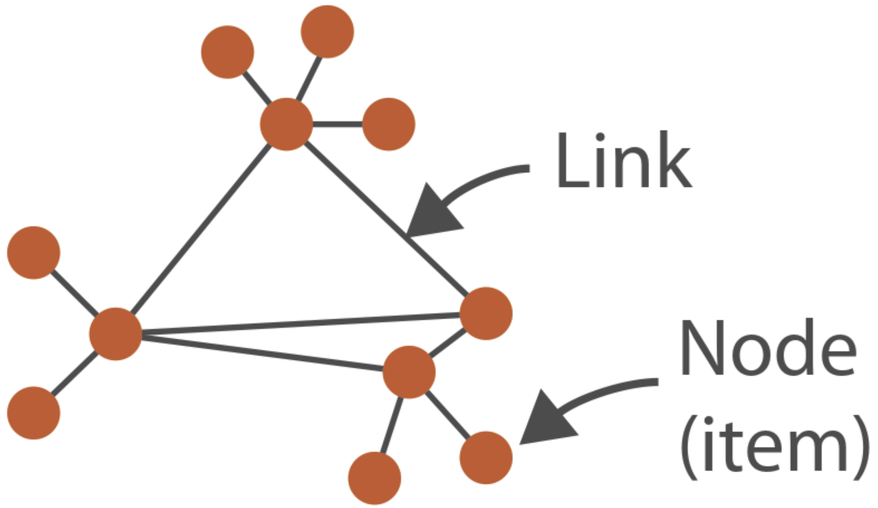
Node-Link Visualizations - Marks & Channels



Gestalt Principles: Grouping, Proximity, Connectedness

Node-Link Visualizations

- Nodes are distributed in space, connected by straight or curved lines
- Typical approach is to use 2D space to break apart breadth and depth
- Often space is used to communicate hierarchical orientation

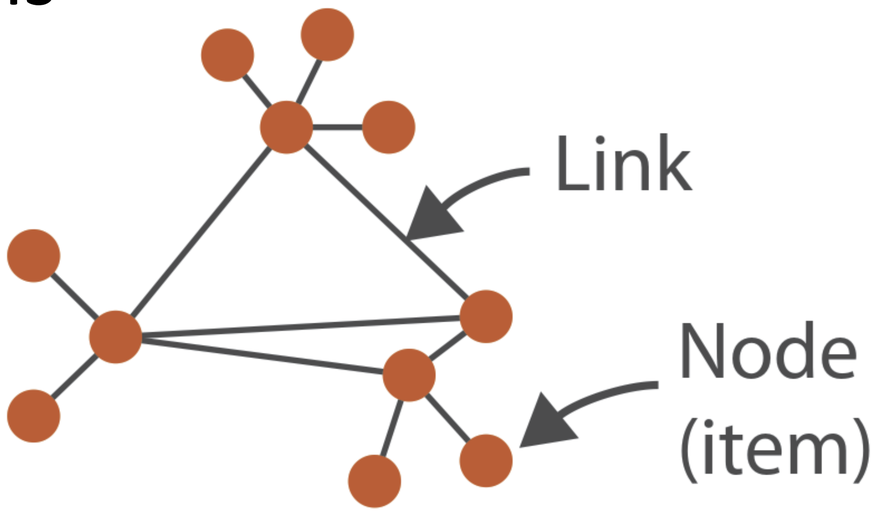


Slide based on Miriah Meyer

Node-Link Visualizations

Pros:

- understandable visual mapping
- can show overall structure, clusters, paths
- flexible, many variations



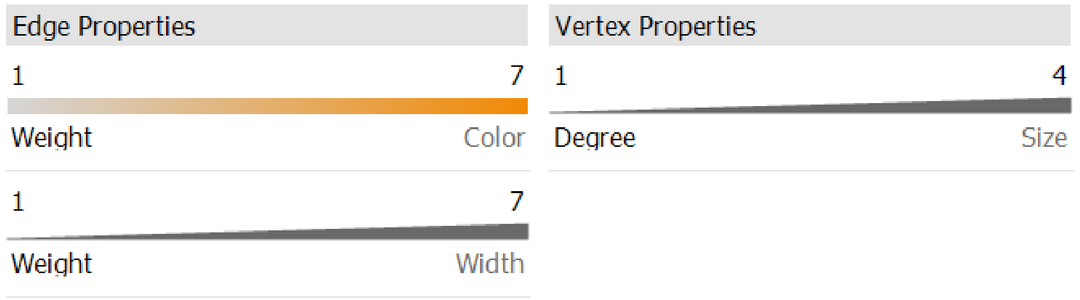
Cons:

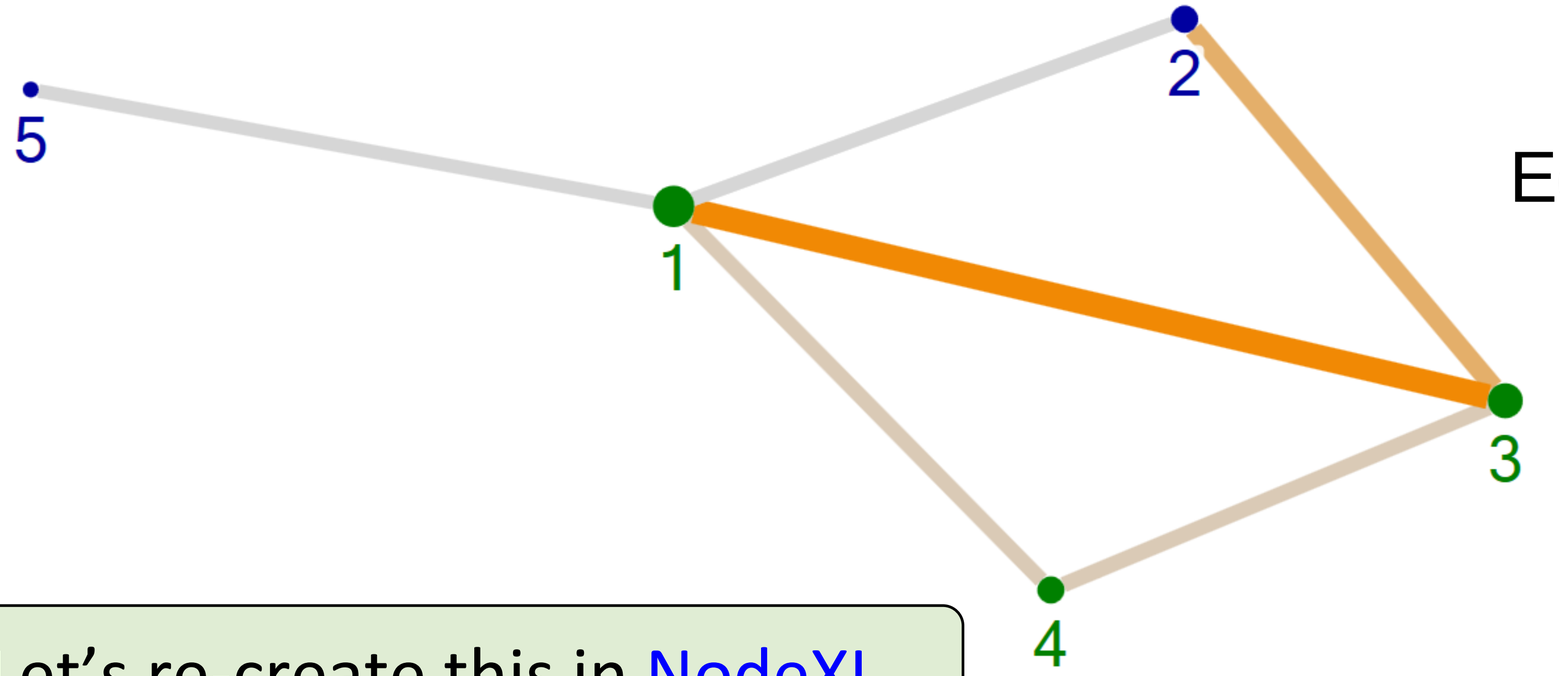
- automatic layout algorithm deficiencies
  - -time consuming to run
  - -non-deterministic results
  - -heuristics with sometimes poor results
- not good for dense graphs - hairball problem!

