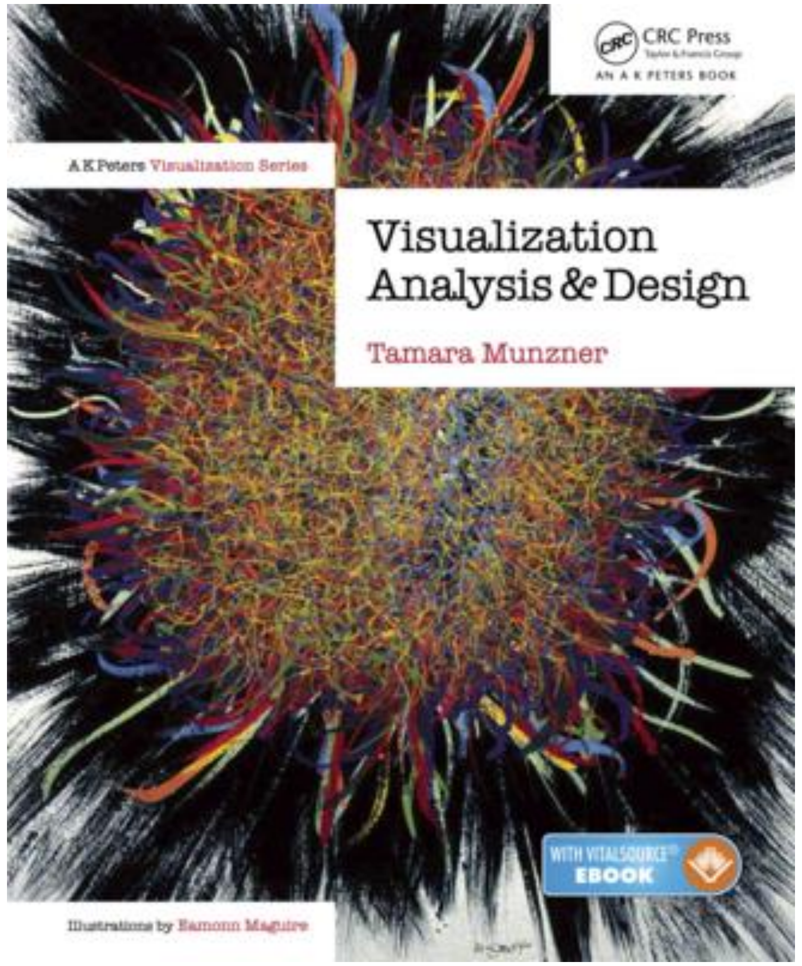
Information Visualization

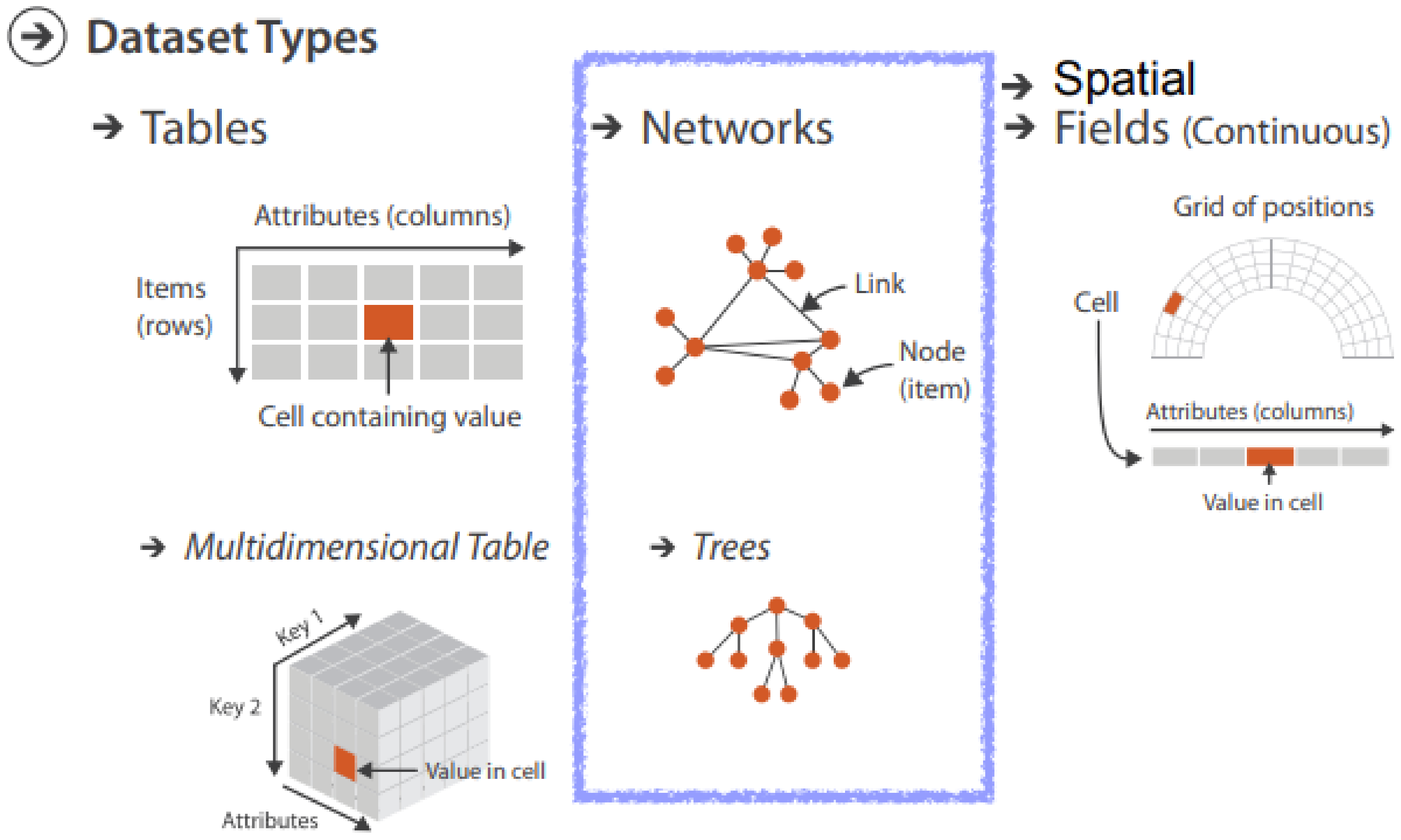
Spatial Layout (c)
Networks and Trees



- Slides refer to https://www.cs.ubc.ca/~tmm/

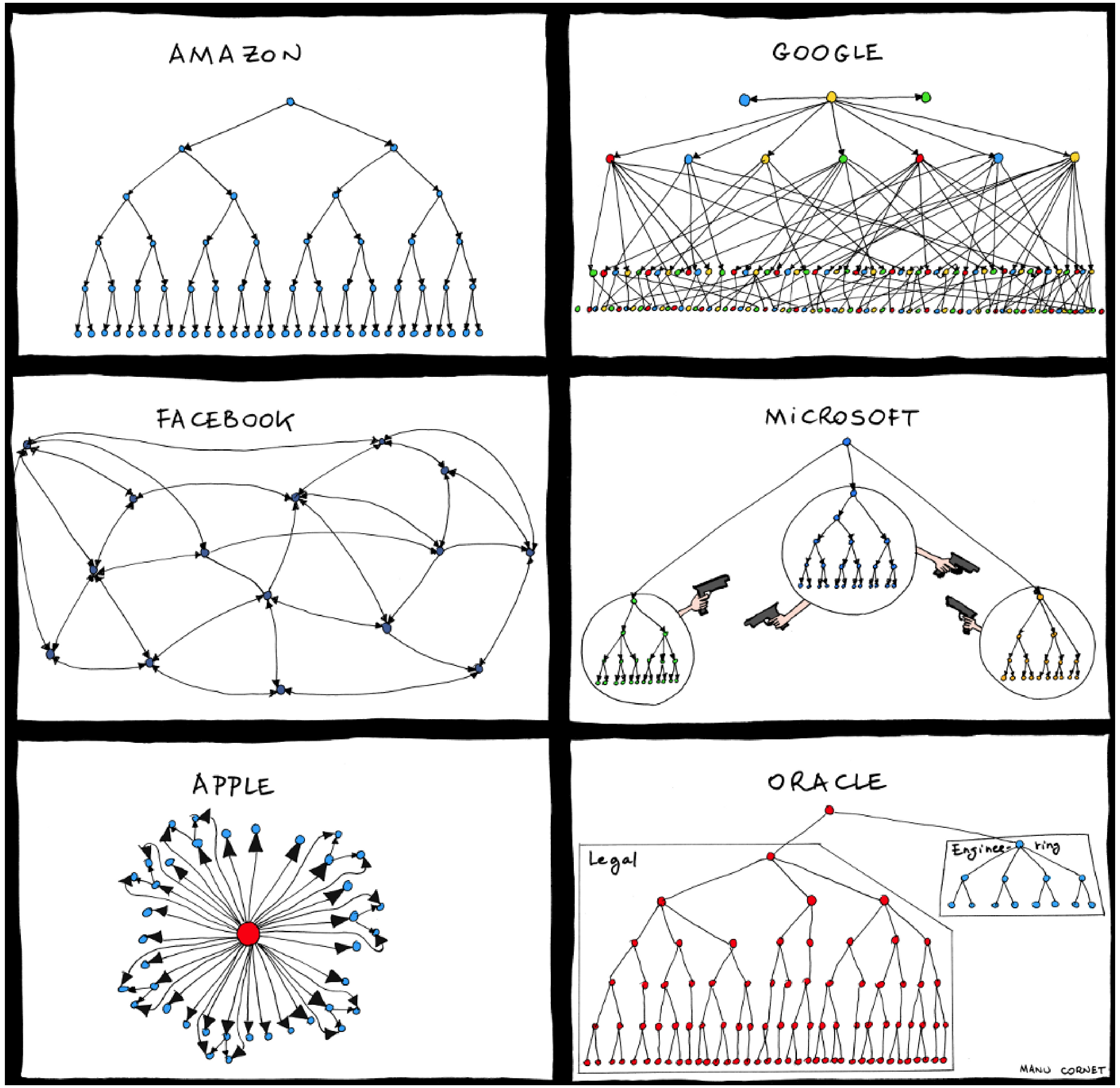
Spatial Layout

- Arrange Tables (ch. 7)
- Arrange Spatial Data (ch. 8)
- Arrange Networks and Trees (ch. 9)



Arrange Networks and Trees

Organization



https://bonkersworld.net/organizational-charts

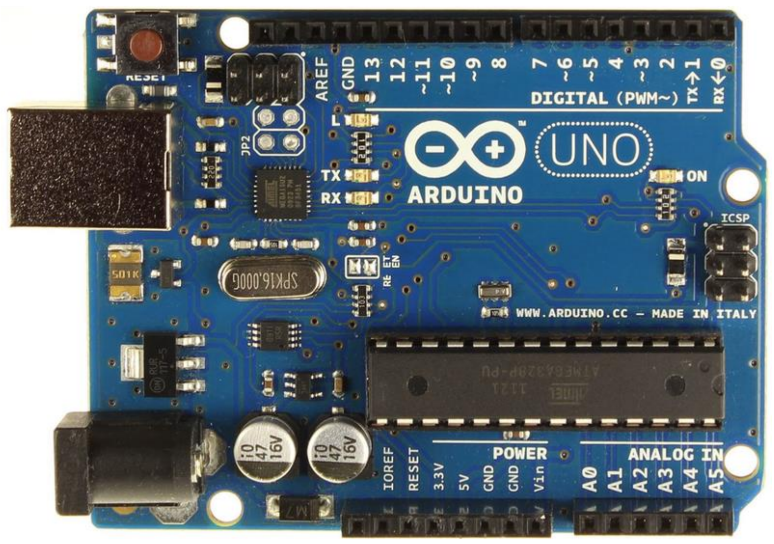
2011 version

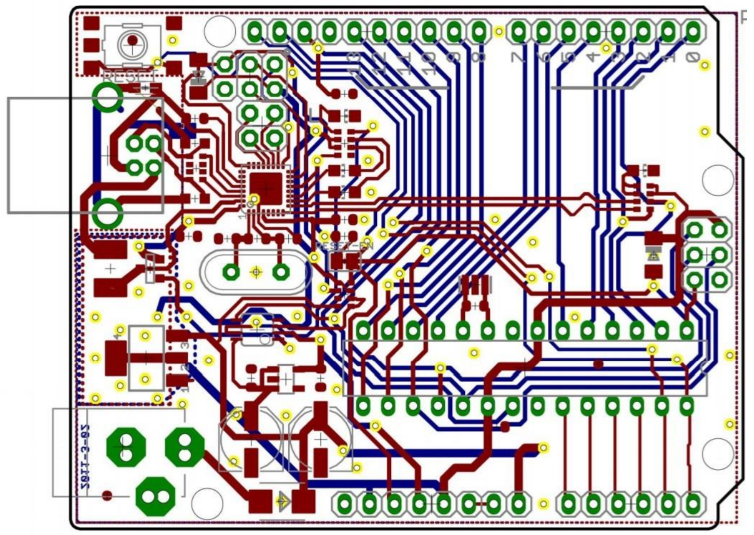
Networks

Applications of networks

- without networks, couldn't have any of these:

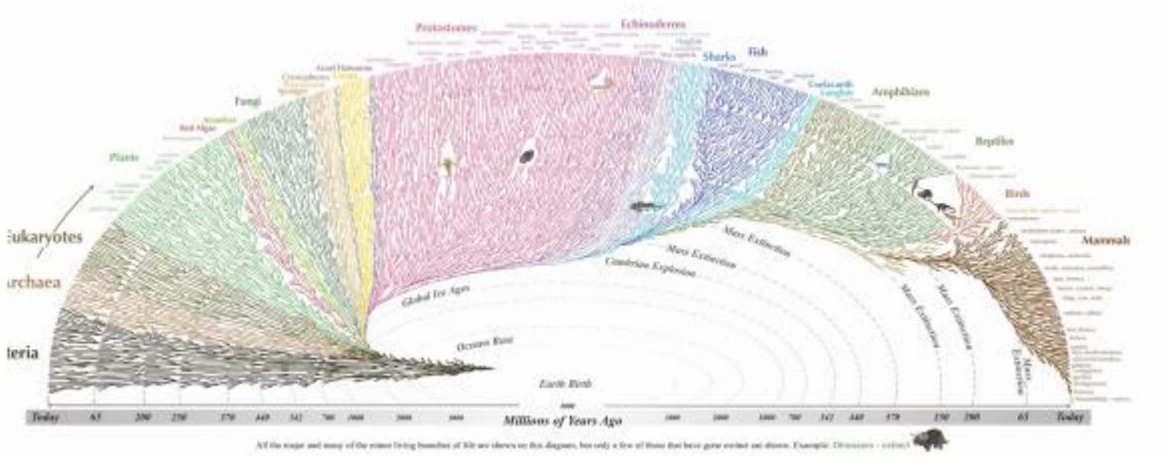
Taipei Rapid Transit Corporation

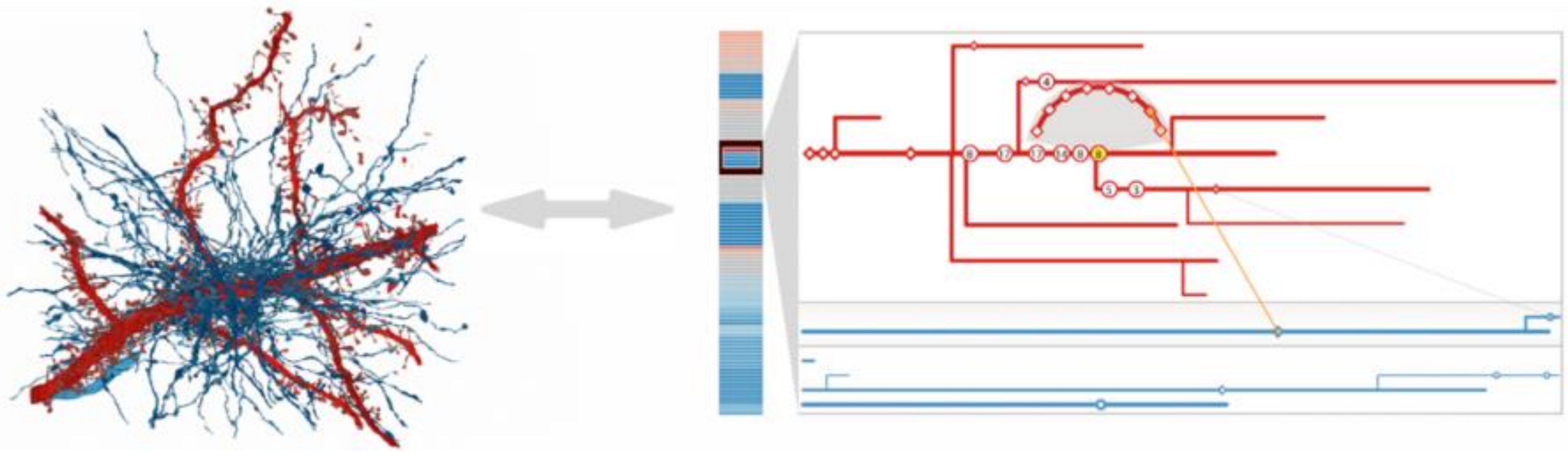




Applications of networks: biological network

- interactions between genes, proteins, and chemical products
- the brain: connections between neurons
- your ancestry: the relations between you and your family
- phylogeny: the evolutionary relationships of life

