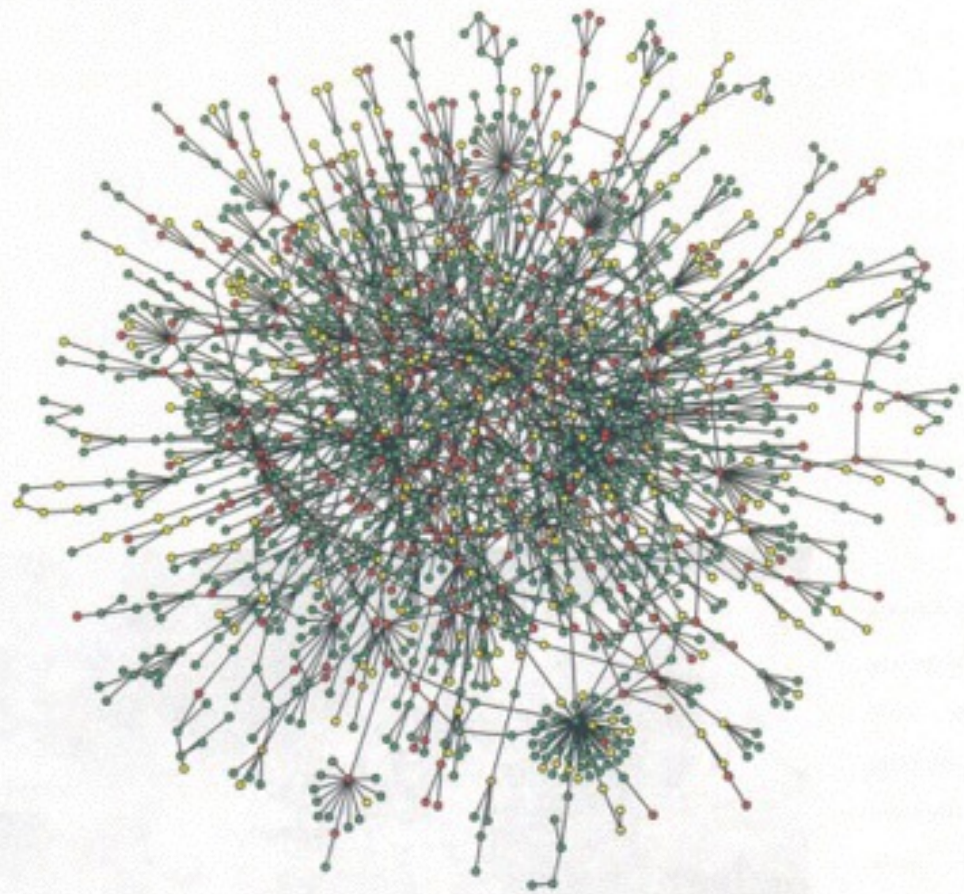
Visualizing Networks

Lynn Cherny
@arnicas

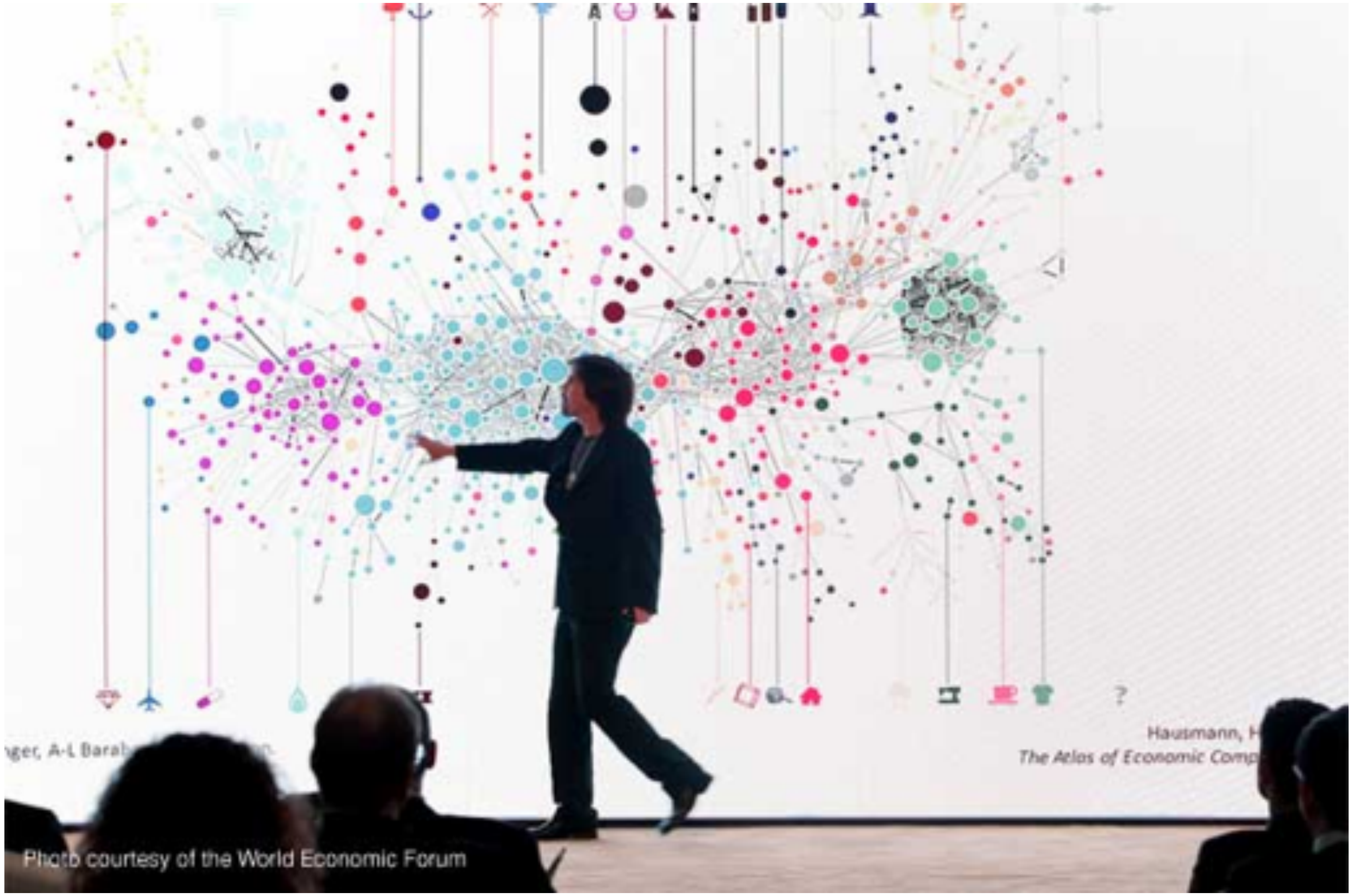
4 Nov 2014

The Hairball: A Metaphor for Complexity

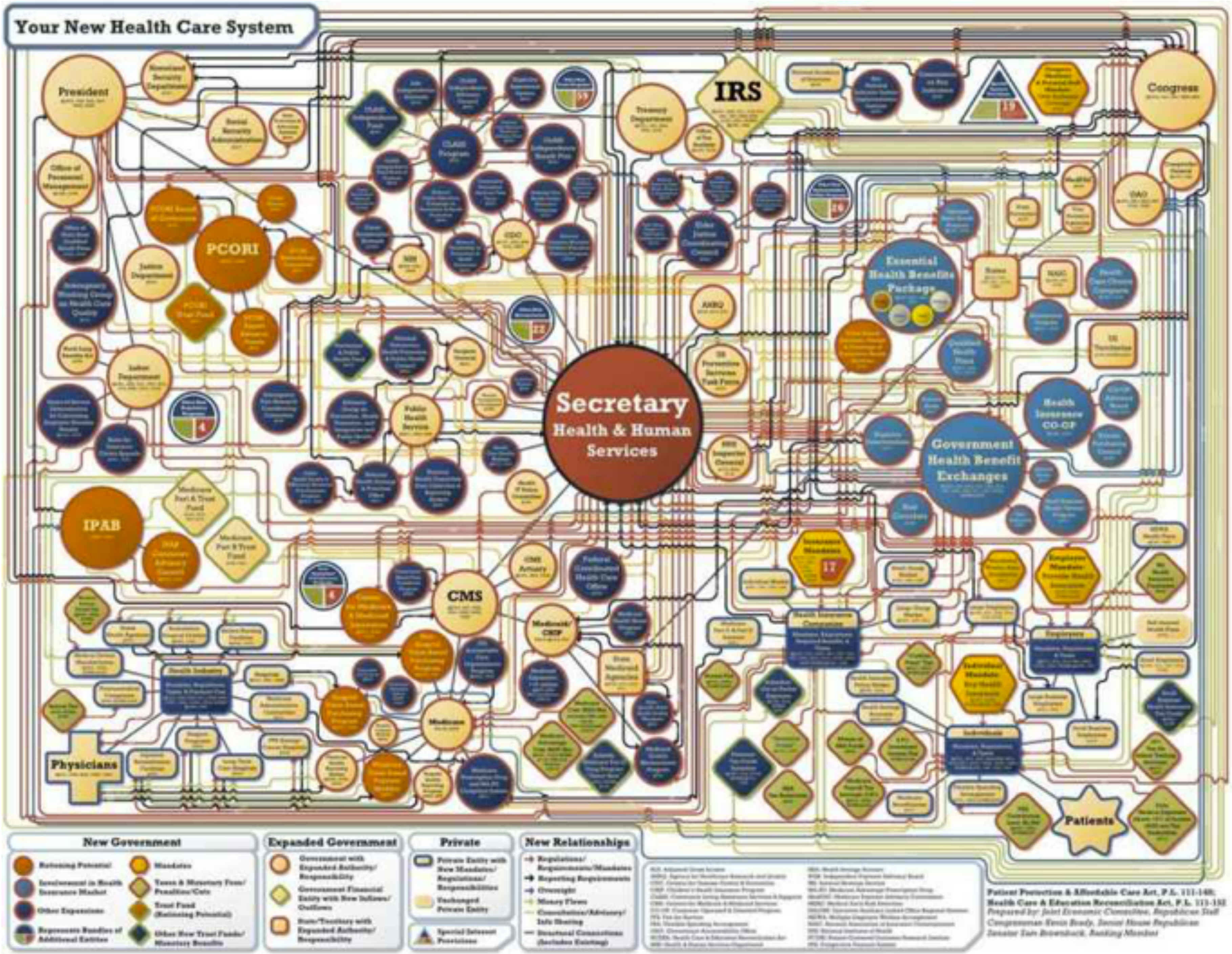


MAP OF INTERACTING PROTEINS in yeast highlights the discovery that highly linked, or hub, proteins tend to be crucial for a cell's survival. Red denotes essential proteins [their removal will cause the cell to die]. Orange represents proteins of some importance [their removal will slow cell growth]. Green and yellow represent proteins of lesser or unknown significance, respectively.

http://www.nd.edu/~networks/Publication%20Categories/01%20Review%20Articles/ScaleFree_Scientific%20Ameri%20288,%2060-69%20(2003).pdf

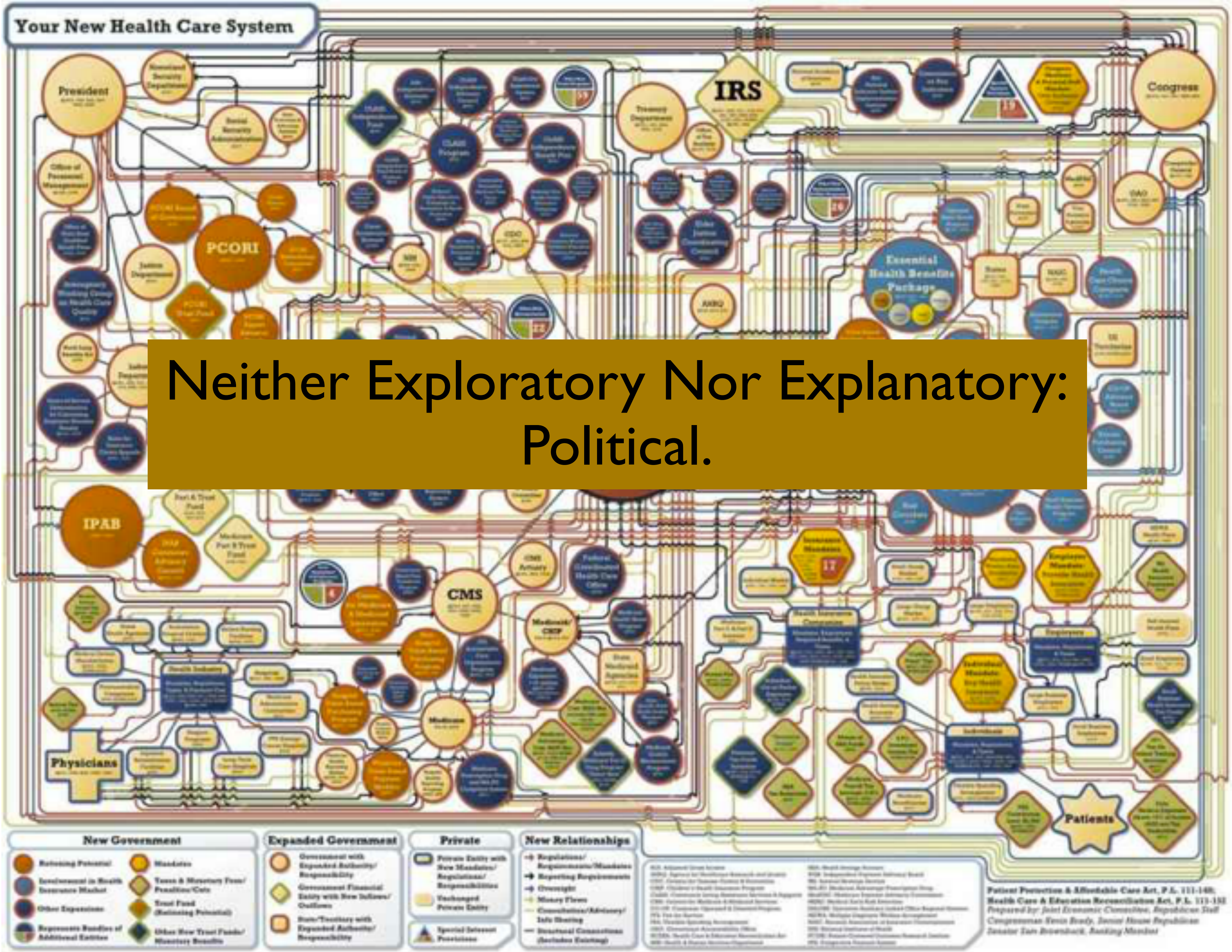


http://www.linkedin.com/today/post/article/20121016185655-10842349-the-hidden-power-of-networks



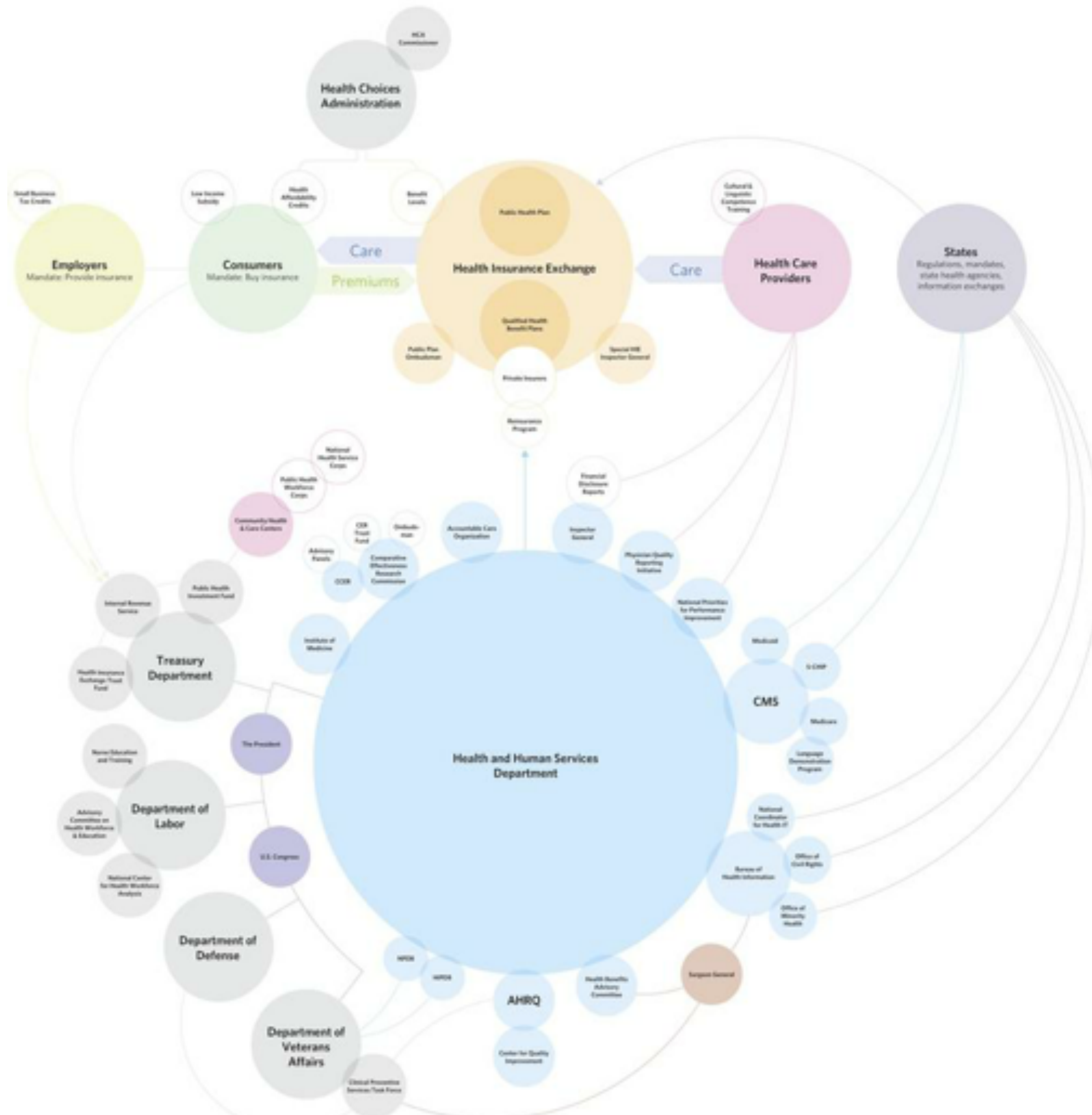
Release by Joint Economic Committee minority Republicans led by Boehner (2010), Iliinsky & Steele fig 4-14

Neither Exploratory Nor Explanatory: Political.

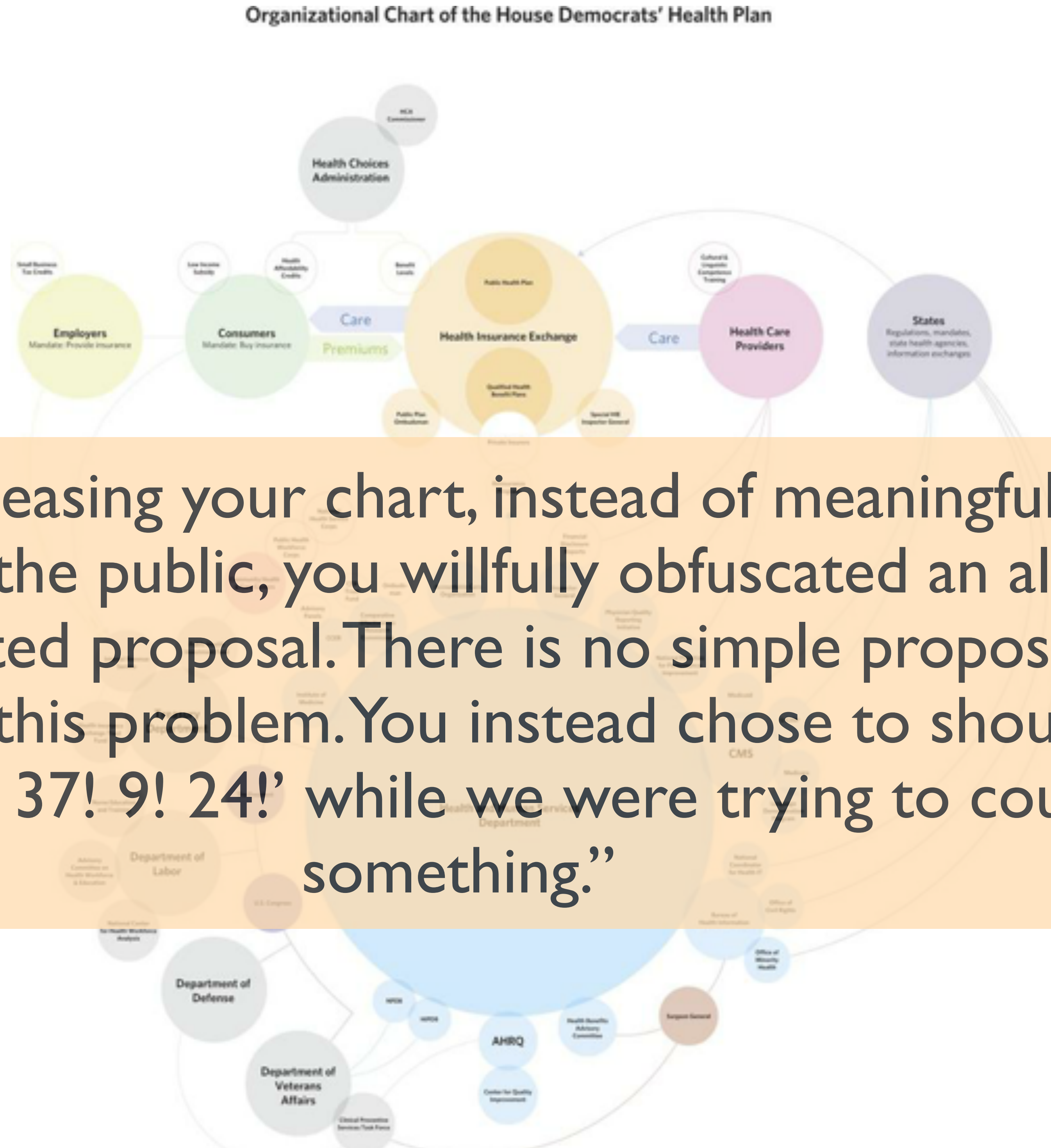


Release by Joint Economic Committee minority Republicans led by Boehner (2010), Iliinsky & Steele fig 4-14

Organizational Chart of the House Democrats' Health Plan



Redesign by Robert Palmer
(Iliisnky & Steele fig 4-15)



"By releasing your chart, instead of meaningfully educating the public, you willfully obfuscated an already complicated proposal. There is no simple proposal to solve this problem. You instead chose to shout '12! 16! 37! 9! 24!' while we were trying to count something."

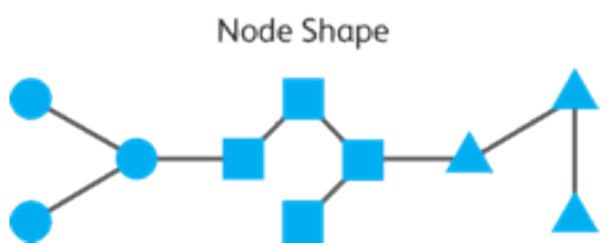
Redesign by Robert Palmer
(Iliisnky & Steele fig 4-15)

WHAT IS A NETWORK?

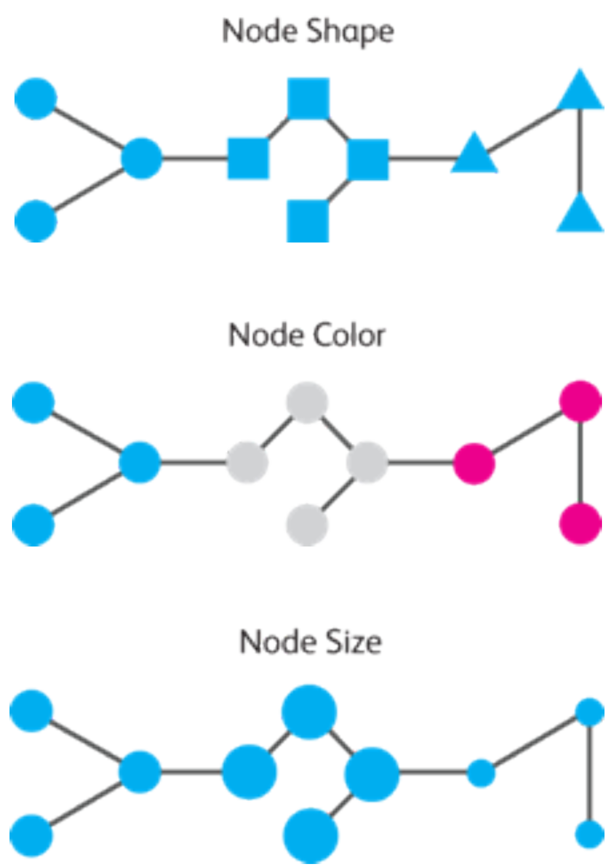
It’s not a visualization. Think of it as a data structure.

- Data relationship: entities + relationships to other objects (node/edge, vertex/link)

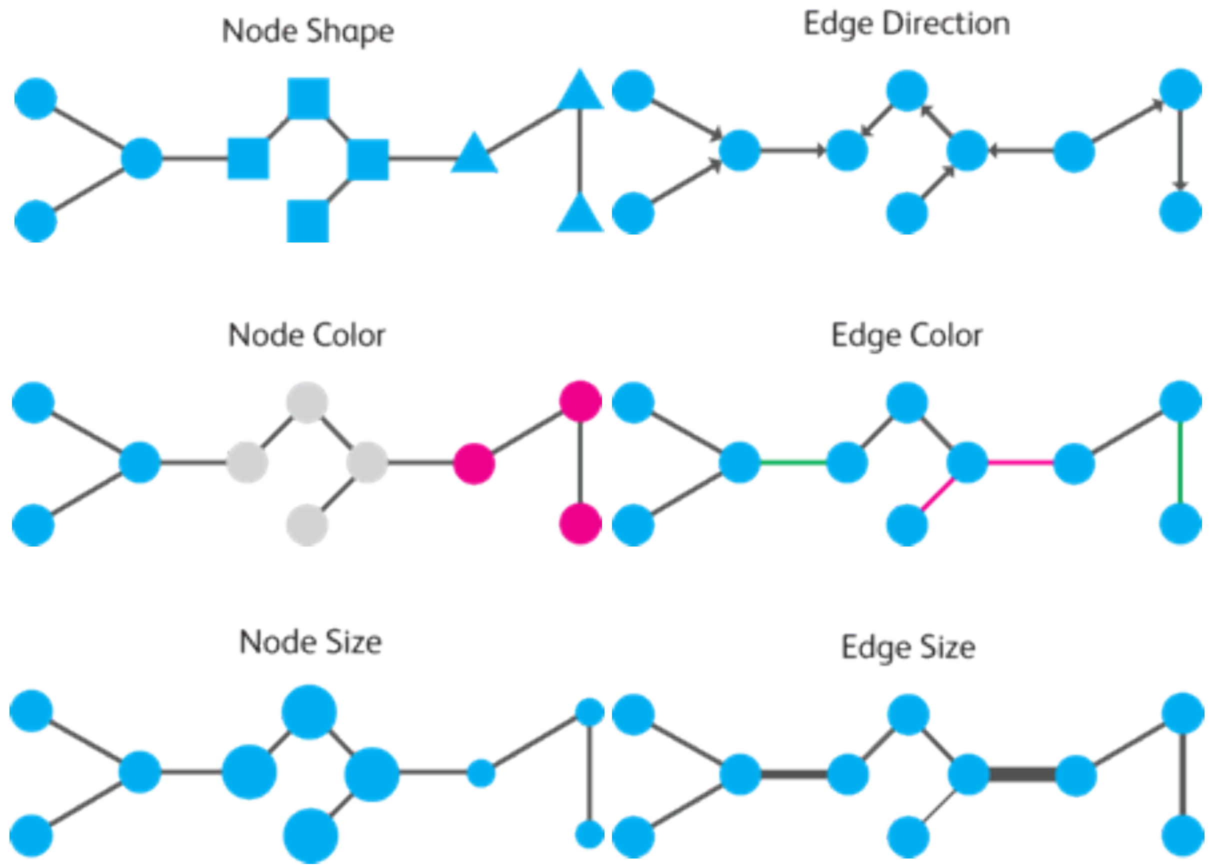
- Nodes and Edges may have attributes, eg.
  - gender, age, weight, tv prefs
  - connection date, frequency of contact, type of exchange, directionality of relationship
  - attributes may be calculated from network relations too



Lane Harrison: http://blog.visual.ly/network-visualizations/



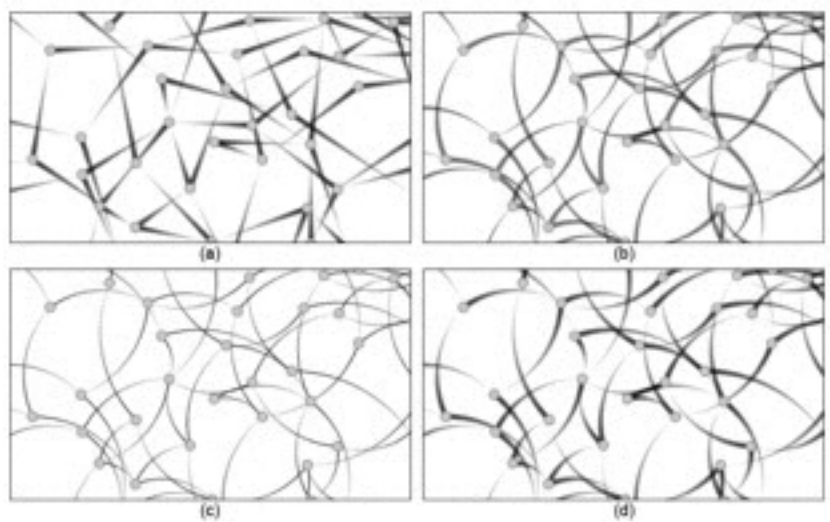
Lane Harrison: http://blog.visual.ly/network-visualizations/



Lane Harrison: http://blog.visual.ly/network-visualizations/

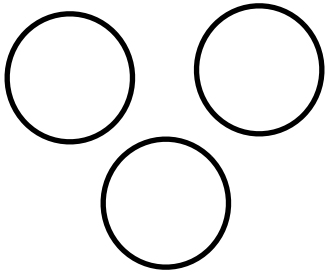
Edge Tapering

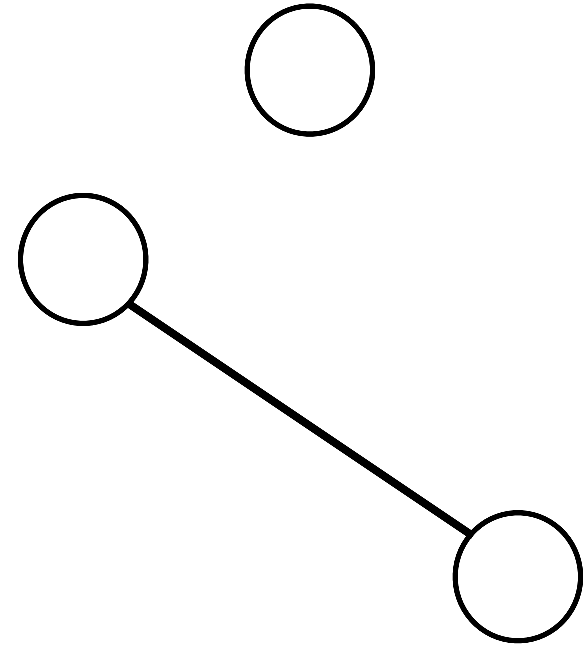
Best!