In-Class Drawing: Node-Link Visualization

~15 min

In-Class Drawing: Node-Link Visualization





Let's re-create this in NodeXL…

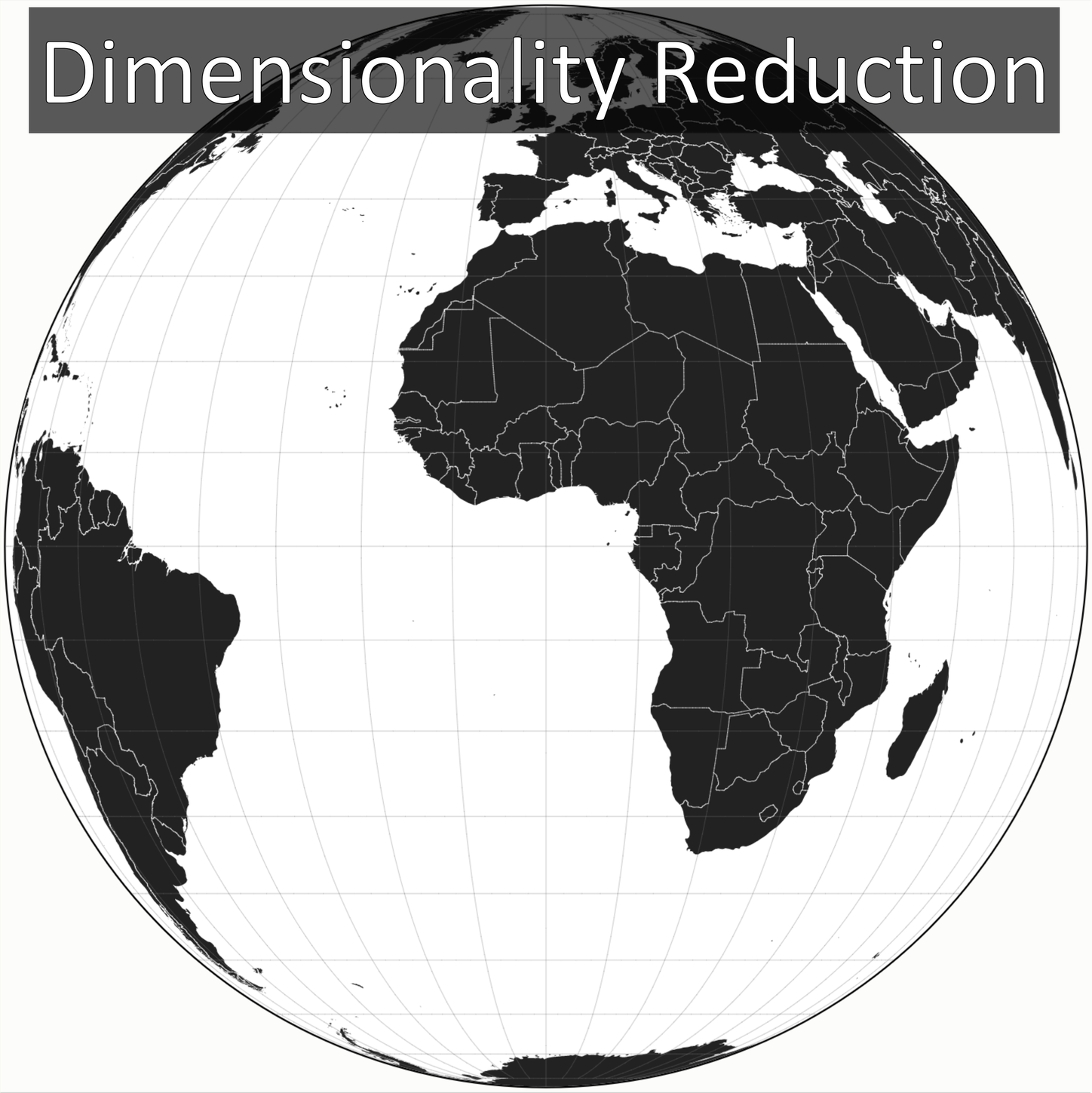
Nodes:

ID	Type
1	A
2	B
3	A
4	A
5	B

Edges:

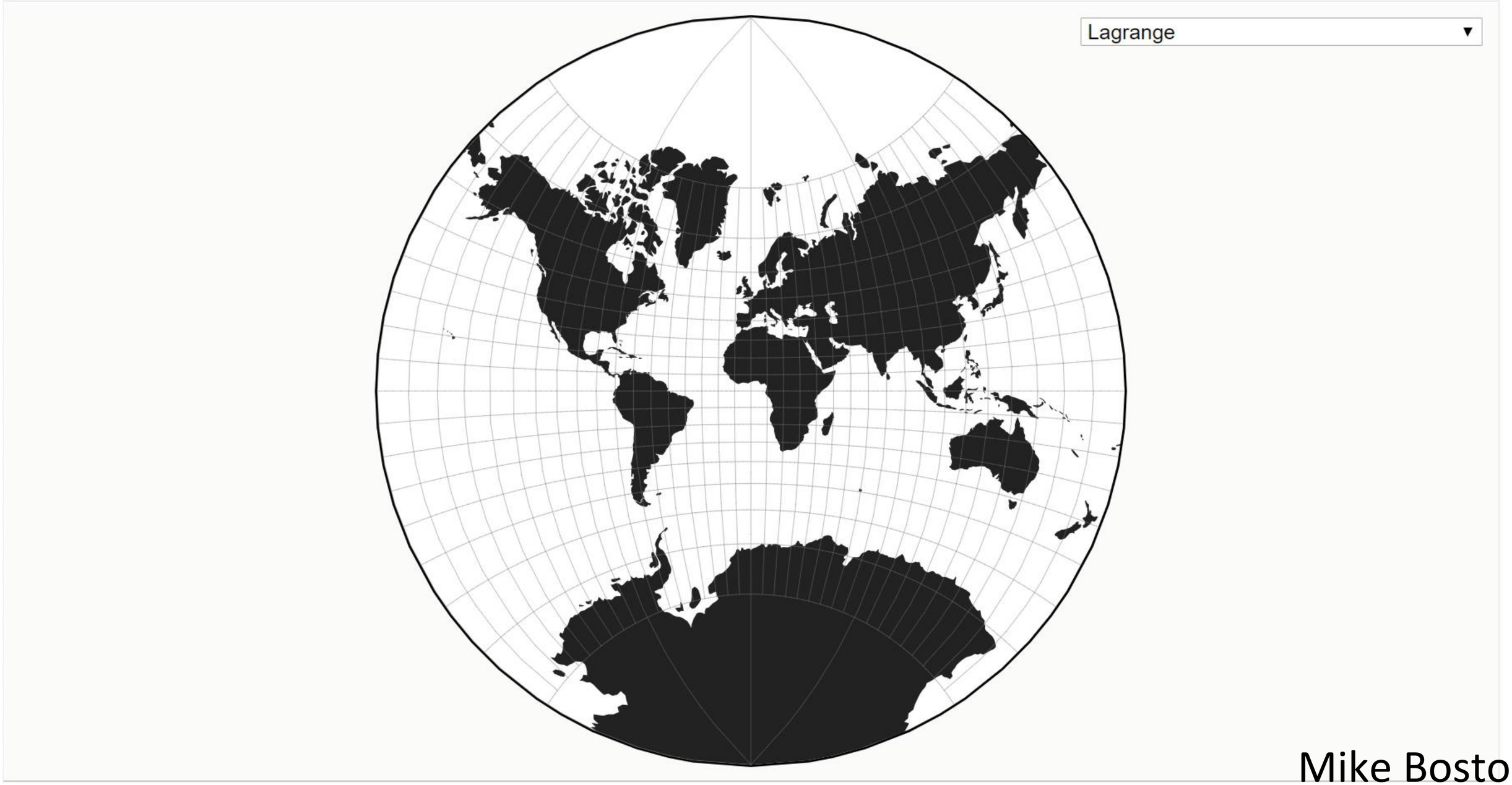
Source	Target	Weight
1	2	1
1	3	7
2	3	4
3	4	2
4	1	2
5	1	1