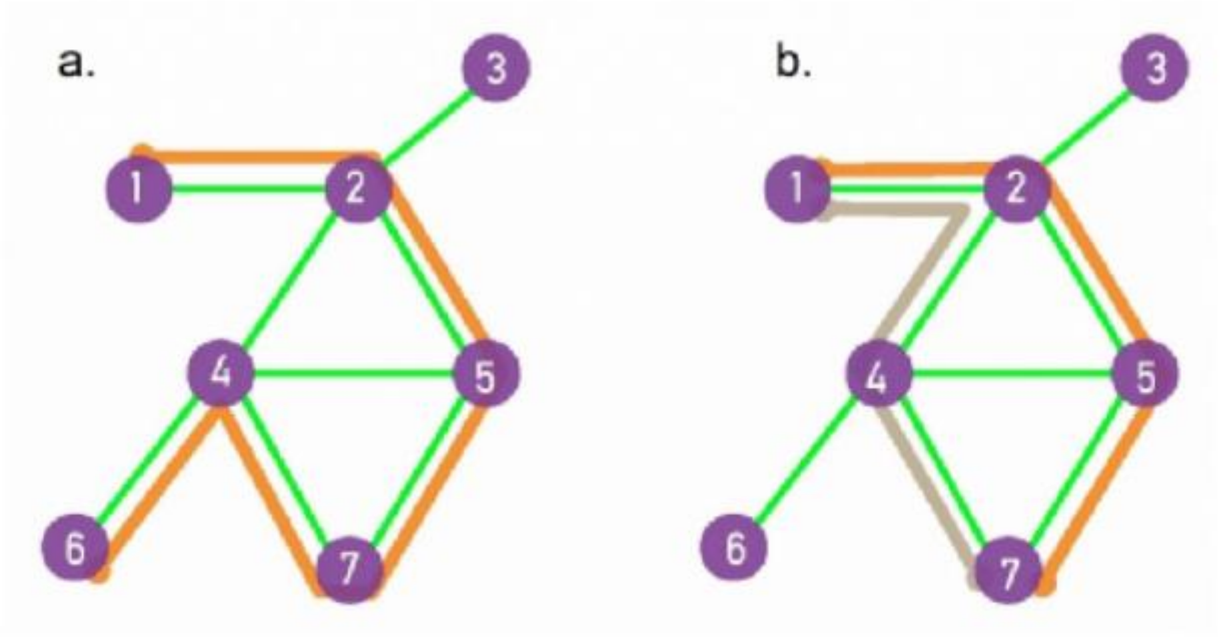


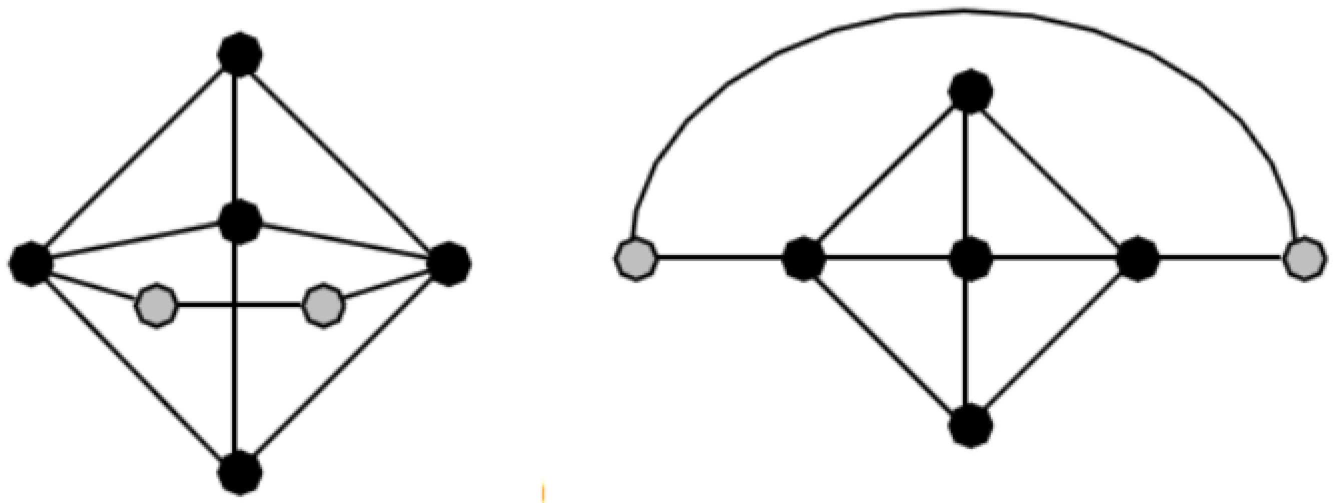


[Beyer 2014]

Network tasks: topology-based (拓樸或位相幾何)

- topological structure of network
  - path following
    - path is route along links
      - hops from one node to another
    - path length is number of links along route
      - shortest path connects nodes i & j with smallest # of hops
- topology vs geometry
  - topological hops different from geometric distance given specific layout
    - topology does not depend on layout
    - geometry does





Network: data

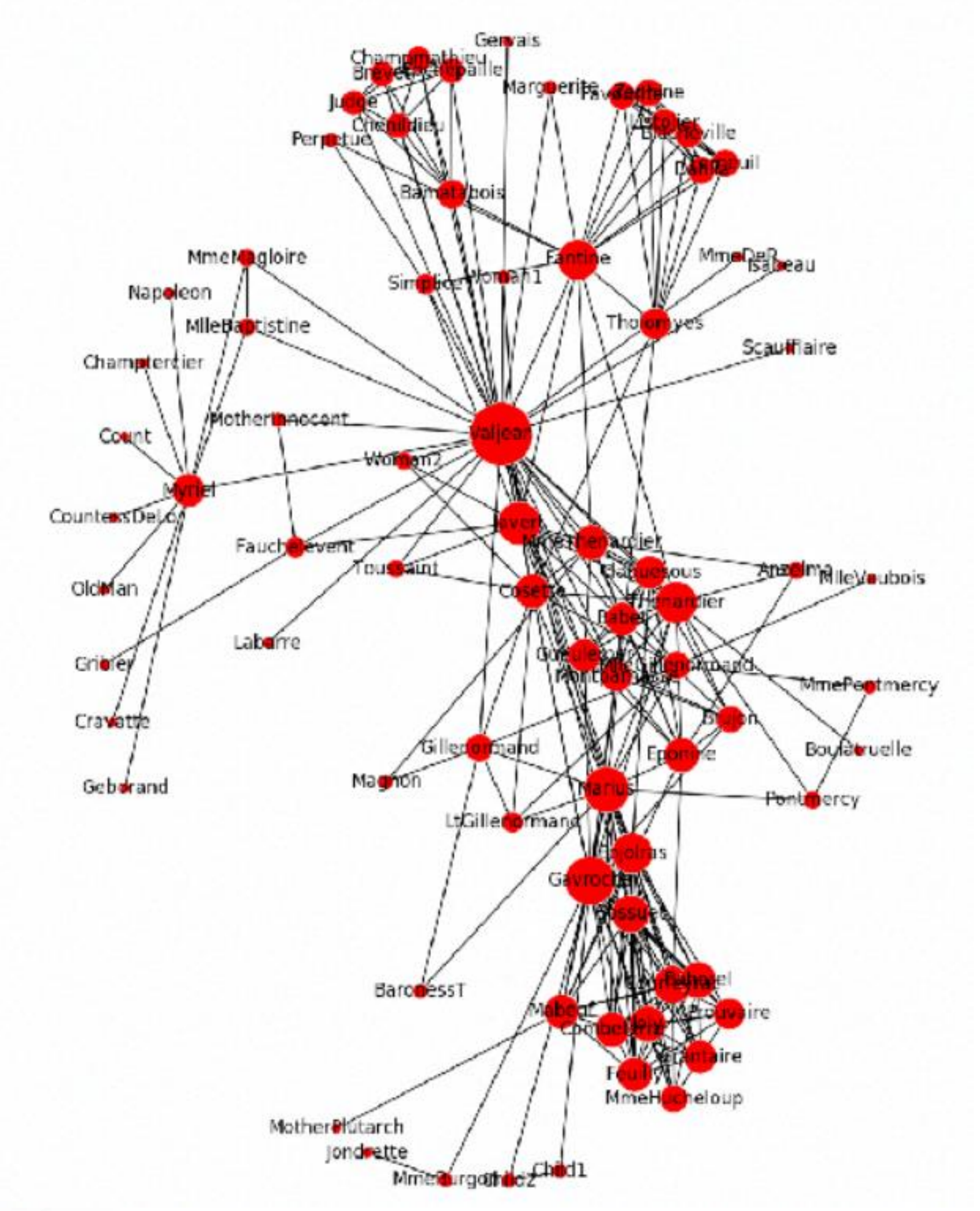
- Example: miserables.json

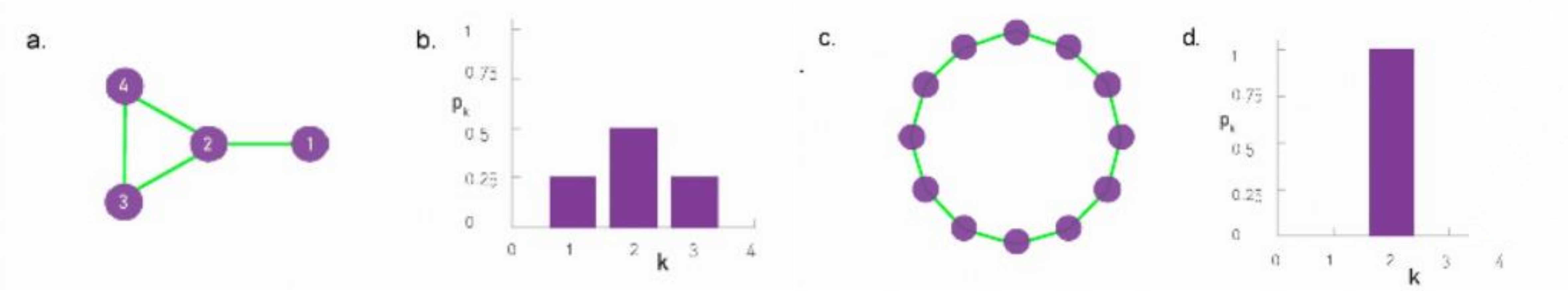
```
{
  "nodes": [
    {"id": "Myriel", "group": 1},
    {"id": "Napoleon", "group": 1},
    {"id": "Mlle.Baptistine", "group": 1},
    {"id": "Mme.Magloire", "group": 1},
    {"id": "CountessdeLo", "group": 1},
    {"id": "Geborand", "group": 1},
    {"id": "Champtercier", "group": 1},
    {"id": "Cravatte", "group": 1},
    {"id": "Count", "group": 1},
    {"id": "OldMan", "group": 1},
    {"id": "Labarre", "group": 2},
    {"id": "Valjean", "group": 2},
    {"id": "Marguerite", "group": 3},
    {"id": "Mme.deR", "group": 2},
    {"id": "Isabeau", "group": 2},
    {"id": "Gervais", "group": 2},
    {"id": "Tholomyes", "group": 3},
    {"id": "Listolier", "group": 3},
    {"id": "Fameuil", "group": 3},
    {"id": "Blacheville", "group": 3},
    {"id": "Favourite", "group": 3},
    {"id": "Dahlia", "group": 3},
    {"id": "Zephine", "group": 3},
    {"id": "Fantine", "group": 3},
    {"id": "Mme.Thenardier", "group": 4},
```

```
  "links": [
    {"source": "Napoleon", "target": "Myriel", "value": 1},
    {"source": "Mlle.Baptistine", "target": "Myriel", "value": 8},
    {"source": "Mme.Magloire", "target": "Myriel", "value": 10},
    {"source": "Mme.Magloire", "target": "Mlle.Baptistine", "value": 6},
    {"source": "CountessdeLo", "target": "Myriel", "value": 1},
    {"source": "Geborand", "target": "Myriel", "value": 1},
    {"source": "Champtercier", "target": "Myriel", "value": 1},
    {"source": "Cravatte", "target": "Myriel", "value": 1},
    {"source": "Count", "target": "Myriel", "value": 2},
    {"source": "OldMan", "target": "Myriel", "value": 1},
    {"source": "Valjean", "target": "Labarre", "value": 1},
    {"source": "Valjean", "target": "Mme.Magloire", "value": 3},
    {"source": "Valjean", "target": "Mlle.Baptistine", "value": 3},
    {"source": "Valjean", "target": "Myriel", "value": 5},
    {"source": "Marguerite", "target": "Valjean", "value": 1},
    {"source": "Mme.deR", "target": "Valjean", "value": 1},
    {"source": "Isabeau", "target": "Valjean", "value": 1},
    {"source": "Gervais", "target": "Valjean", "value": 1},
    {"source": "Listolier", "target": "Tholomyes", "value": 4},
    {"source": "Fameuil", "target": "Tholomyes", "value": 4},
    {"source": "Fameuil", "target": "Listolier", "value": 4},
    {"source": "Blacheville", "target": "Tholomyes", "value": 4},
    {"source": "Blacheville", "target": "Listolier", "value": 4},
    {"source": "Blacheville", "target": "Fameuil", "value": 4},
    {"source": "Favourite", "target": "Tholomyes", "value": 3},
    {"source": "Favourite", "target": "Listolier", "value": 3}
```

Network tasks: topology-based

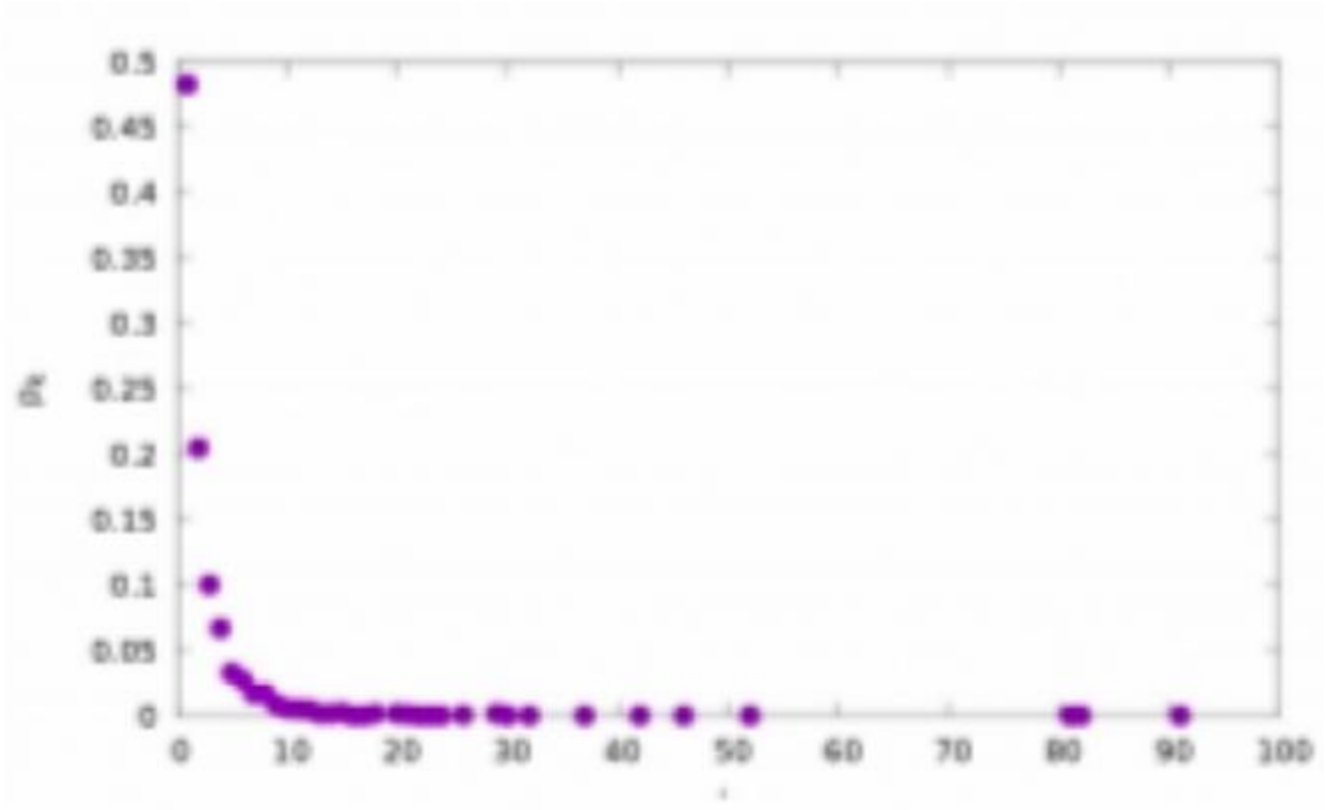
- topological structure of network
  - node importance metrics
    - **node degree**: attribute on nodes
      - number of links connected to a node
      - local measure of importance
      - average degree, degree distribution