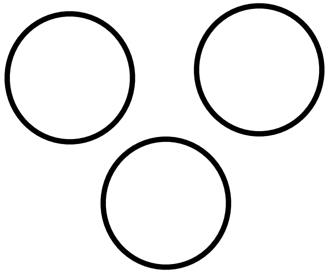


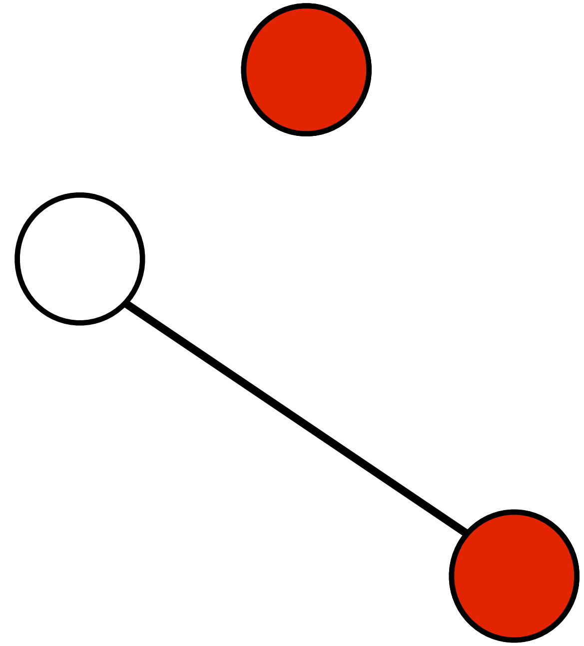
A User Study on Visualizing Directed Edges in Graphs"
Danny Holten and Jarke J. van Wijk, 27th SIGCHI Conference on Human Factors in Computing Systems (Proceedings of CHI 2009),

http://blog.visual.ly/network-visualizations/





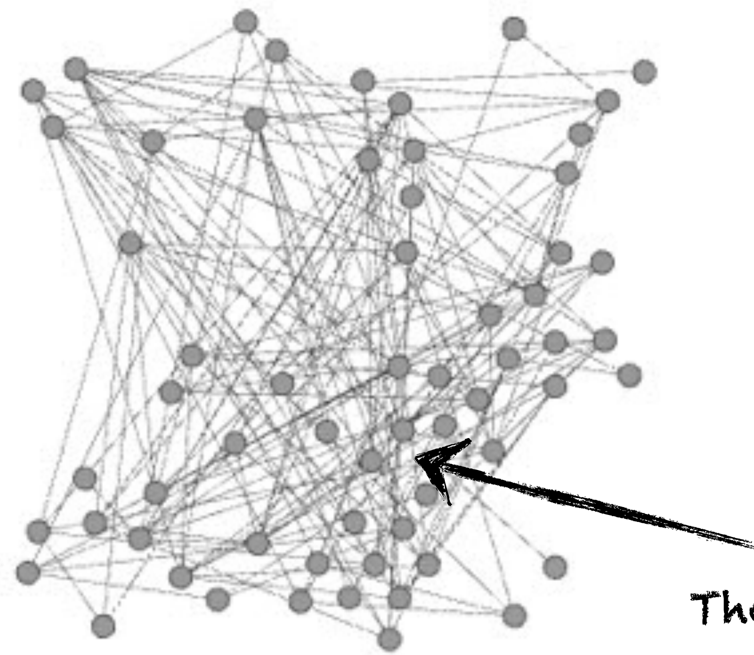




It’s a natural human trait to see visual similarity and proximity as meaningful.

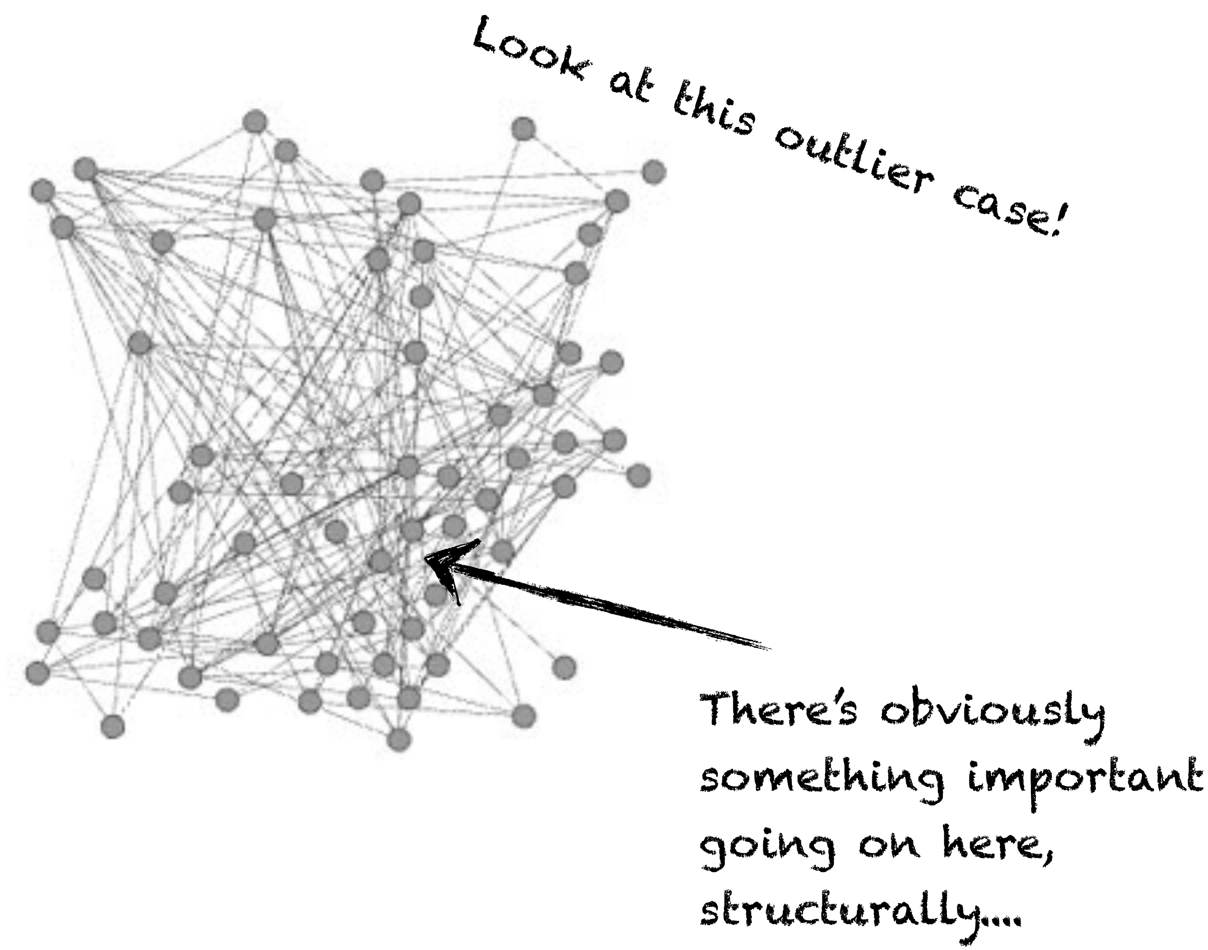
Be very careful about your display choices and layout methods!

Reading a network visualization

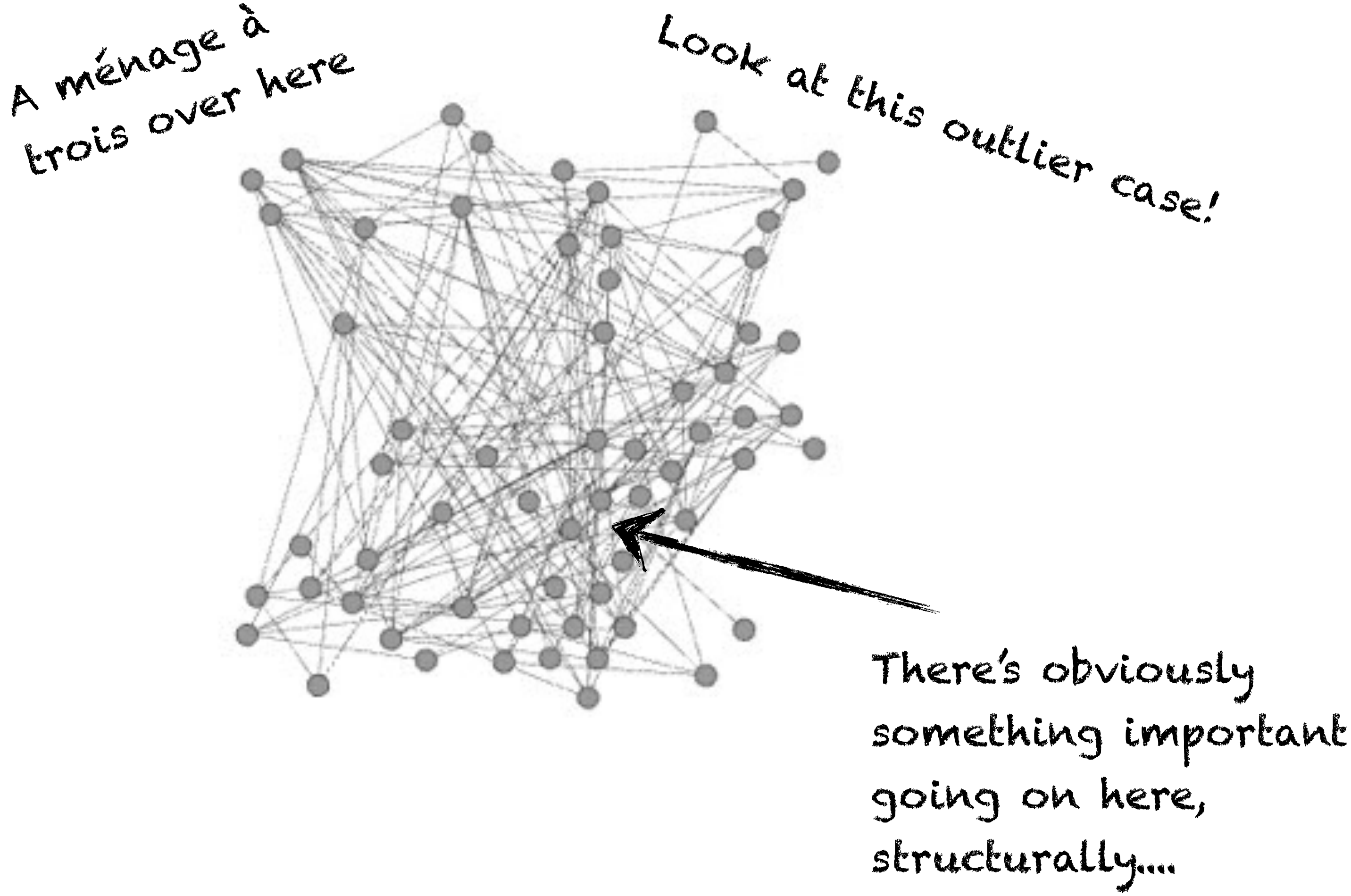


There's obviously something important going on here, structurally....

Reading a network visualization

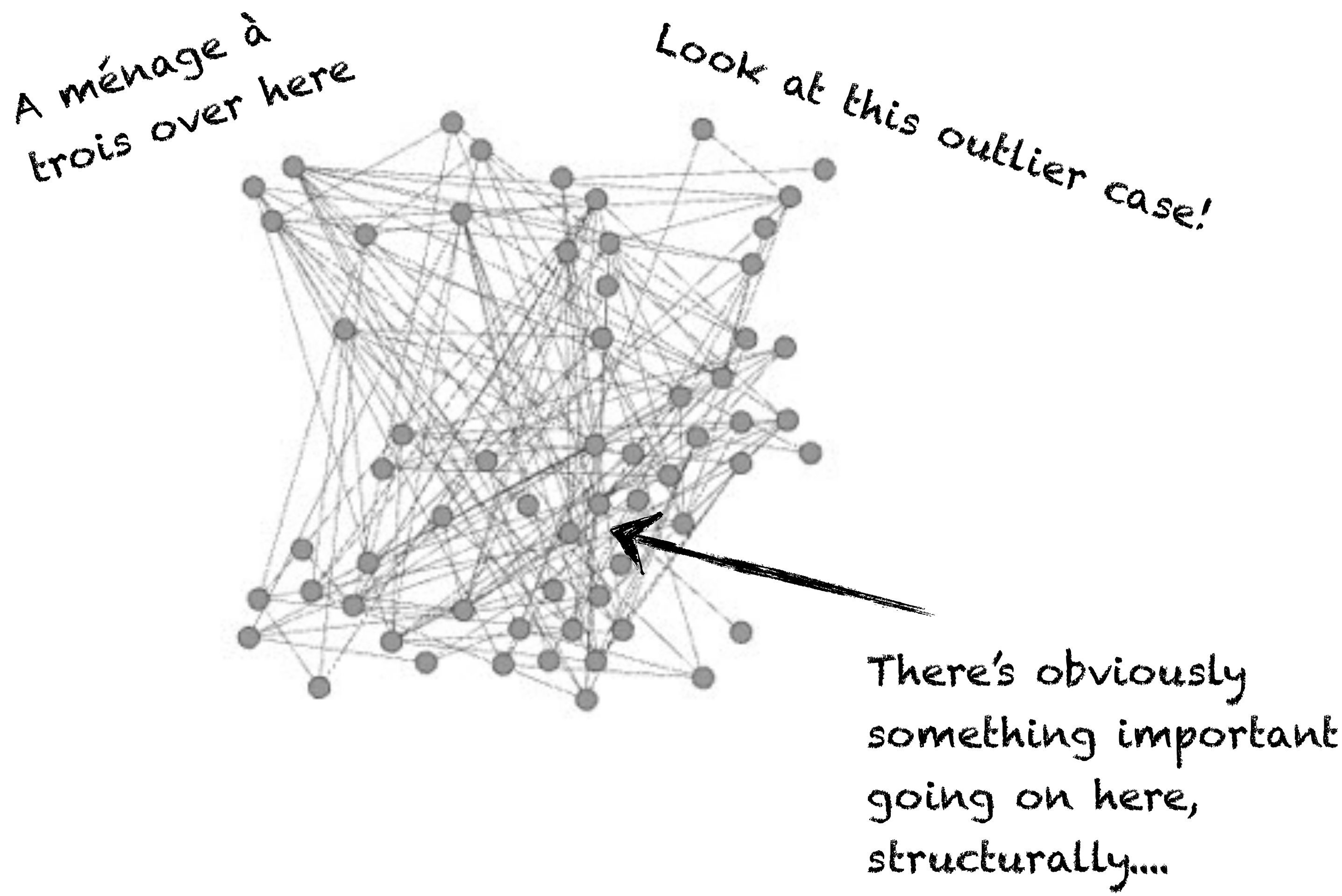


Reading a network visualization



MIS?

Reading a network visualization



Using a "random" Gephi layout on the dolphins

MIS?

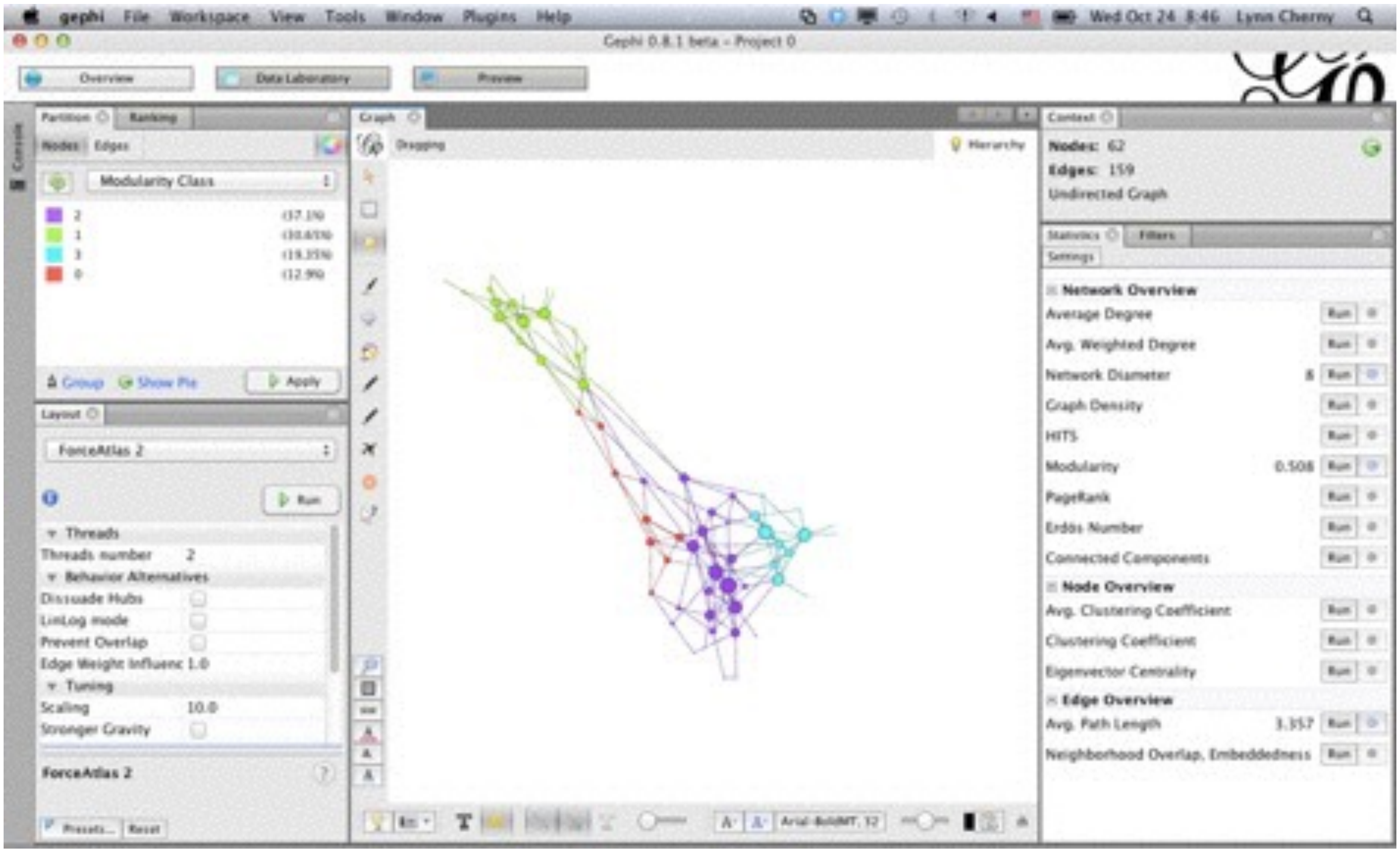
Reading a network visualization

Using a "random" Gephi layout on the dolphins

TOOLS AND CONCEPTS FOR TODAY

Creating network layouts...

Gephi (and a little D3)



J Bertin: Semiology of Graphics

Linear

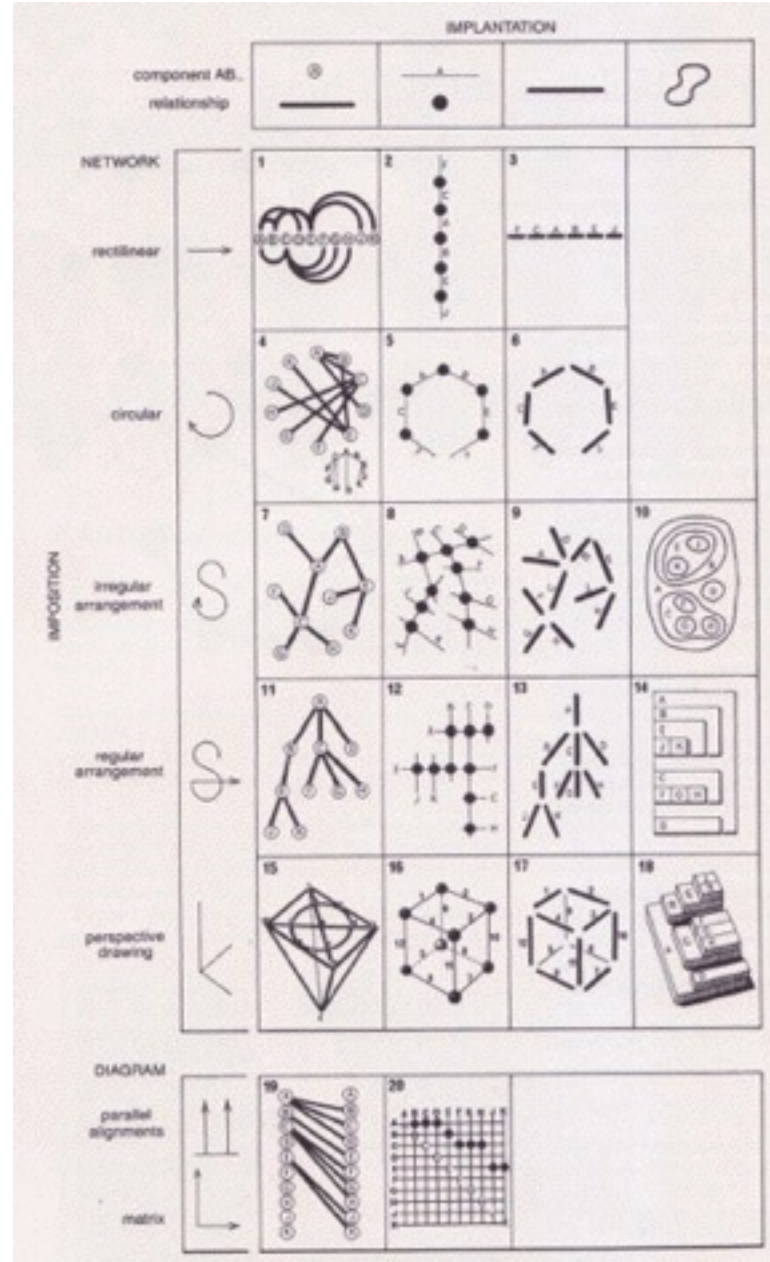
Circular

Irregular

Regular (Tree)

3D

Matrix / Bipartite



Bertin, J. Semiology of Graphics: Diagrams, Networks, Maps (1967)

J Bertin: Semiology of Graphics

Linear

Circular

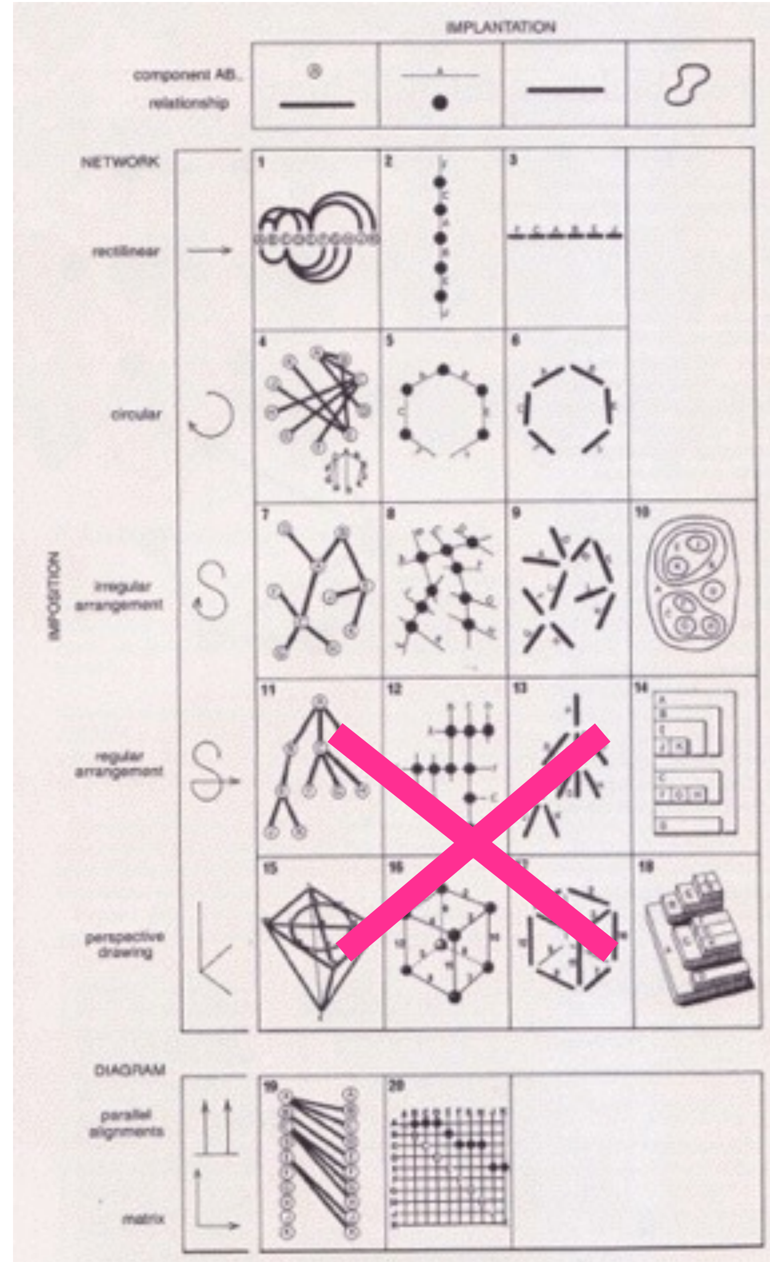
Irregular

Regular (Tree)

3D

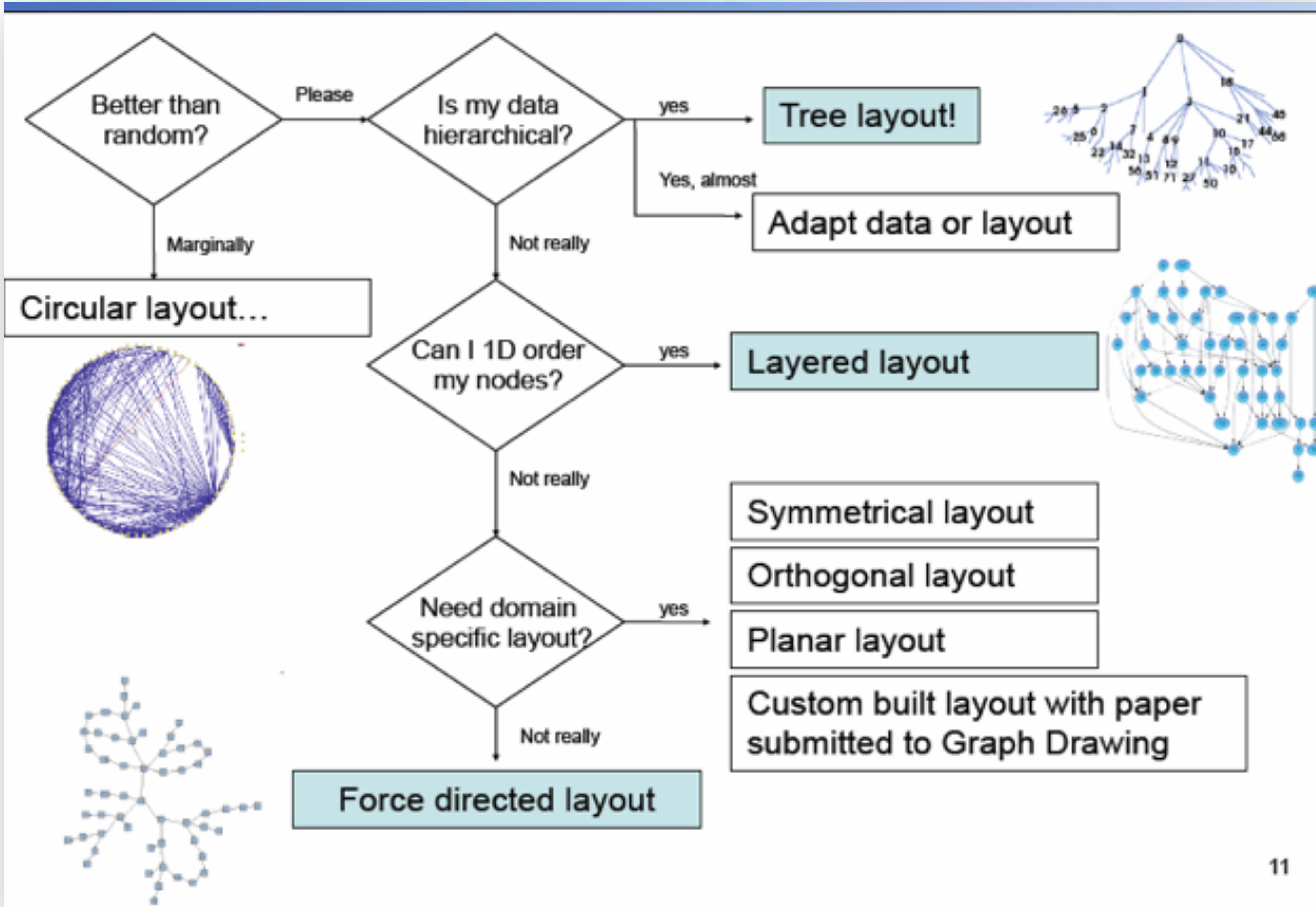
Matrix /

Bipartite



Bertin, J. Semiology of Graphics: Diagrams, Networks, Maps (1967)

Algorithmic Approaches



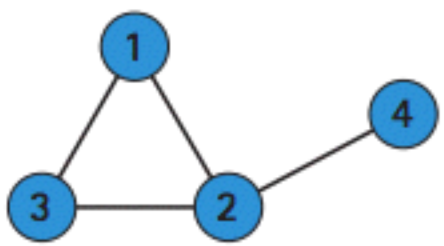
Frank van Ham talk slides: http://bit.ly/s6udpy

MATRIX LAYOUTS / REPRESENTATIONS

Adjacency matrix

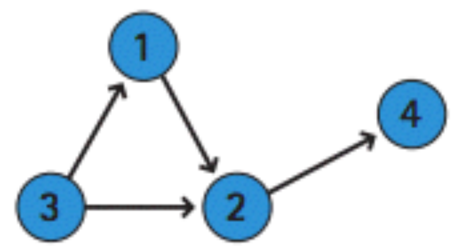
$$A_{ij} = \begin{pmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ A_{21} & A_{22} & A_{23} & A_{24} \\ A_{31} & A_{32} & A_{33} & A_{34} \\ A_{41} & A_{42} & A_{43} & A_{44} \end{pmatrix}$$

Undirected network



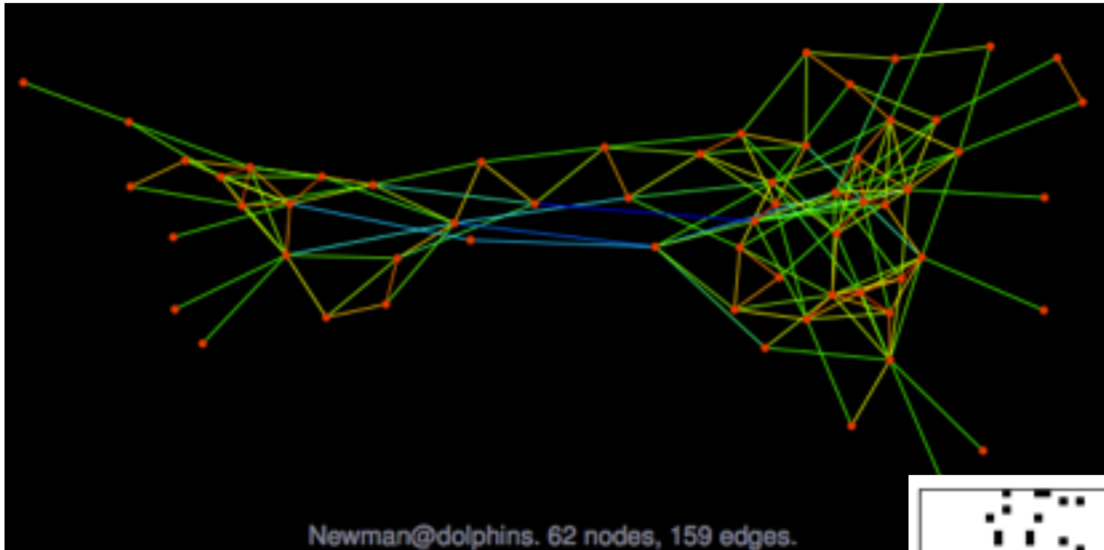
$$A_{ij} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