Dimensionality Reduction



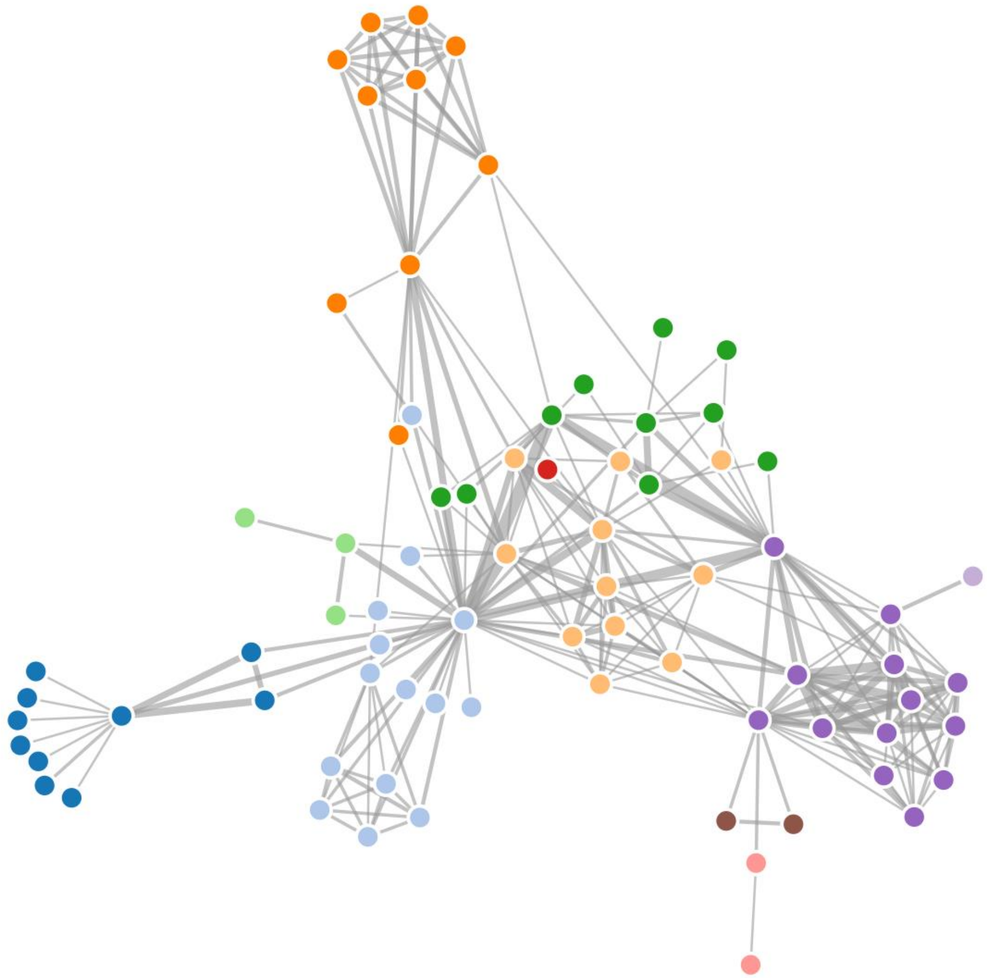
Mike Bostock

Projection Transitions



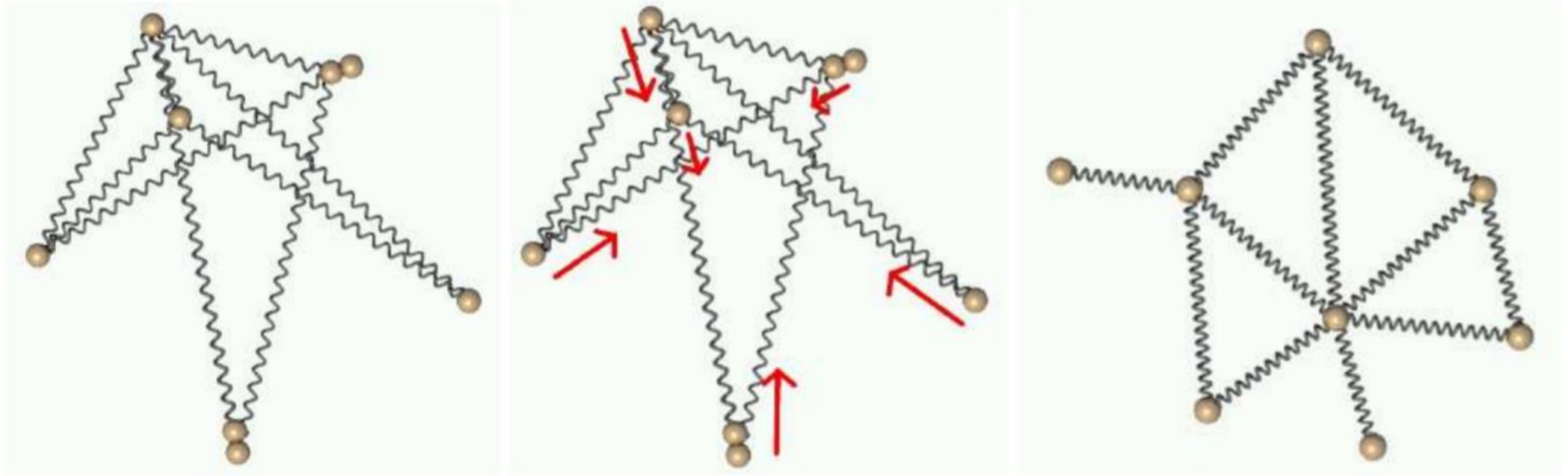
Mike Bostock