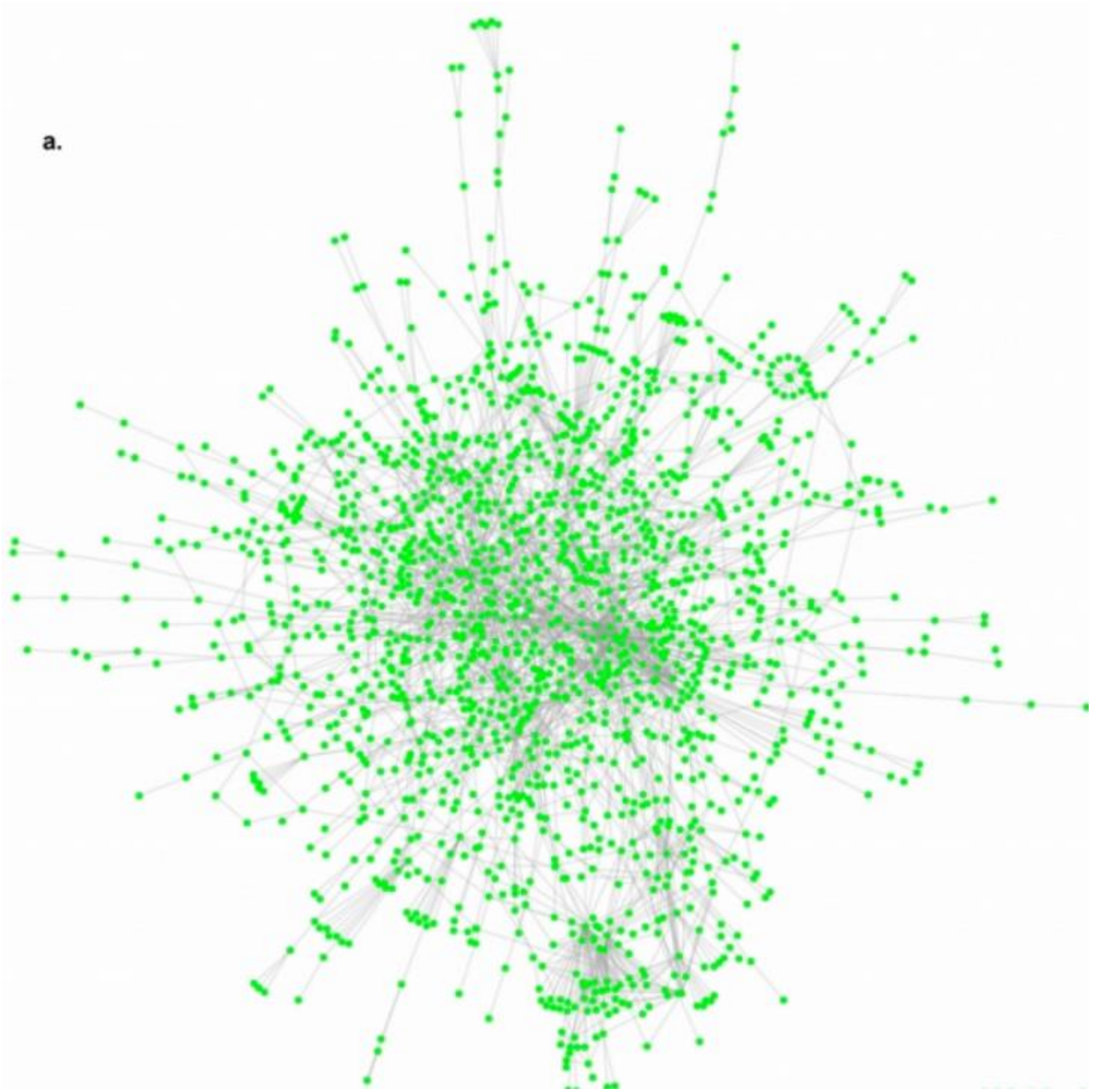




Degree distribution

- **real network**
  - power law distributions are common

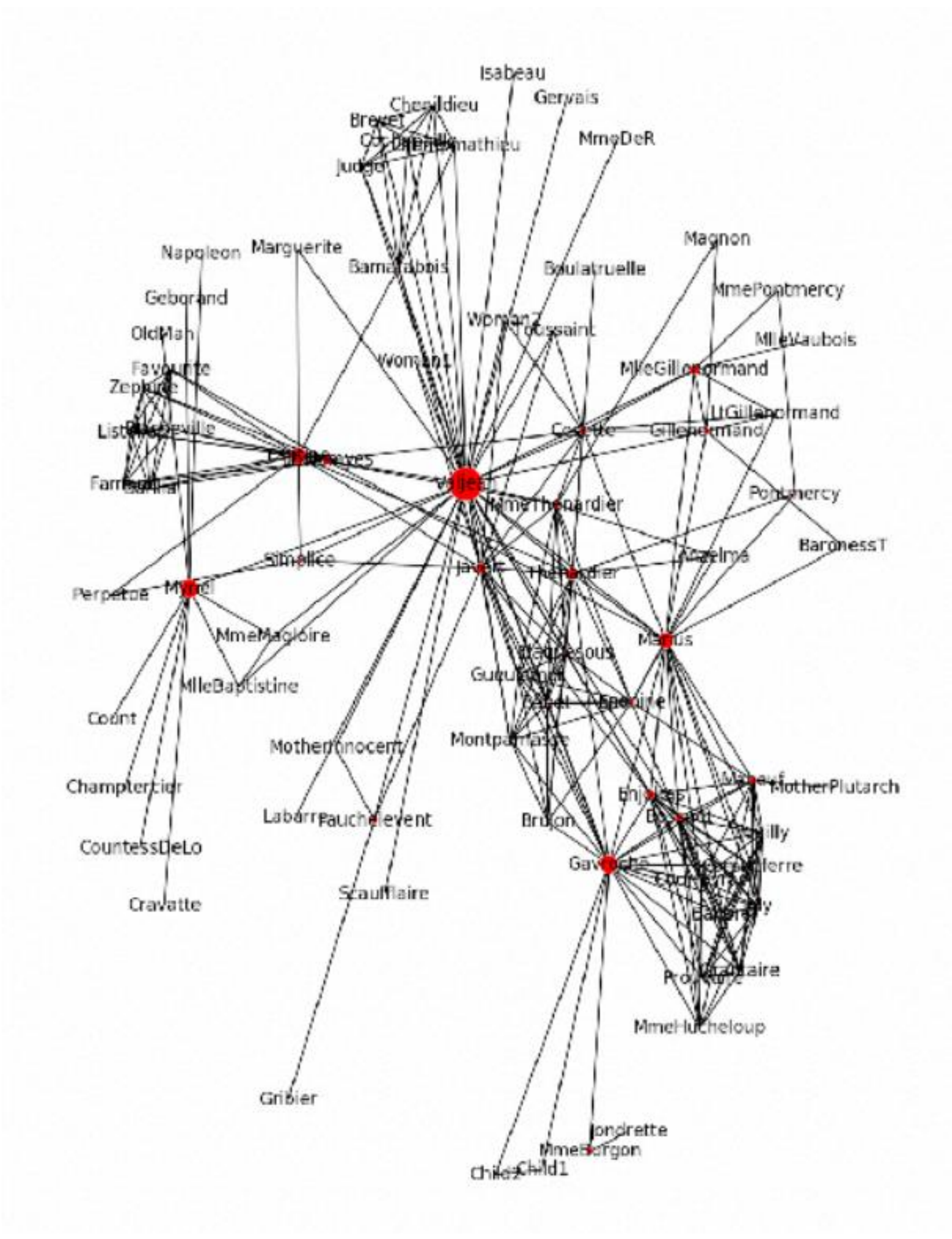




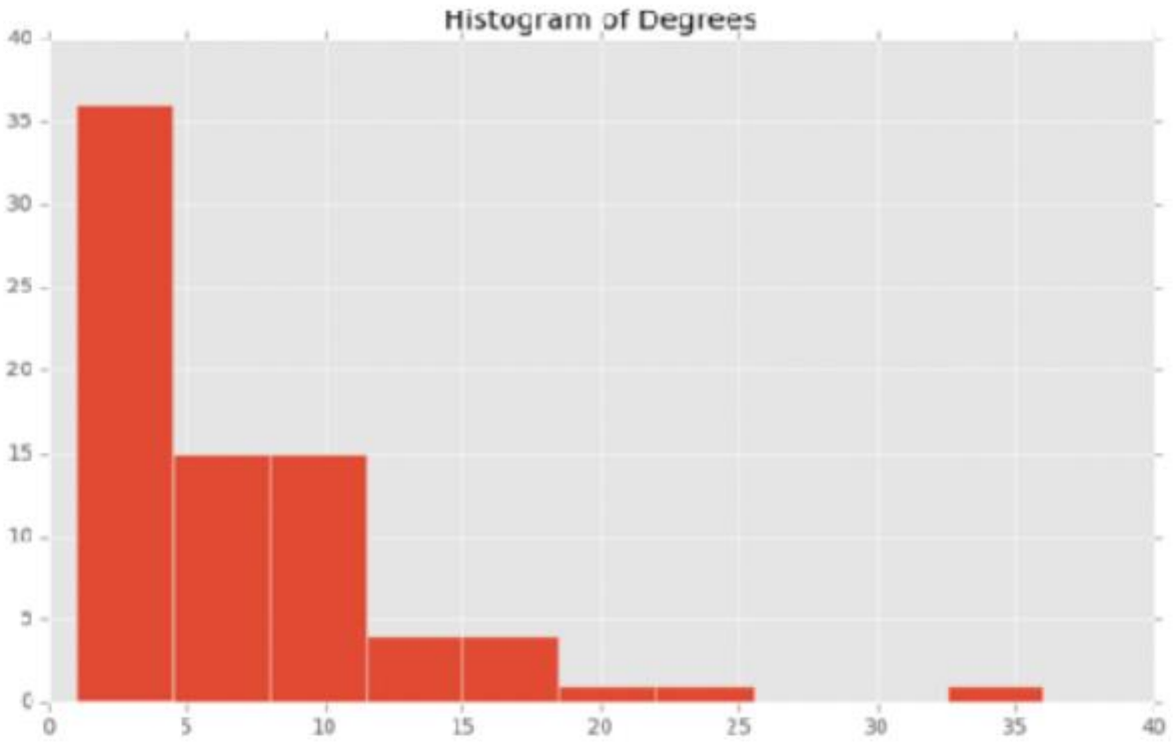
Protein interaction network, Barabasi

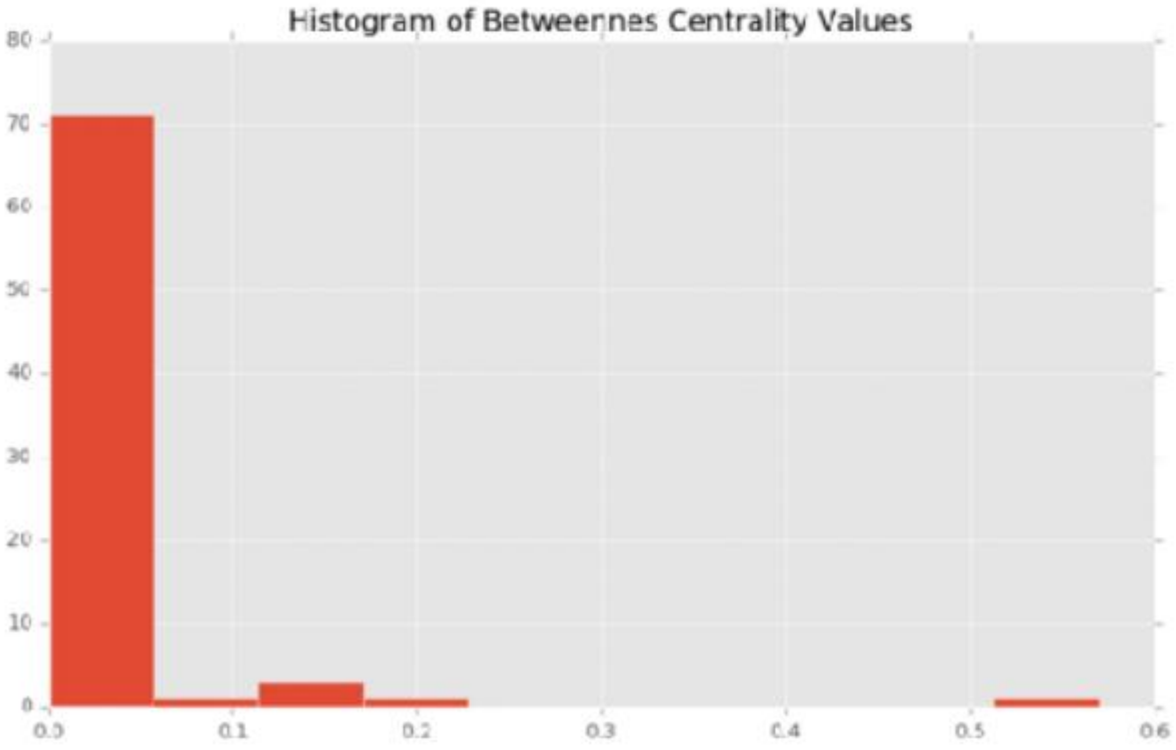
Network tasks: topology-based

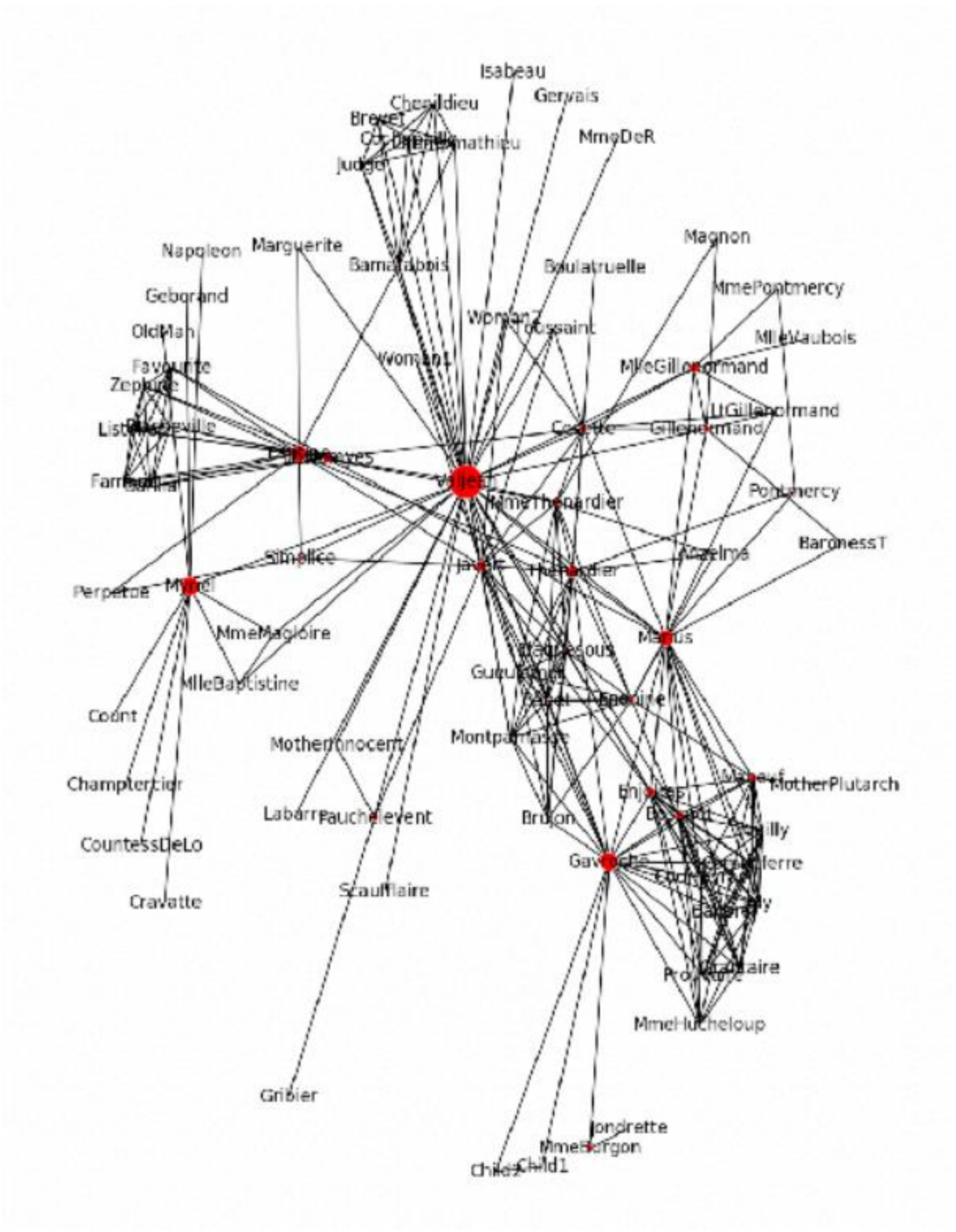
- topological structure of network
  - node importance metrics
    - **betweenness centrality**: attribute on nodes
      - how many shortest paths pass through a node
      - global measure of importance
      - good measure for overall relevance of node in network