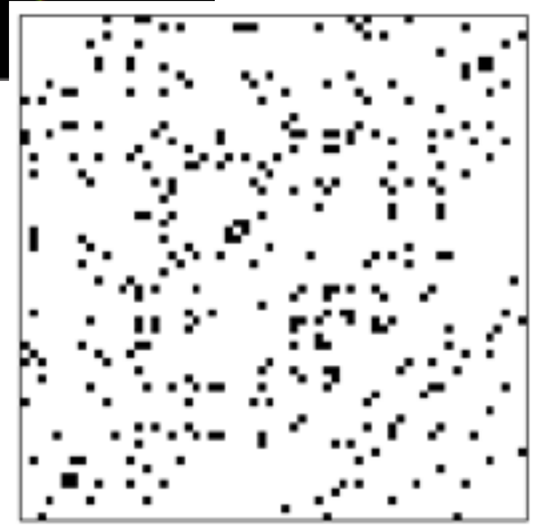
Directed network



$$A_{ij} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

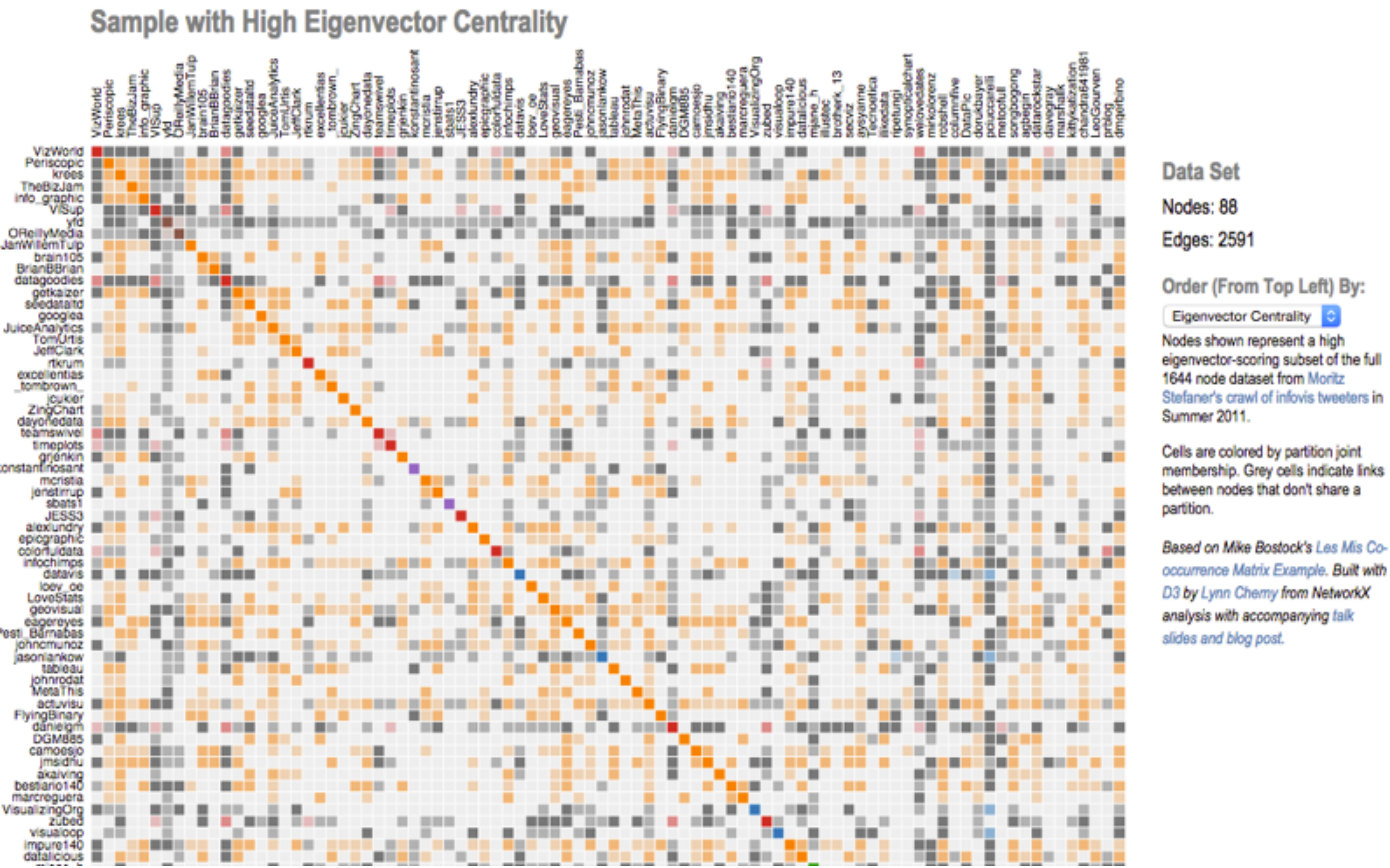
http://barabasilab.neu.edu/networksciencebook/download/network_science_October_2012.pdf



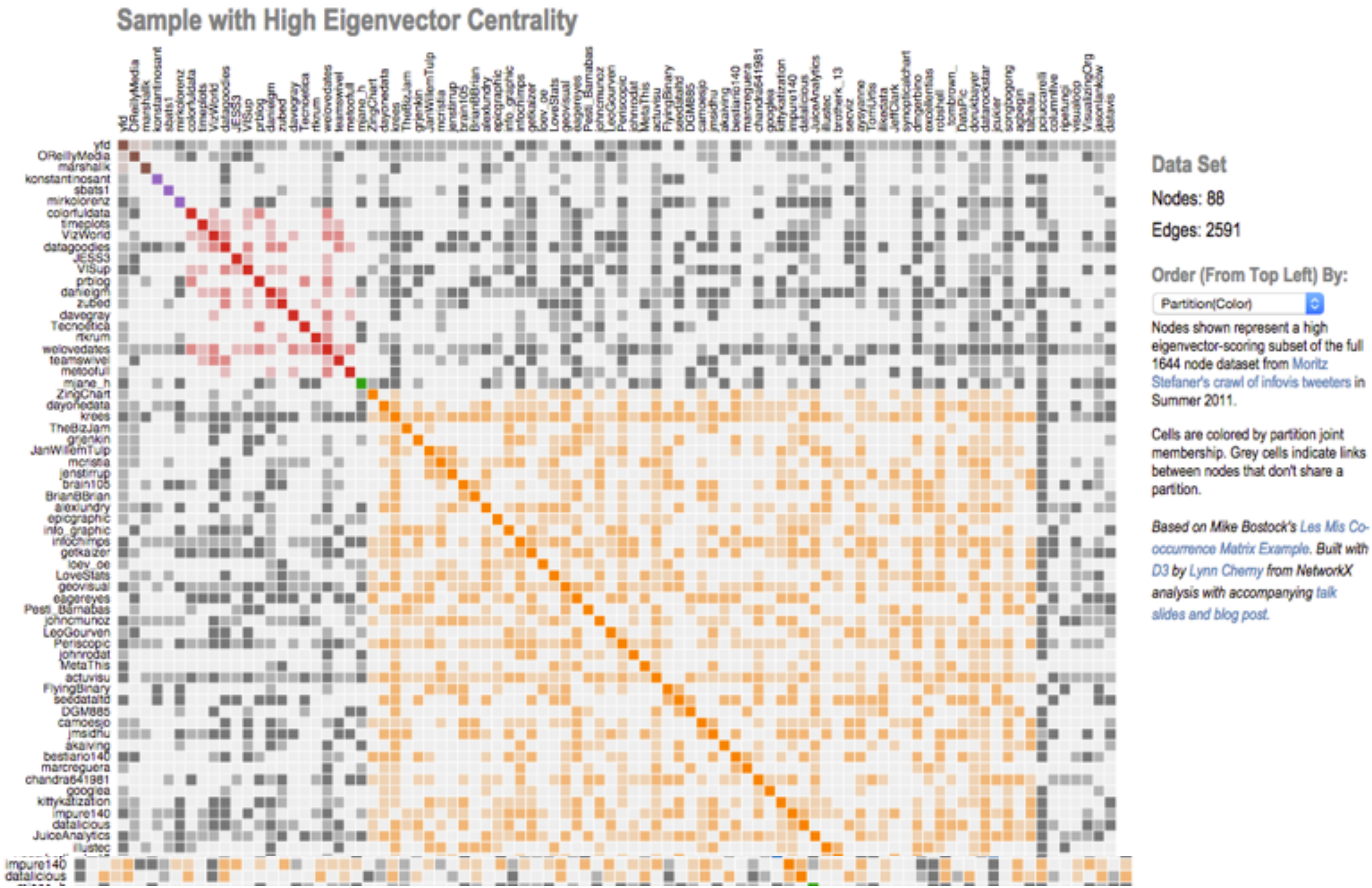


Real social networks are generally quite sparse.

http://www.cise.ufl.edu/research/sparse/matrices/Newman/dolphins.html

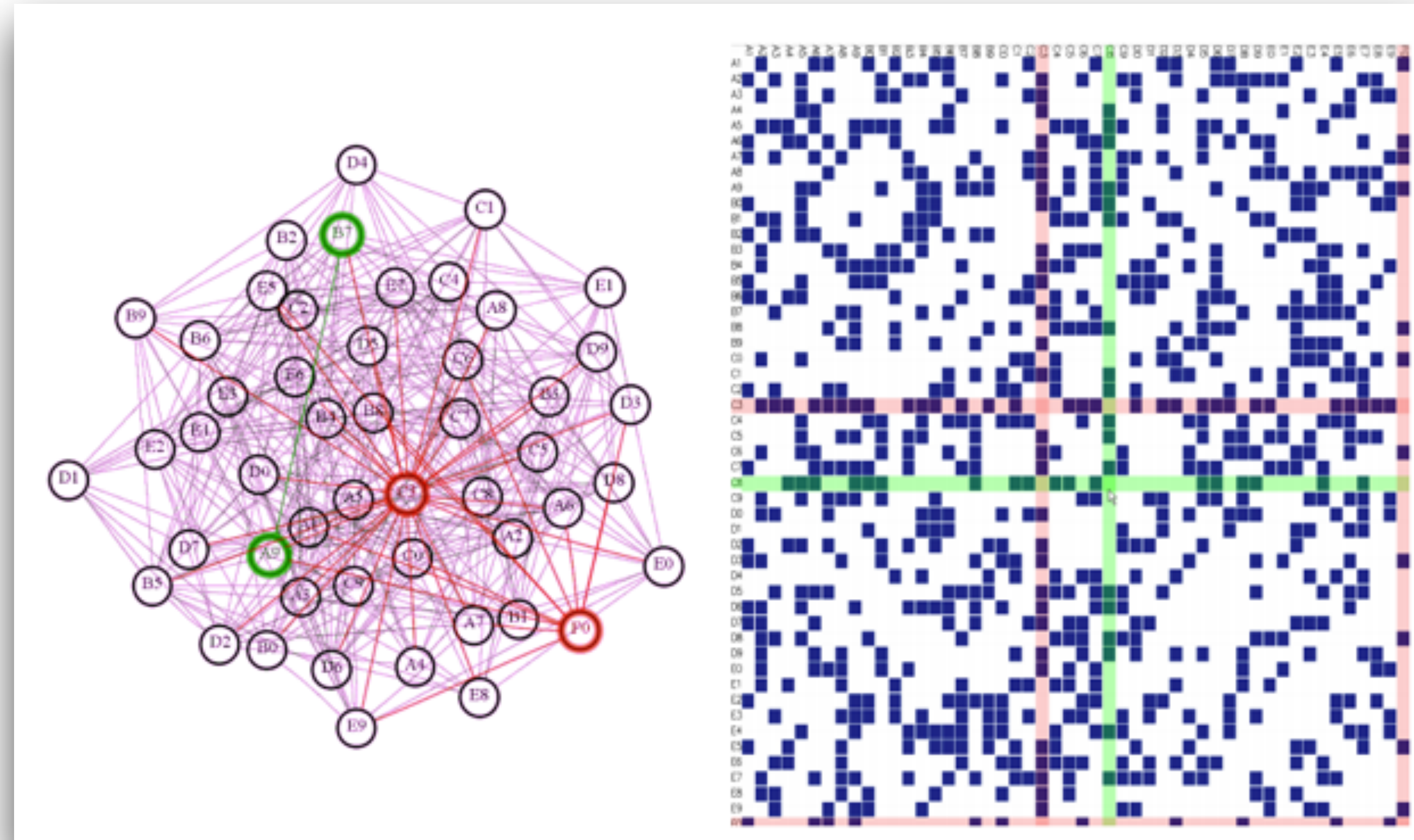


D3 demo by me http://www.ghostweather.com/essays/talks/networkx/adjacency.html



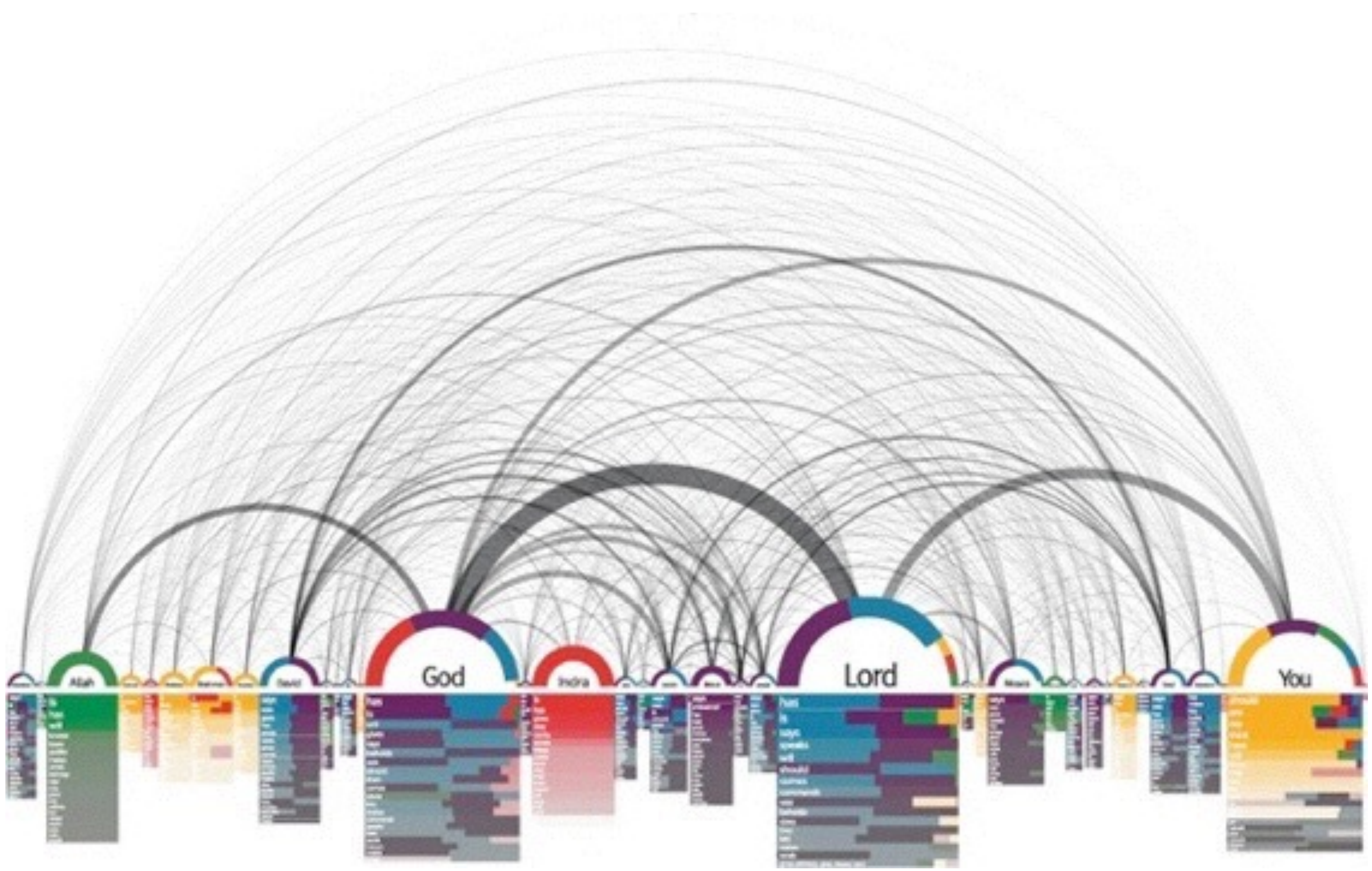
D3 demo by me http://www.ghostweather.com/essays/talks/networkx/adjacency.html

Readability: Matrix Better (except for path finding)



Jean-Daniel Fekete http://lamut.informatik.uni-wuerzburg.de/gd2014/data/uploads/gd2014-fekete-keynote.pdf

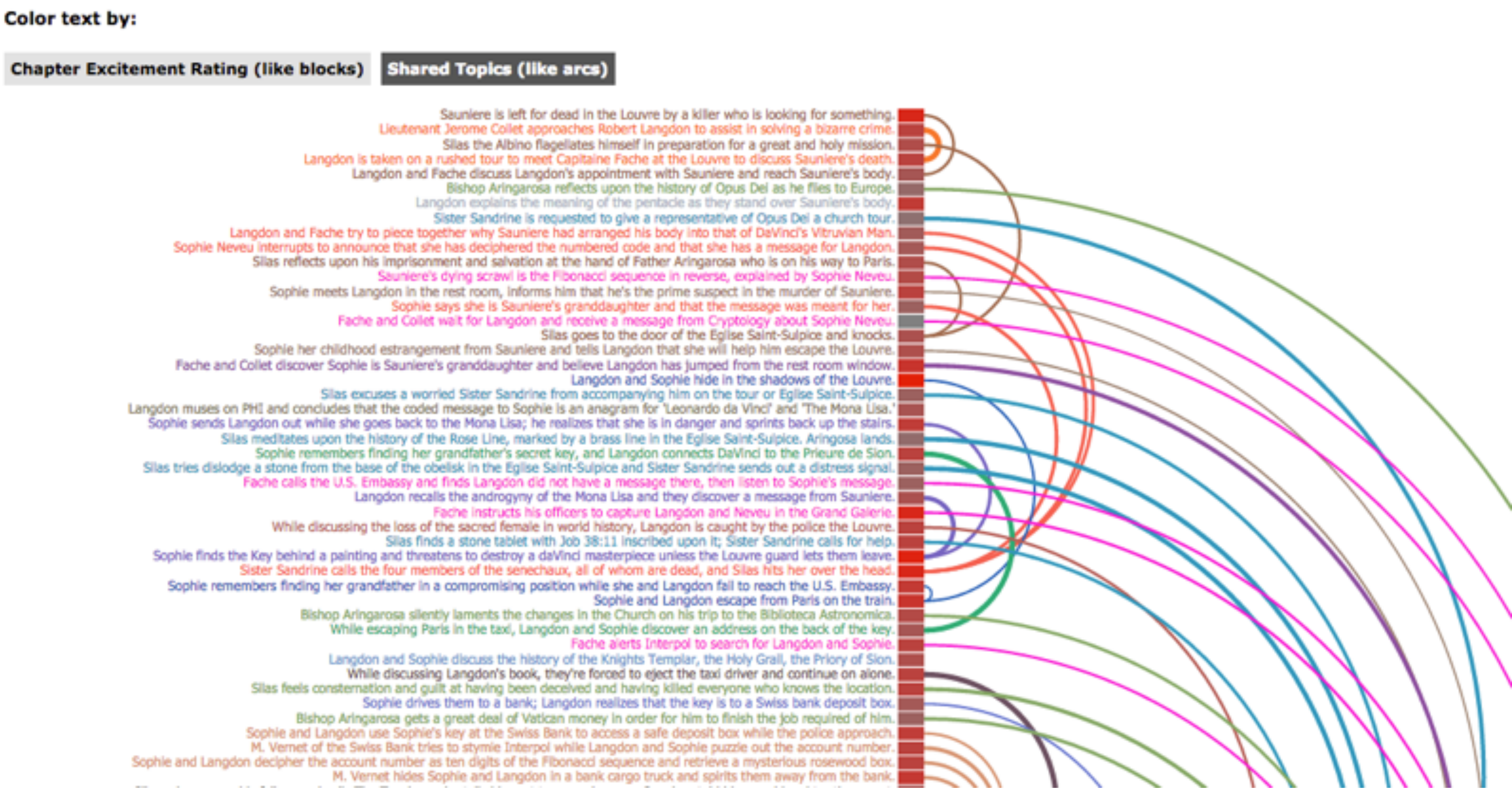
ARC / LINEAR LAYOUTS



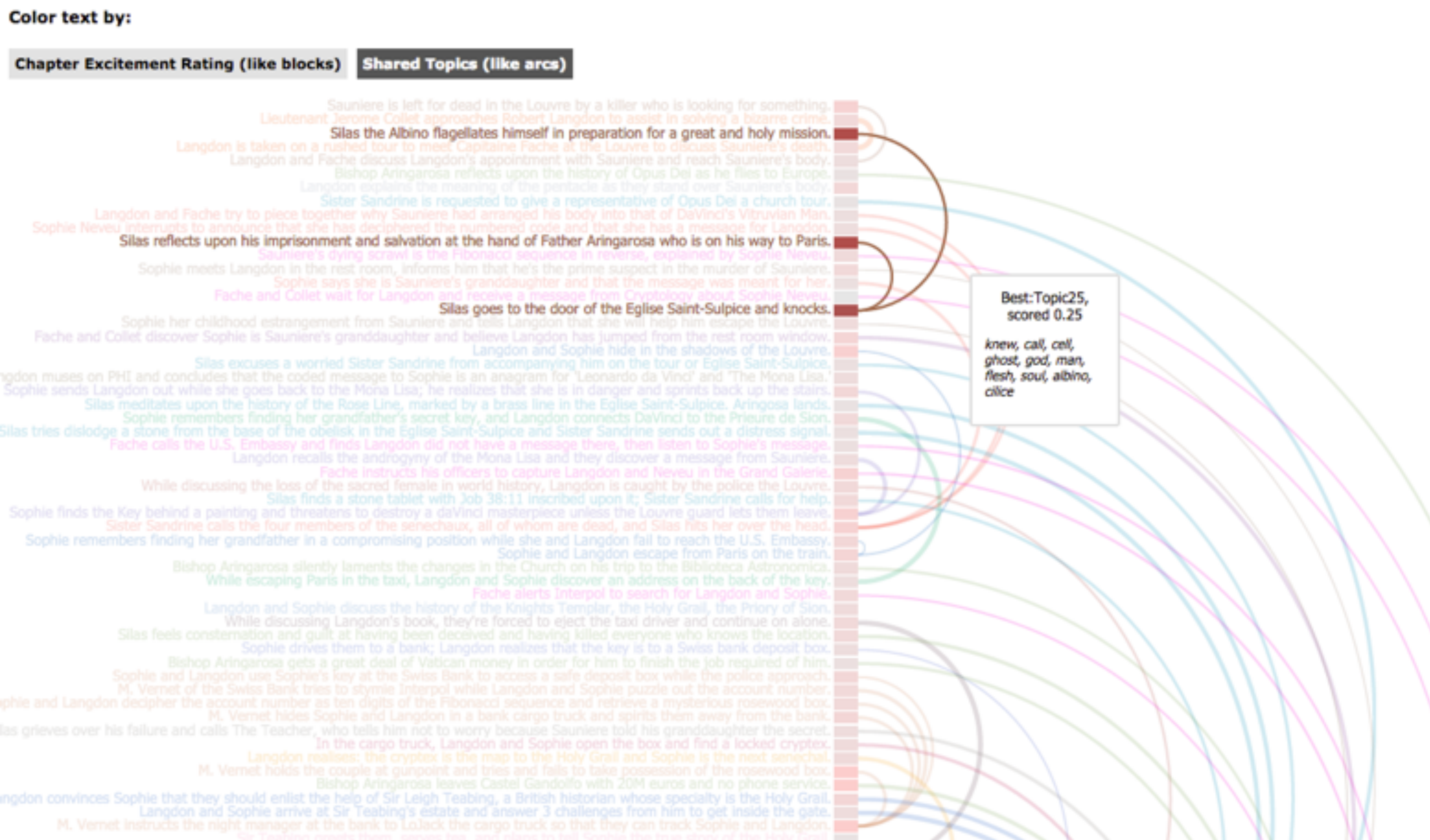
Philipp Steinweber and Andreas Koller
Similar Diversity, 2007

For a D3 example in another domain: http://tradearc.laserdeathstehr.com/

Topics in *Da Vinci Code* (by me)

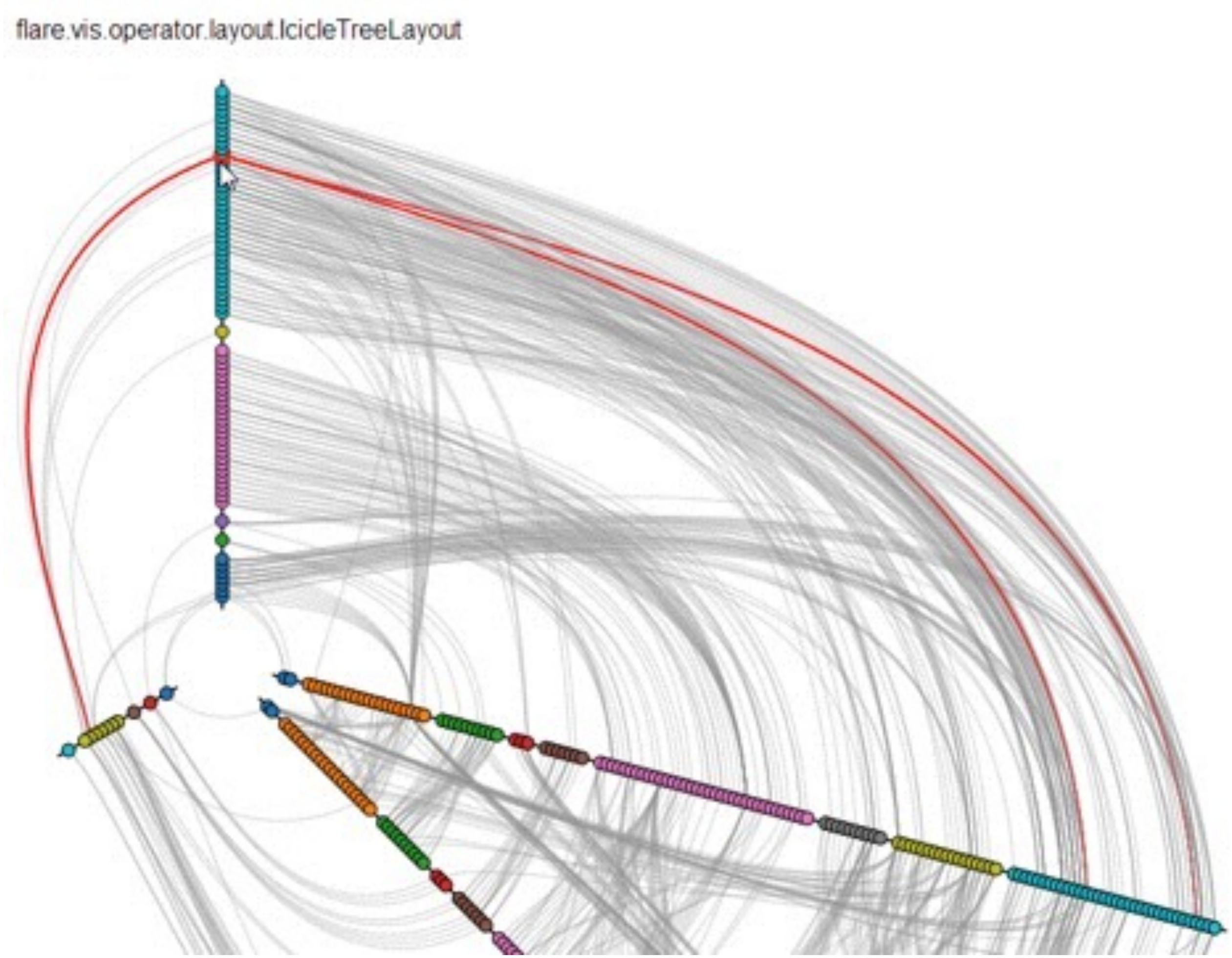


Topics in *Da Vinci Code* (by me)



Lynn Cherny - http://www.ghostweather.com/essays/talks/openvisconf/topic_arc_diagram/TopicArc.html

Hive Plots

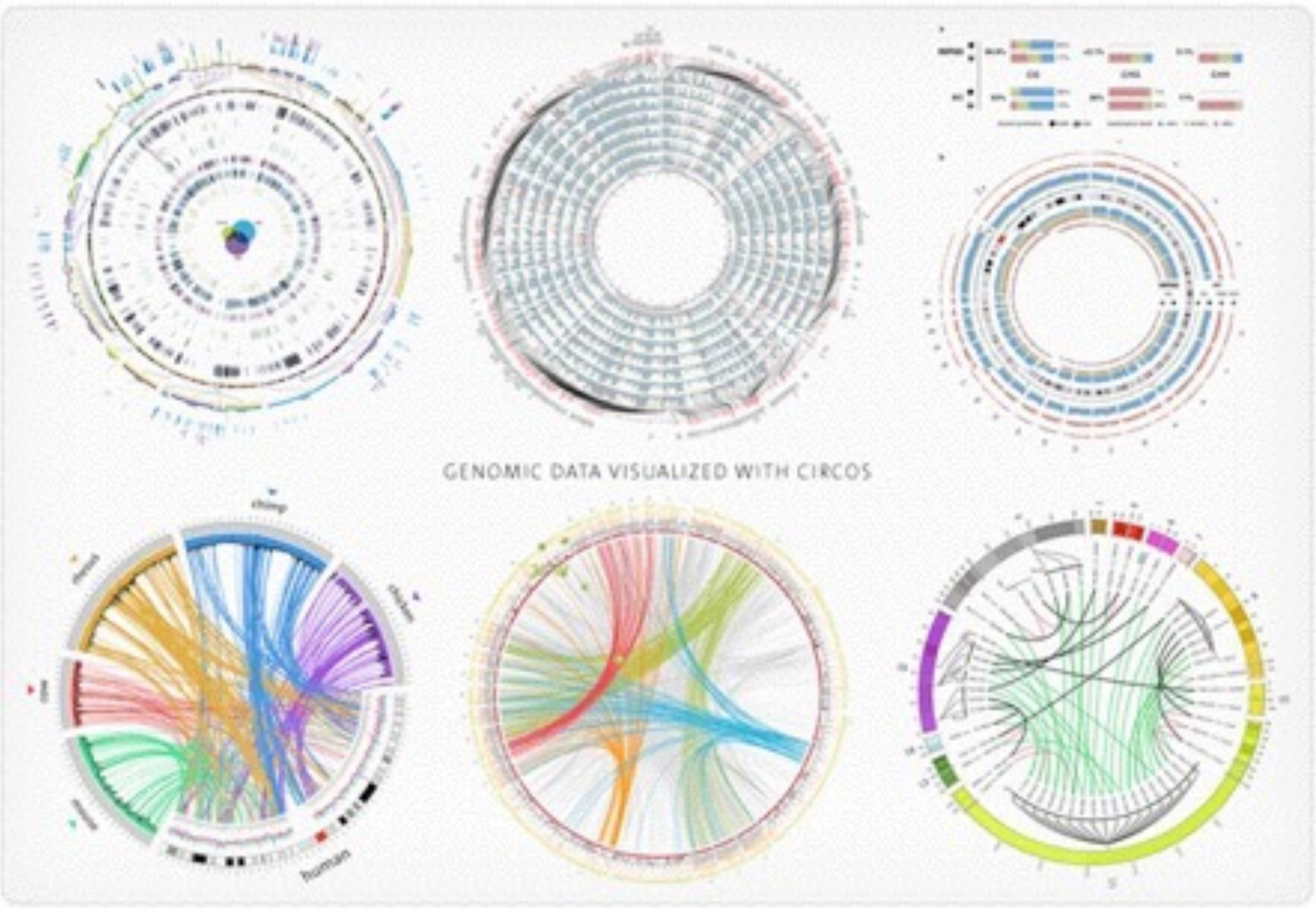


D3: http://bost.ocks.org/mike/hive/

CIRCULAR / CHORD LAYOUTS

Circos

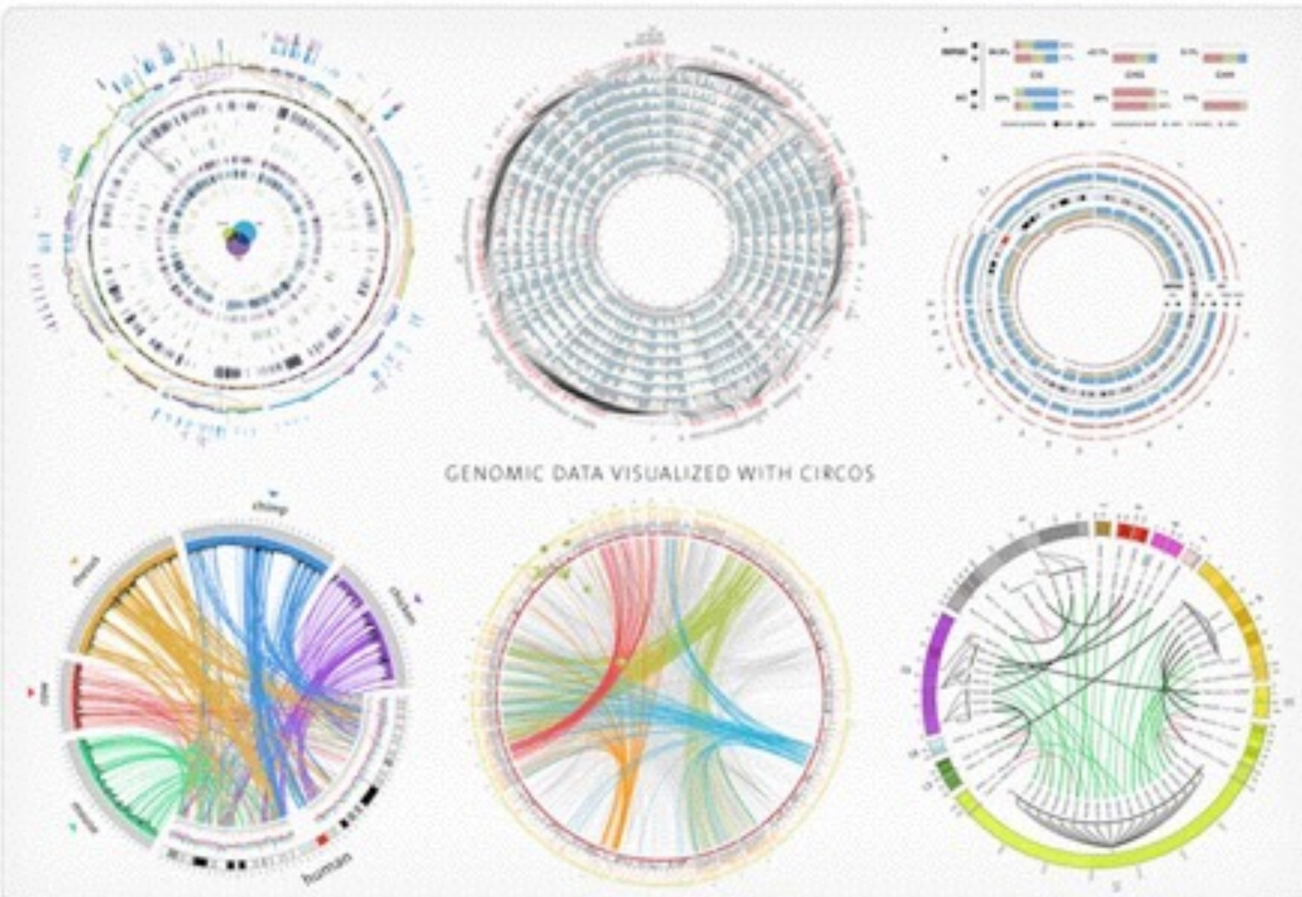
"If it's round, Circos can probably do it"



http://circos.ca/images/

Circos

"If it's round, Circos can probably do it"