Layout Algorithm: D3 Force-Directed

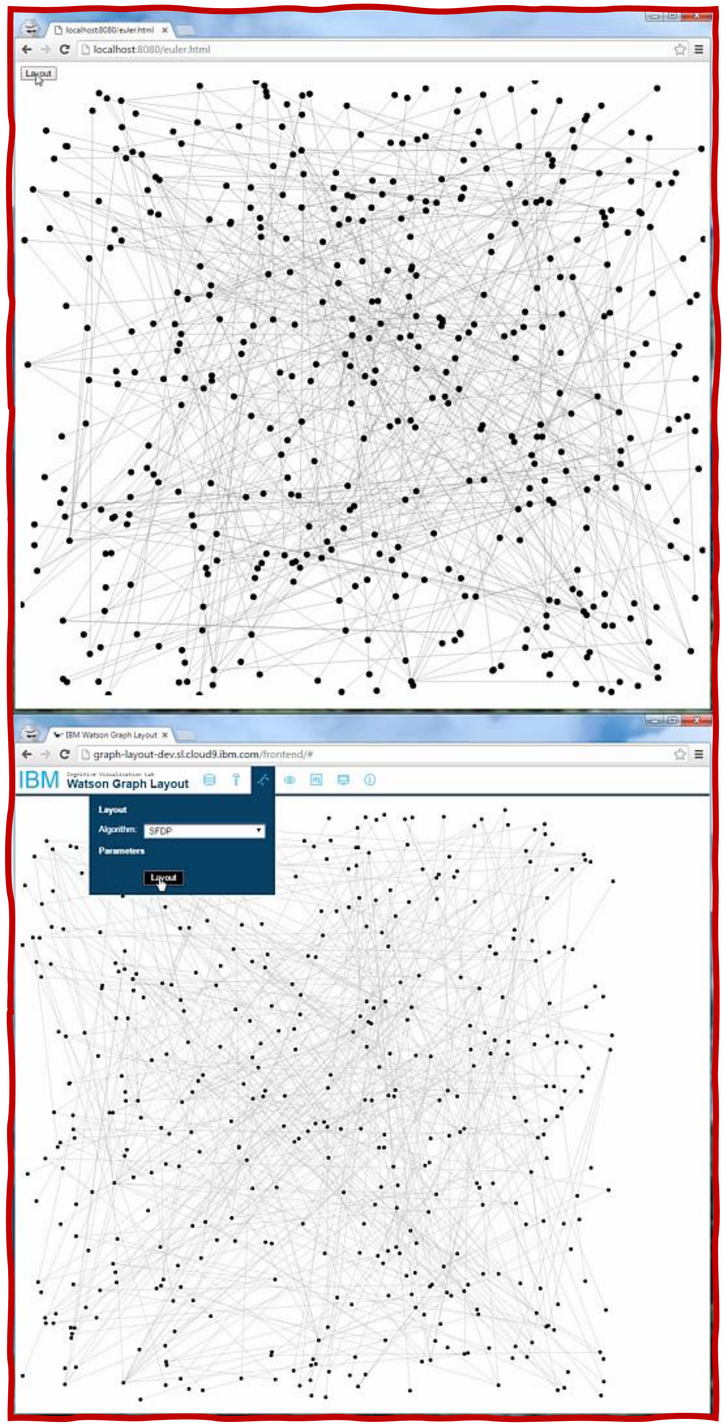


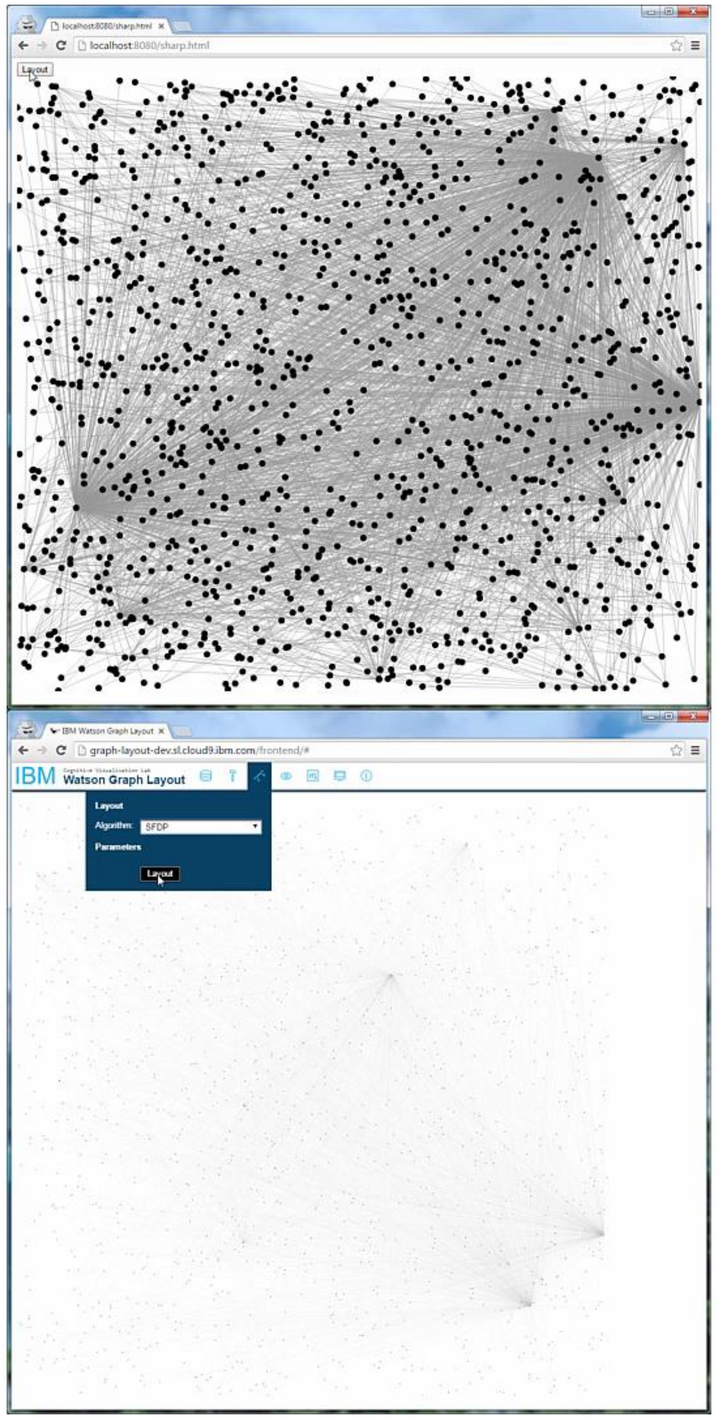
https://observablehq.com/@d3/force-directed-graph