Centrality measures: Degree vs betweenness centrality

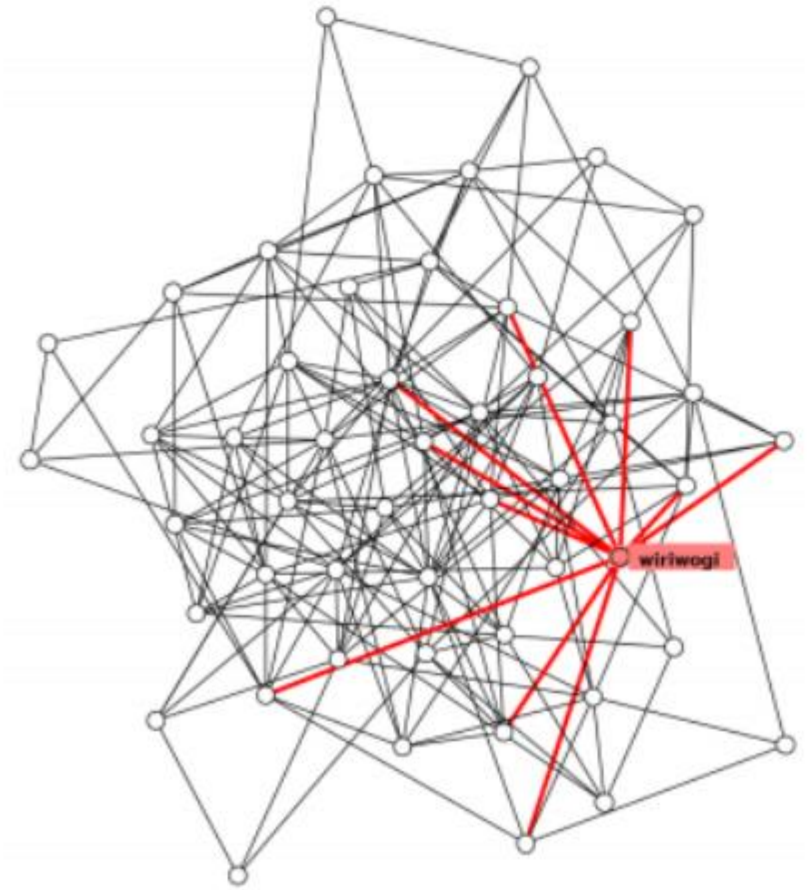






Network tasks: attribute-based vs topology-based

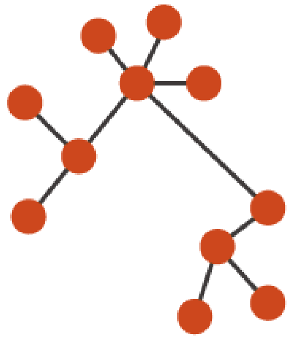
- topology based tasks
  - find paths
  - find (topological) neighbors
  - compare centrality/importance measures
  - identify clusters / communities
- attribute based tasks (similar to table data)
  - find extreme values, ...
- combination tasks - incorporating both
  - example: locate - find single or multiple nodes/links with a given property
    - topology: find all adjacent nodes of given node
    - attributes: find edges with maximum edge weight



Arrange networks and trees

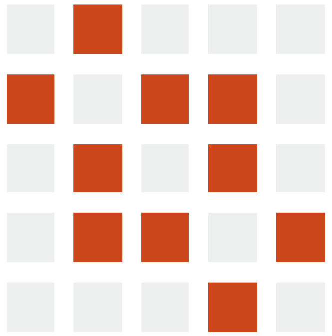
Node–Link Diagrams
Connection Marks

✔ NETWORKS ✔ TREES



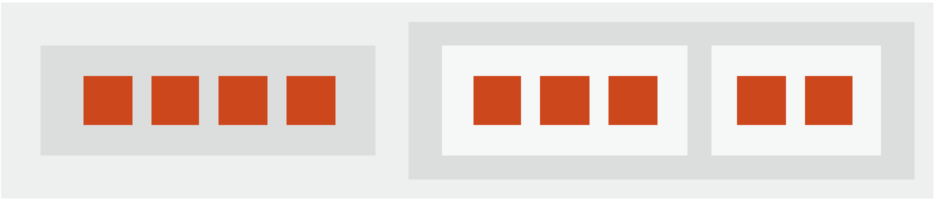
Adjacency Matrix
Derived Table

✔ NETWORKS ✔ TREES



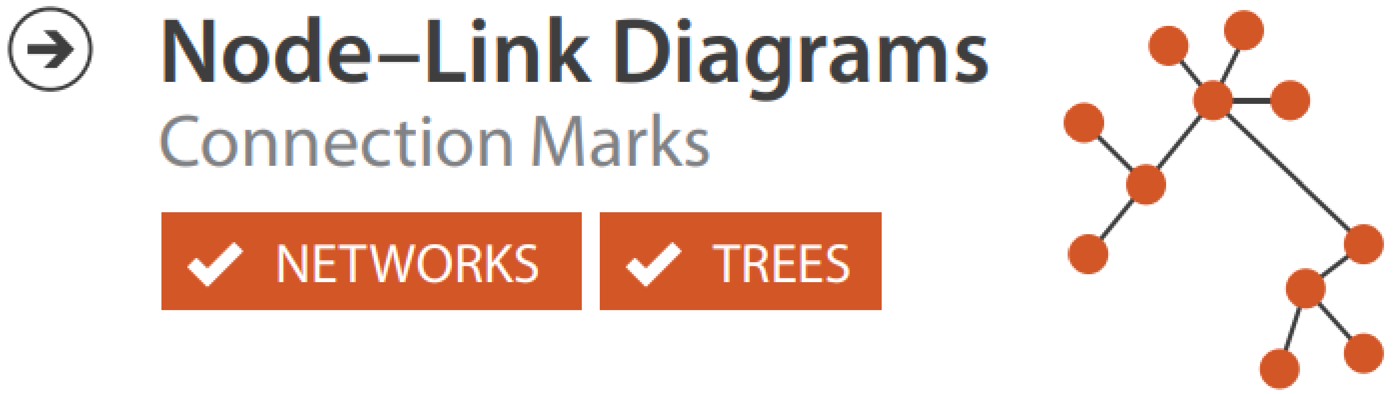
Enclosure
Containment Marks

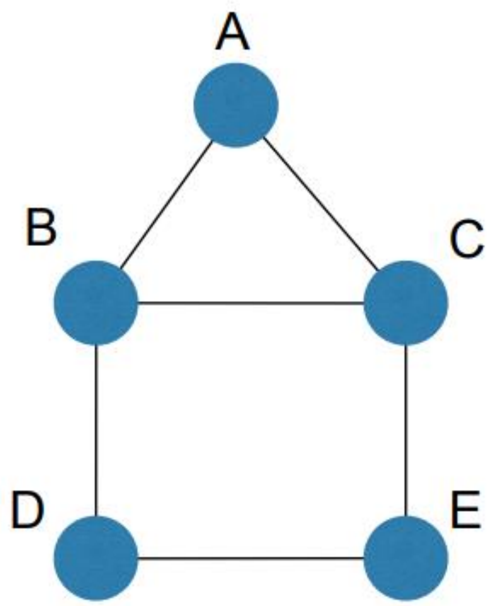
✖ NETWORKS ✔ TREES

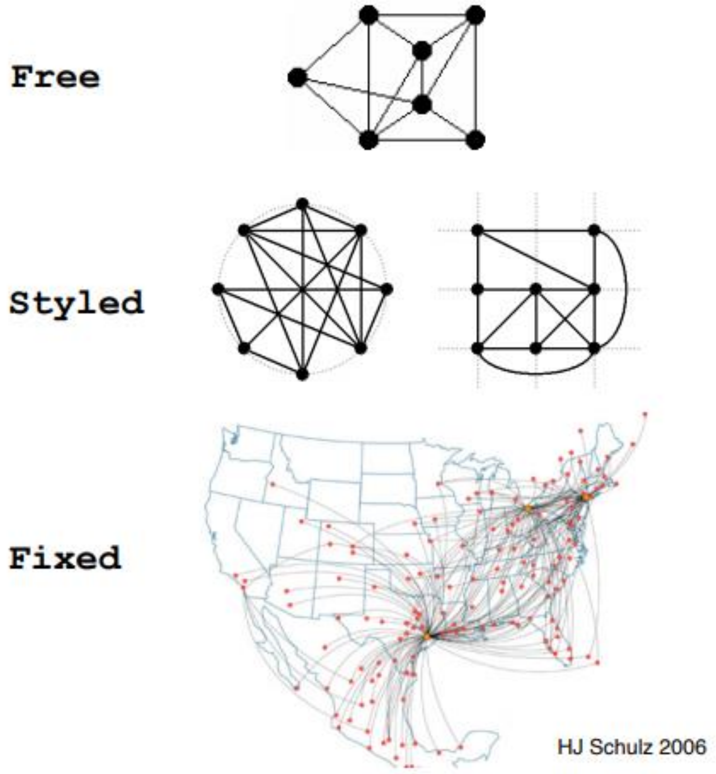


Node-link diagrams

- **nodes**: point marks
- **links**: line marks
  - straight lines or arcs
  - connections between nodes
- intuitive & familiar
  - most common
  - many, many variants

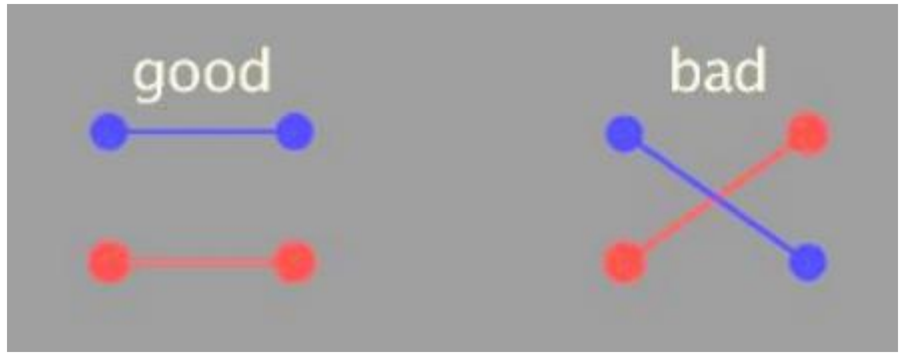






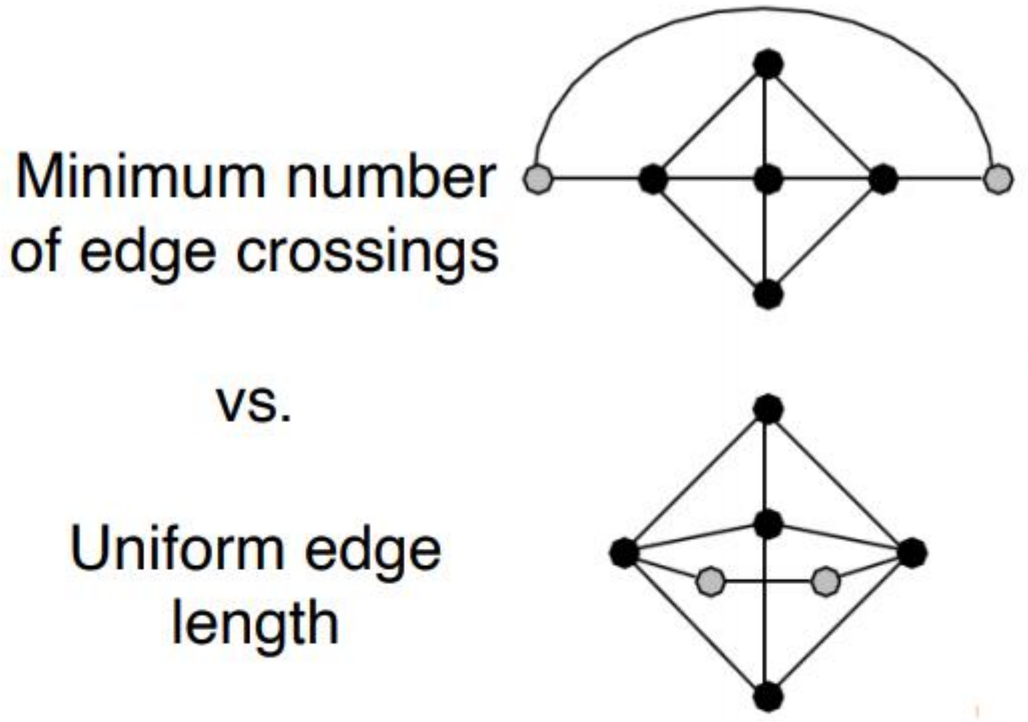
Criteria for good node-link layouts

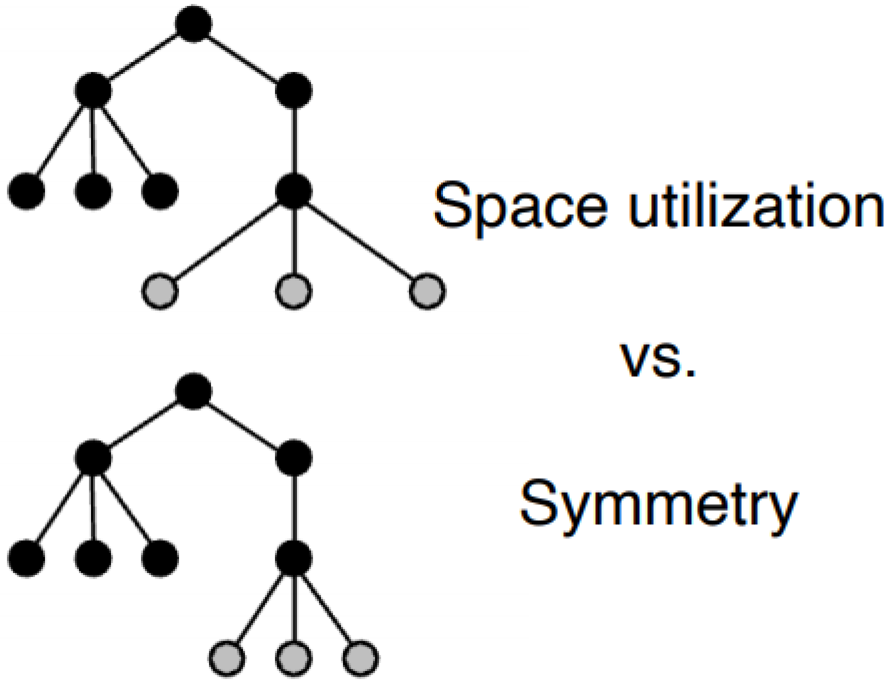
- minimize
  - edge crossings
  - distances between topological neighbor nodes
  - total drawing area
  - edge bends
  - edge length disparities (sometimes)
- maximize
  - angular distance between different edges
  - aspect ratio disparities
- emphasize symmetry
  - similar graph structures should look similar in layout



Criteria conflict

- most criteria NP-hard individually
- many criteria directly conflict with each other