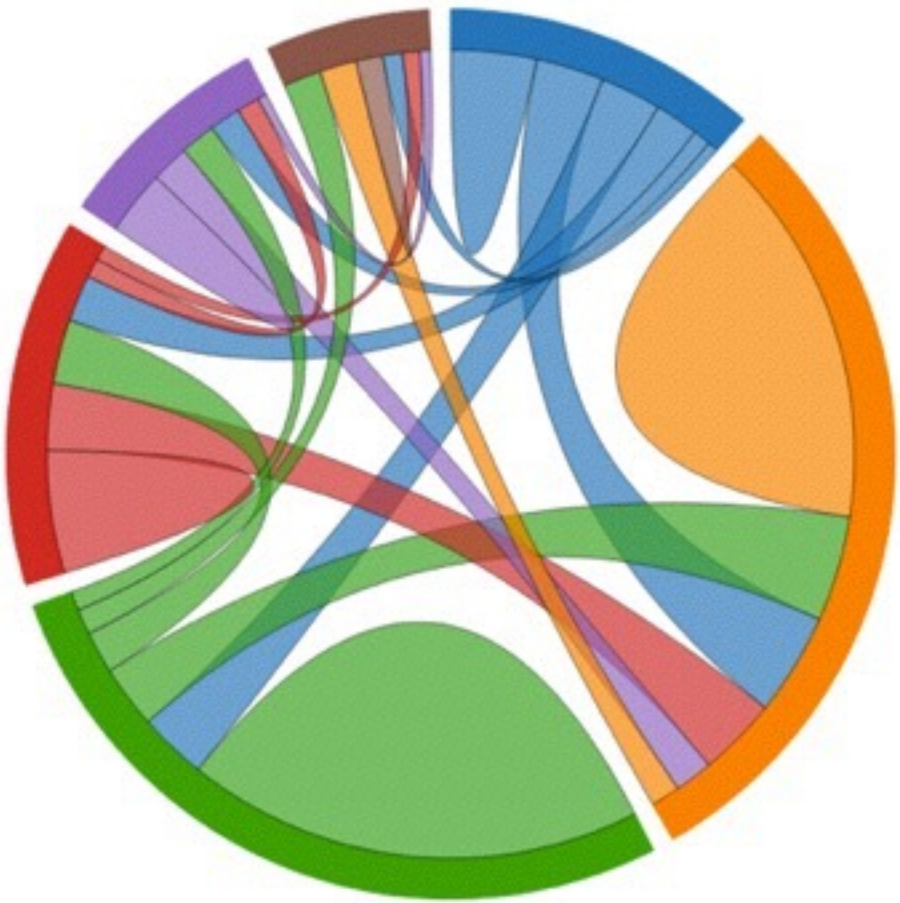


TIL (shocking): It's all perl code?!

http://circos.ca/images/

Computed Relations Between Nodes as Chord Diagram

Nodes shown represent a high-scoring subset of the full 1644 node dataset from Moritz Stefaner's crawl of infovis tweeters in Summer 2011.



Select Data Set:

Full Data (1644)

Nodes: 1644

Edges: 73922

Computed Source/Target Matrix

	0	1	2	3	4	5
0	2663	2107	1896	1360	544	332
1	3182	11823	3263	2456	1192	911
2	2111	1958	14203	1302	597	784
3	1452	2071	1983	3803	598	461
4	964	1183	1096	712	2009	282
5	617	1151	1162	528	333	833

Color chords by larger:

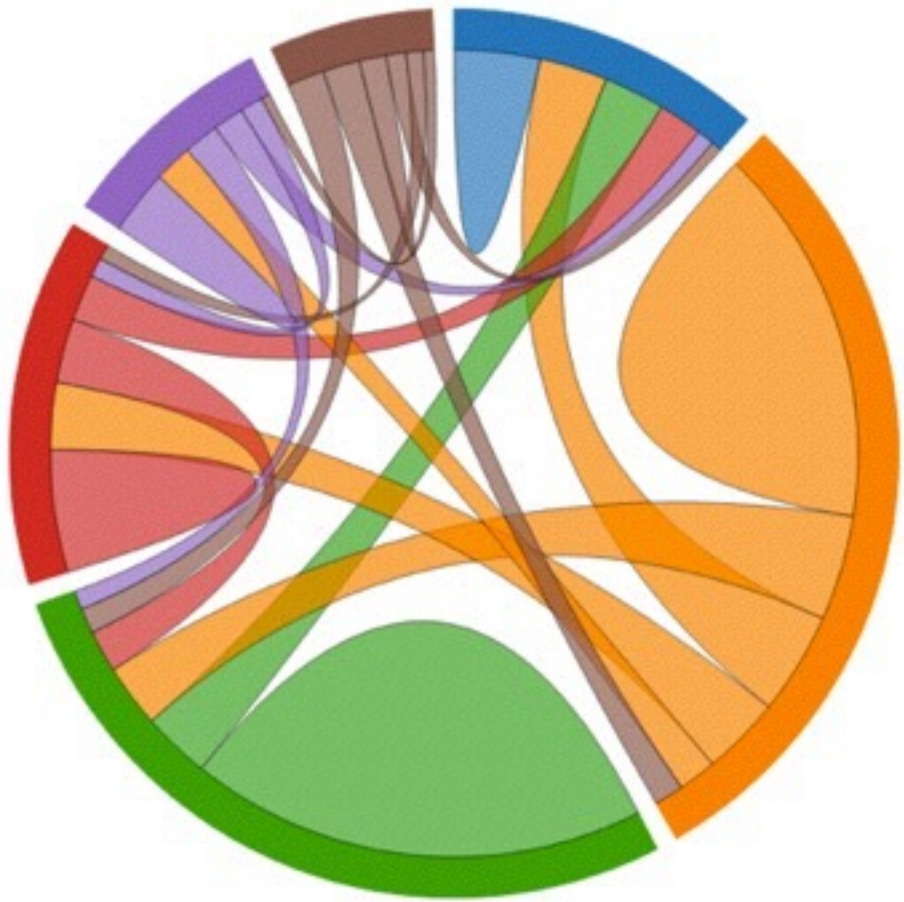
○ Source ⦿ Target

Built with D3 by Lynn Cherny from NetworkX analysis with accompanying talk slides and blog post.

http://www.ghostweather.com/essays/talks/networkx/chord.html

Computed Relations Between Nodes as Chord Diagram

Nodes shown represent a high-scoring subset of the full 1644 node dataset from Moritz Stefaner's crawl of infovis tweeters in Summer 2011.



Select Data Set:

Full Data (1644)

Nodes: 1644

Edges: 73922

Computed Source/Target Matrix

	0	1	2	3	4	5
0	2663	2107	1896	1360	544	332
1	3182	11823	3263	2456	1192	911
2	2111	1958	14203	1302	597	784
3	1452	2071	1983	3803	598	461
4	964	1183	1096	712	2009	282
5	617	1151	1162	528	333	833

Color chords by larger:

(•) Source ( ) Target

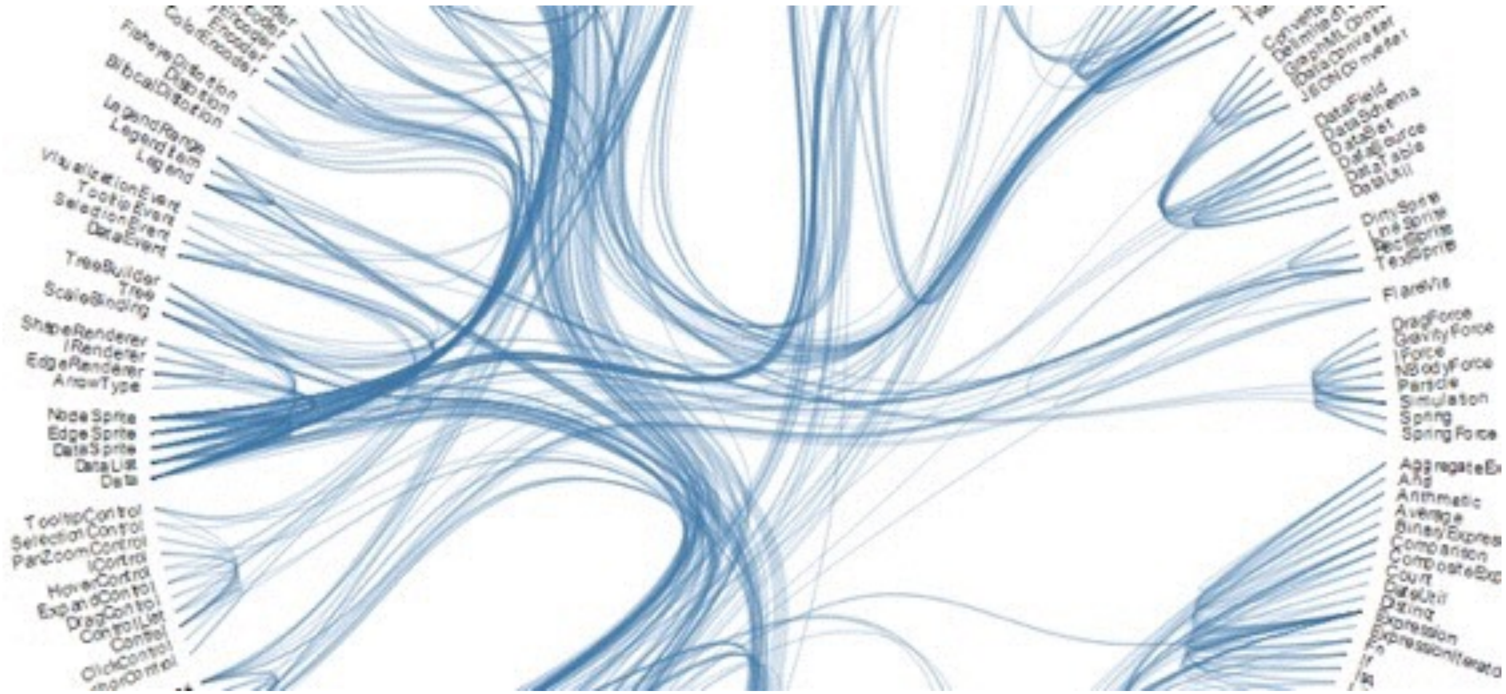
Built with D3 by Lynn Cherny from NetworkX analysis with accompanying talk slides and blog post.

http://www.ghostweather.com/essays/talks/networkx/chord.html

A Nice UI Improvement - Differentiate Source/ Destination

Nikola Sander, http://www.global-migration.info/

Hierarchical Edge Bundling demo by @mbostock



D3: http://bl.ocks.org/1044242

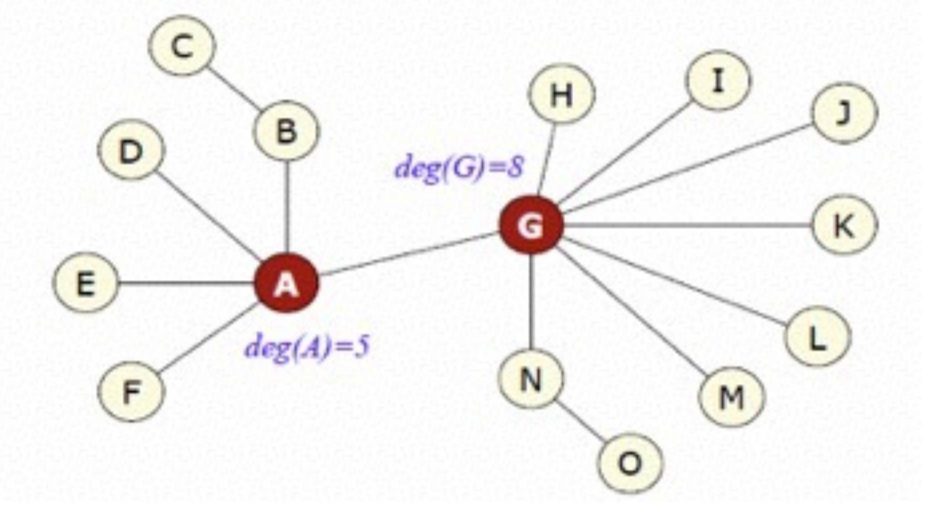
"Hierarchical Edge Bundles: Visualization of Adjacency Relations in Hierarchical Data", Danny Holten, IEEE Transactions on Visualization and Computer Graphics (TVCG; Proceedings of Vis/InfoVis 2006), Vol. 12, No. 5, Pages 741 - 748, 2006.

SIMPLE CALCULATIONS ON NETWORKS CAN TELL YOU LOADS

Often you need to visualize the structure/role of the graph elements as part of the visualization: So, do some simple math.

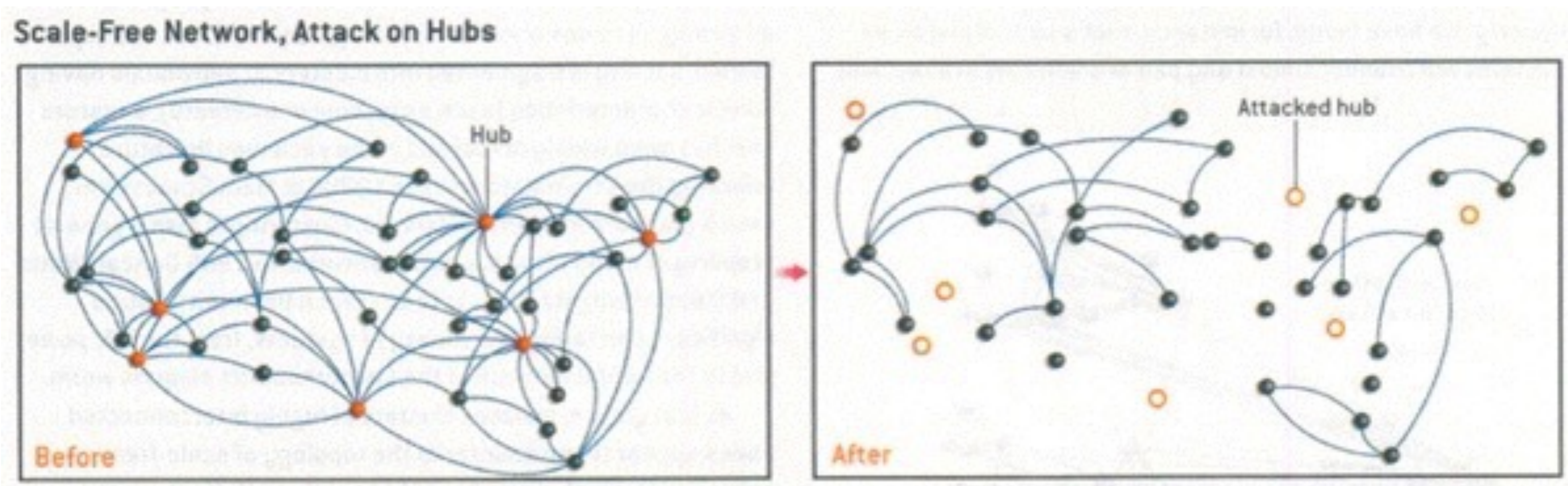
Degree (In, Out)

- "Degree" is a measure of the edges in (directed), out (directed), or total (in directed or undirected graphs) to a node
- "Hub" nodes have high in-degree. In scale-free networks, we see preferential attachment to the popular kids.



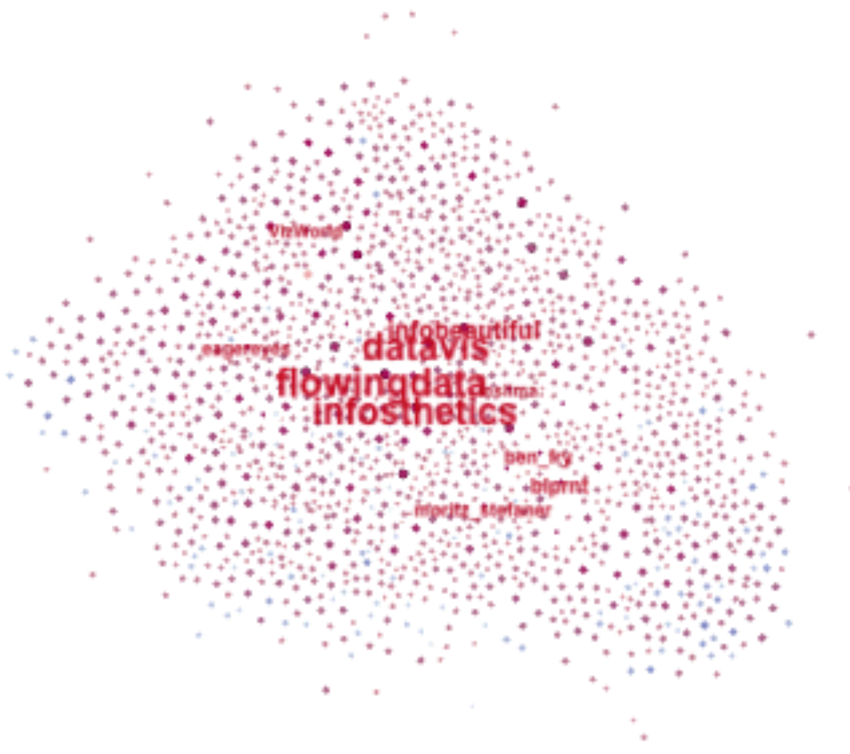
http://mlg.ucd.ie/files/summer/tutorial.pdf

The Threat of Hub-Loss



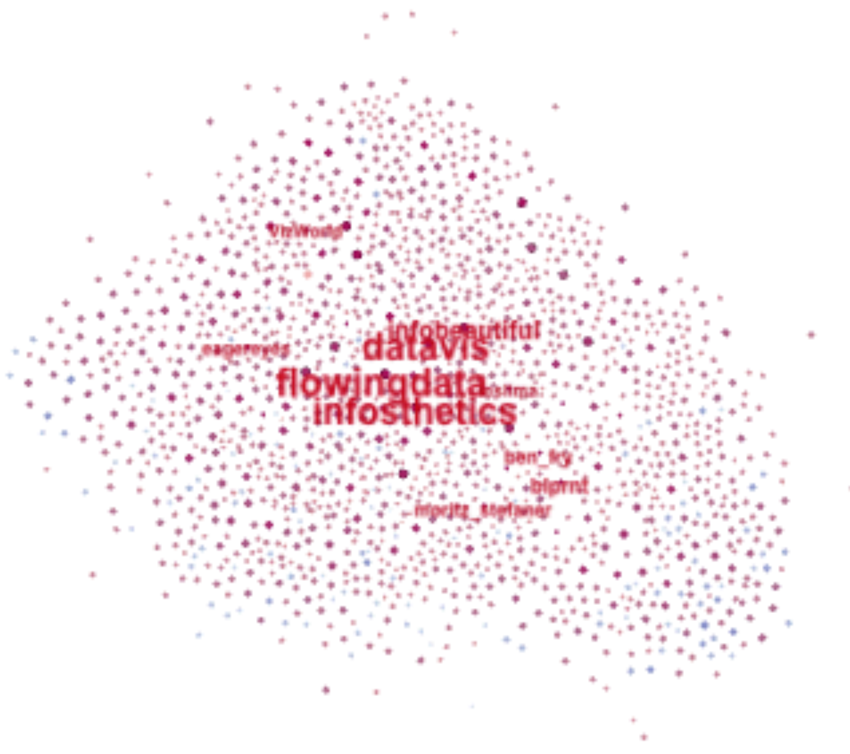
Albert-László Barabási and Eric Bonabeau, Scale-Free Networks, 2003.http://www.scientificamerican.com/article.cfm?id=scale-free-networks

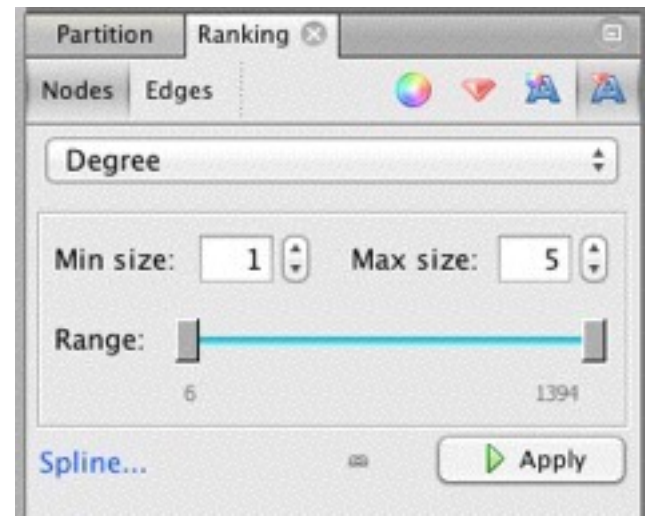
Visualization Aside: If Some Names are Huge, the Others are Invisible-?



Correcting for text size by degree display issue

Visualization Aside: If Some Names are Huge, the Others are Invisible-?

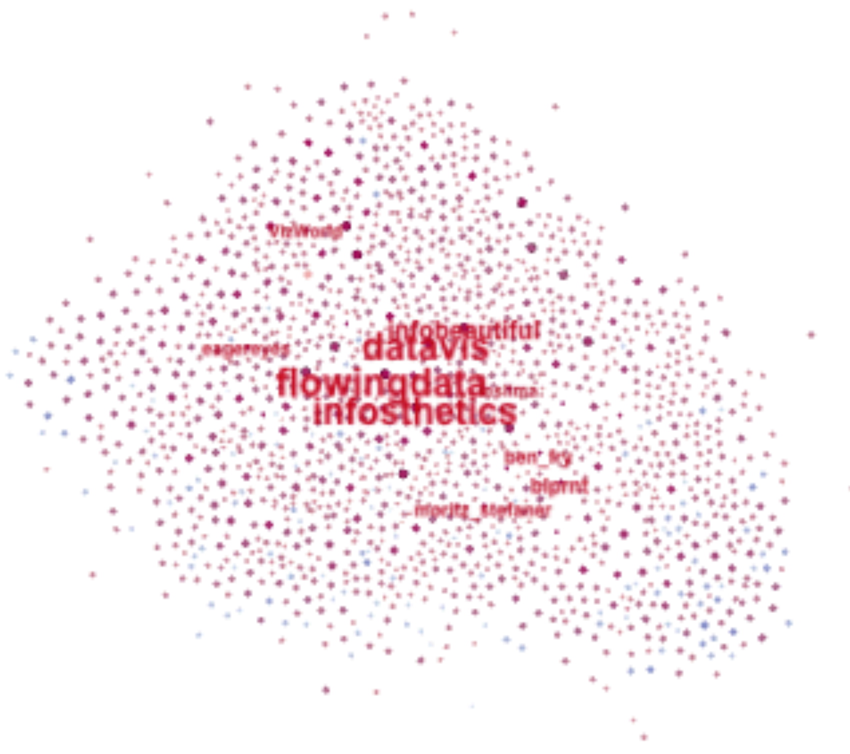


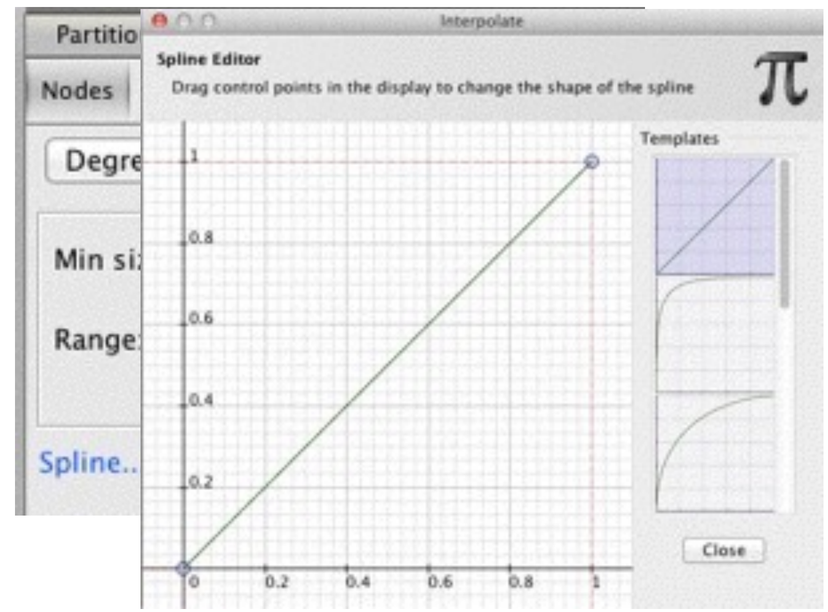


Gephi Panel

Correcting for text size by degree display issue

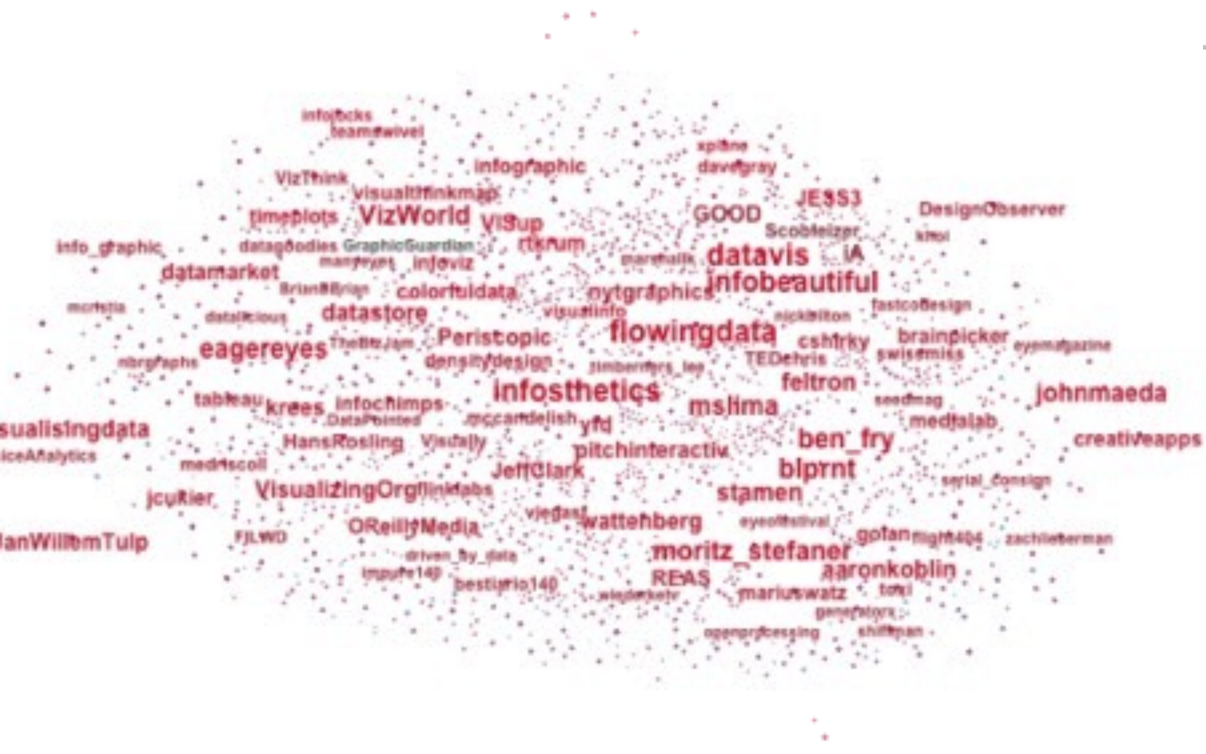
Visualization Aside: If Some Names are Huge, the Others are Invisible-?

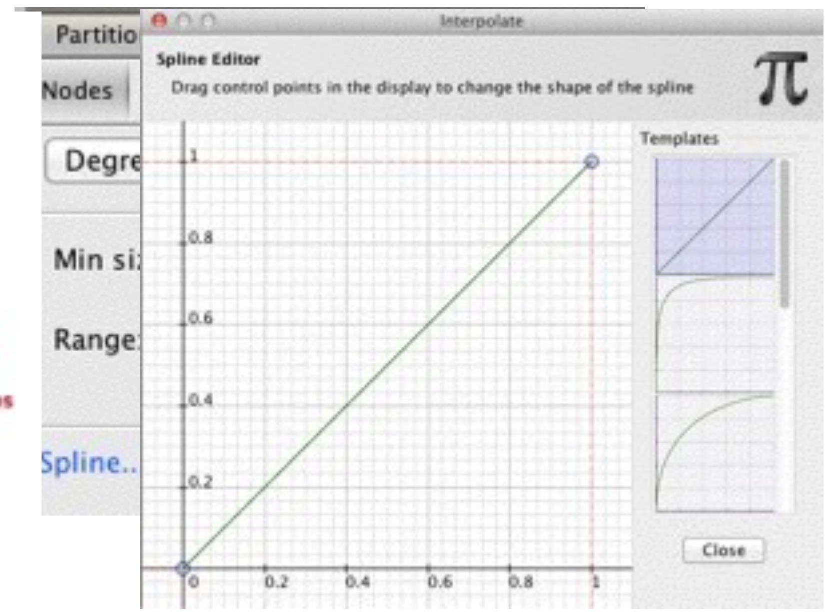




Correcting for text size by degree display issue

Visualization Aside: If Some Names are Huge, the Others are Invisible-?



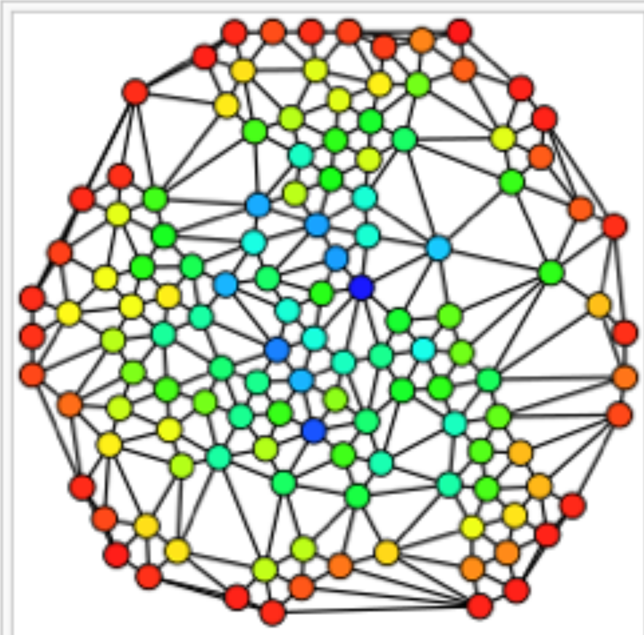


Correcting for text size by degree display issue

Betweenness

A measure of connectedness between (sub)components of the graph

“Betweenness centrality thus tends to pick out boundary individuals who play the role of brokers between communities.”