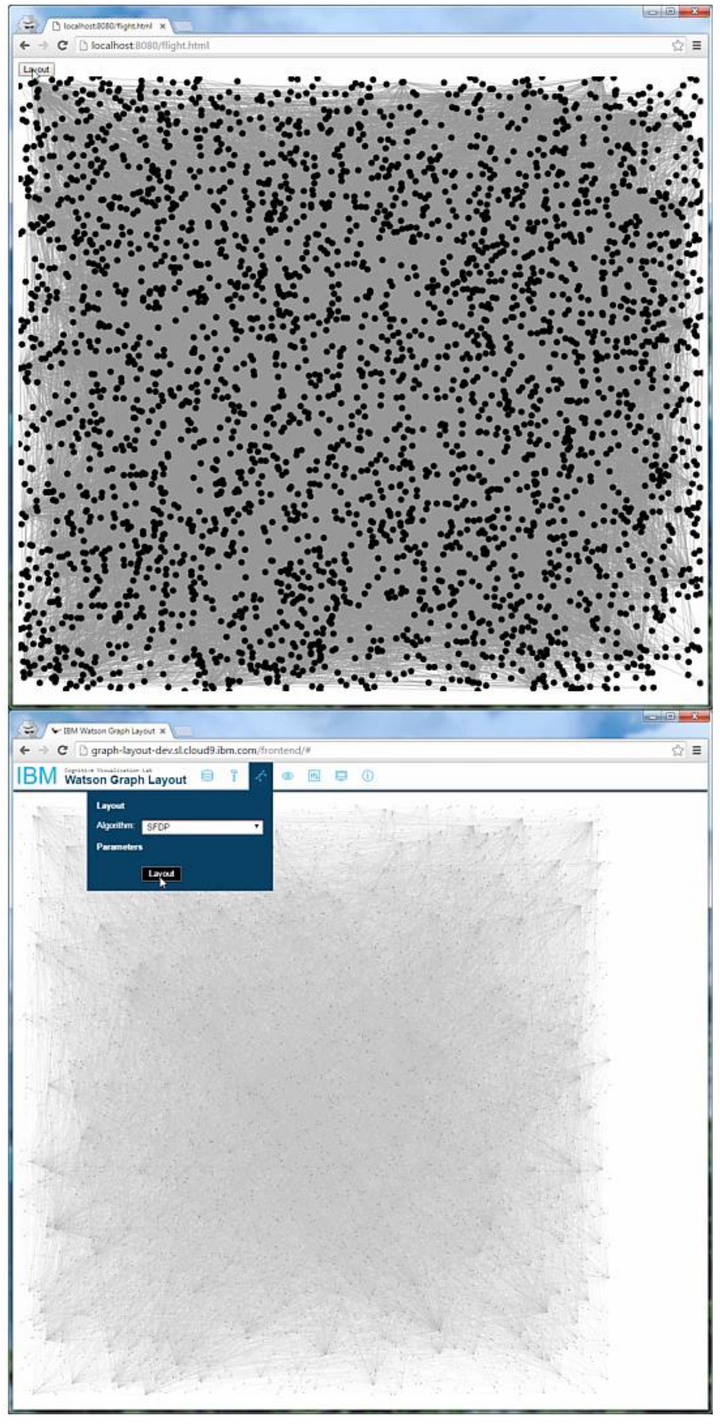
Force-Directed Layout Algorithms

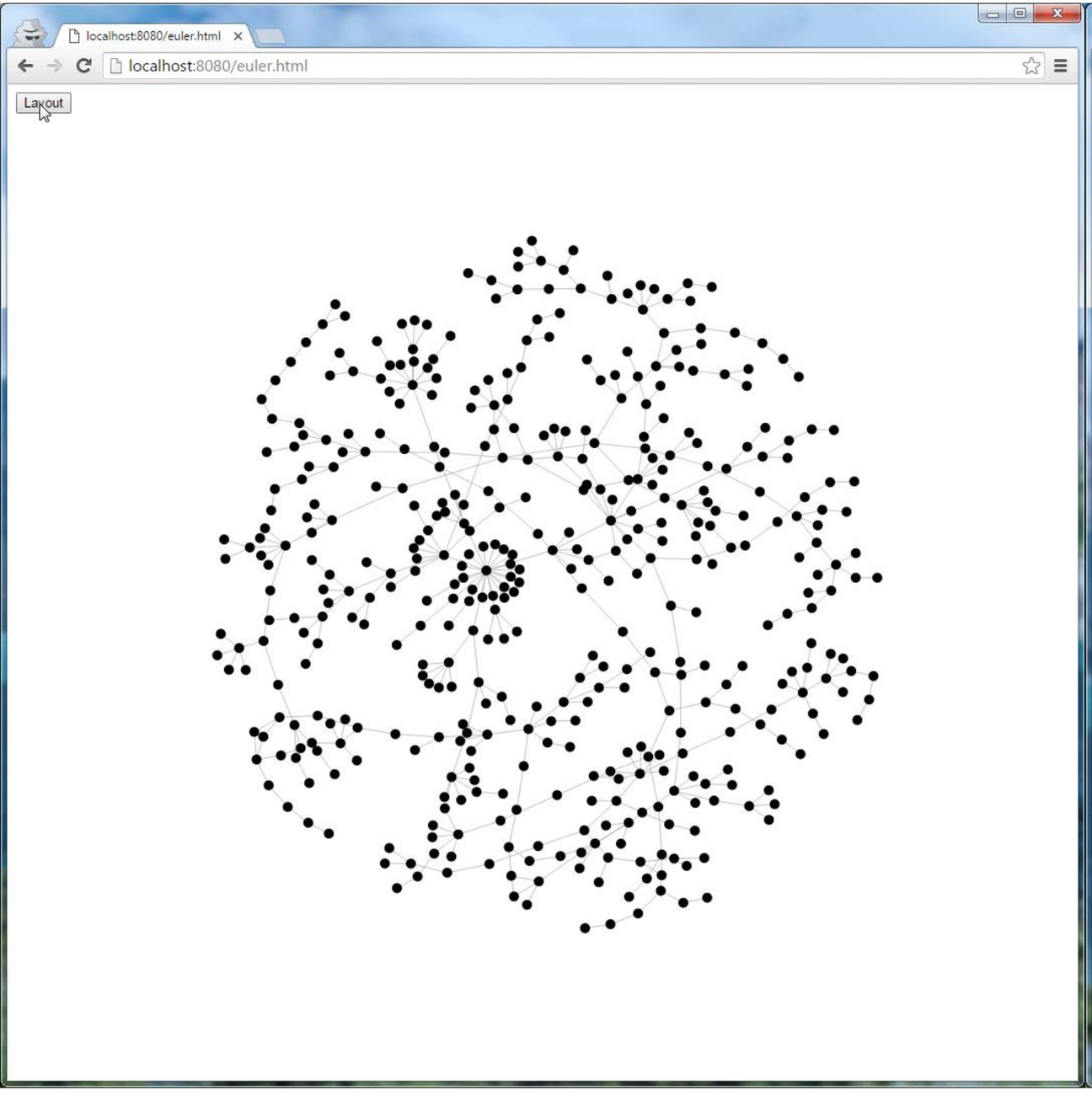


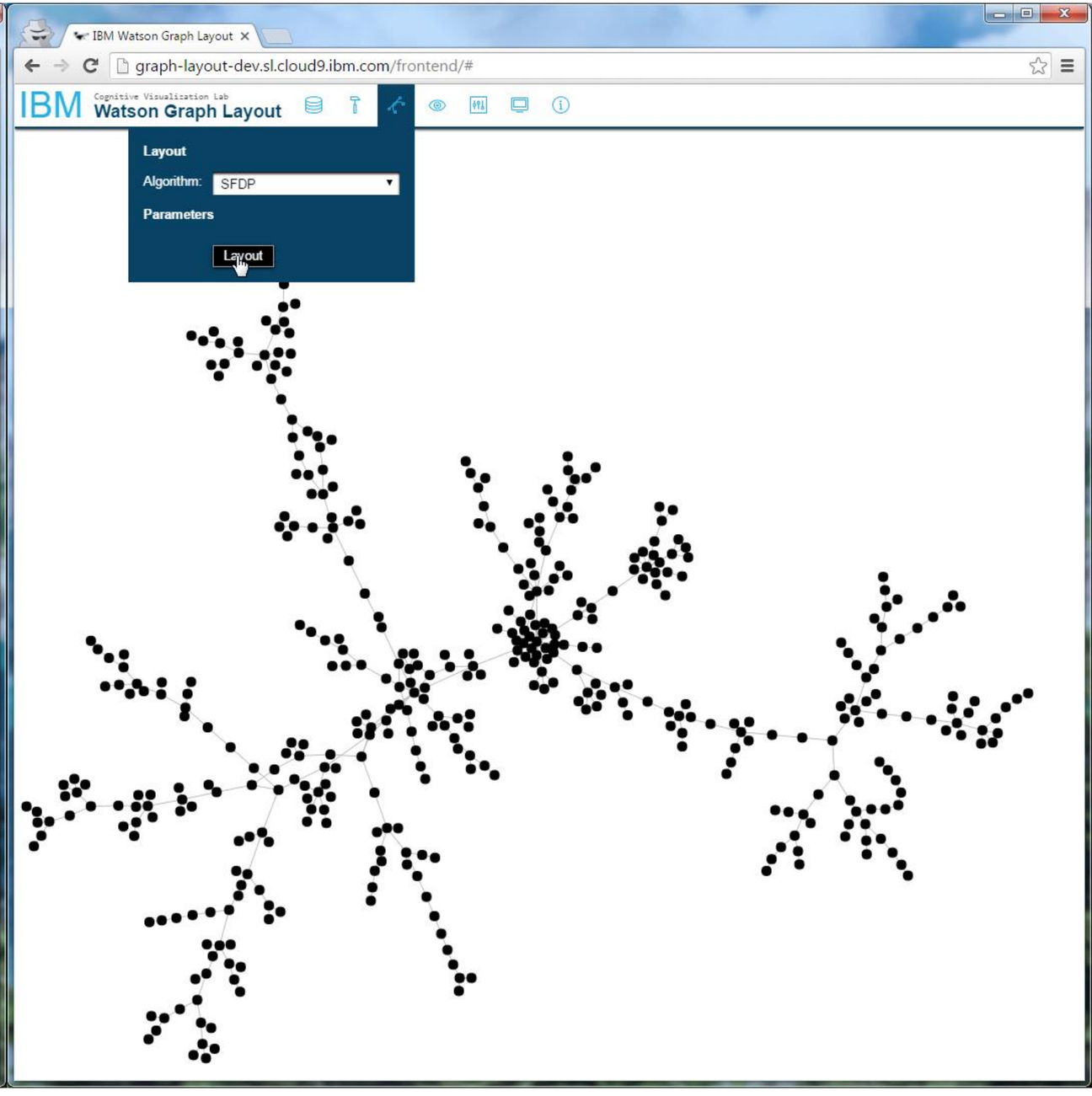
Kobourov, 2012


localhost:8080/euler.html
localhost:8080/euler.html
Layout
IBM Watson Graph Layout
graph-layout-dev.sl.cloud9.ibm.com/frontend/#
IBM
Watson Graph Layout
Layout
Algorithm:
SFDP
Parameters
Layout