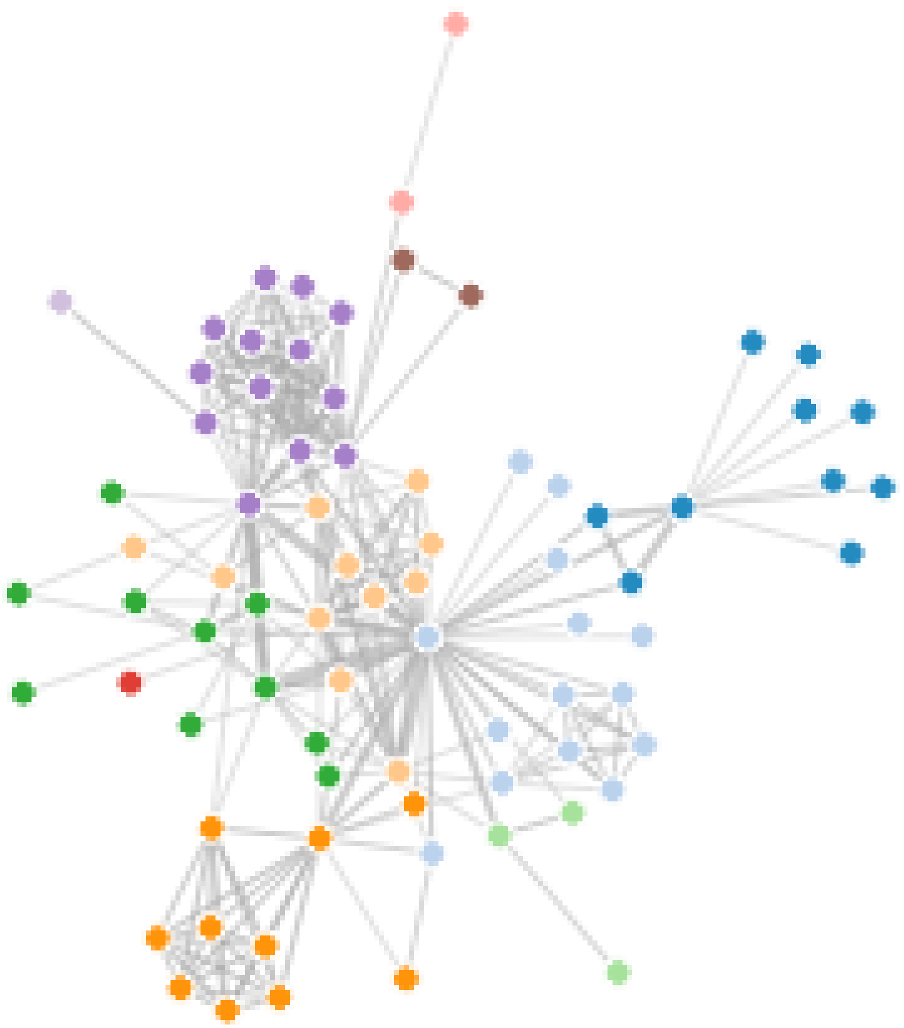
Schulz 2004

Optimization-based layouts

- formulate layout problem as optimization problem
- convert criteria into weighted cost function
  - F(layout) = a*[crossing counts] + b*[drawing space used] + …
- use known optimization techniques to find layout at minimal cost
  - energy-based physics models
  - force-directed placement
  - spring embedders

Force-directed placement

- physics model
  - links = springs pull together
  - nodes = magnets repulse apart
- algorithm
  - place vertices in random locations
  - while not equilibrium
    - calculate force on vertex
      - sum of
        - » pairwise repulsion of all nodes
        - » attraction between connected nodes
    - move vertex by **c * vertex_force**



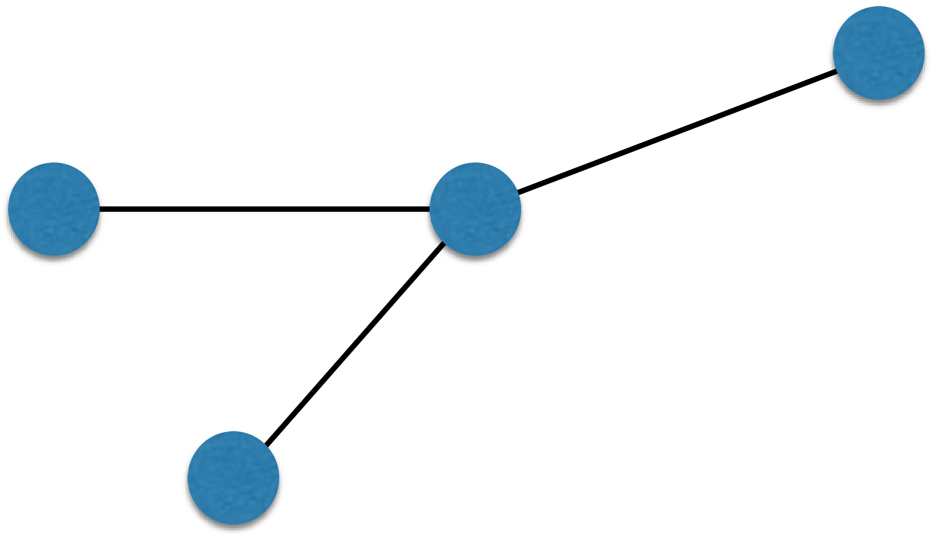
[character co-occurence in *Les Misérables*](https://homes.cs.washington.edu/~jheer/files/zoo/ex/networks/force.html)

https://homes.cs.washington.edu/~jheer/files/zoo/ex/networks/force.html

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

A heuristic for graph drawing [1984]

- *Spring force*
  - $F_{spring} = c_1 \log(\frac{d}{c_2})$
    - *force = 0, when* $d = c_2$
- *Repel force for non-adjacent vertex*
  - $F_{repeling} = c_3/d^2$
- *Move*
  - $c_4 F_{total}$



```
Spring (G: graph)
--------------------------------
Place vertices of G in random locations;

Repeat M times
    calculate the force on each vertex;
    move the vertex;

Draw the graph;
```

Force-Directed Graph

- Fruchterman and Reingold[1991]

Even vertex distribution $$k = \sqrt{\frac{area}{number\ of\ vertices}}$$

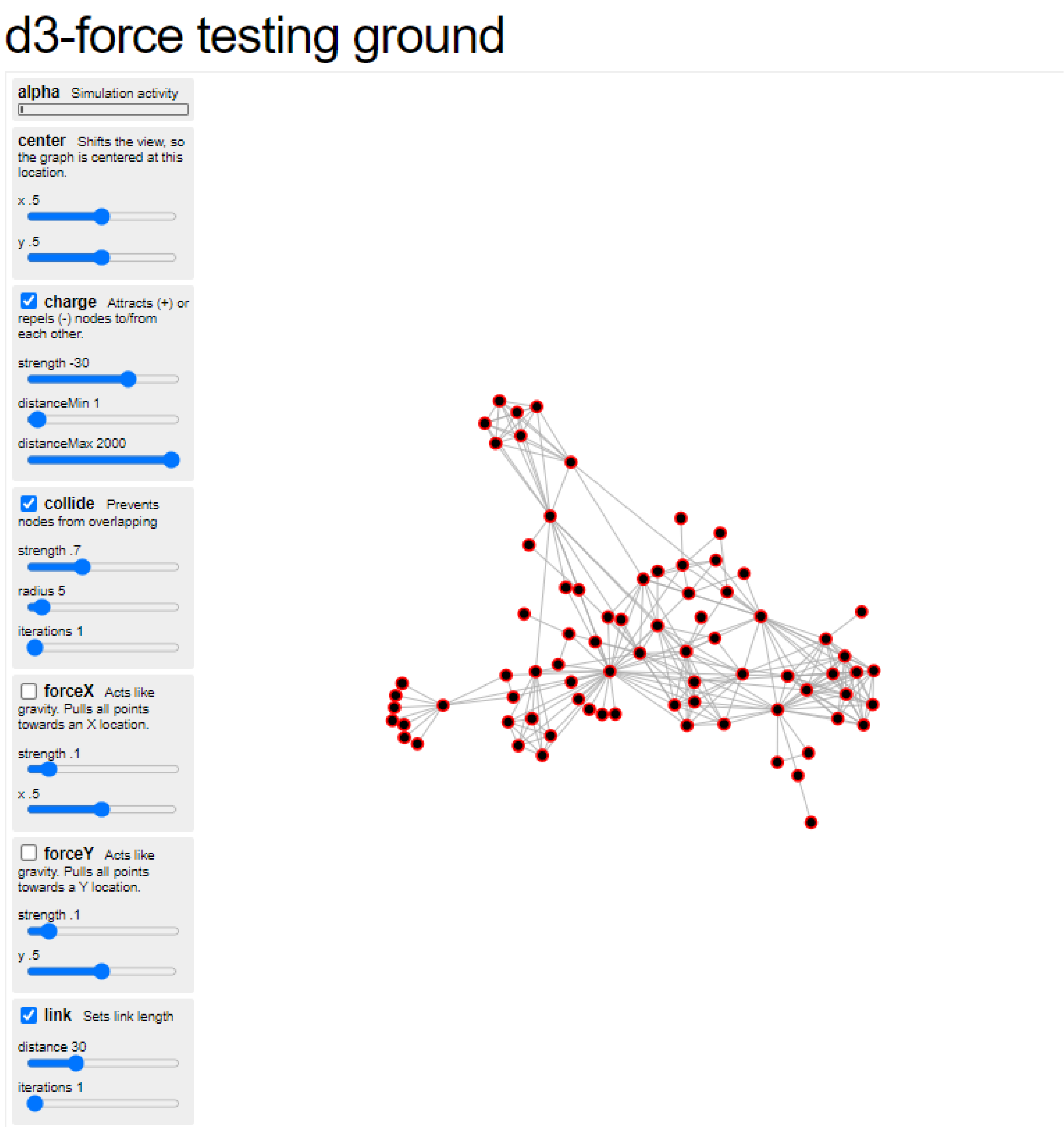
$$\Delta = v.pos - u.pos, \ \ for\ each\ node\ u, v$$

attractive $$f_a(\Delta) = \frac{\Delta^2}{k}$$

repulsive $$f_r(\Delta) = \frac{k^2}{\Delta}$$

Force-directed placement properties

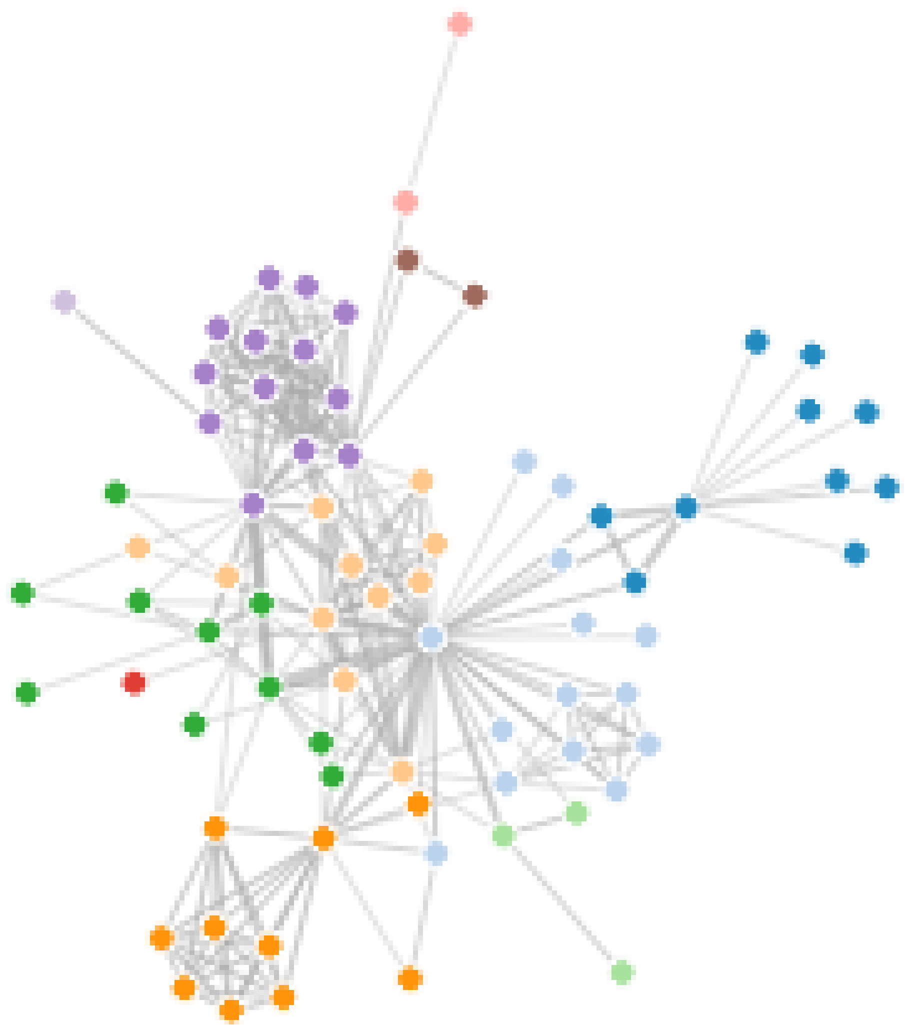
- strengths
  - reasonable layout for small, sparse graphs
  - clusters typically visible
  - edge length uniformity
- weaknesses
  - nondeterministic
  - computationally expensive: O(n^3) for n nodes
- each step is n^2, takes ~n cycles to reach equilibrium
  - naive FD doesn't scale well beyond 1K nodes
  - iterative progress: engaging but distracting



https://bl.ocks.org/steveharoz/8c3e2524079a8c440df60c1ab72b5d03

Idiom: **force-directed placement**

- visual encoding
  - link connection marks, node point marks
- considerations
  - spatial position: no meaning directly encoded
    - left free to minimize crossings
  - proximity semantics?
    - sometimes meaningful
    - sometimes arbitrary, artifact of layout algorithm
    - tension with length
      - long edges more visually salient than short
- tasks
  - explore topology; locate paths, clusters
- scalability
  - node/edge density E < 4N