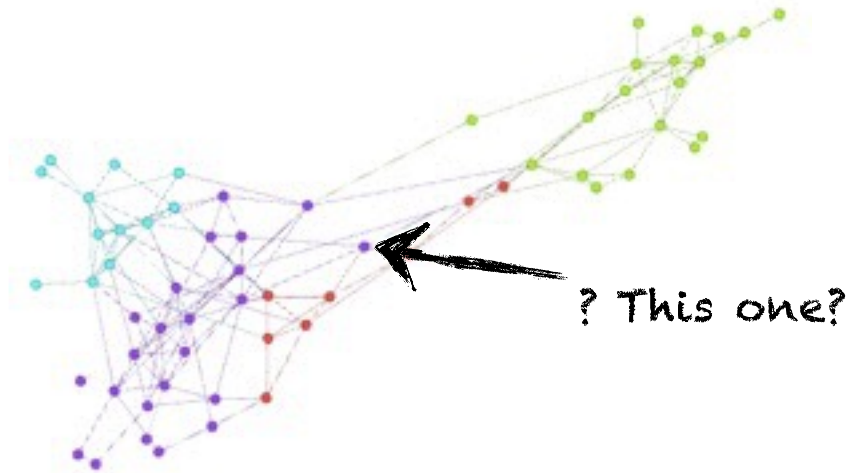
Hue (from red=0 to blue=max) shows the node betweenness.

Lusseau and Newman. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1810112/pdf/15801609.pdf

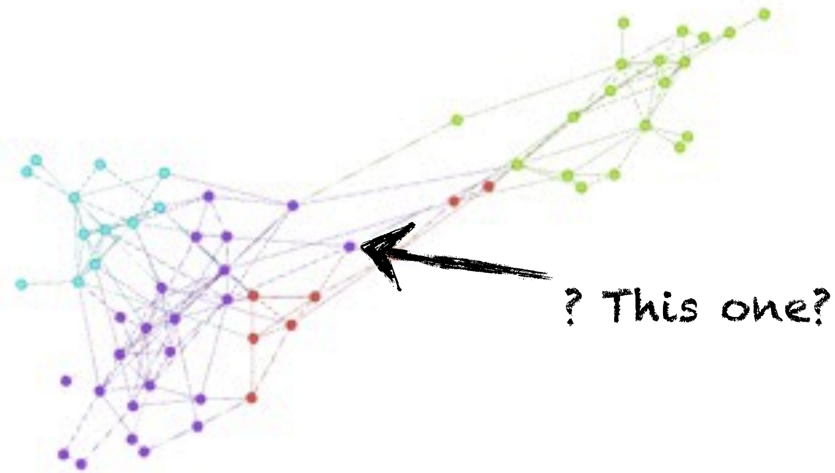
http://en.wikipedia.org/wiki/Centrality#Betweenness_centrality

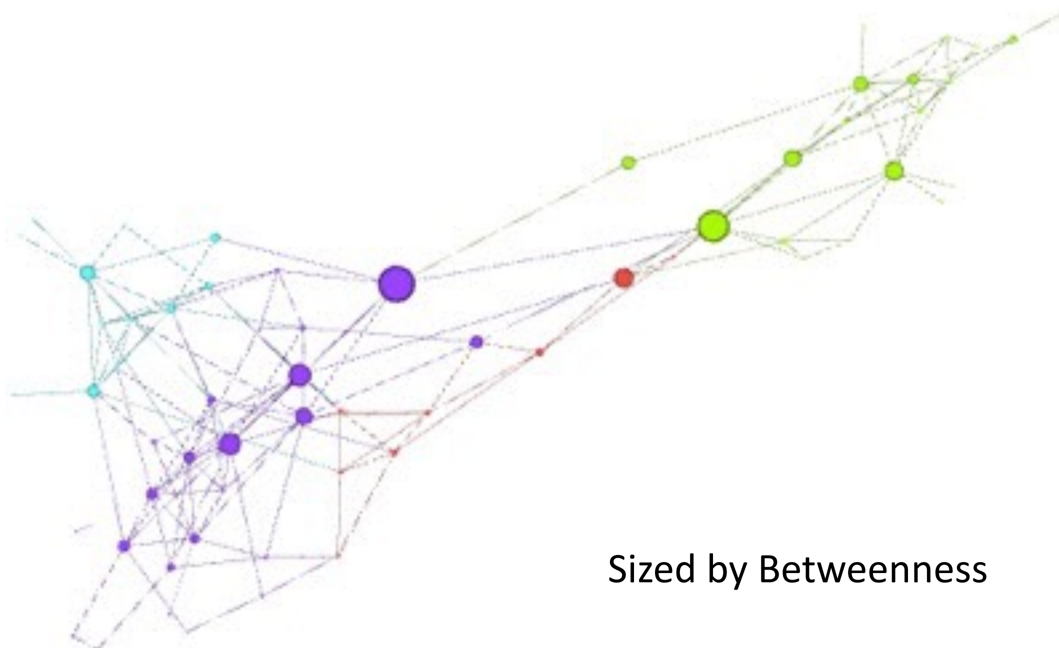
Judging By Eye Will Probably Be Wrong...

Judging By Eye Will Probably Be Wrong...



Judging By Eye Will Probably Be Wrong...





Sized by Betweenness

Eigenvector Centrality

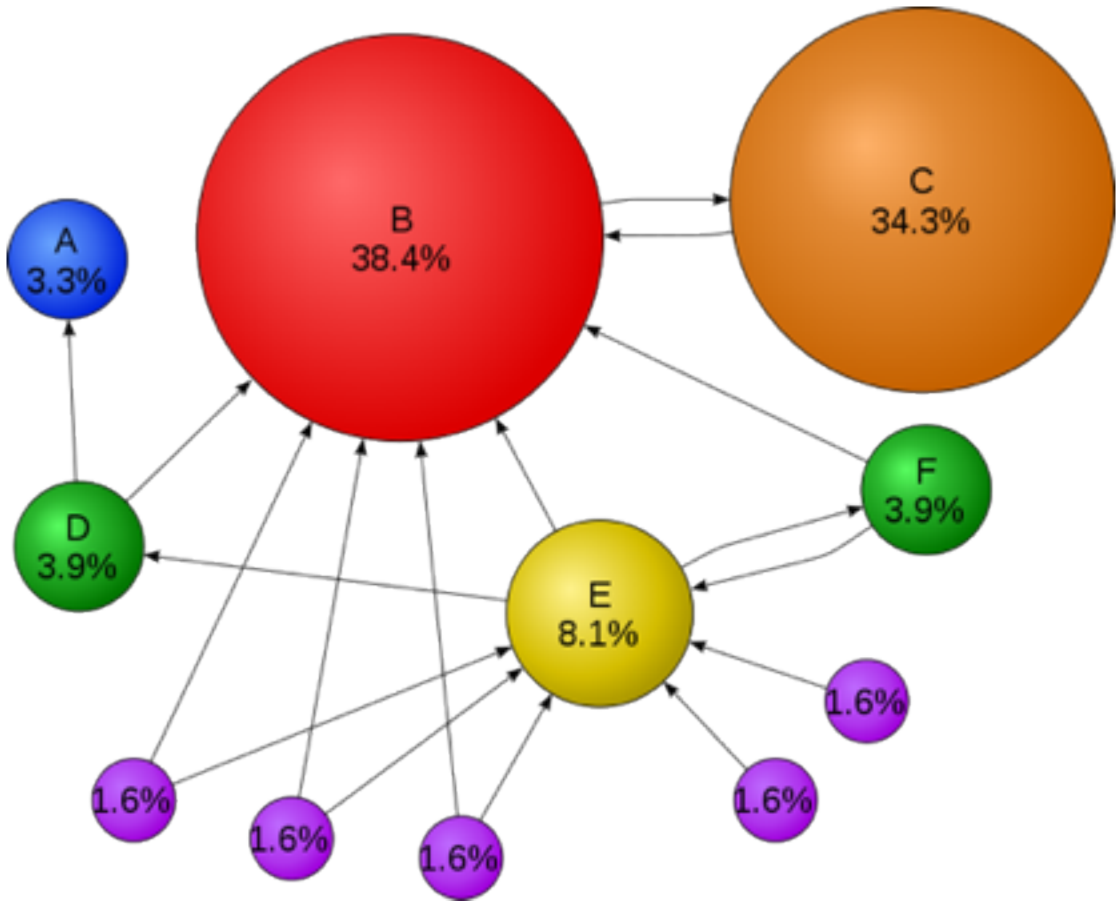
- Intuition: A node is important if it is connected to other important nodes
- A node with a small number of influential contacts may outrank one with a larger number of mediocre contacts

$$\text{adjacency matrix } A: \begin{pmatrix} 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \end{pmatrix} \qquad \text{centrality matrix } x: \begin{pmatrix} 0.38374 \\ 0.833698 \\ 0.730143 \\ 0.833698 \\ 0.30813 \\ 1. \end{pmatrix}$$

http://mlg.ucd.ie/files/summer/tutorial.pdf

http://demonstrations.wolfram.com/NetworkCentralityUsingEigenvectors/

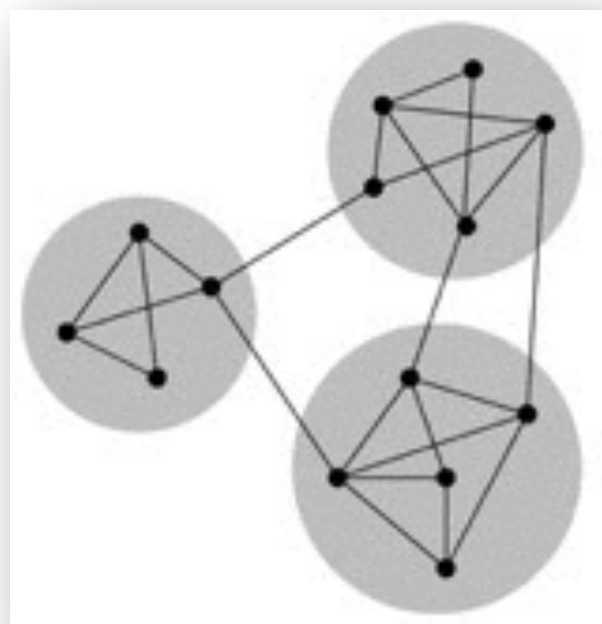
Pagerank



Wikipedia image

Community Detection Algorithms

E.g., the Louvain method, in Gephi as "Modularity." Many layout algorithms help you intuit these structures, but don't rely on perception of layout!



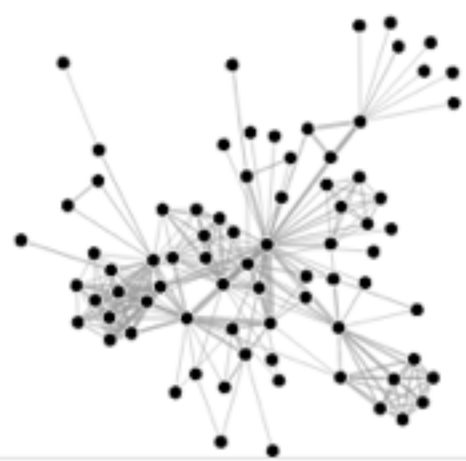
http://en.wikipedia.org/wiki/File:Network_Community_Structure.png

You can even do it in your browser now…

john-guerra's block #ecdde32ab4ad91a1a240 October 31, 2014

Javascript Network Clustering Library

Cluster



netClustering allows you to detect clusters in networks using the Clauset, Newman and Moore community detection algorithm directly from the browser, as simple as:

Open in a new window.

netClustering.cluster(nodes, edges);

http://bl.ocks.org/john-guerra/ecdde32ab4ad91a1a240

You can even do it in your browser now…

john-guerra's block #ecdde32ab4ad91a1a240 October 31, 2014

Javascript Network Clustering Library

Cluster

netClustering allows you to detect clusters in networks using the Clauset, Newman and Moore community detection algorithm directly from the browser, as simple as:

Open in a new window.

netClustering.cluster(nodes, edges);

http://bl.ocks.org/john-guerra/ecdde32ab4ad91a1a240

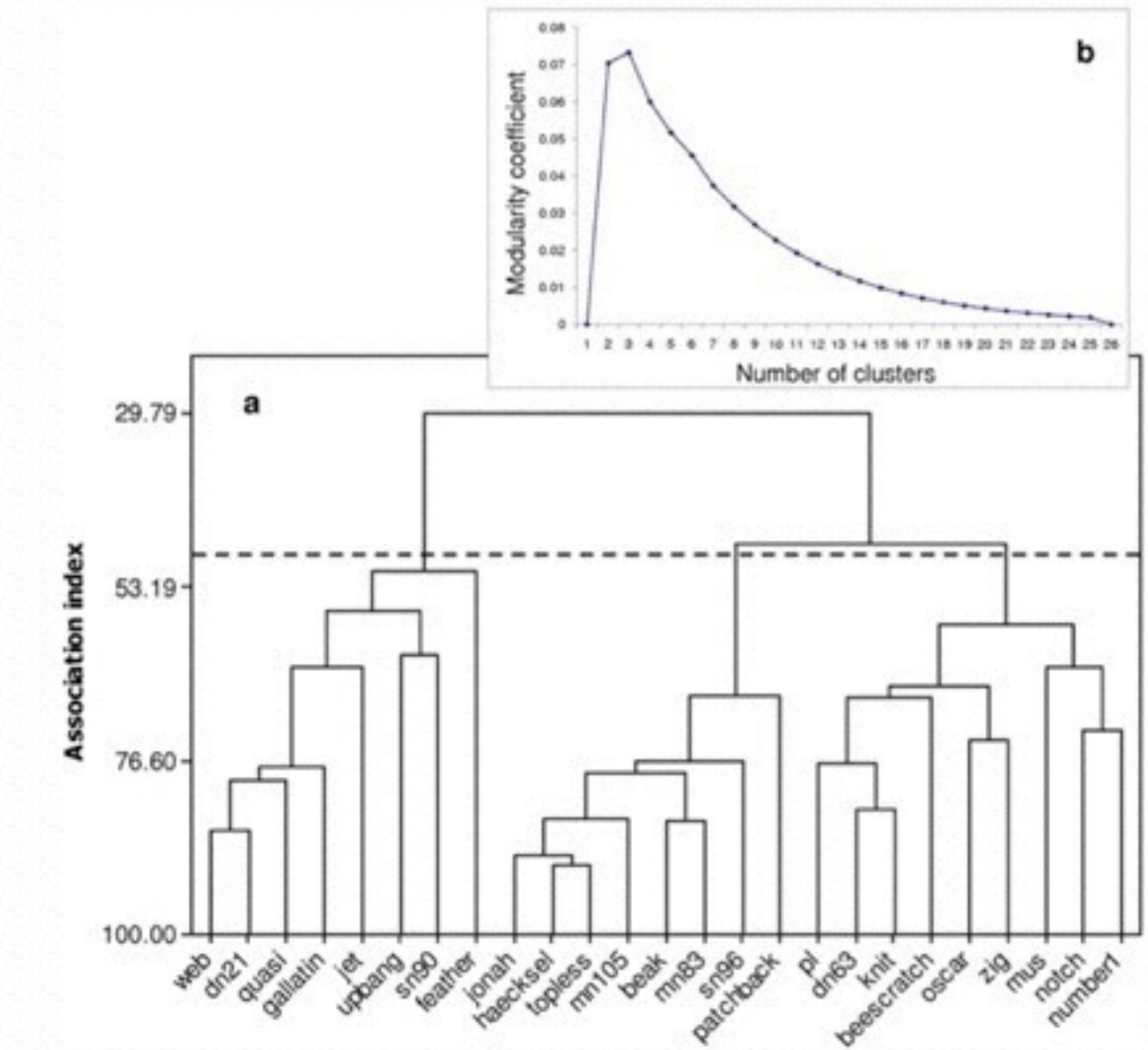


Figure 1. (a) Average linkage cluster analysis of the male association matrix, based on half-weight index (the association index). There appears to be three clusters of individuals spending more time together, the dashed line represents the more parsimonious split in the network which is given by the peak in the modularity coefficient (b).

Citation: Lusseau D (2007) Why Are Male Social Relationships Complex in the Doubtful Sound Bottlenose Dolphin Population?. PLoS ONE 2(4): e348. doi:10.1371/journal.pone.0000348

Identifying the role that animals play in their social networks (2004)
D Lusseau, MEJ Newman
Proceedings of the Royal Society of London. Series B: Biological Sciences

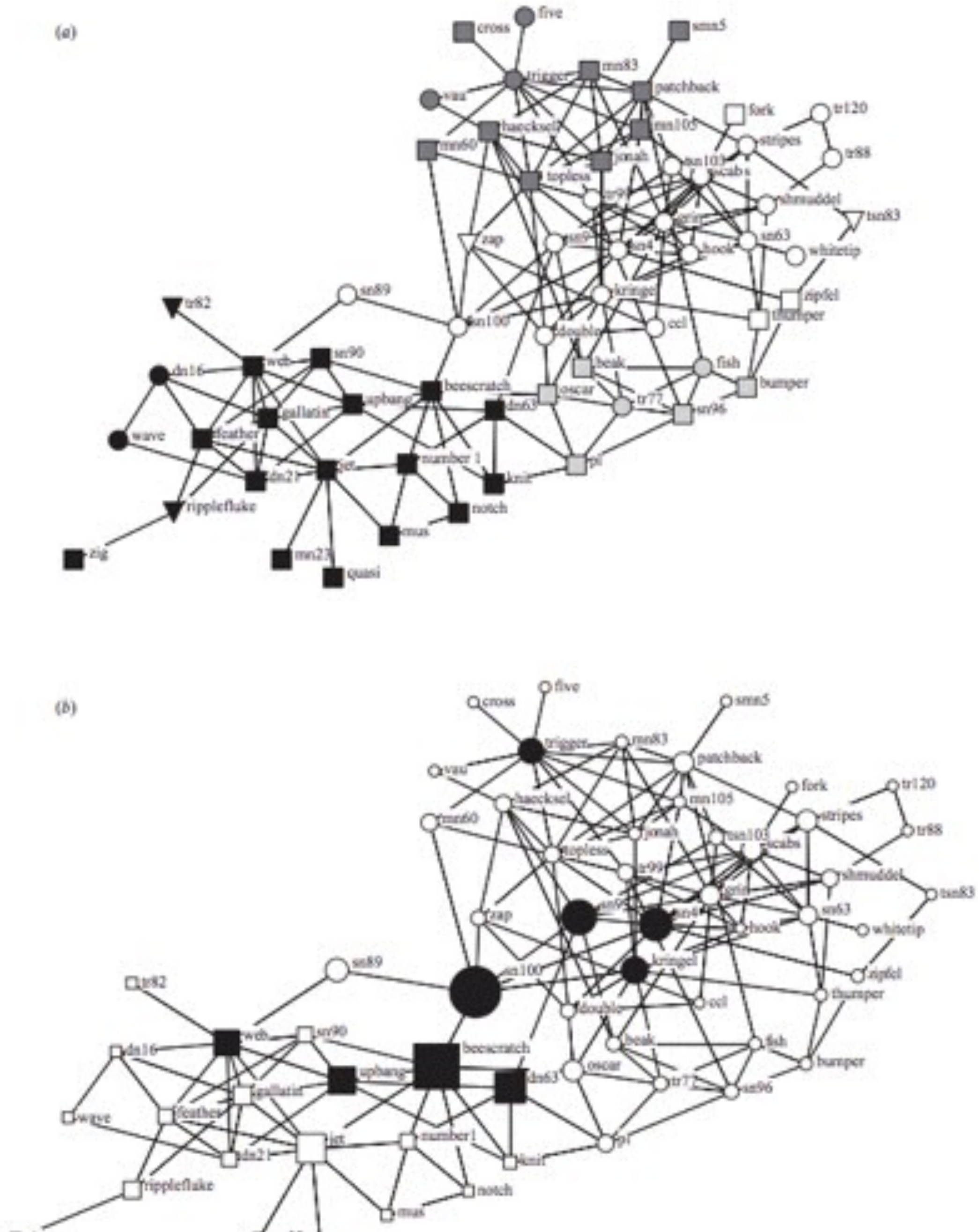


Figure 1. Communities and sub-communities identified in the dolphin social network using the betweenness-based algorithm of Girvan & Newman (2002). (*a*) Vertex shading indicates community membership: vertices in black are all part of one community, while all other vertices are part of the second community. This second community is further divided into three sub-communities by the algorithm; these are represented by the shading (white, light grey and dark grey). Females are represented as circles, males as squares and individuals with unknown gender as triangles. (*b*) Individuals with high betweenness values (greater than 7.32) are represented by black symbols. The size of the black symbol is directly related to the betweenness of the vertex.

Identifying the role that animals play in their social networks (2004)
D Lusseau, MEJ Newman
Proceedings of the Royal Society of London. Series B: Biological Sciences

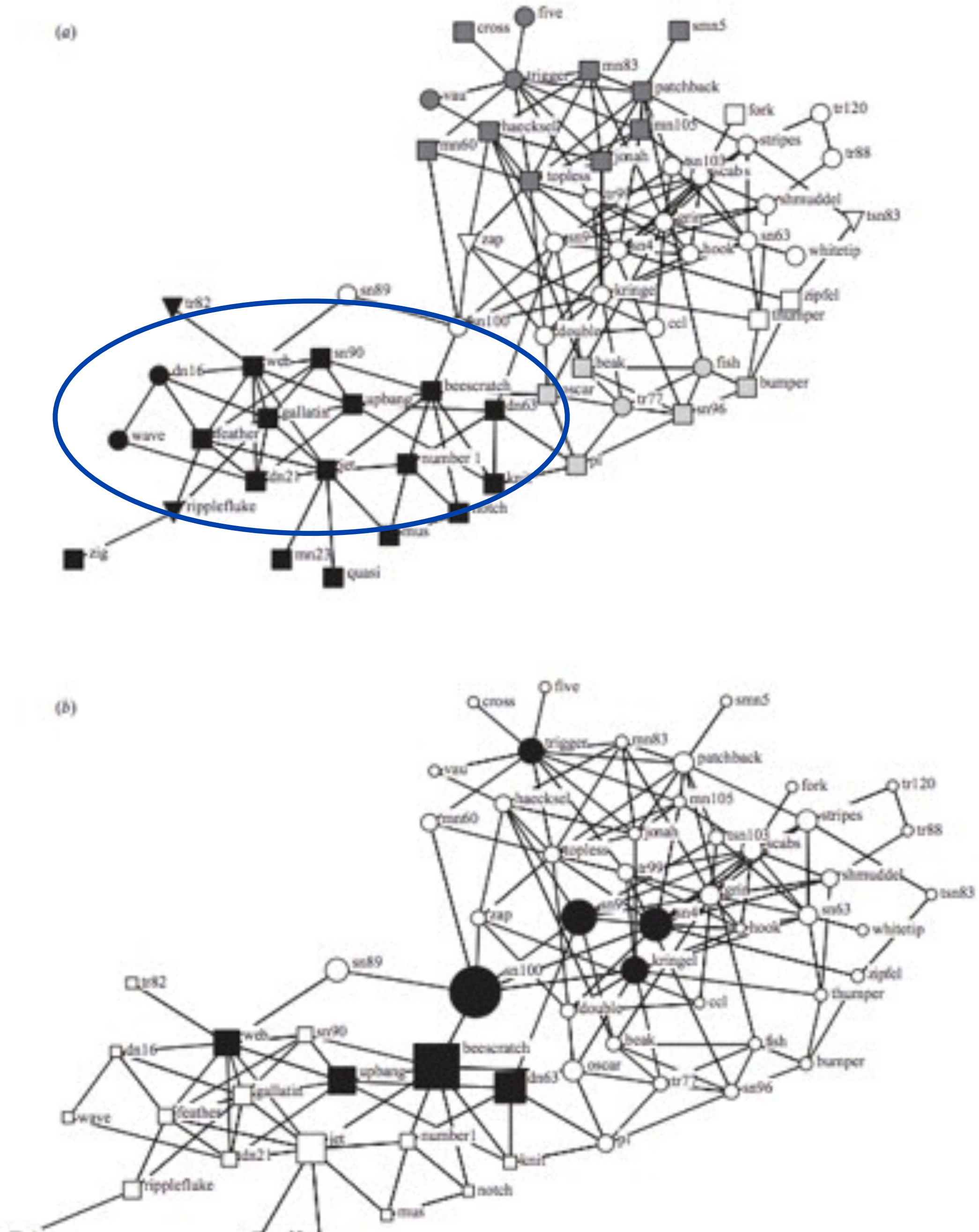


Figure 1. Communities and sub-communities identified in the dolphin social network using the betweenness-based algorithm of Girvan & Newman (2002). (*a*) Vertex shading indicates community membership: vertices in black are all part of one community, while all other vertices are part of the second community. This second community is further divided into three sub-communities by the algorithm; these are represented by the shading (white, light grey and dark grey). Females are represented as circles, males as squares and individuals with unknown gender as triangles. (*b*) Individuals with high betweenness values (greater than 7.32) are represented by black symbols. The size of the black symbol is directly related to the betweenness of the vertex.

Identifying the role that animals play in their social networks (2004)
D Lusseau, MEJ Newman
Proceedings of the Royal Society of London. Series B: Biological Sciences

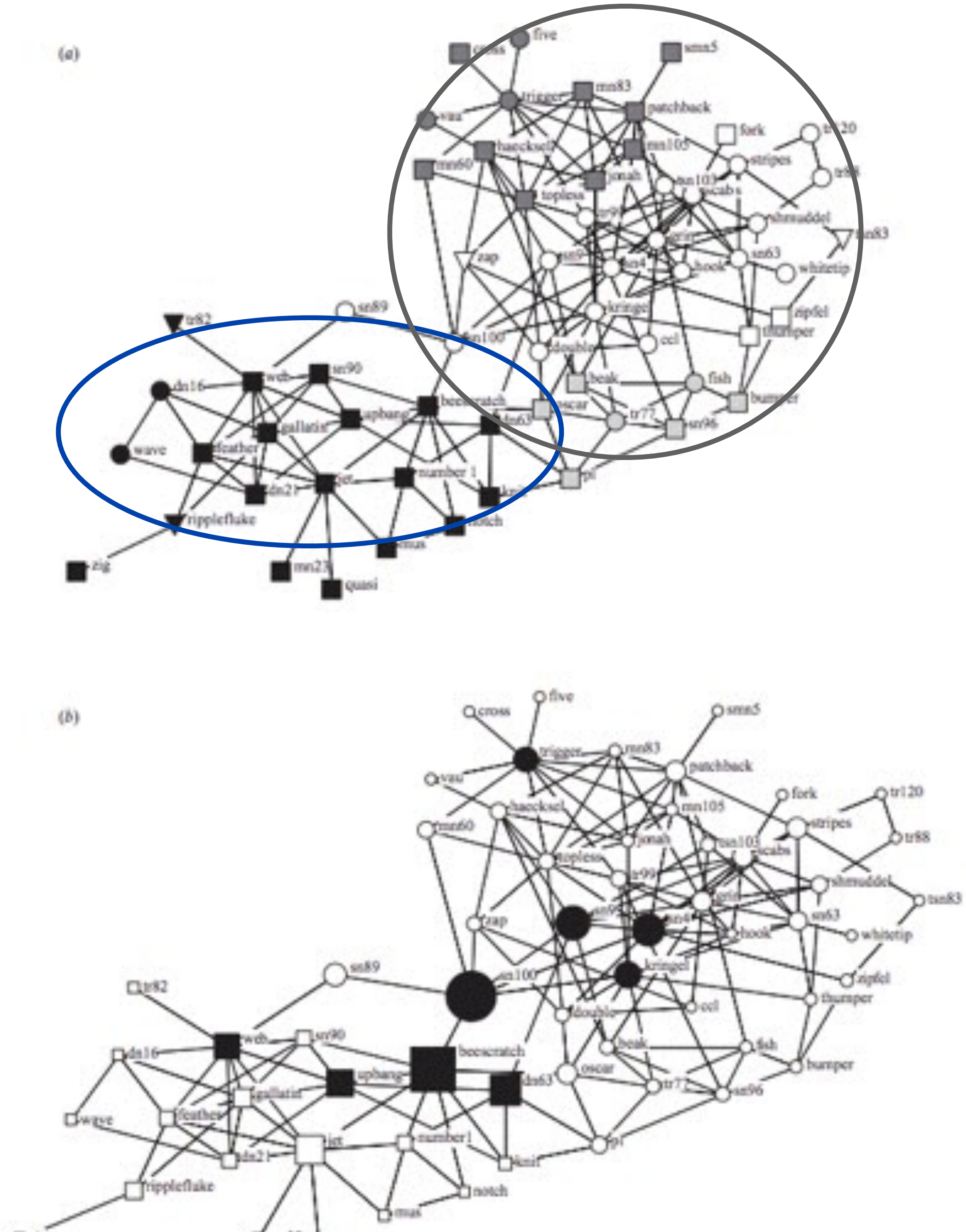


Figure 1. Communities and sub-communities identified in the dolphin social network using the betweenness-based algorithm of Girvan & Newman (2002). (*a*) Vertex shading indicates community membership: vertices in black are all part of one community, while all other vertices are part of the second community. This second community is further divided into three sub-communities by the algorithm; these are represented by the shading (white, light grey and dark grey). Females are represented as circles, males as squares and individuals with unknown gender as triangles. (*b*) Individuals with high betweenness values (greater than 7.32) are represented by black symbols. The size of the black symbol is directly related to the betweenness of the vertex.

Identifying the role that animals play in their social networks (2004)
D Lusseau, MEJ Newman
Proceedings of the Royal Society of London. Series B: Biological Sciences

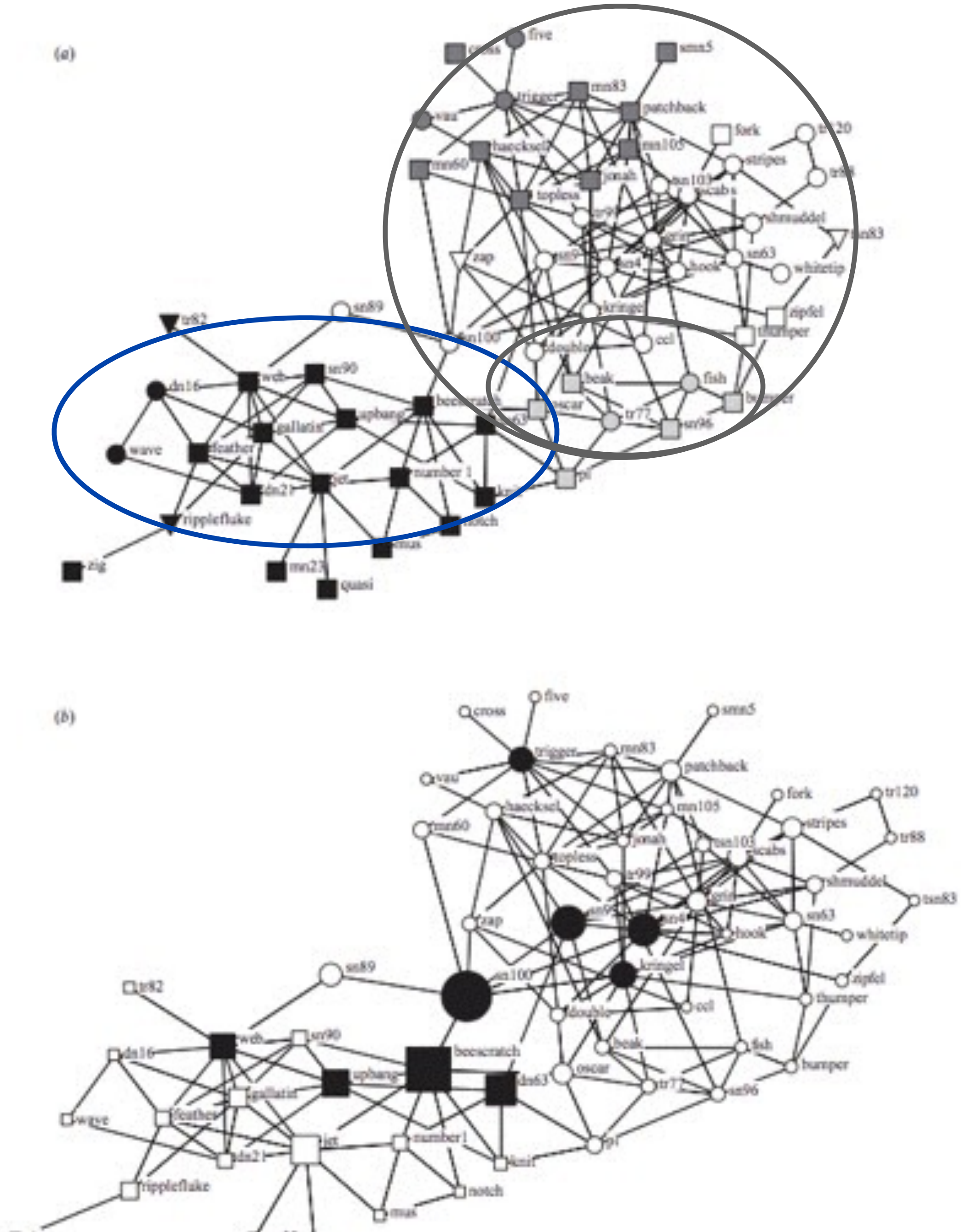


Figure 1. Communities and sub-communities identified in the dolphin social network using the betweeness-based algorithm of Girvan & Newman (2002). (*a*) Vertex shading indicates community membership: vertices in black are all part of one community, while all other vertices are part of the second community. This second community is further divided into three sub-communities by the algorithm; these are represented by the shading (white, light grey and dark grey). Females are represented as circles, males as squares and individuals with unknown gender as triangles. (*b*) Individuals with high betweenness values (greater than 7.32) are represented by black symbols. The size of the black symbol is directly related to the betweenness of the vertex.