localhost:8080/sharp.html
localhost:8080/sharp.html
Layout
IBM Watson Graph Layout
graph-layout-dev.sl.cloud9.ibm.com/frontend/#
IBM
Watson Graph Layout
Layout
Algorithm:
SFDP
Parameters
Layout

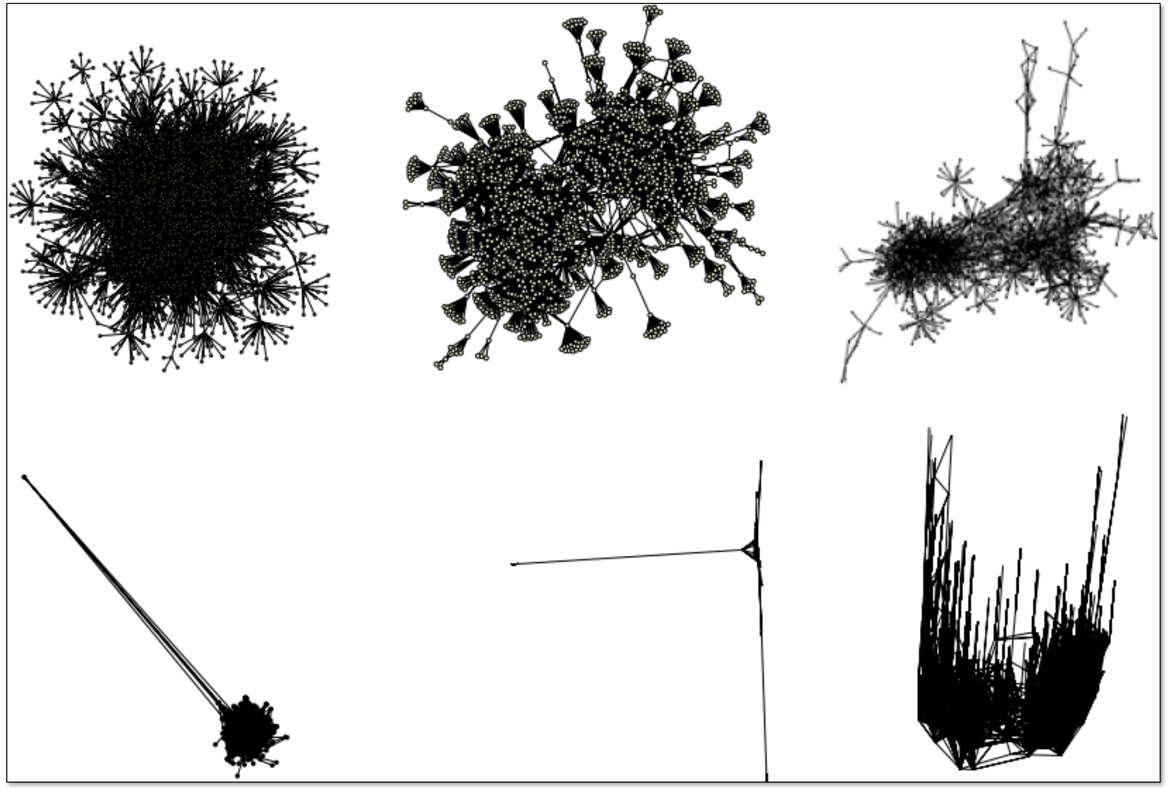

localhost:8080/flight.html
localhost:8080/flight.html
Layout
IBM Watson Graph Layout
graph-layout-dev.sl.cloud9.ibm.com/frontend/#
IBM
Watson Graph Layout
Layout
Algorithm:
SFDP
Parameters
Layout


localhost:8080/euler.html
localhost:8080/euler.html
Layout

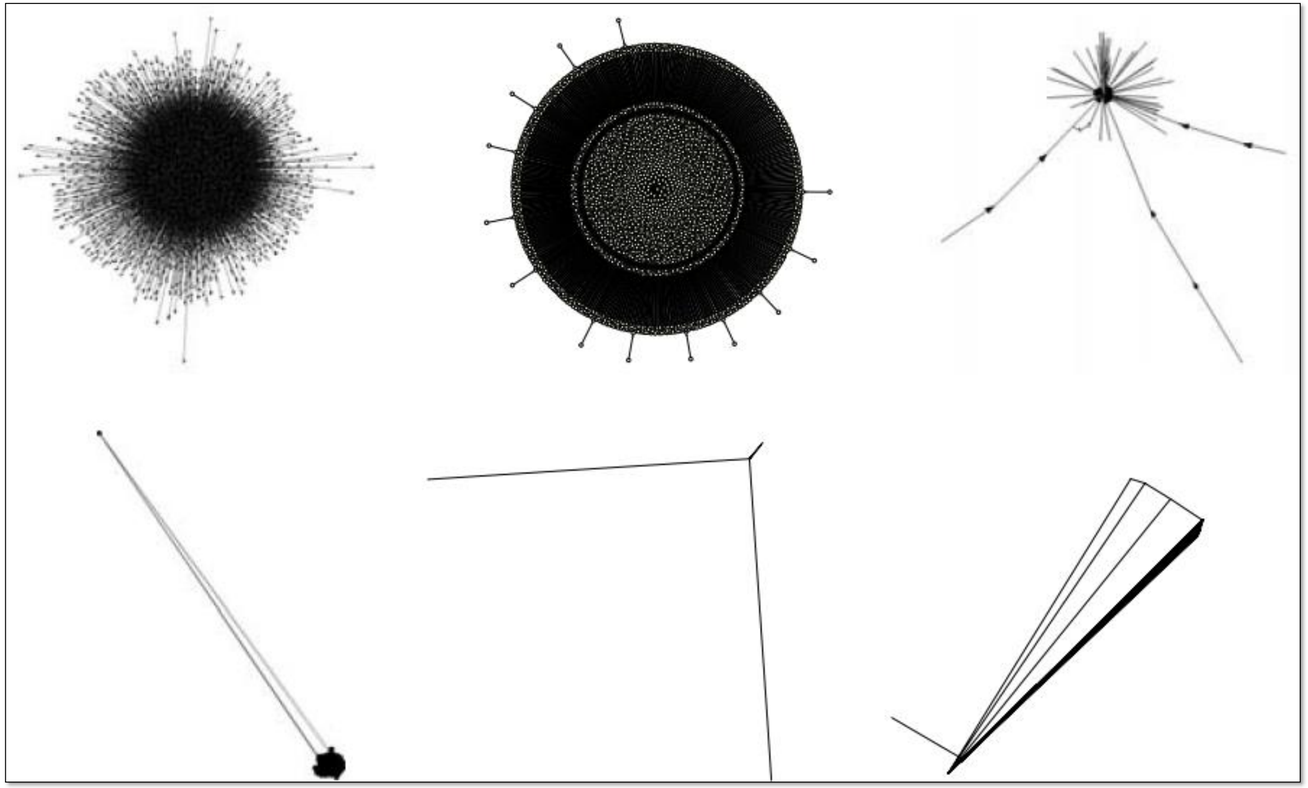

IBM Watson Graph Layout
graph-layout-dev.sl.cloud9.ibm.com/frontend/#
IBM
Cognitive Visualization Lab
Watson Graph Layout
Layout
Algorithm:
SFDP
Parameters
Layout

Layout Algorithm Comparisons

Graph A



Graph B



Hachul & Jünger, 2006

How to compare?