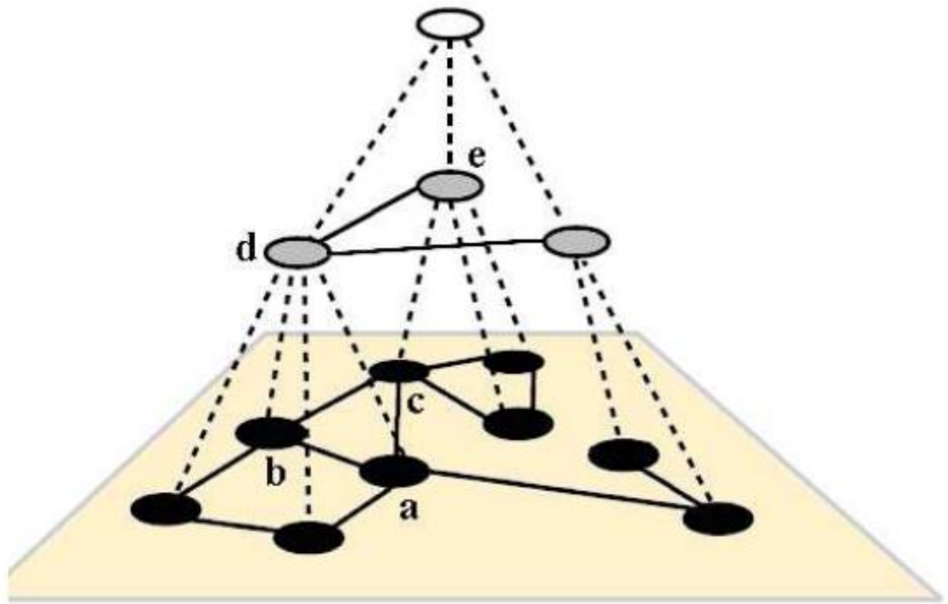


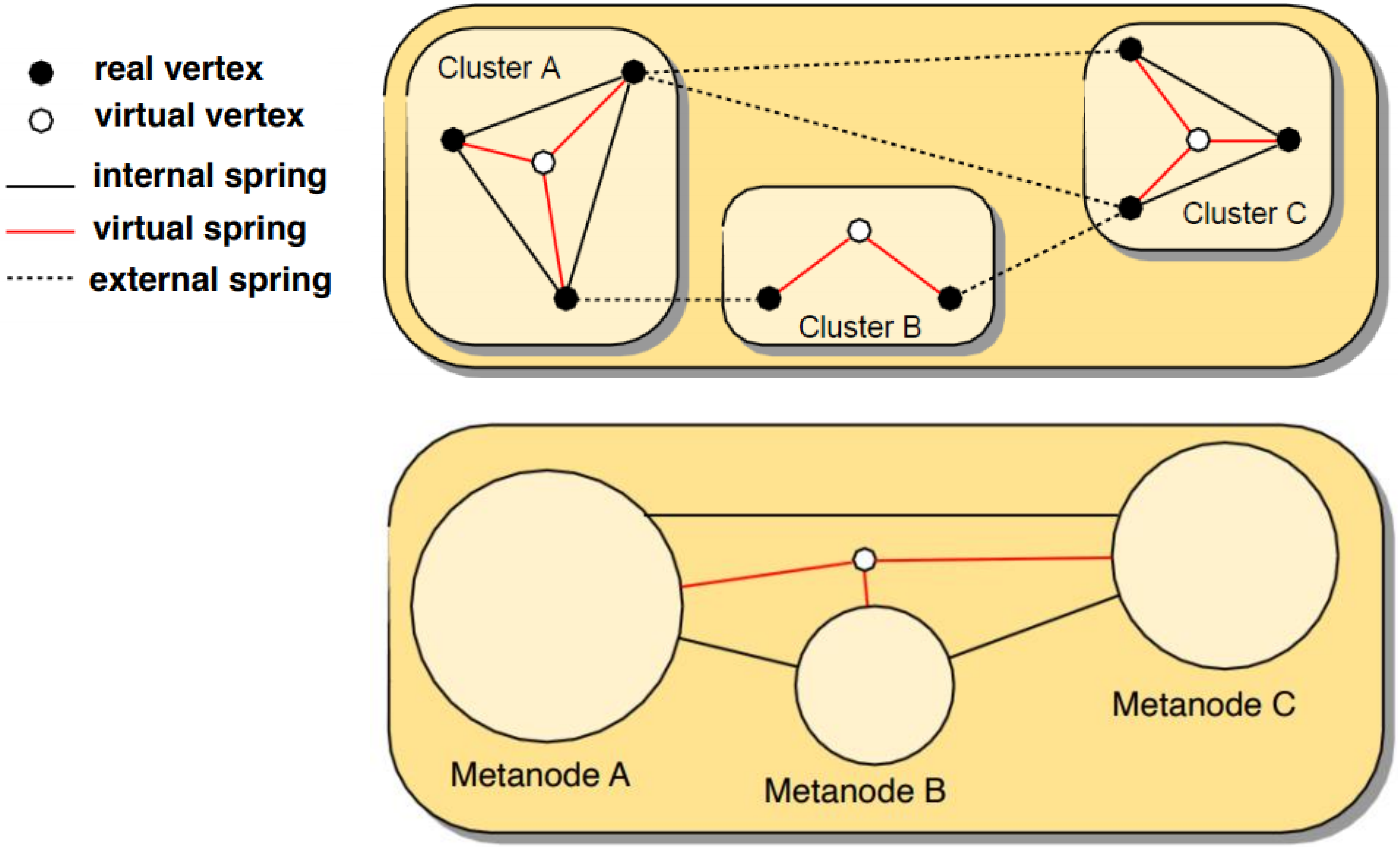
[character co-occurence in *Les Misérables*](http://mbostock.github.io/d3/ex/force.html)

http://mbostock.github.com/d3/ex/force.html

Multilevel approaches

- derive cluster hierarchy of metanodes on top of original graph nodes

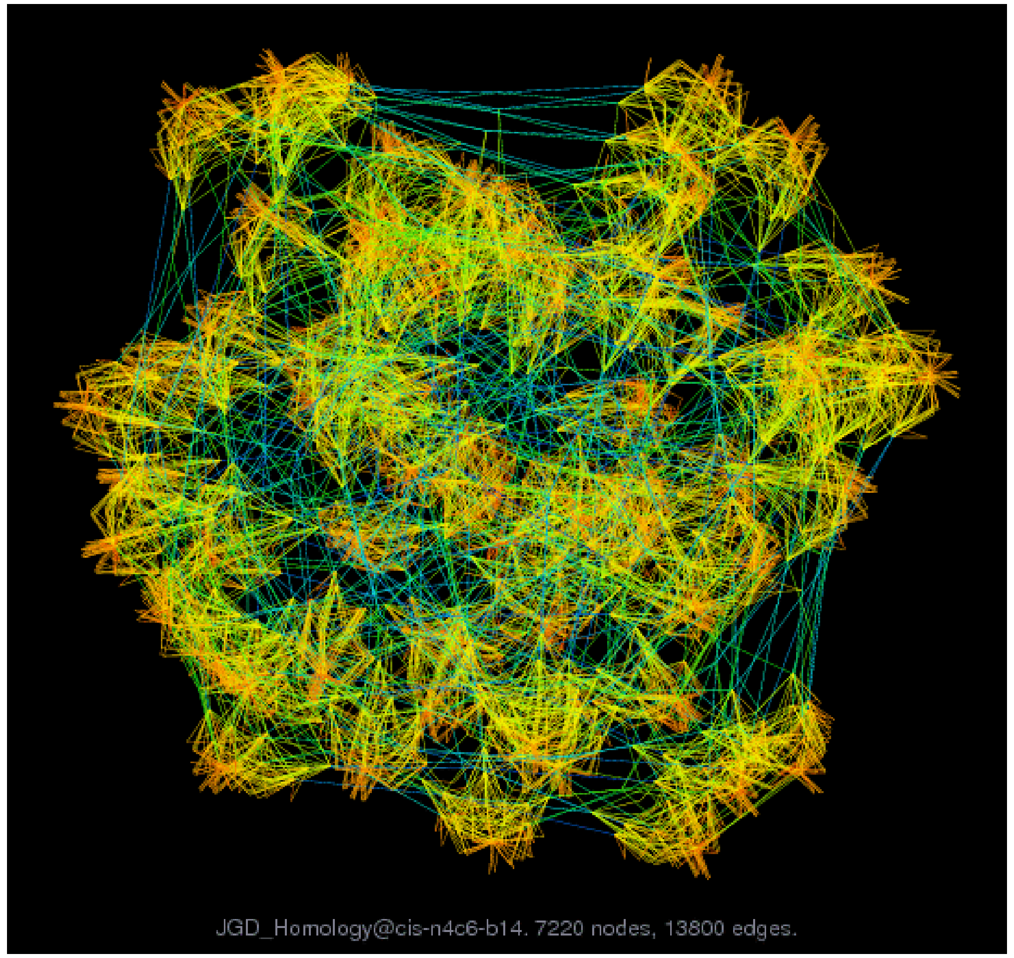




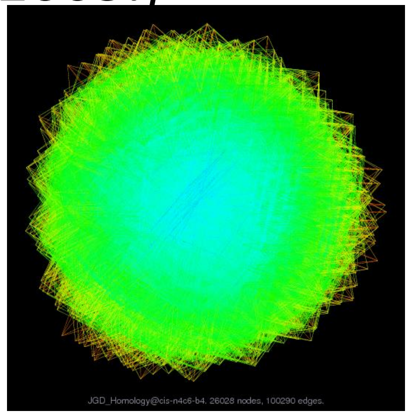
[Schulz 2004]

Idiom: **sfdp** (multi-level scalable force-directed placement)

- data
  - original: network
  - derived: cluster hierarchy atop it
- considerations
  - better algorithm for same encoding technique
    - same: fundamental use of space
    - hierarchy used for algorithm speed/quality but not shown explicitly
- scalability
  - nodes, edges: 1K-10K
  - hairball problem eventually hits



[Efficient and high quality force-directed graph drawing. Hu. The Mathematica Journal 10:37–71, 2005.]

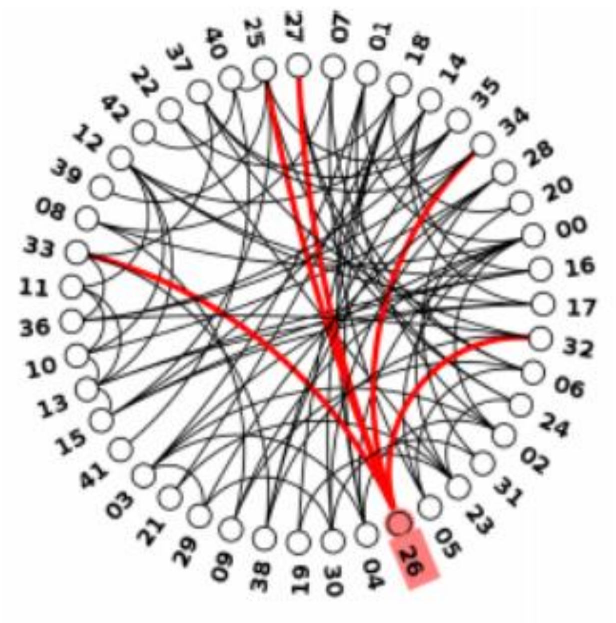


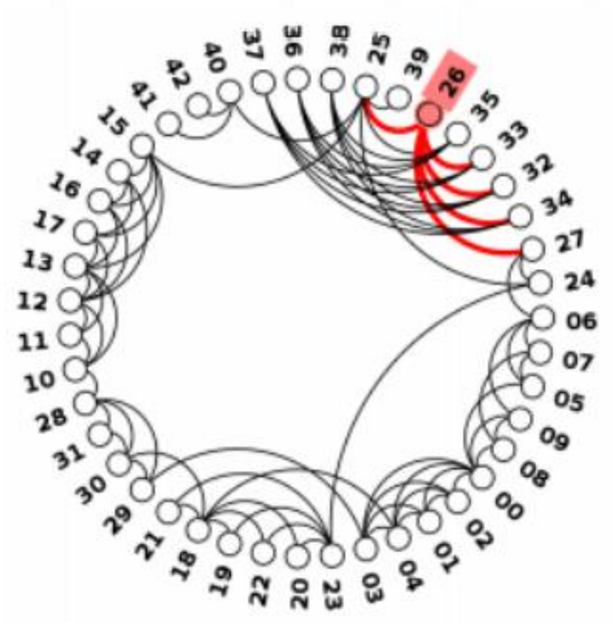
26,020 nodes and 100,290 edges

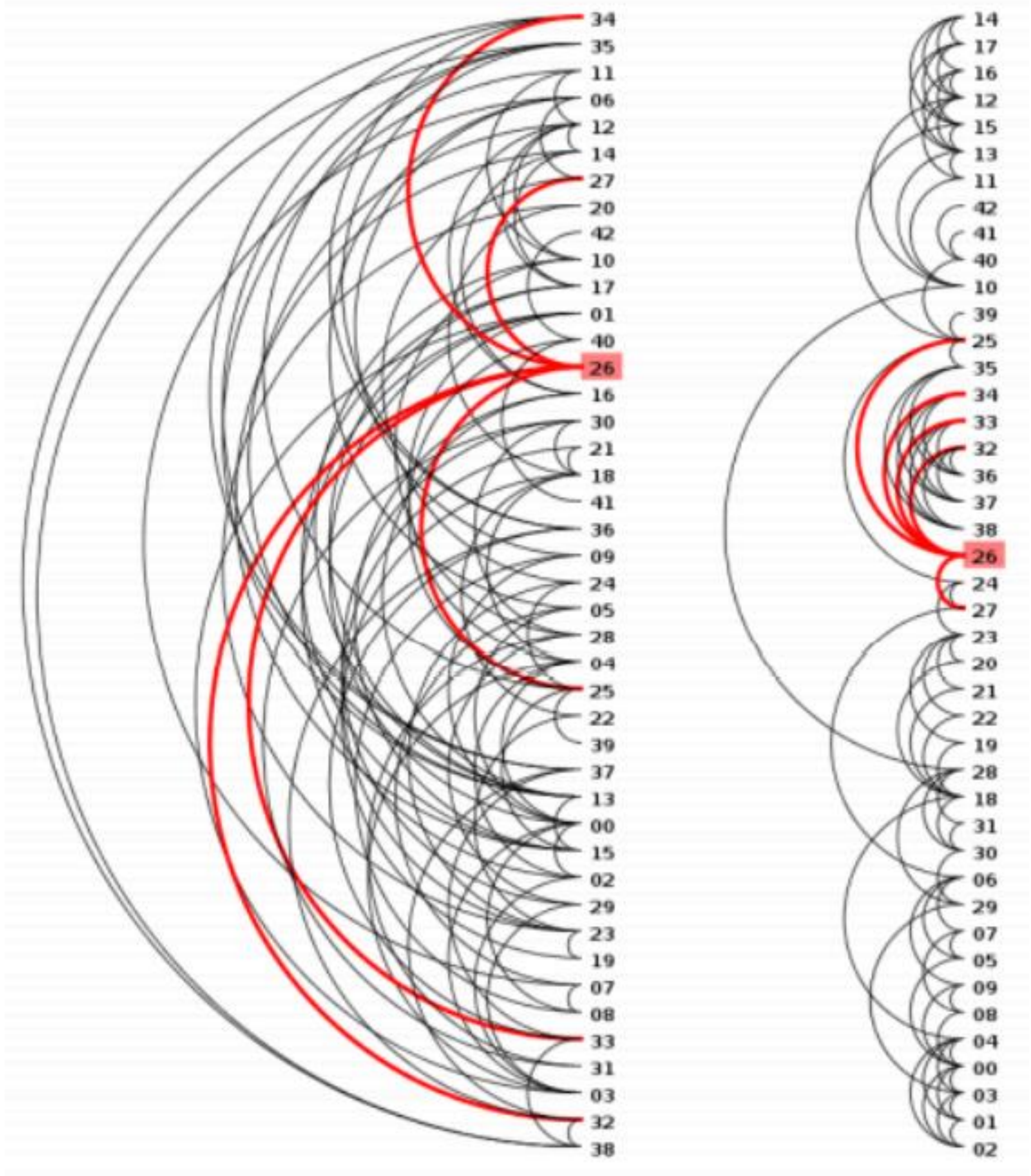
http://www.research.att.com/yifanhu/GALLERY/GRAPHS/index1.html

Restricted layouts: Circular, arc

- lay out nodes around **circle** or along **line**
  - circular layouts
  - arc diagrams
- **node ordering** crucial to avoid excessive clutter from edge crossings
  - barycentric ordering before & after
  - derived attribute: global computation