Identifying the role that animals play in their social networks (2004)
D Lusseau, MEJ Newman
Proceedings of the Royal Society of London. Series B: Biological Sciences

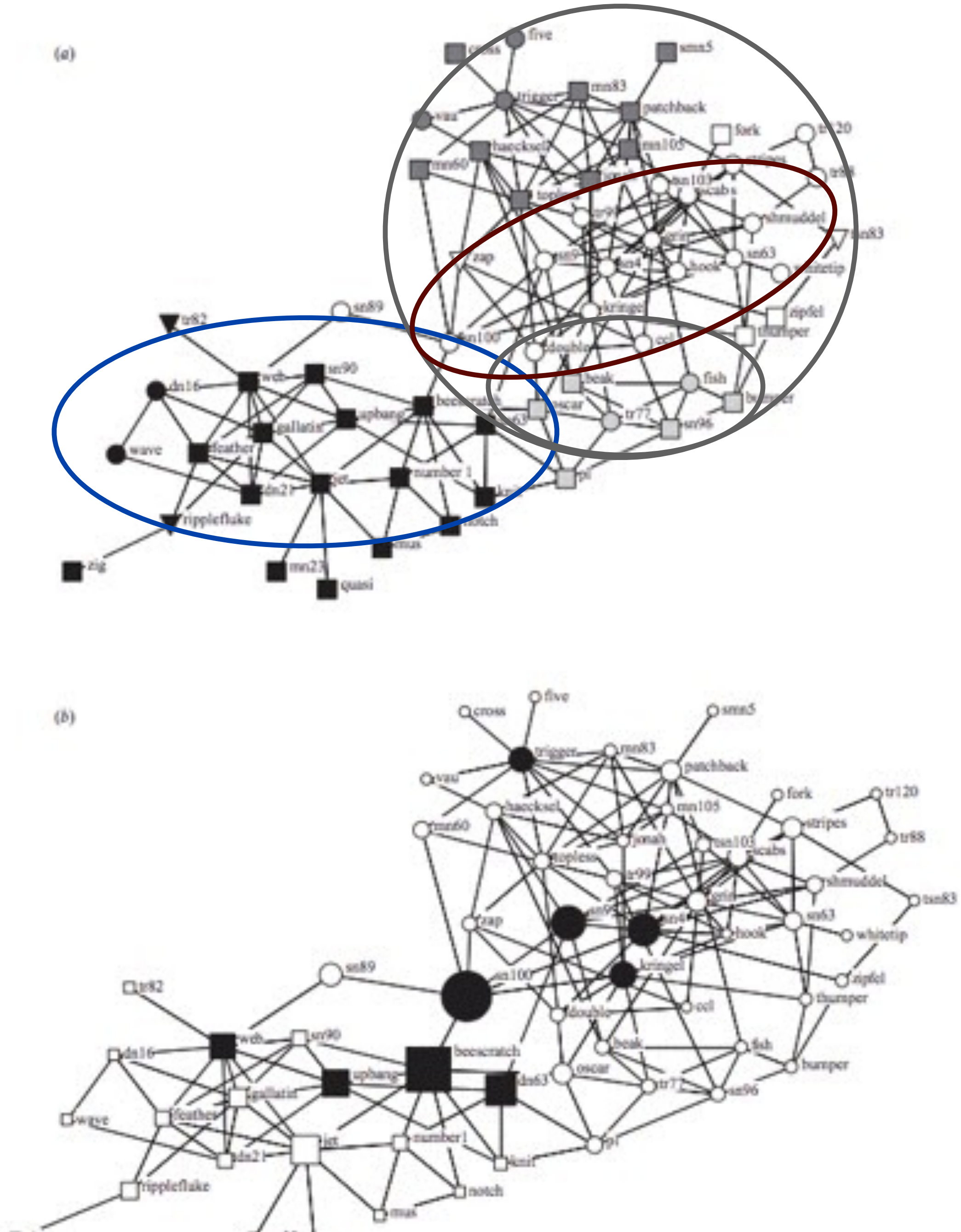


Figure 1. Communities and sub-communities identified in the dolphin social network using the betweenness-based algorithm of Girvan & Newman (2002). (a) Vertex shading indicates community membership: vertices in black are all part of one community, while all other vertices are part of the second community. This second community is further divided into three sub-communities by the algorithm; these are represented by the shading (white, light grey and dark grey). Females are represented as circles, males as squares and individuals with unknown gender as triangles. (b) Individuals with high betweenness values (greater than 7.32) are represented by black symbols. The size of the black symbol is directly related to the betweenness of the vertex.

Identifying the role that animals play in their social networks (2004)
D Lusseau, MEJ Newman
Proceedings of the Royal Society of London. Series B: Biological Sciences

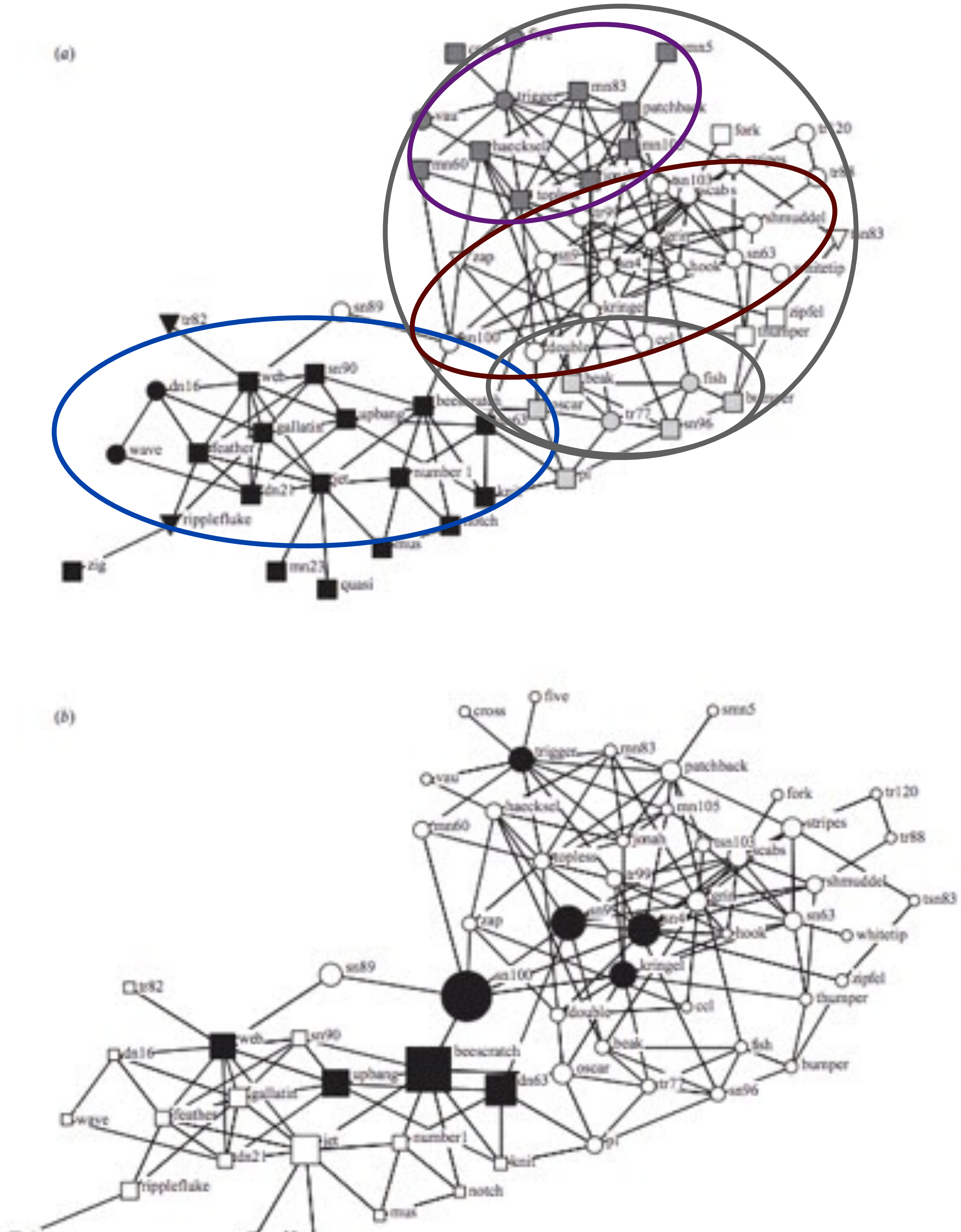
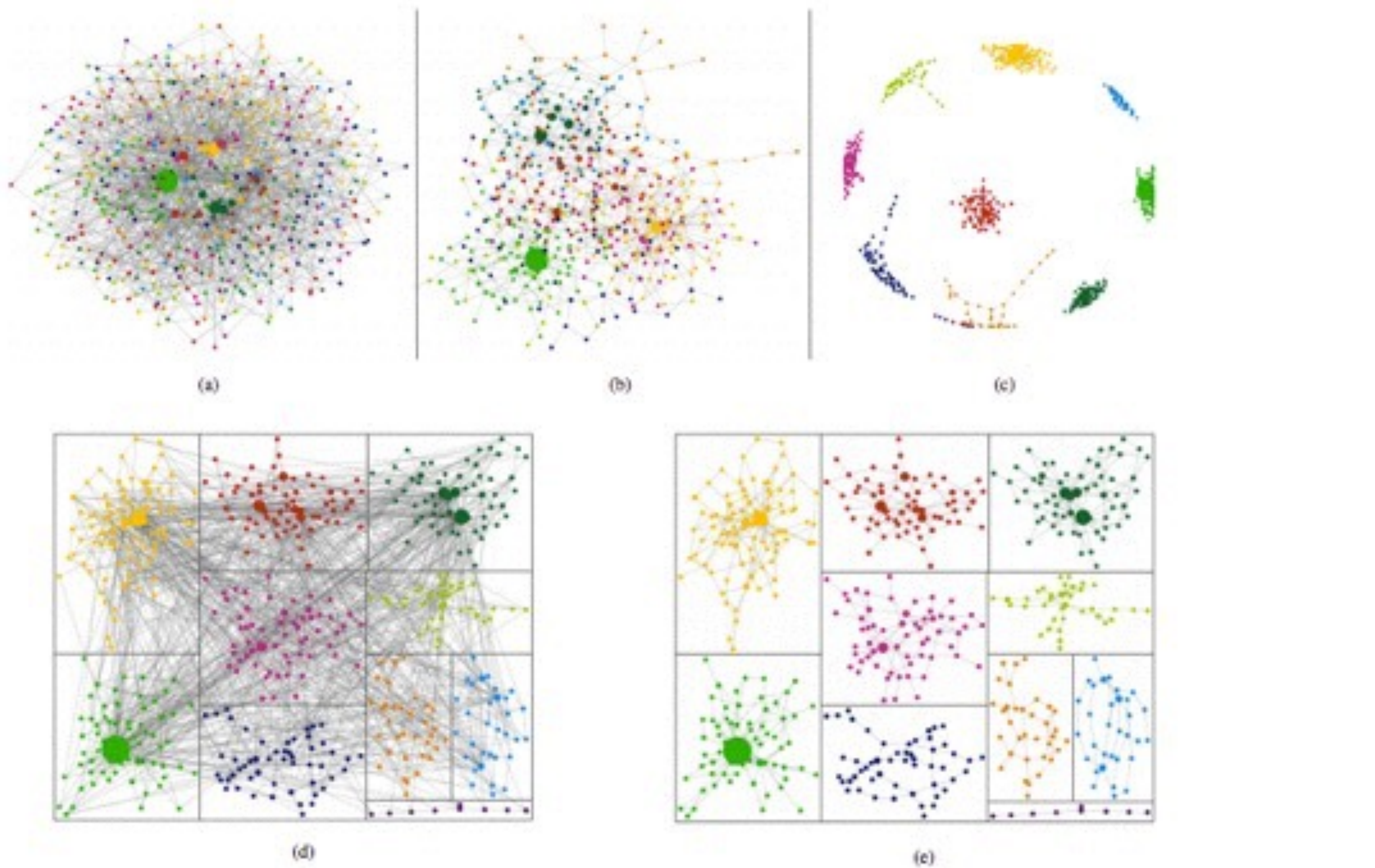


Figure 1. Communities and sub-communities identified in the dolphin social network using the betweeness-based algorithm of Girvan & Newman (2002). (*a*) Vertex shading indicates community membership: vertices in black are all part of one community, while all other vertices are part of the second community. This second community is further divided into three sub-communities by the algorithm; these are represented by the shading (white, light grey and dark grey). Females are represented as circles, males as squares and individuals with unknown gender as triangles. (*b*) Individuals with high betweenness values (greater than 7.32) are represented by black symbols. The size of the black symbol is directly related to the betweenness of the vertex.

Figure 6. Scale-free network with 10 clusters detected by the Clauset-Newman-Moore algorithm. Vertices are colored by cluster membership and sized by betweeness centrality. (a) Harel–Koren layout of the clustered graph. (b) Harel–Koren layout after removing inter-cluster edges. (c) Fruchterman-Reingold layout after removing inter-cluster edges. (d) GIB showing inter-cluster edges and (e) GIB showing intra-clsuter edges.



Eduarda Mendes Rodrigues, Natasa Milic-Frayling, Marc Smith, Ben Shneiderman, Derek Hansen, Group-in-a-box Layout for Multi-faceted Analysis of Communities. IEEE Third International Conference on Social Computing, October 9-11, 2011. Boston, MA

http://mbostock.github.com/d3/talk/20111116/force-collapsible.html

Also see Ger Hobbelt D3: http://bl.ocks.org/3616279

http://mbostock.github.com/d3/talk/20111116/force-collapsible.html

Also see Ger Hobbelt D3: http://bl.ocks.org/3616279

ALGORITHMIC LAYOUTS

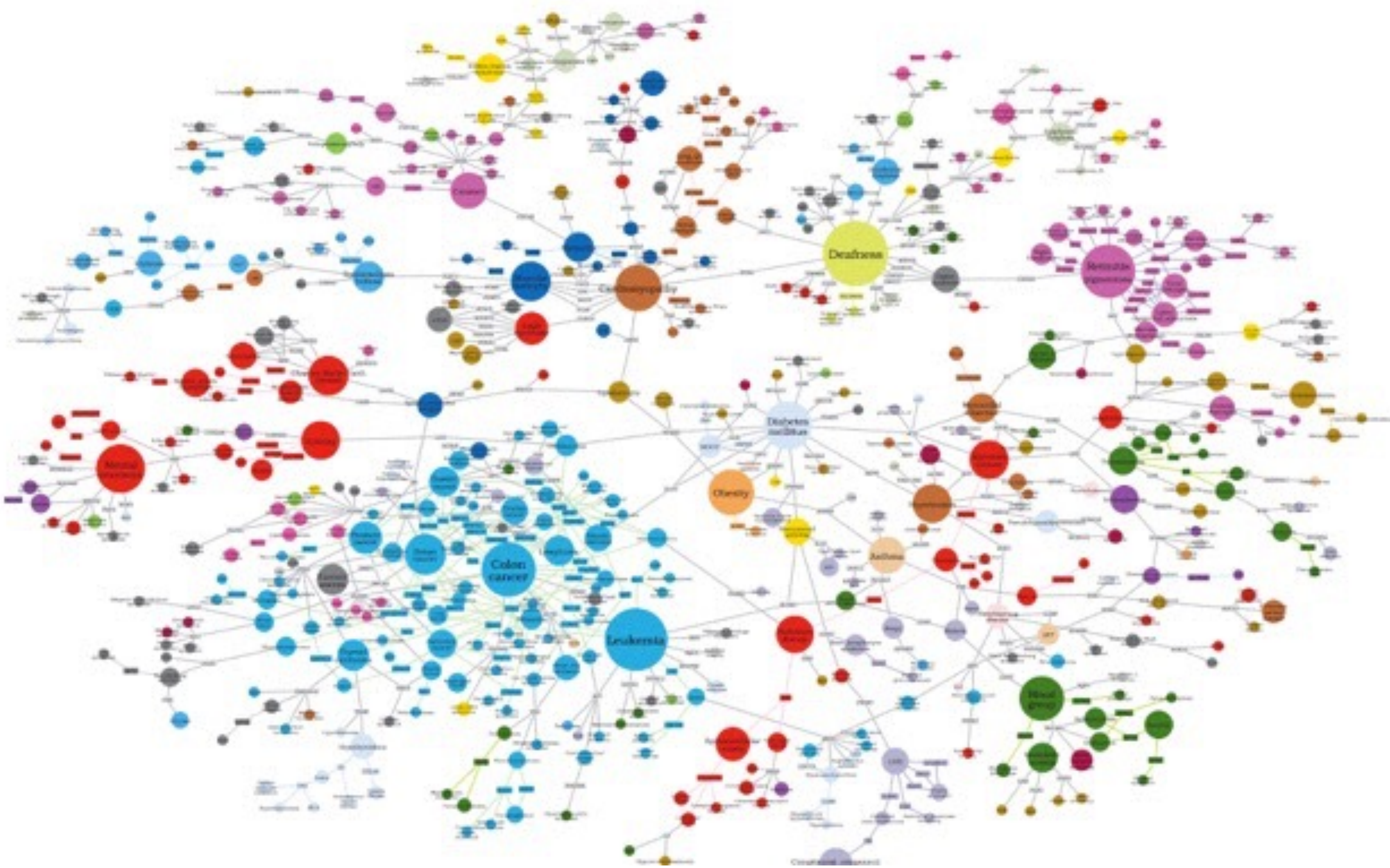
Gephi / D3.js / Other tools

Gephi → Sigma.js / D3

- Gephi.org: Open source, runs on Mac, Linux, PC
  - Can be run from a python-esque console plugin or UI
  - Can be run "headless" for layouts with Java or Python (I have code for this)
  - Plugins include a Neo4j graph db access, and streaming support
- Sigma.js :
  - Will display a gexf gephi layout file with minimal work, using a plugin interpreter for sigma site export
  - Also offers a force-directed layout plugin for graphs without x&y coords
  - Does CANVAS drawing, not SVG
- Export JSON from Gephi and load into D3 network layout

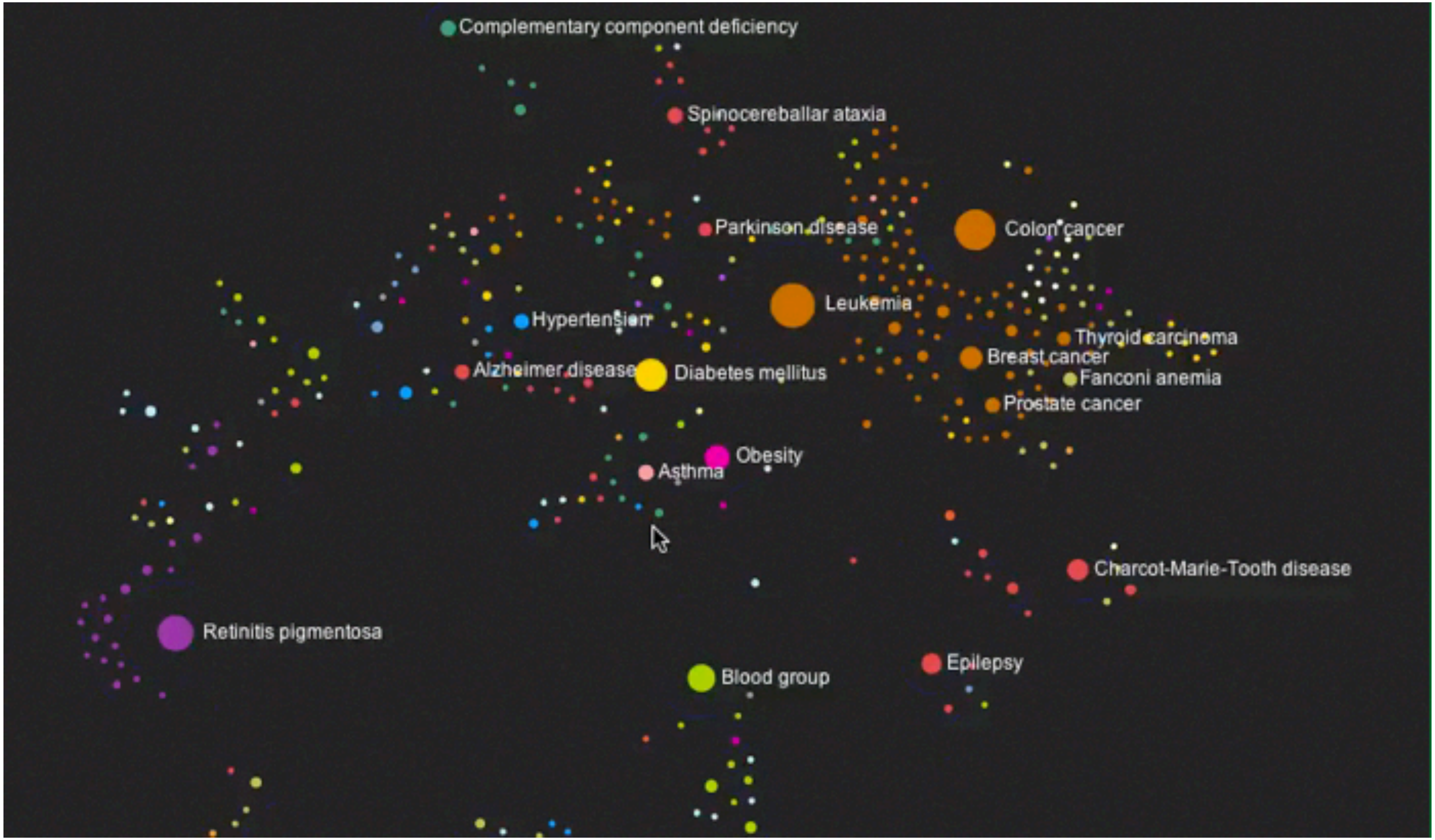
The human disease network

Goh K-I, Cusick ME, Valle D, Childs B, Vidal M, Barabási A-L (2007) *Proc Natl Acad Sci USA* 104:8685-8690



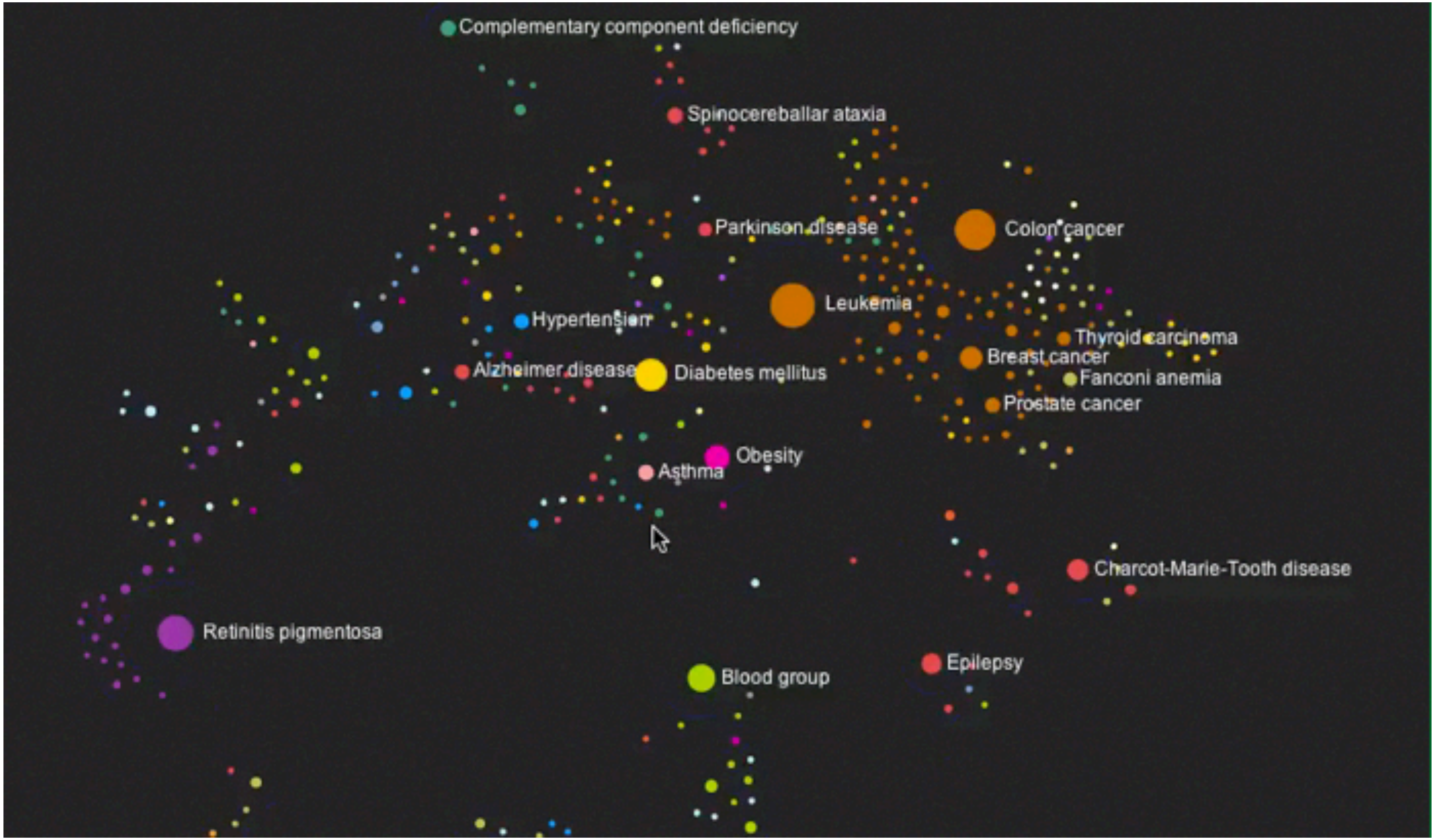
http://www.barabasilab.com/pubs/CCNR-ALB_Publications/200705-14_PNAS-HumanDisease/Suppl/Goh_etal_poster.pdf

Movie:



Sigma.js version of the Gephi export http://exploringdata.github.com/vis/human-disease-network/

Movie:



Sigma.js version of the Gephi export http://exploringdata.github.com/vis/human-disease-network/

Sample Layout Plugins in Gephi

emphasis DIVISIONS OpenOrd	*emphasis* COMPLEMENTARITIES ForceAtlas, Yifan Hu, Frushterman-Reingold
emphasis RANKING Circular, Radial Axis	*emphasis* GEOGRAPHIC REPARTITION GeoLayout

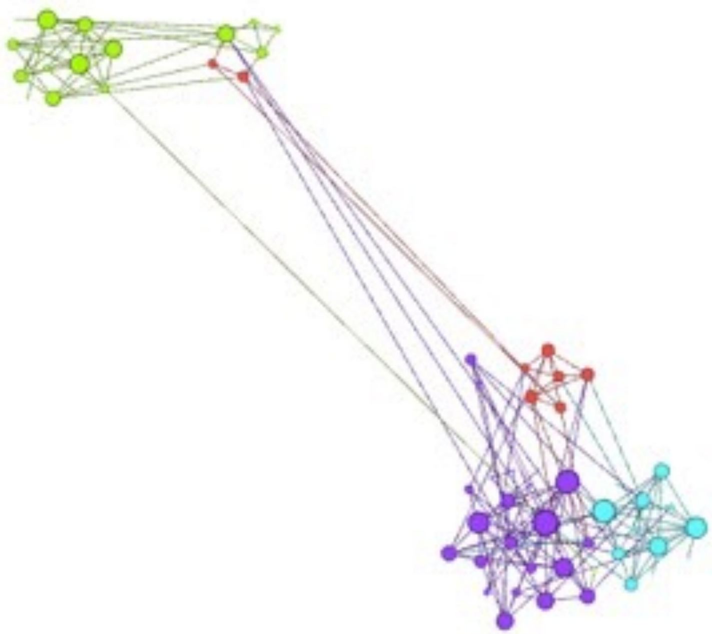
https://gephi.org/tutorials/gephi-tutorial-layouts.pdf

Gephi Plugin Layout Details

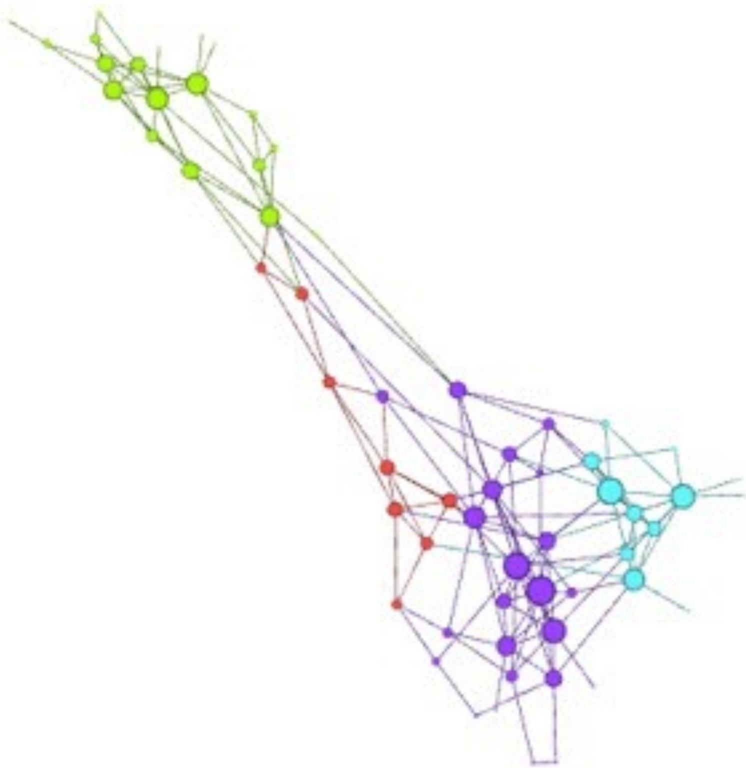
Layout	Complexity	Graph Size	Author	Comment
Circular	O(N)	1 to 1M nodes	Matt Groeninger	Used to show distribution, ordered layout
Radial Axis	O(N)	1 to 1M nodes	Matt Groeninger	Show ordered groups (homophily)
Force Atlas	$O(N^2)$	1 to 10K nodes	Mathieu Jacomy	Slow, but uses edge weight and few biases
Force Atlas 2	O(N*log(N))	1 to 1M nodes	Mathieu Jacomy	(does use weight)
OpenOrd	O(N*log(N))	100 to 1M nodes	S. Martin, W. M. Brown, R. Klavans, and K. Boyack	Focus on clustering (uses edge weight)
Yifan Hu	O(N*log(N))	100 to 100K nodes	Yifan Hu	(no edge weight)
Fruchterman-Rheingold	$O(N^2)$	1 to 1K nodes	Fruchterman & Rheingold!	Particle system, slow (no edge weight)
GeoLayout	O(N)	1 to 1M nodes	Alexis Jacomy	Uses Lat/Long for layout

https://gephi.org/2011/new-tutorial-layouts-in-gephi/

Dolphins Again

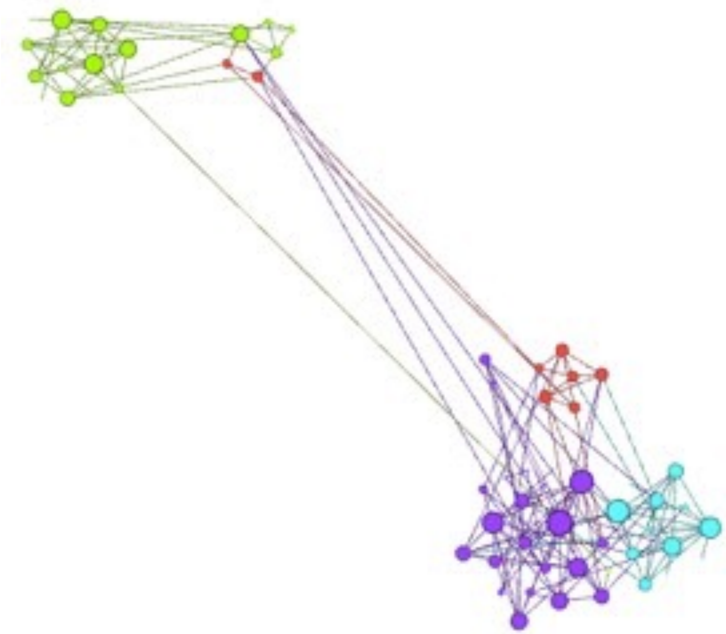


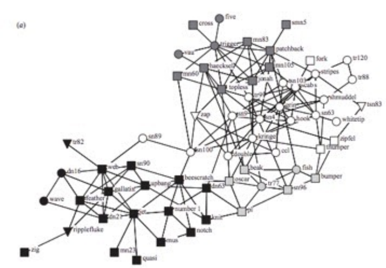
OpenOrd + “No Overlap”