User performance

Huang et al., 2007, etc.

Simple rules or heuristics

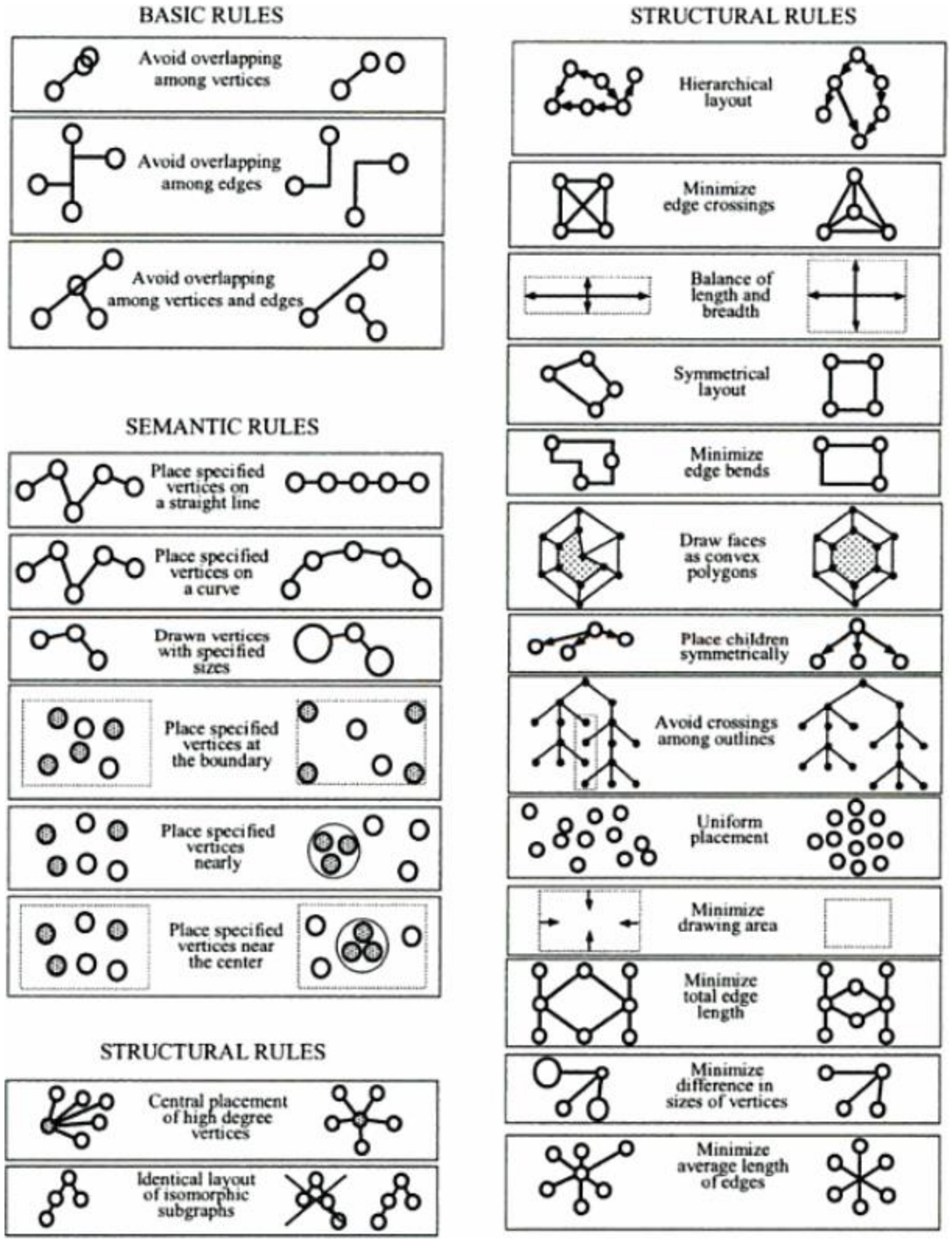
Davidson & Harel, 1996

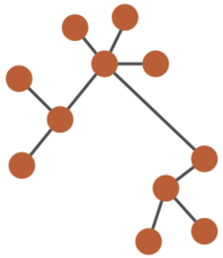
Global and local readability metrics

Purchase et al., 2002

Dunne et al., 2015

Sugiyama, 2002, p. 14





Scale Problems...

- Quickly run out of space!
- Tree breadth often grows exponentially
- Layout algorithms are slow and heuristics
- Solutions:

- scrolling or panning
- filtering or zooming
- aggregation & simplification

Graph Viewer Test

www.yasiv.com/graphs#HB/blckhole

Yasiv About Contact... More Visualizations... vivagraph.js Tweet

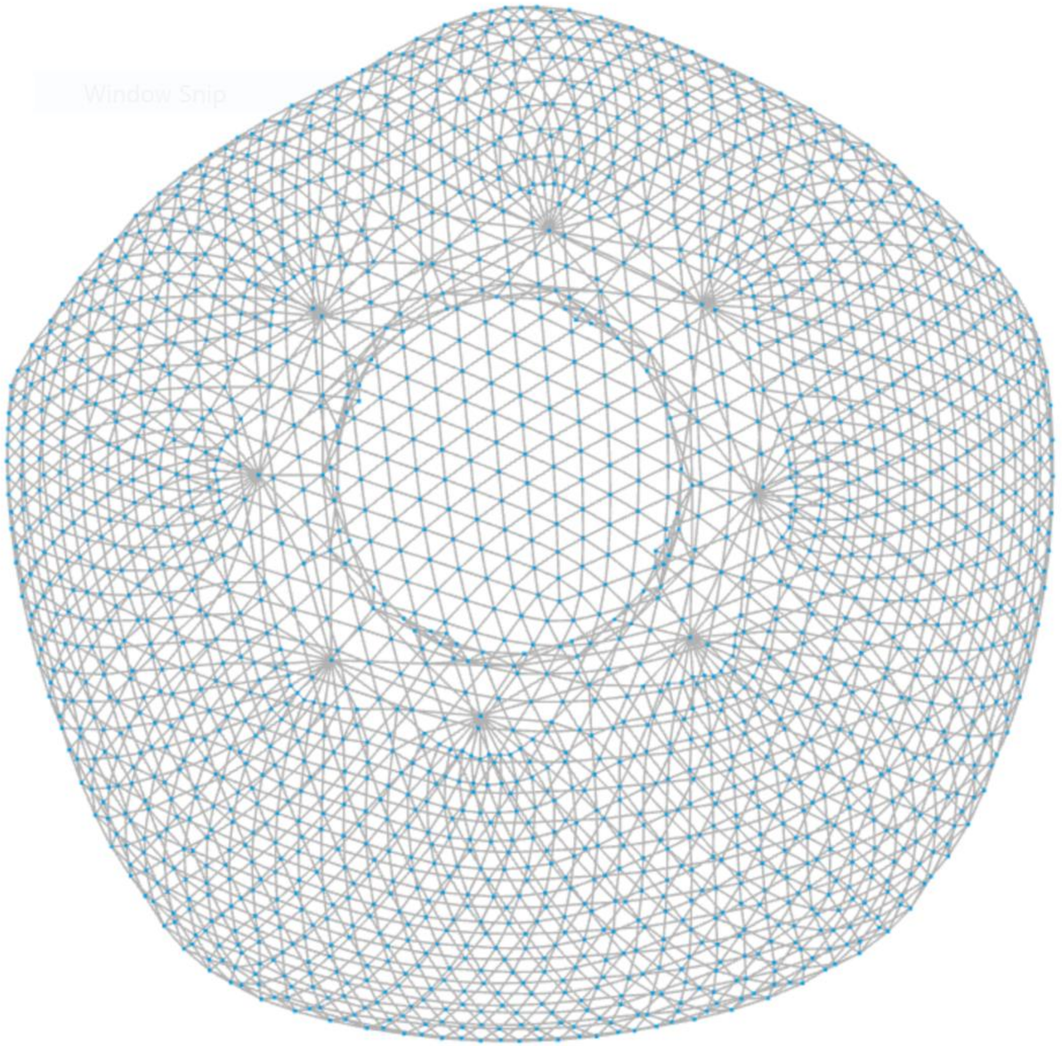
Choose Graph:

FAVORITE GRAPHS

- HB/blckhole
- Bai/rw5151
- HB/bcsstm13
- HB/jagmesh6
- HB/watt_1
- HB/lshp1882
- HB/plat1919
- HB/bcsstk26
- Bai/dw256A
- Bai/tols2000
- Bai/dw1024
- Bai/rdb2048
- Pajek/CSphd
- GHS_indef/laser

BAI

- bfwa398
- bfwa62
- bfwb398
- bfwb62
- bfwb782
- bwm200
- cdde1
- cdde2
- cdde3
- cdde4
- cdde5
- cdde6
- ck104
- ck400
- ck656



Layout Settings

Spring Coeff: 0.0008

Spring Length: 30

Gravity Coeff: -1.2

Drag Coeff: 0.009

Theta Coeff: 0.8

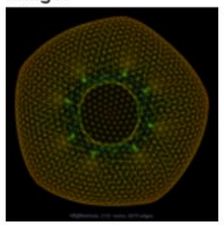
Reset to default

HB/blckhole

Nodes: 2121

Edges: 6370

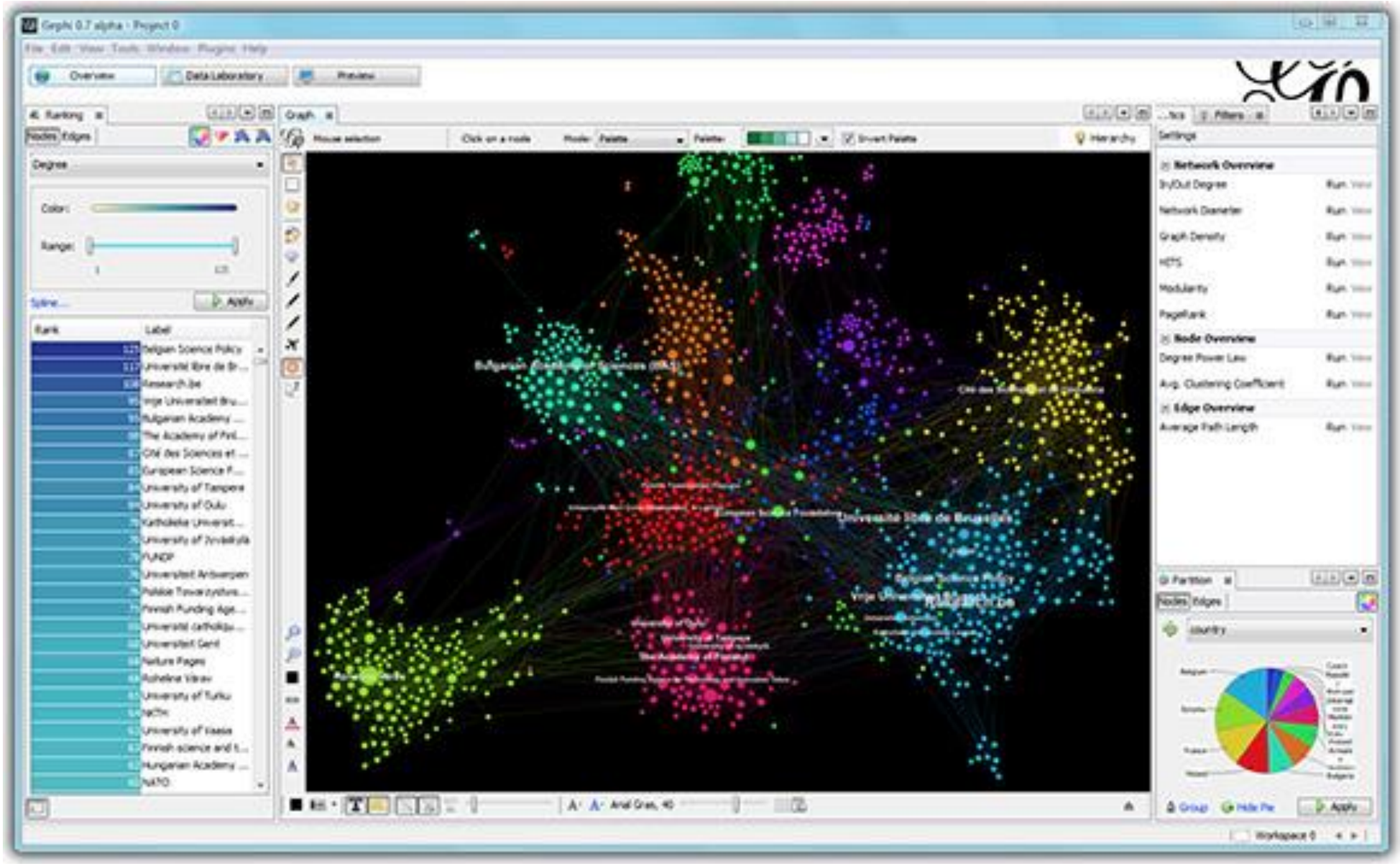
Image:



Pause layout

Use mouse wheel to zoom in/zoom out

http://www.yasiv.com/graphs#HB/blckhole



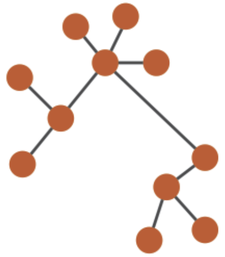
https://gephi.org/

Arrange Networks and Trees

Node–Link Diagrams

Connection Marks

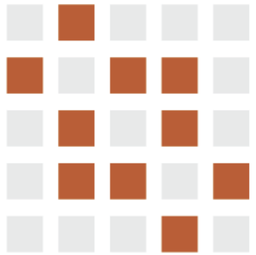
✔ NETWORKS ✔ TREES



Adjacency Matrix

Derived Table

✔ NETWORKS ✔ TREES



Enclosure

Containment Marks

✘ NETWORKS ✔ TREES

"Treemap"

Upcoming Assignments & Communication

Look at the upcoming assignments and deadlines regularly!

- Textbook, Readings, & Reading Quizzes — Variable days
- In-Class Activities — 11:59pm same day as class
  F: Lecture, T: In-Class Project Feedback Meetings & Work, F/T: Lecture/TBD
- Assignments & Projects— Generally due R 11:59pm
  R (2 days):
    Assignment 8 — Brushing and Linking in D3
    Project 5 — Final "Interactive" Visualization Sketch, Implementation Plan, & Group Charter
  Next R (9 days):
    Project 6 — Sprint 1
  Next-Next R (16 days):
    Project 7 — Sprint 2 & Paper Draft

Use Canvas Discussions for general questions, email the TAs/S-LTA/instructor for questions specific to you: codydunne-and-tas@ccs.neu.edu. Include links!

If you're emailing about a particular assignment, please **include the URL of the Submission Details page**. (Canvas documentation.)

If you have a project question, **give us your group number**. E.g., include: `Group ## — Topic` with '##' replaced by your group number and 'Topic' replaced by your topic.