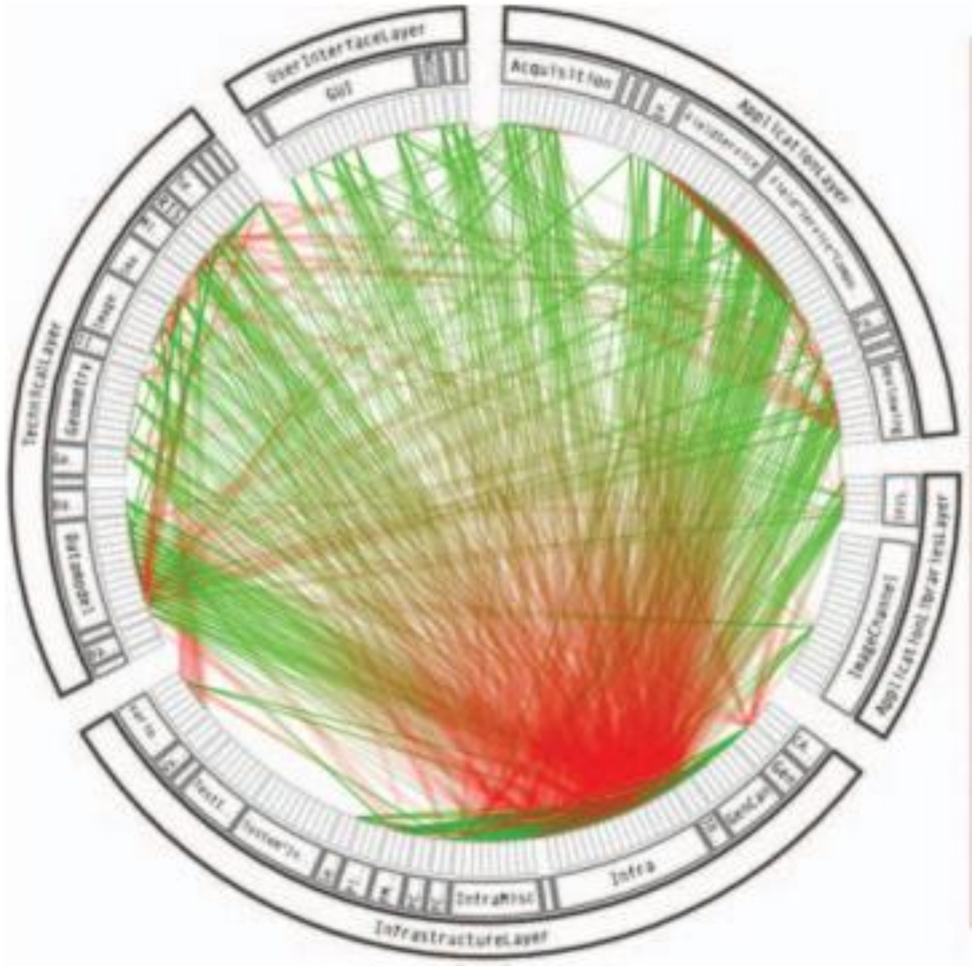




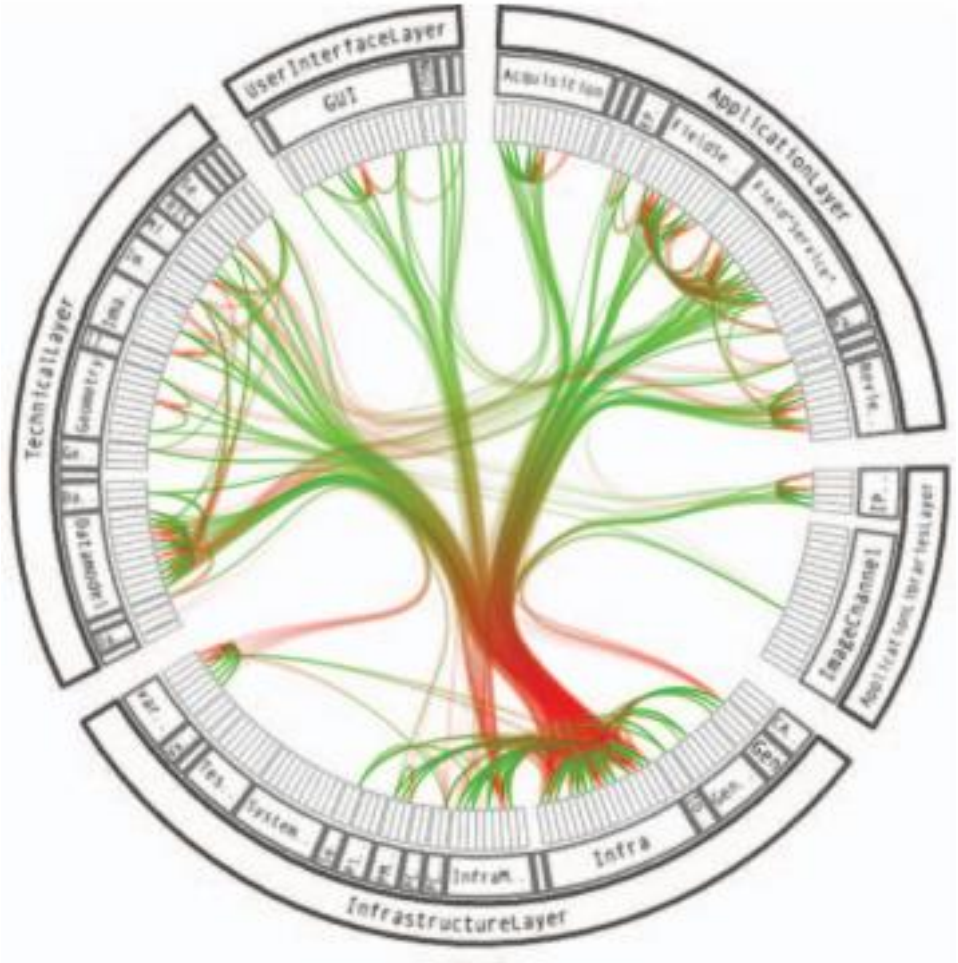


[Simple algorithms for network visualization: A tutorial 2012]

Edge clutter reduction: hierarchical edge bundling



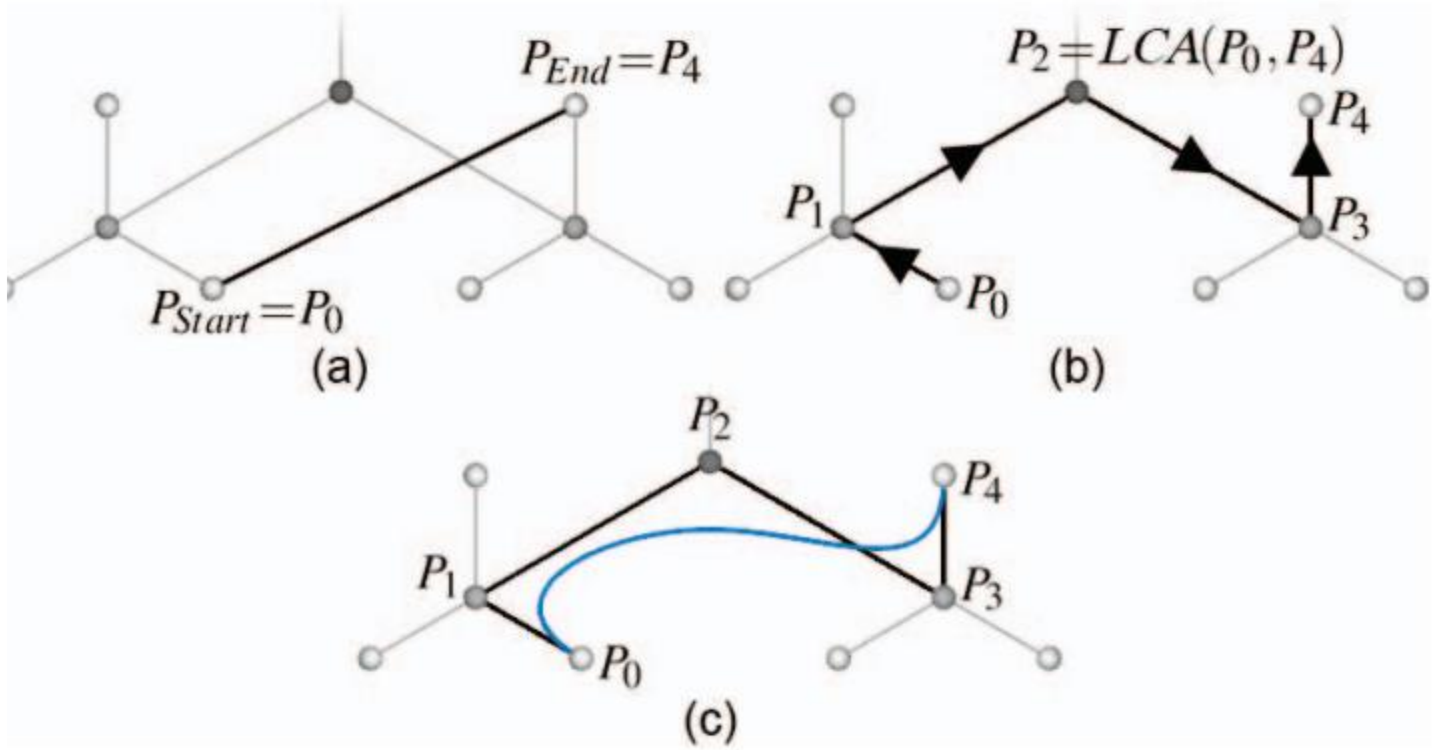
(a)

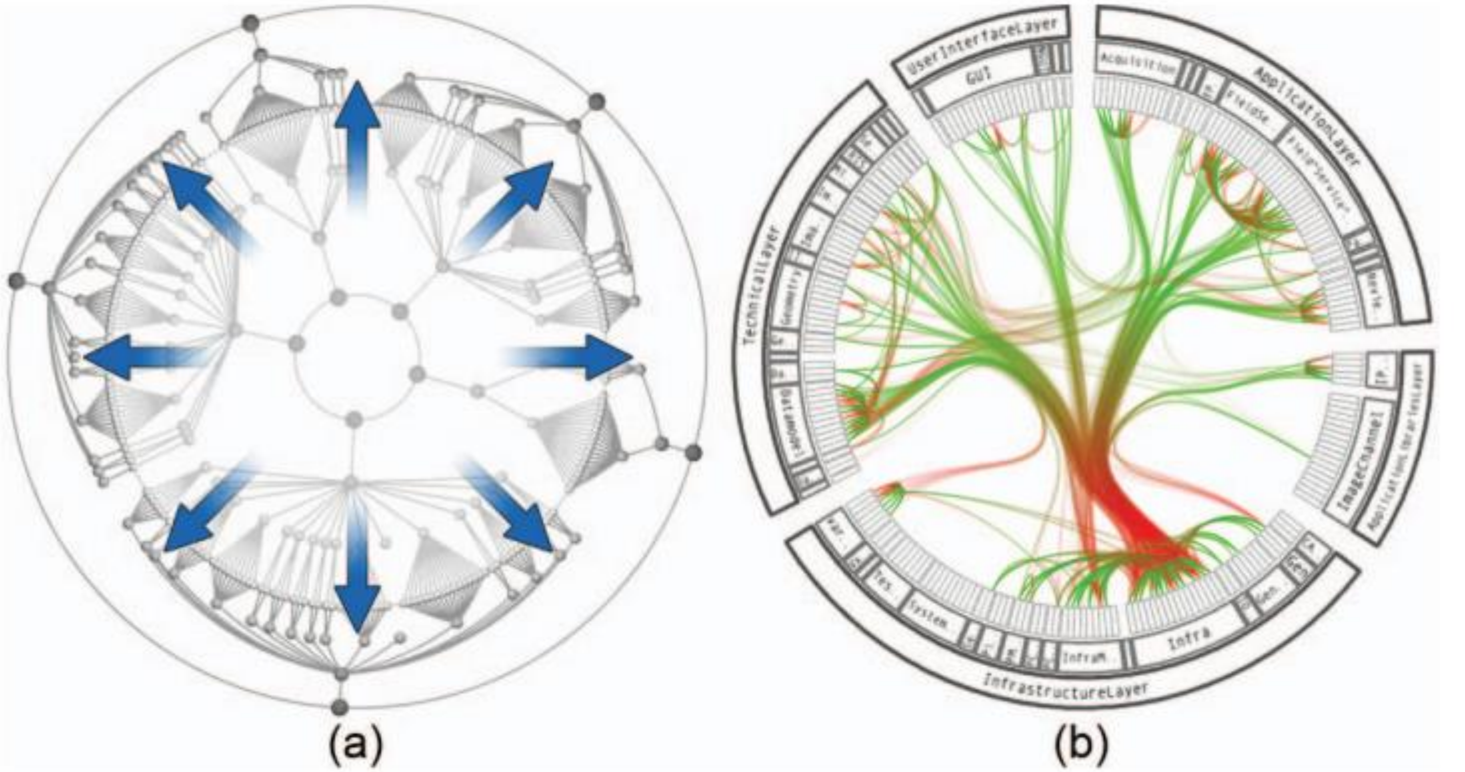


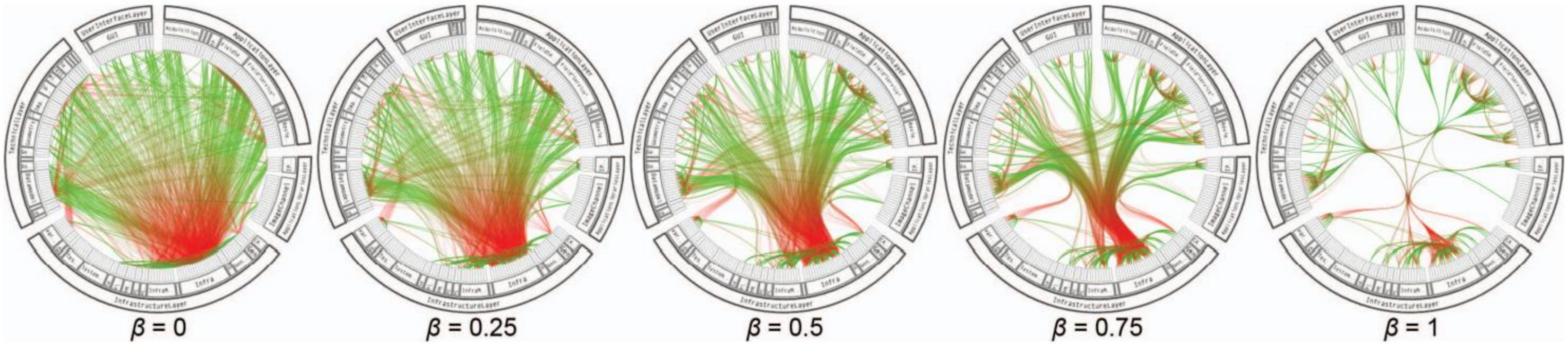
(b)

[Hierarchical Edge Bundles: Visualization of Adjacency Relations in Hierarchical Data. Danny Holten. TVCG 12(5):741-748 2006]