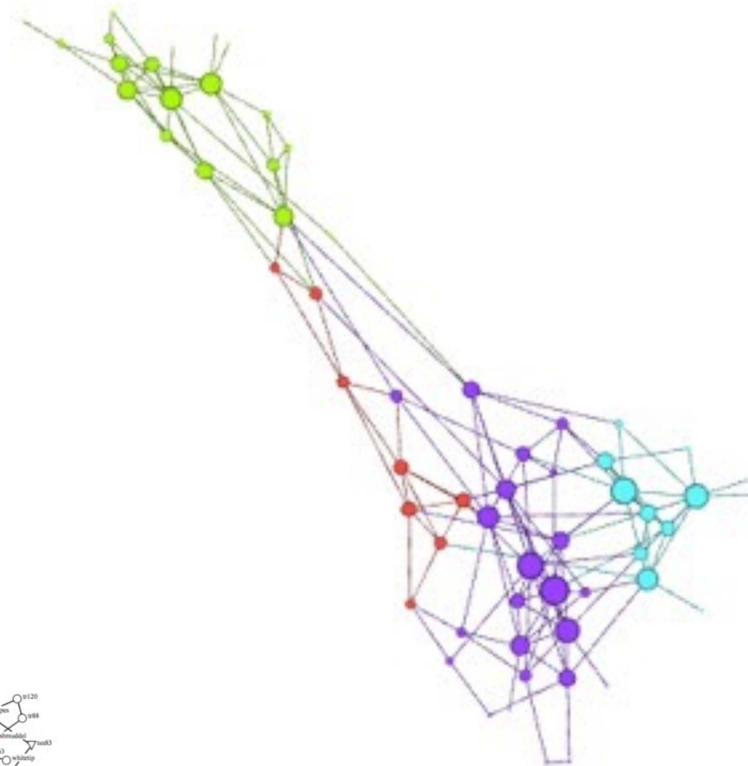


ForceAtlas2

Dolphins Again



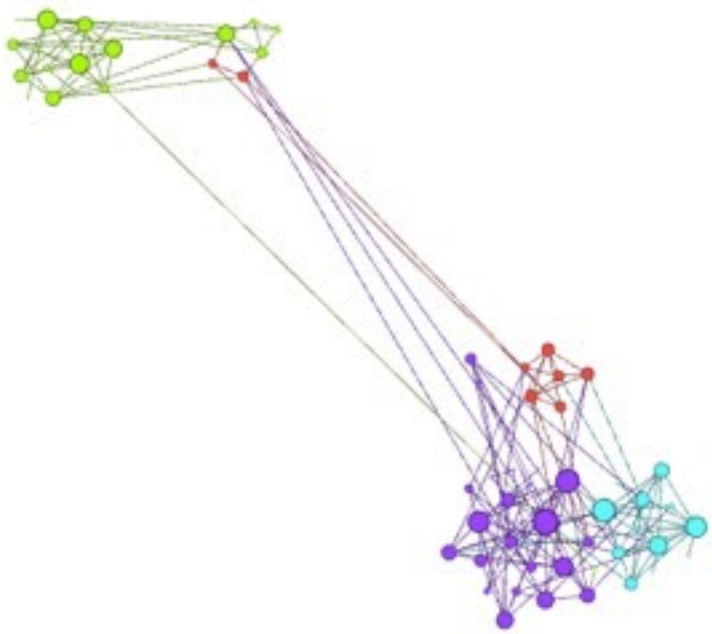




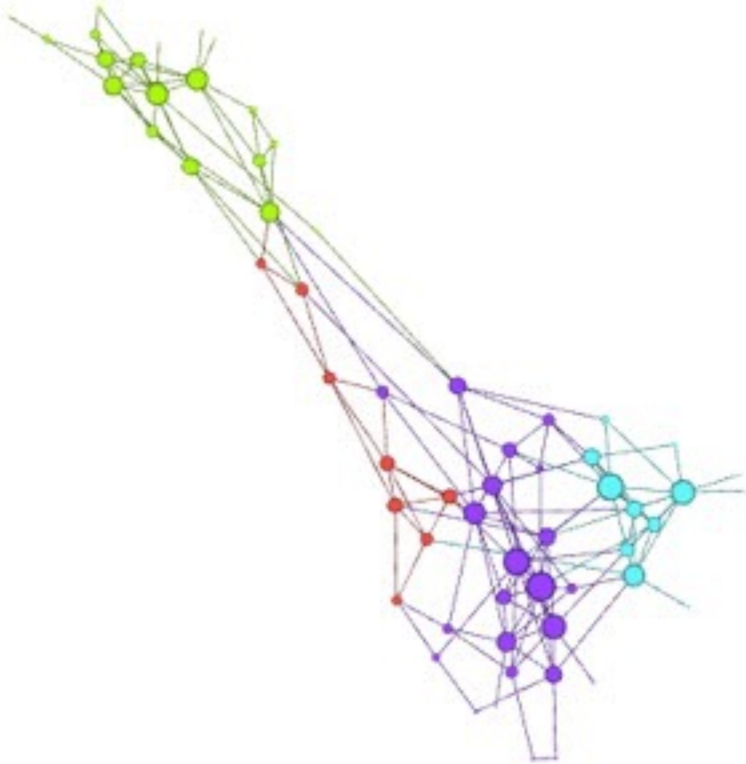
OpenOrd + “No Overlap”

ForceAtlas2

Dolphins Again

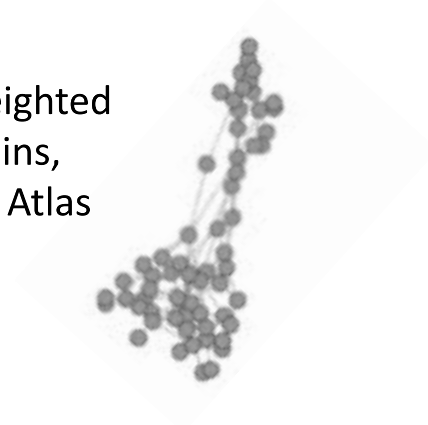


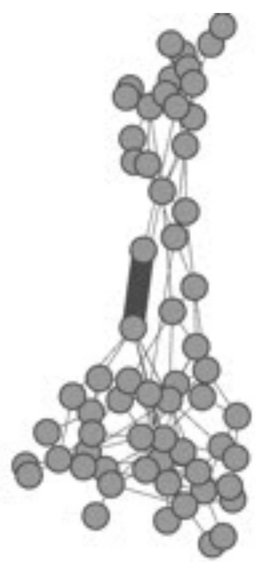
OpenOrd + “No Overlap”

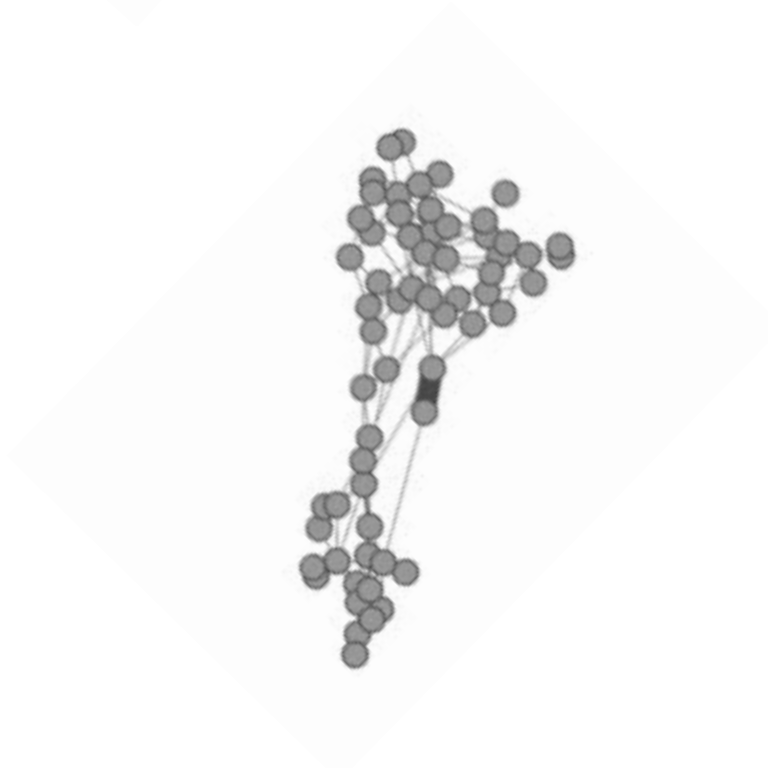


ForceAtlas2

Unweighted dolphins, Force Atlas



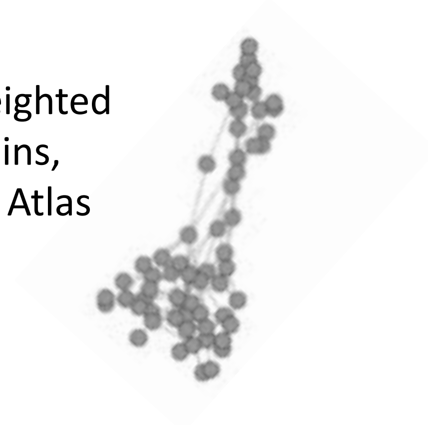


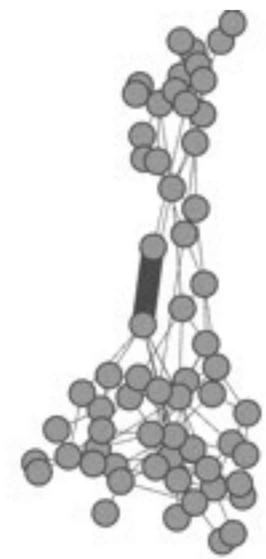


Weight 2: Force Atlas

Weight 4: Force Atlas

Unweighted dolphins, Force Atlas





Weight 2: Force Atlas

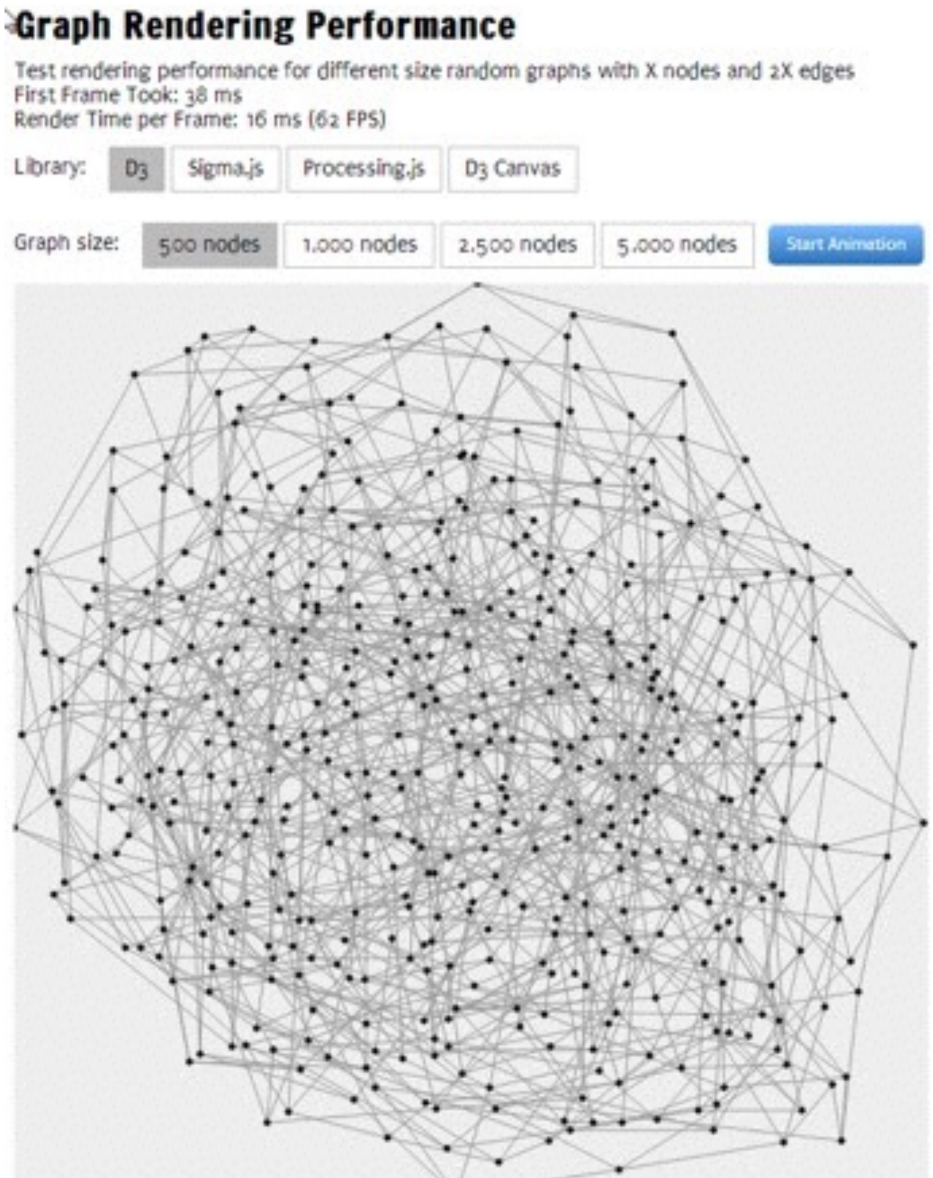
Weight 4: Force Atlas

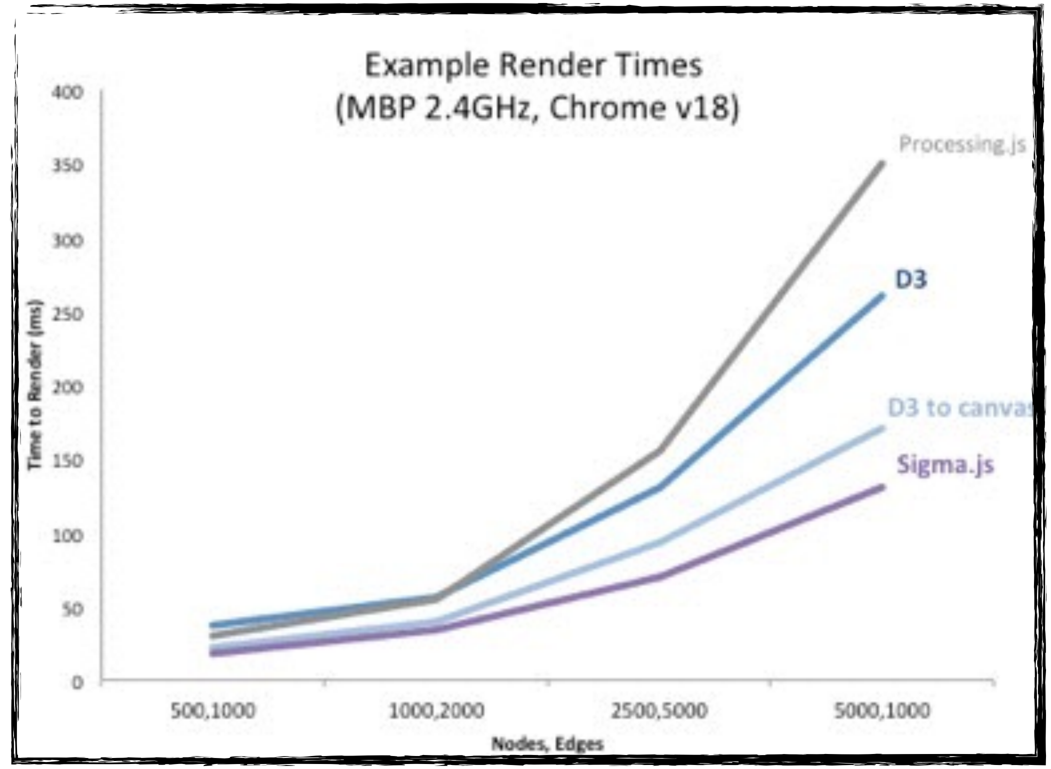
Weight 4: Yifan Hu

Canvas/SVG benchmarks from the d3.js group:

https://docs.google.com/spreadsheet/ccc?key=0AtvlFoSBUC5kdEZJNVFySG9wSHZka0NsOTZDSkt3Nnc#gid=0

Nick Diakapolous: http://nad.webfactional.com/ntap/graphscale/



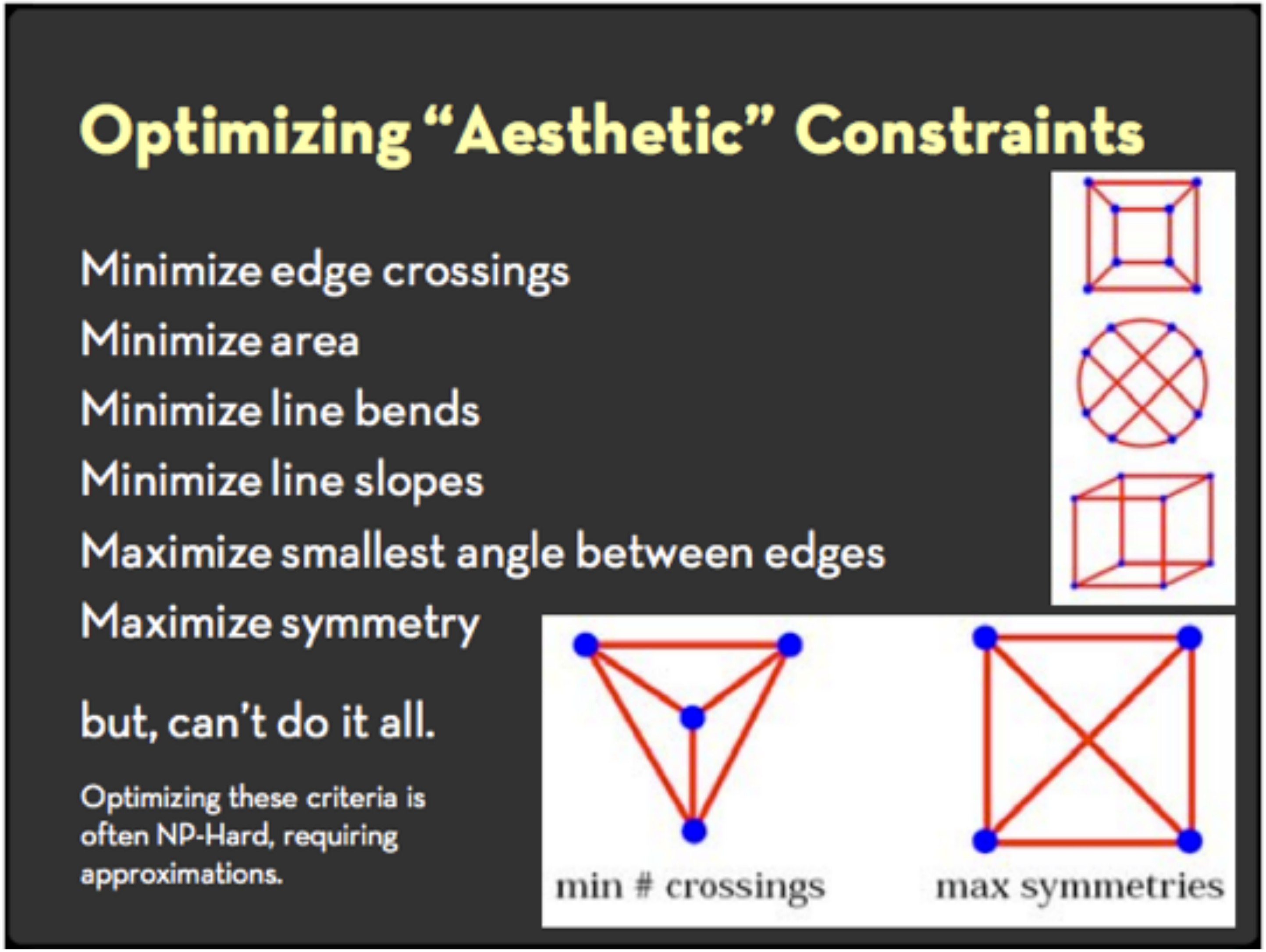


Canvas/SVG benchmarks from the d3.js group:

https://docs.google.com/spreadsheet/ccc?key=0AtvlFoSBUC5kdEZJNVFySG9wSHZka0NsOTZDSkt3Nnc#gid=0

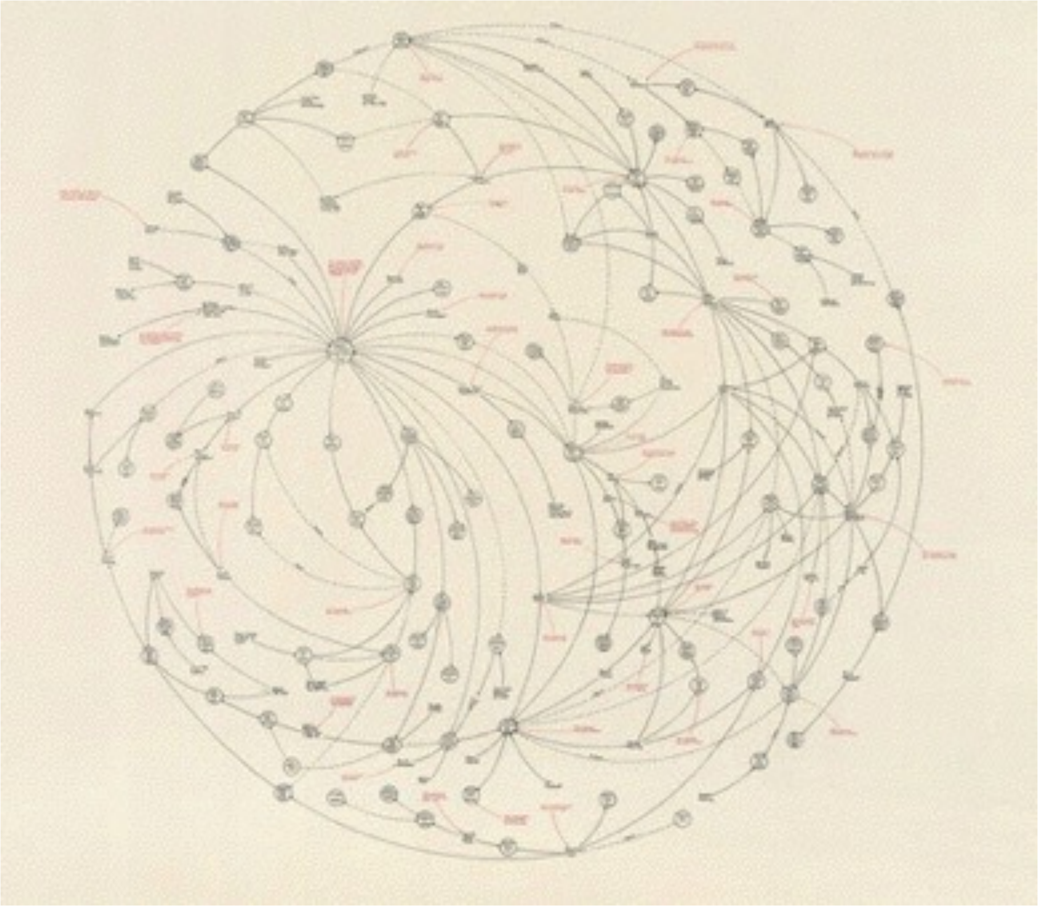
Nick Diakapolous: http://nad.webfactional.com/ntap/graphscale/

FINAL DESIGN THOUGHTS…
BY HAND? BY ALGORITHM-?



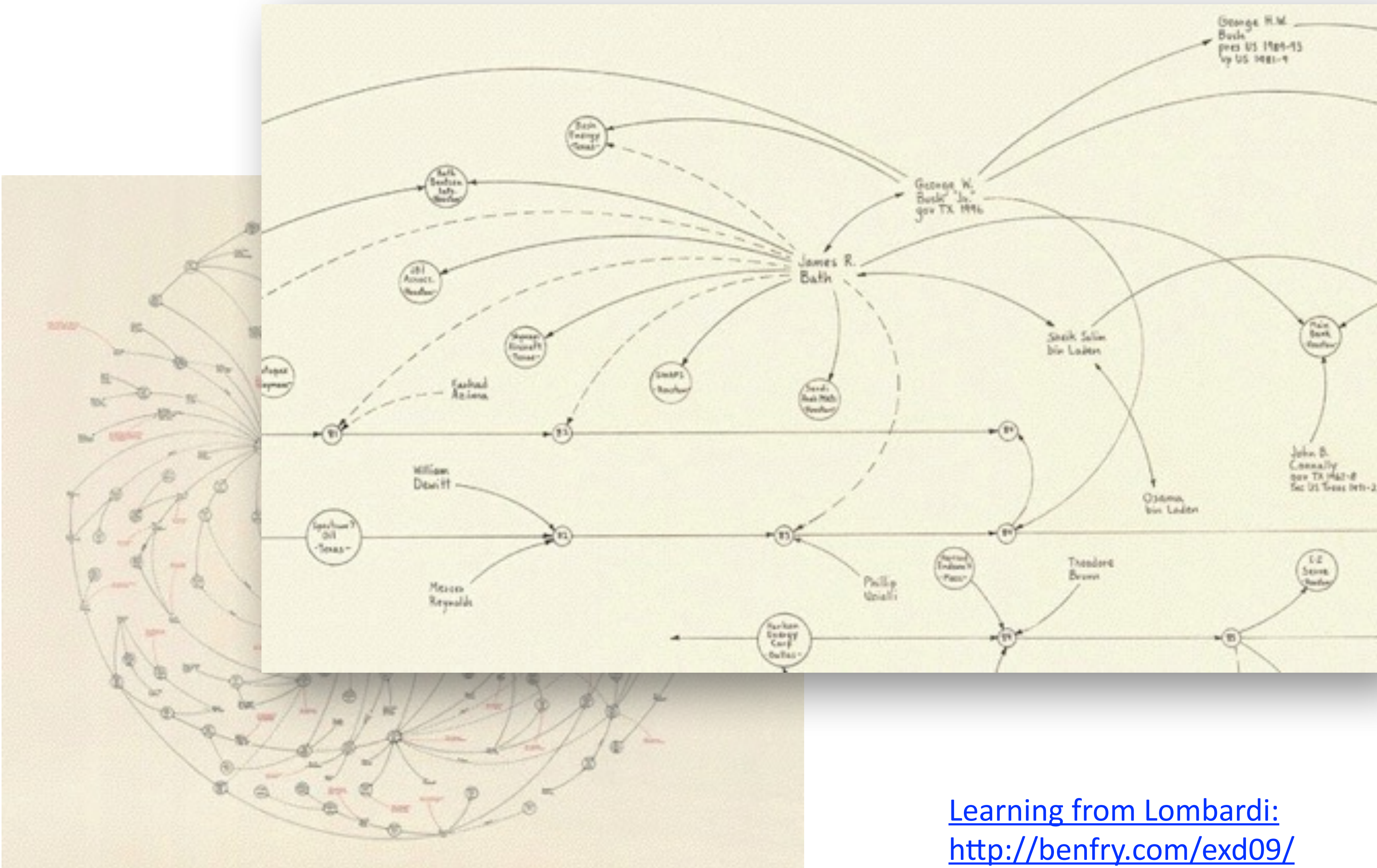
Jeff Heer: http://hci.stanford.edu/courses/cs448b/f11/lectures/CS448B-20111110-GraphsAndTrees.pdf

Conspiracy Theorist Mark Lombardi



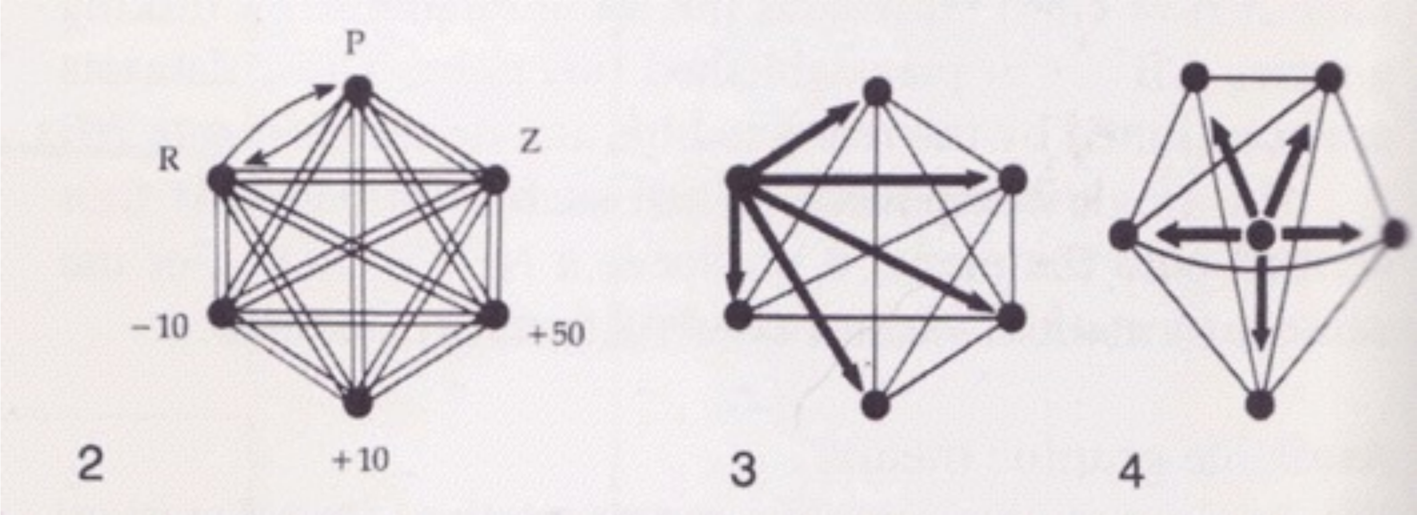
[Learning from Lombardi: http://benfry.com/exd09/](http://benfry.com/exd09/)

Conspiracy Theorist Mark Lombardi



[Learning from Lombardi: http://benfry.com/exd09/](http://benfry.com/exd09/)

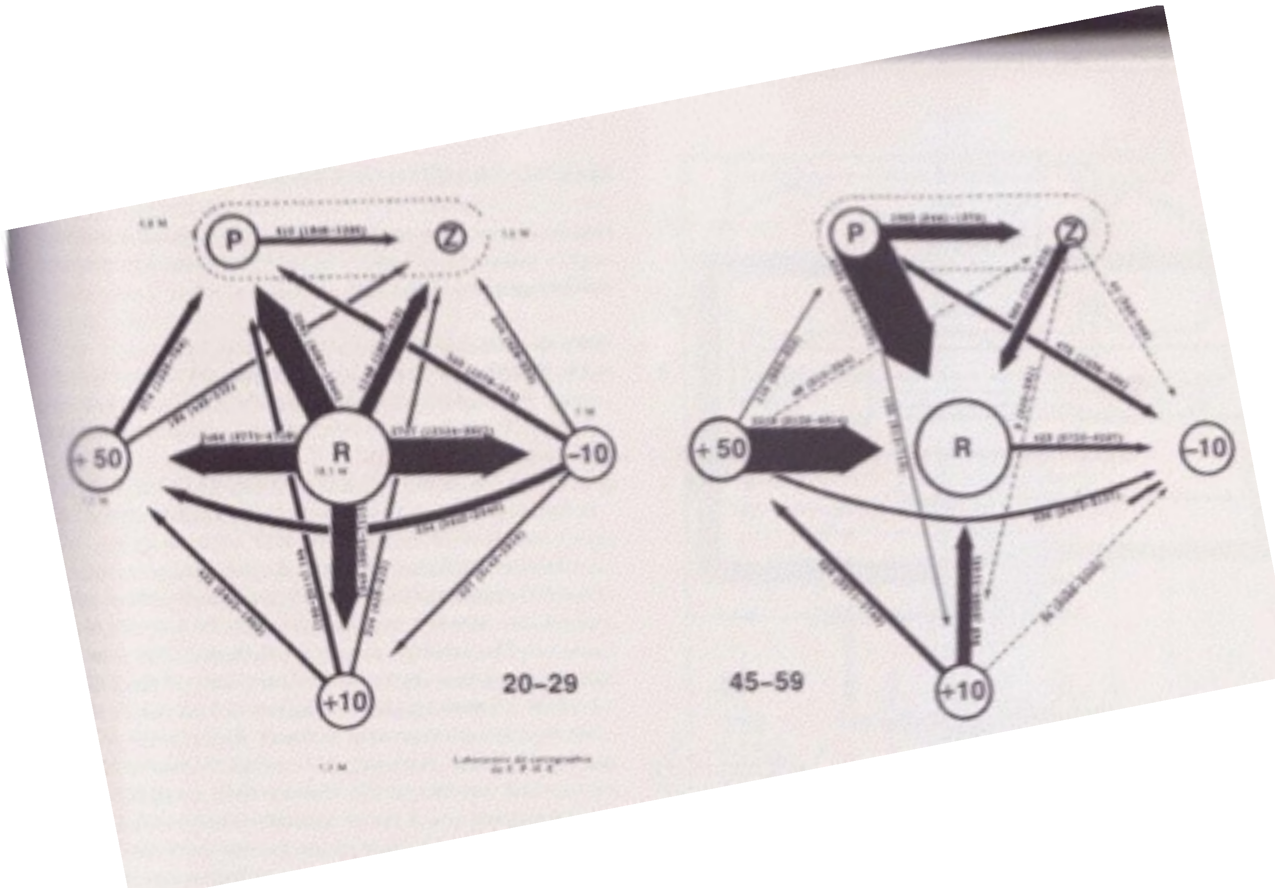
J. Bertin Again…

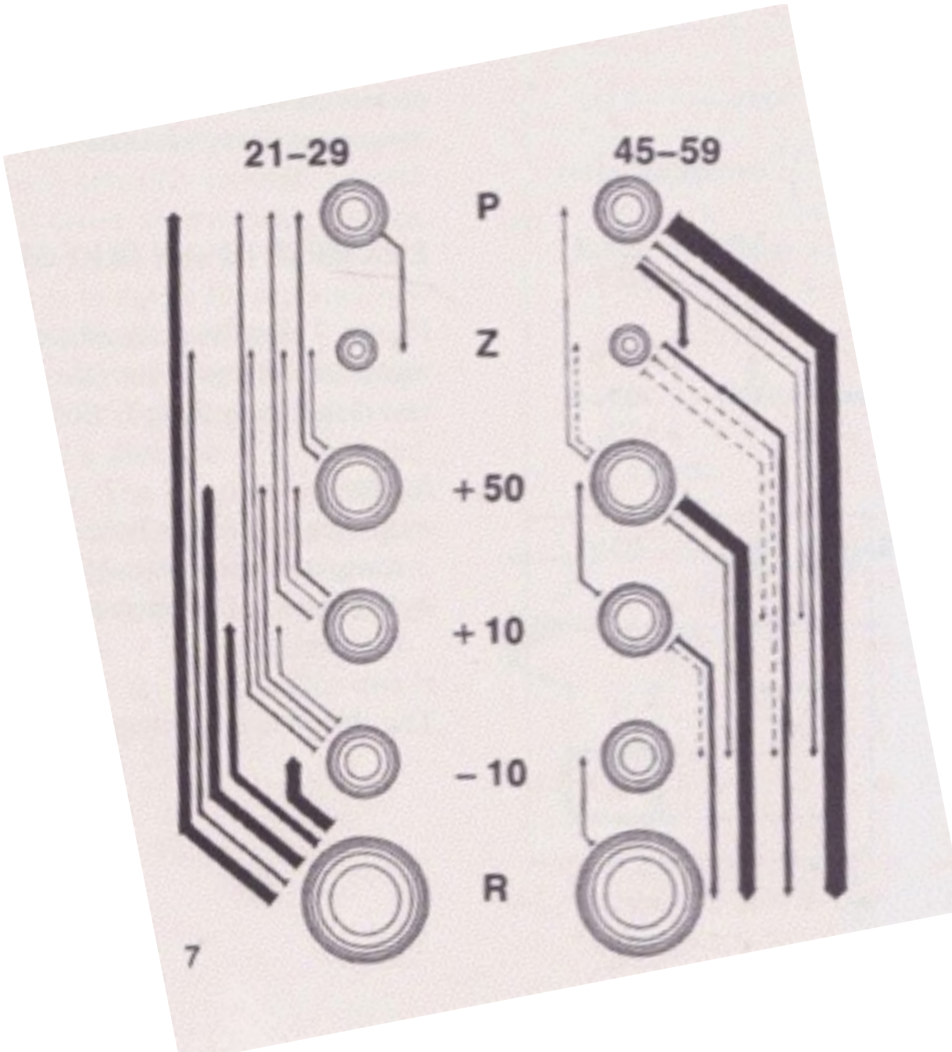


2 3 4

Bertin, J. Semiology of Graphics: Diagrams, Networks, Maps (1967)

Age Groups and Movement to/from Paris to Rural Spots





(P) Paris

(Z) Paris Suburbs

(+50) Communes of >50K

(+10) Communes of >10K

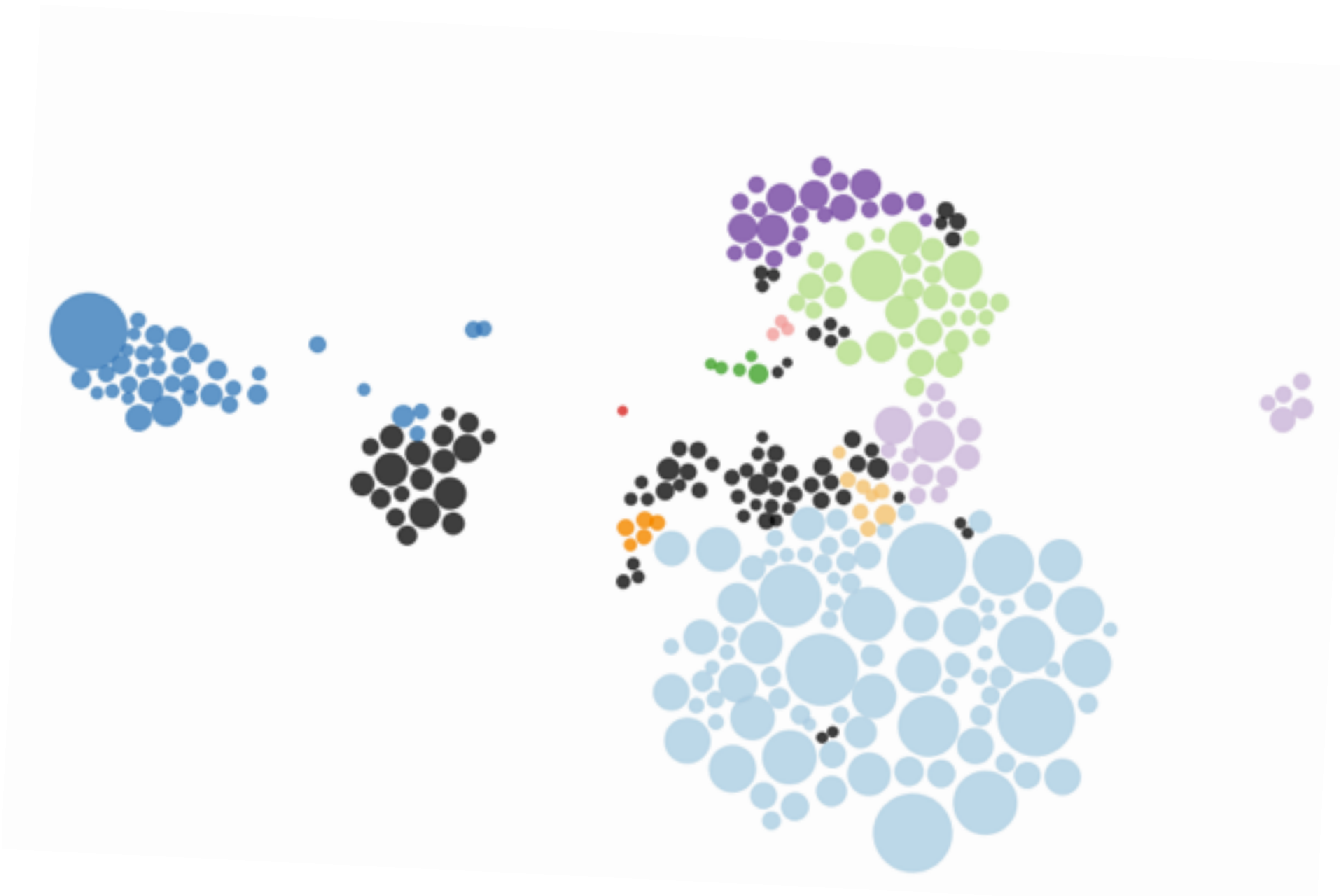
(-10) Communes of <10K

(R) Rural

Hybrid Method: Use algorithmic layout, and then adjust nodes by hand.

Ger Hobbelt in D3: http://bl.ocks.org/3637711

Minimize Info: Less is More



Color by important data value, sized by degree, no edges

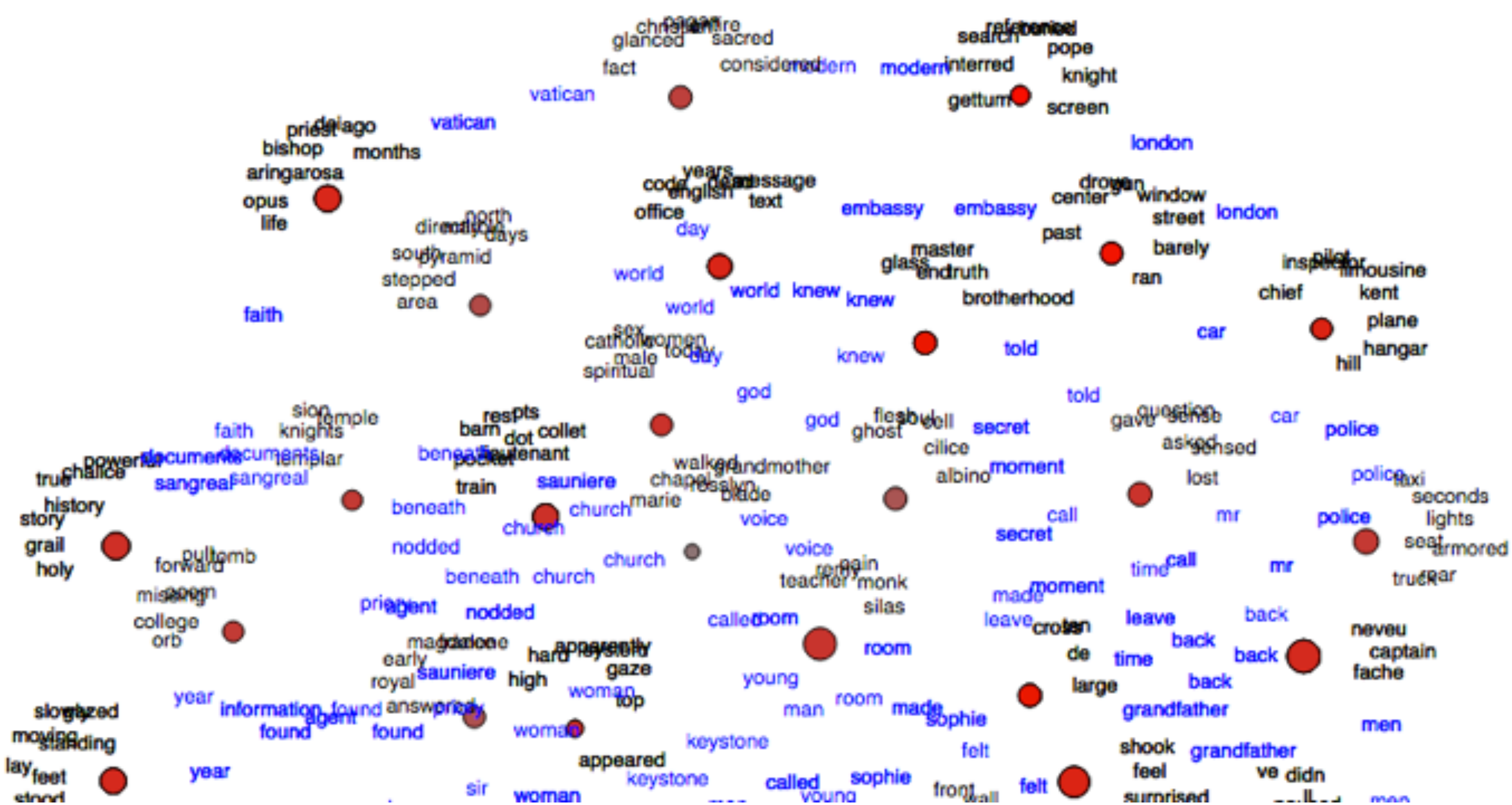
'DaVinci Code' LDA Topic Words

Topic node sized by number of topic assignments, colored by relative excitement score as tallied off Mechanical Turk ratings by chapter. Mouse over to see details and network relations.

Blue words are found in more than one topic. NB: This layout has a lot of overlaps because there is no collision detection for the labels, but I abandonned it when I realized it wasn't very illuminating anyway. Blog post/slides by Lynn Cherny.

Filter Topics By...

All | Most Exciting | Less Exciting



http://www.ghostweather.com/essays/talks/openvisconf/topic_docs_network/

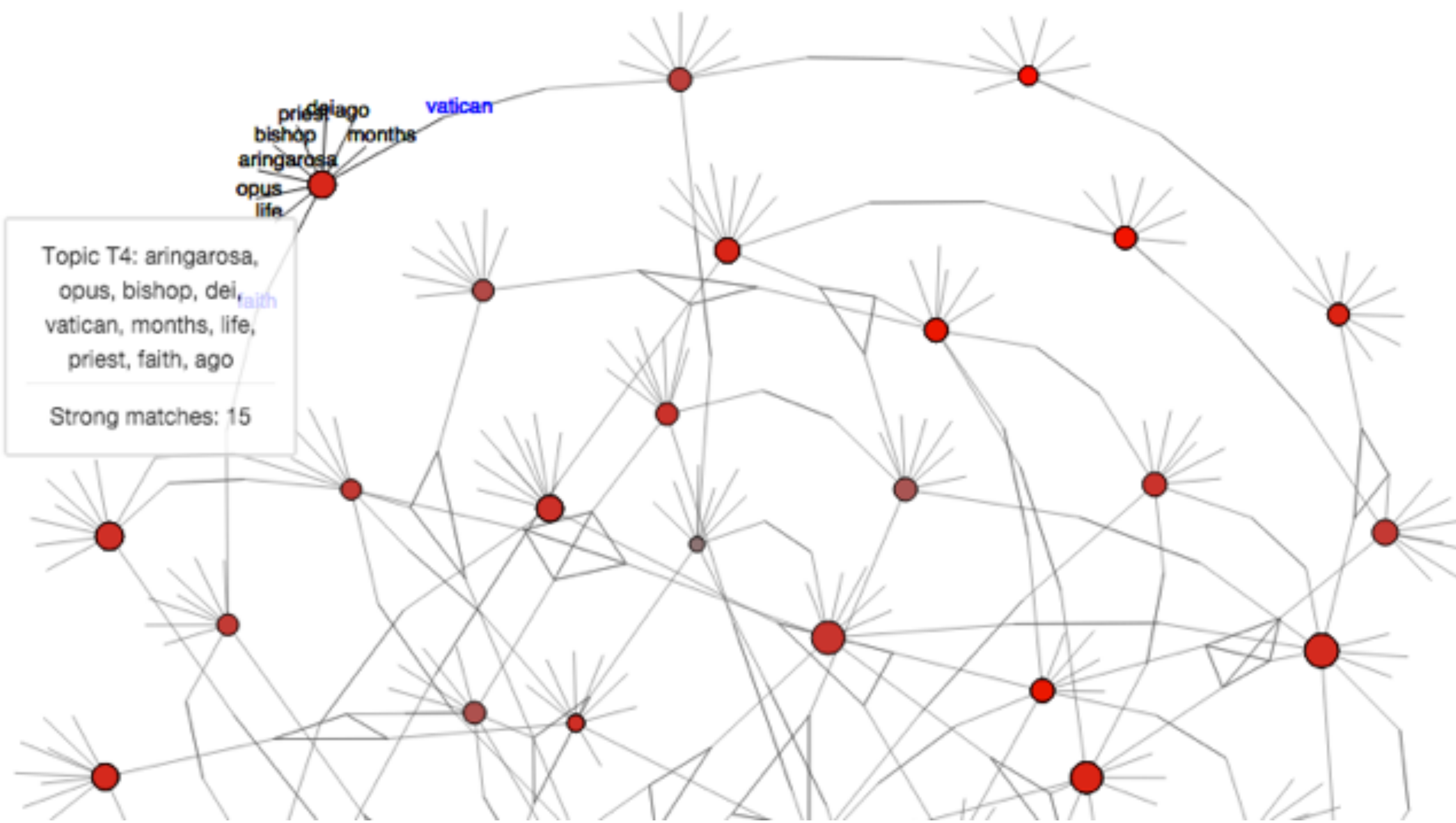
'DaVinci Code' LDA Topic Words

Topic node sized by number of topic assignments, colored by relative excitement score as tallied off Mechanical Turk ratings by chapter. Mouse over to see details and network relations.

Blue words are found in more than one topic. NB: This layout has a lot of overlaps because there is no collision detection for the labels, but I abandonned it when I realized it wasn't very illuminating anyway. Blog post/slides by Lynn Cherny.

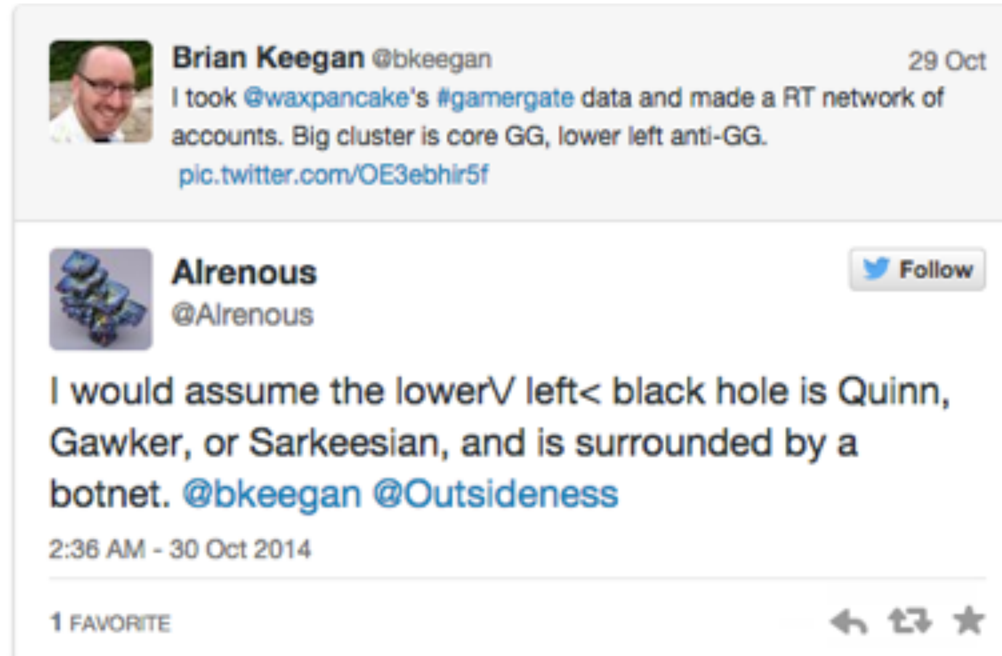
Filter Topics By...

All | Most Exciting | Less Exciting



http://www.ghostweather.com/essays/talks/openvisconf/topic_docs_network/

Brian Keegan's "15 Minutes of Fame as a B-List GamerGate Celebrity"



http://www.brianckeegan.com/2014/10/my-15-minutes-of-fame-as-a-b-list-gamergate-celebrity/

Brian Keegan's "15 Minutes of Fame as a B-List GamerGate Celebrity"

Brian Keegan @bkeegan 29 Oct

I took @waxpancake's #gamergate data and made a RT network of accounts. Big cluster is core GG, lower left anti-GG. pic.twitter.com/OE3ebhir5f

Alrenous @Alrenous Follow

I would assume the lowerV left< black hole is Quinn, Gawker, or Sarkeesian, and is surrounded by a botnet. @bkeegan @Outsideness

2:36 AM - 30 Oct 2014

1 FAVORITE

ROGUE☆ @RogueStarGamez Follow

How diverse is #GamerGate? Lower left is AntiGG forces. ProGG is the center.

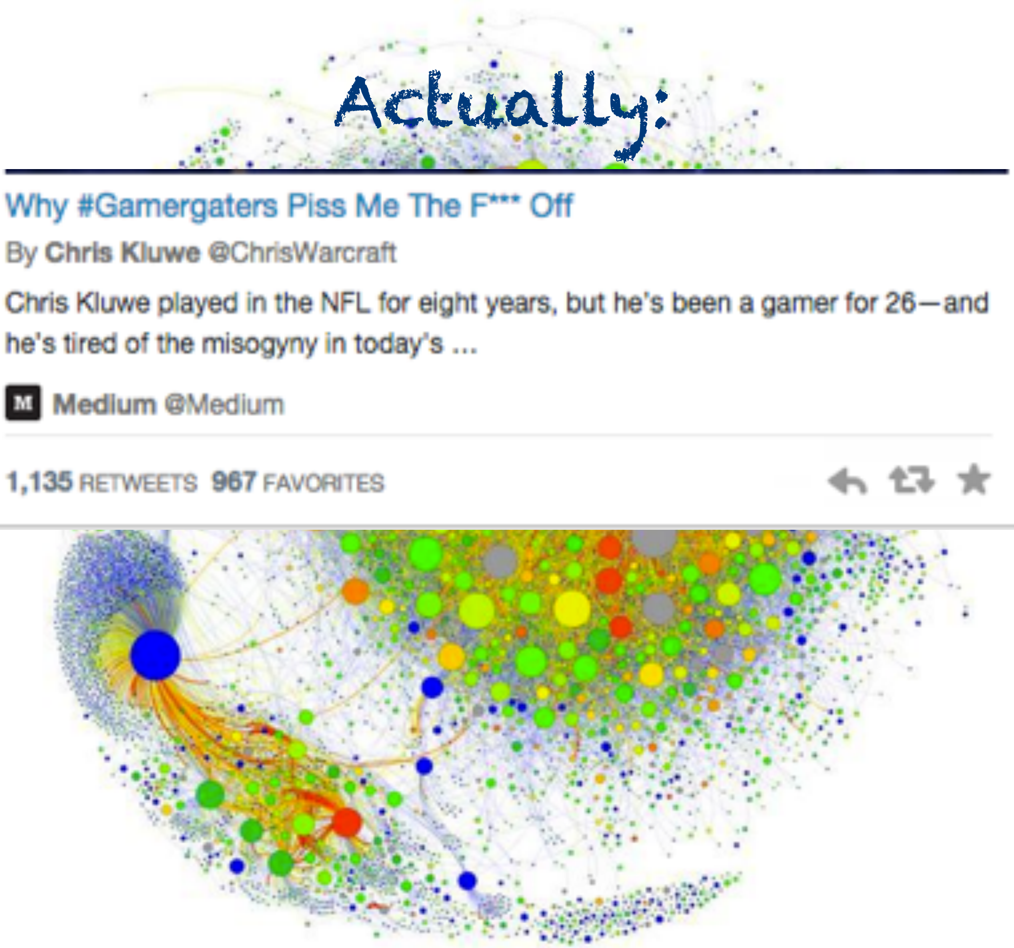
AntiGG is INSULAR. Data does not lie.

2:40 PM - 29 Oct 2014

145 RETWEETS 133 FAVORITES

http://www.brianckeegan.com/2014/10/my-15-minutes-of-fame-as-a-b-list-gamergate-celebrity/

Brian Keegan's "15 Minutes of Fame as a B-List GamerGate Celebrity"



Why #Gamergaters Piss Me The F*** Off

By **Chris Kluwe** @ChrisWarcraft

Chris Kluwe played in the NFL for eight years, but he's been a gamer for 26—and he's tired of the misogyny in today's ...

Medium @Medium

1,135 RETWEETS **967** FAVORITES

Brian Keegan @bkeegan 29 Oct

I took @waxpancake's #gamergate data and made a RT network of accounts. Big cluster is core GG, lower left anti-GG. pic.twitter.com/OE3ebhir5f

Alrenous @Alrenous Follow

I would assume the lowerV left< black hole is Quinn, Gawker, or Sarkeesian, and is surrounded by a botnet. @bkeegan @Outsideness

2:36 AM - 30 Oct 2014

1 FAVORITE

ROGUE☆ @RogueStarGamez Follow

How diverse is #GamerGate? Lower left is AntiGG forces. ProGG is the center.

AntiGG is INSULAR. Data does not lie.

2:40 PM - 29 Oct 2014

145 RETWEETS **133** FAVORITES

http://www.brianckeegan.com/2014/10/my-15-minutes-of-fame-as-a-b-list-gamergate-celebrity/

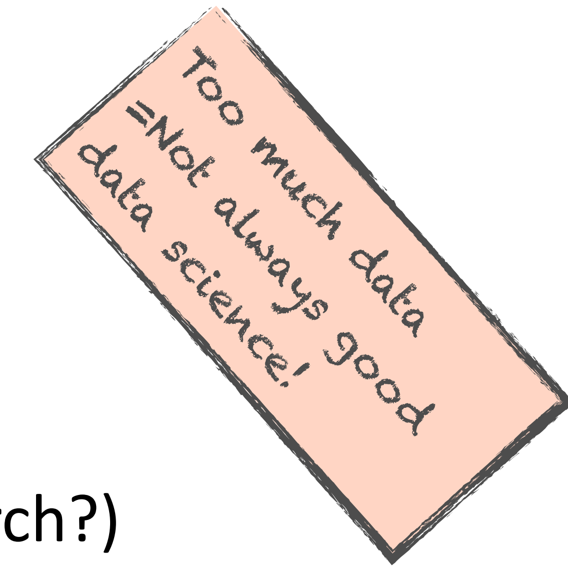
Wrap it up on design...

Reminders

- Do data analysis / reduction - why would you want to show 1T of network data?
- Calculate and reveal important facts about node relationships.
- Consider your visual encodings:
  – visualization, not data vomit.
- Make it interactive - details on demand.
  - Help people find things in your network! (Search?)
  - Show more on hover/click

Reminders

- Do data analysis / reduction - why would you want to show 1T of network data?
- Calculate and reveal important facts about node relationships.
- Consider your visual encodings:
  - visualization, not data vomit.
- Make it interactive - details on demand.
  - Help people find things in your network! (Search?)
  - Show more on hover/click



More Reminders!

- Different layouts communicate different things to your viewer - choose wisely.
  - Take care: people will infer things from proximity/similarity even if it was not intended!
- Consider if it has to be a “network” display at all: Is it the stats you care about? The important nodes who branch sub-communities? The ones with the most in/out links? Graph those instead.

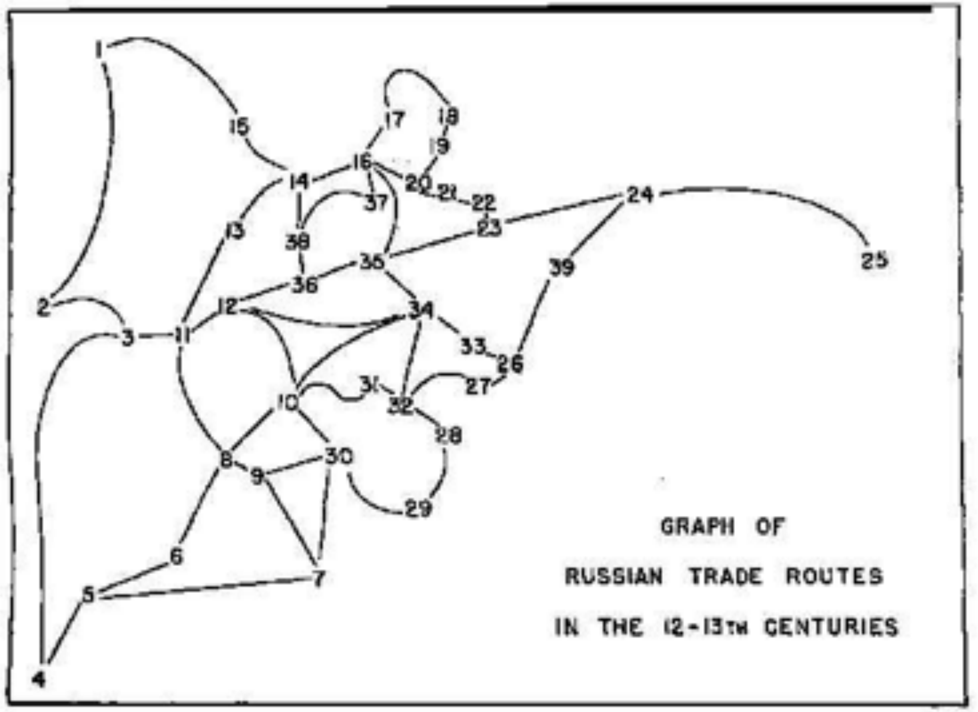
The Map is Not the Territory…

Figure 1. *Russian trade routes in the 12th - 13th centuries.*

Forest Pitts (thanks to Noah Friedkin) http://www.analytictech.com/networks/pitts.htm

The Map is Not the Territory…

Figure 2. *Graph of Russian trade routes in the 12th - 13th centuries.*



Forest Pitts (thanks to Noah Friedkin) http://www.analytictech.com/networks/pitts.htm

Thanks!

@arnicas
lynn@ghostweather.com

Dec 11-12, NYC

http://www.ghostweather.com/workshops/nyc_visdata.html

lynn @ ghostweather.com

A Few More References

Jeff Heer class slides: http://hci.stanford.edu/courses/cs448b/w09/lectures/20090204-GraphsAndTrees.pdf

A great in-progress book on networks: http://barabasilab.neu.edu/networksciencebook/

Mark Newman's new book- NetWorks: An Introduction

Eyeo Festival videos from Moritz Stefaner, Manuel Lima, Stefanie Posavec

Journal of Graph Algorithms and Applications: http://jgaa.info/home.html

Jim Vallandingham's D3 network tutorials: http://flowingdata.com/2012/08/02/how-to-make-an-interactive-network-visualization/, http://vallandingham.me/bubble_charts_in_d3.html

Brian Keegan's post about GamerGate tweets and the dangers of network vis

Robert Kosara's post: http://eagereyes.org/techniques/graphs-hairball

Lane Harrison's post: http://blog.visual.ly/network-visualizations/

MS Lima's book Visual Complexity

Jason Sundram's tool to drive Gephi layout from command line: https://github.com/jsundram/pygephi

A couple articles on community structure:

- Overlapping Community Detection in Networks: State of the Art and Comparative Study by Jierui Xie, Stephen Kelley, Boleslaw K. Szymanski
- Empirical Comparison of Algorithms for Network Community Detection by Leskovec, Lang, Mahoney
- My posts on NetworkX with D3 and Twitter network vis with Gephi