hierarchical edge bundling

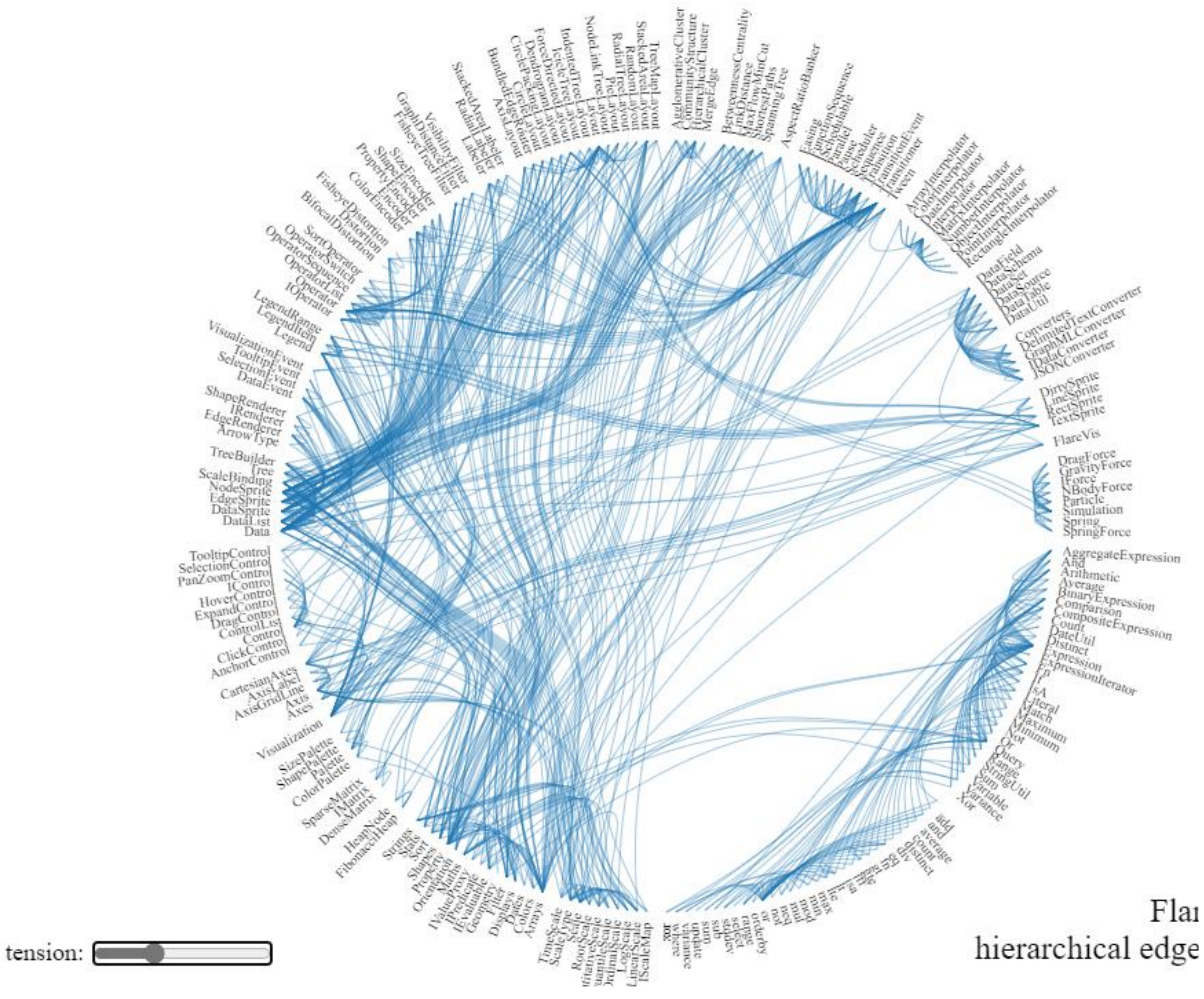


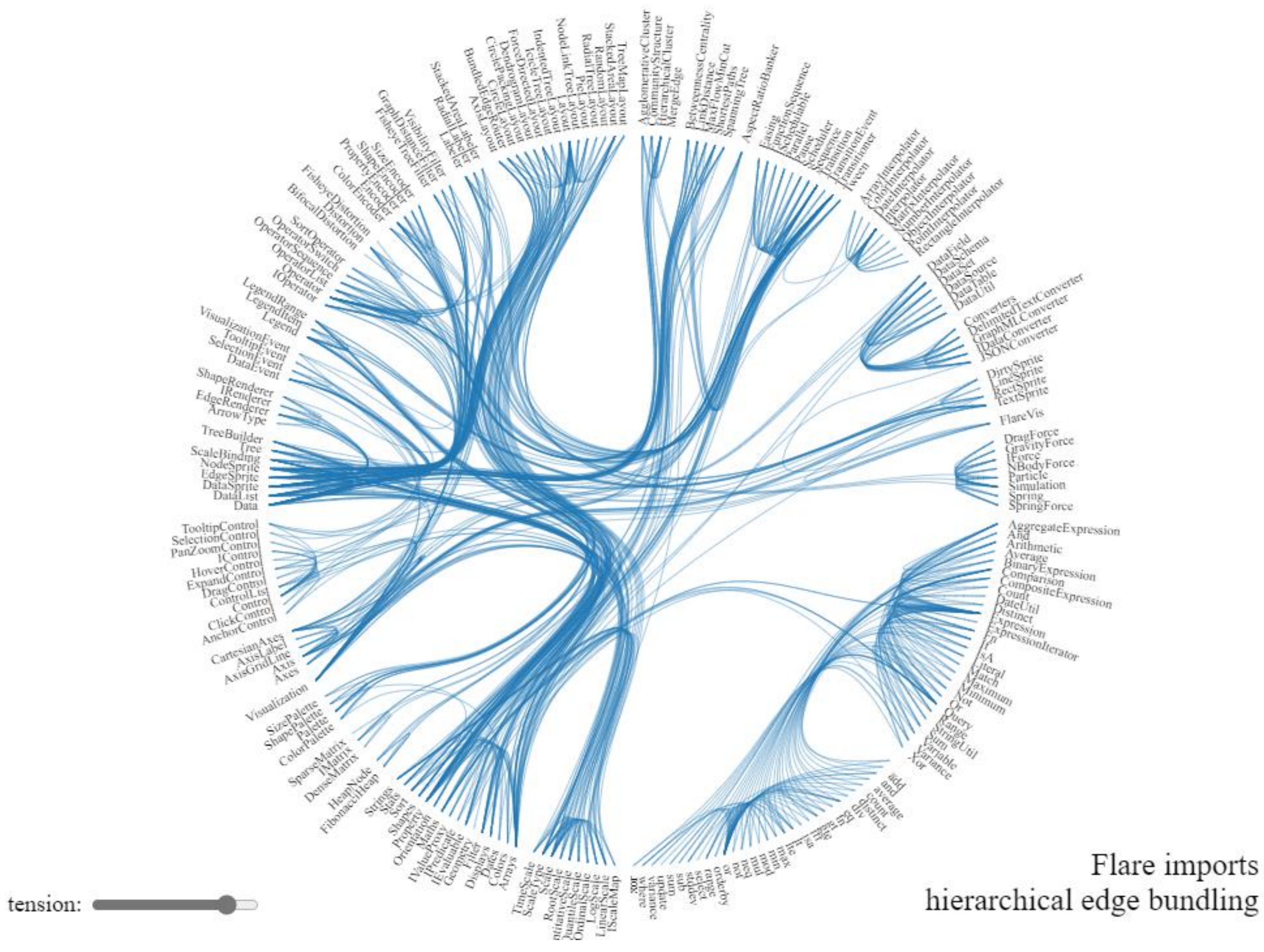




[Hierarchical Edge Bundles: Visualization of Adjacency Relations in Hierarchical Data. Danny Holten. TVCG 12(5):741-748 2006]

Bundle strength





mbostock.github.io/d3/talk/20111116/bundle.html

Fixed layouts: Geographic

- lay out network nodes using given/fixed spatial data
  - route edges accordingly
  - edge bundling also applicable

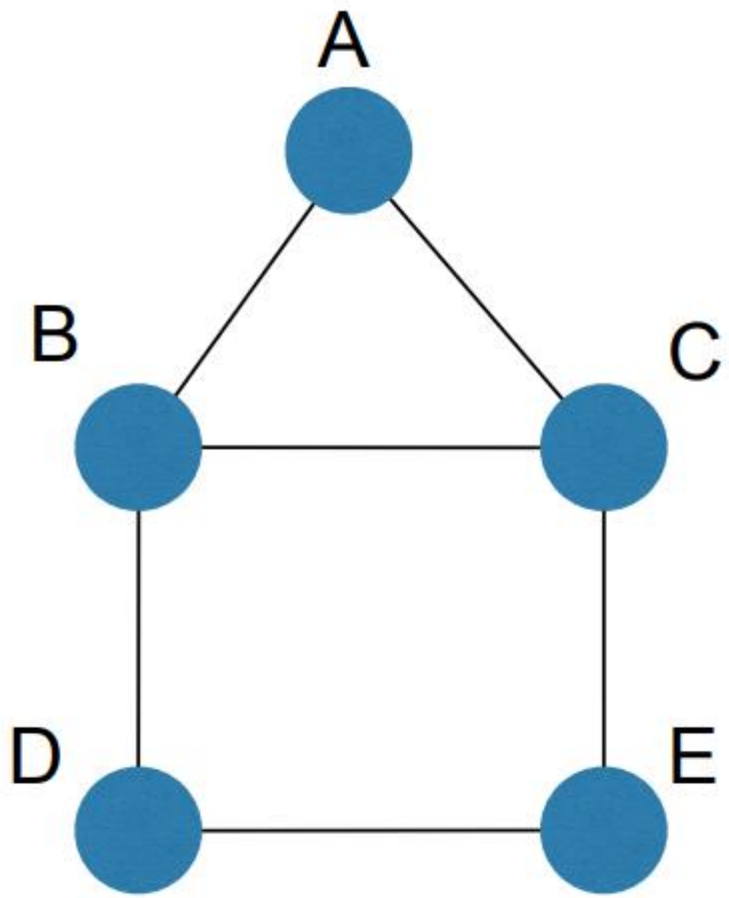
Facebook

(a)

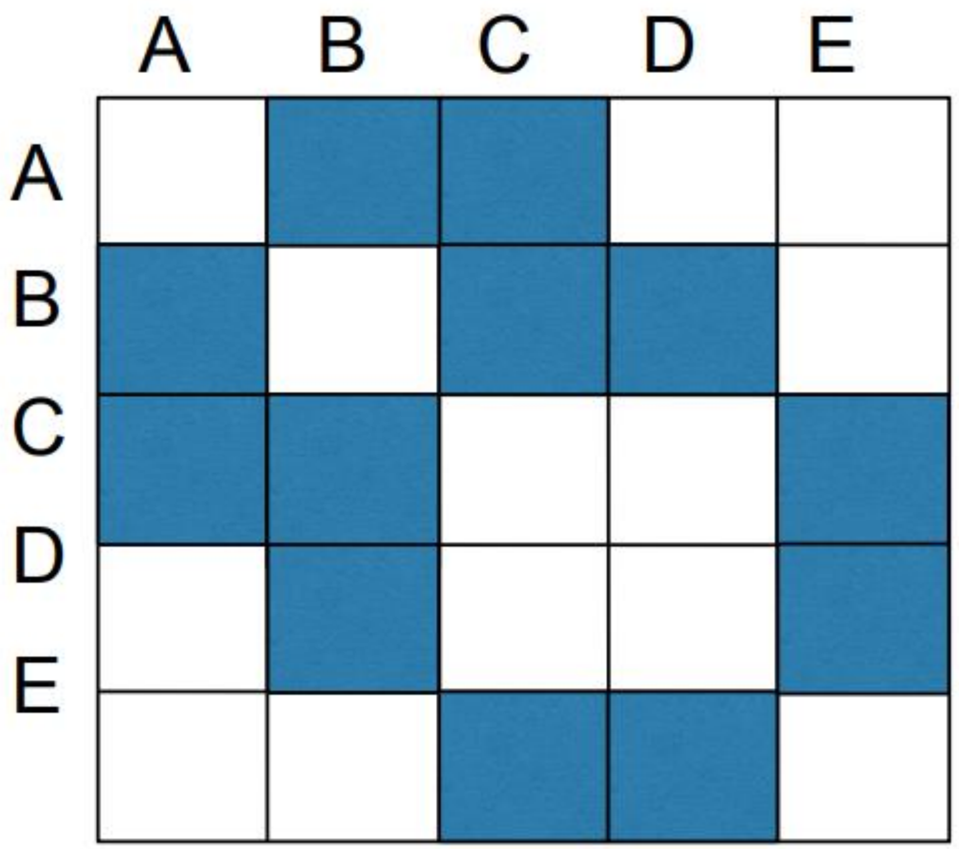
(c)

Adjacency matrix representations

- derive adjacency matrix from network

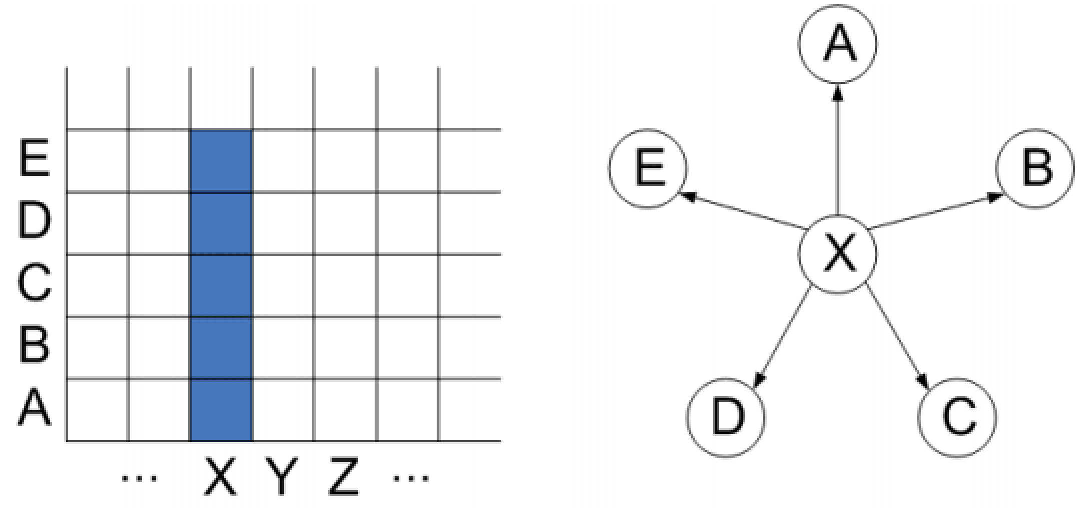


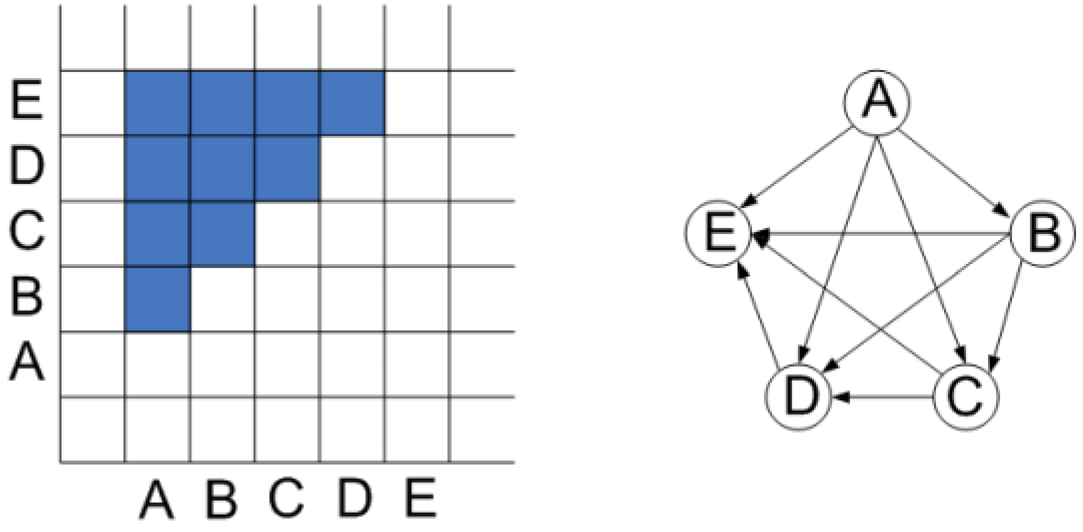
Node link

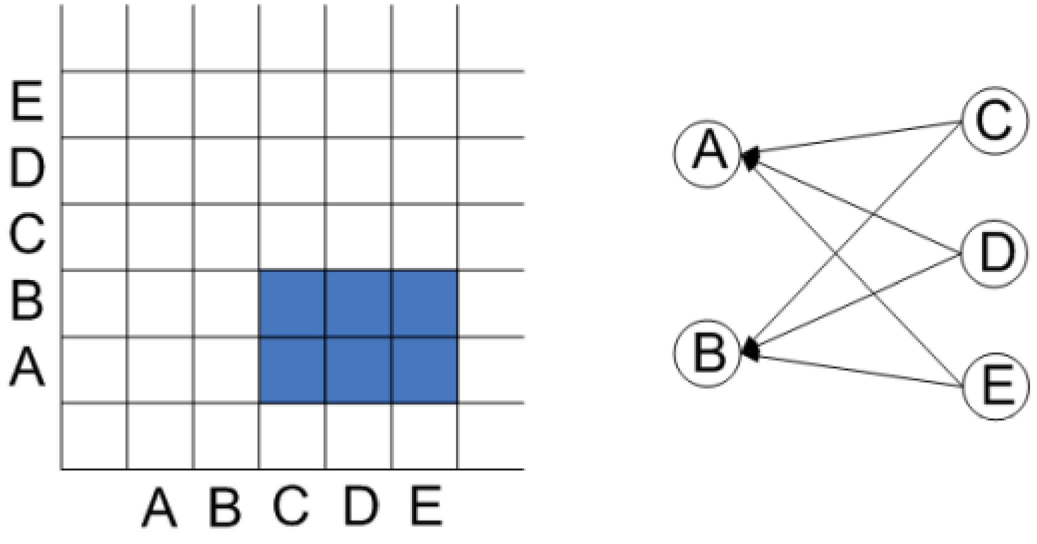


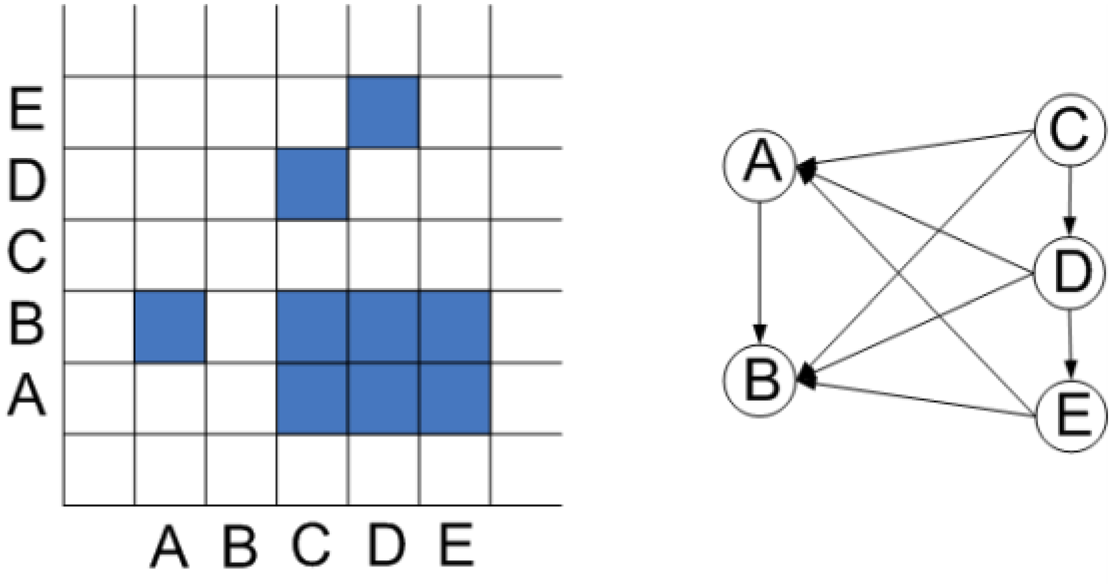
adjacency matrix

Adjacency matrix example

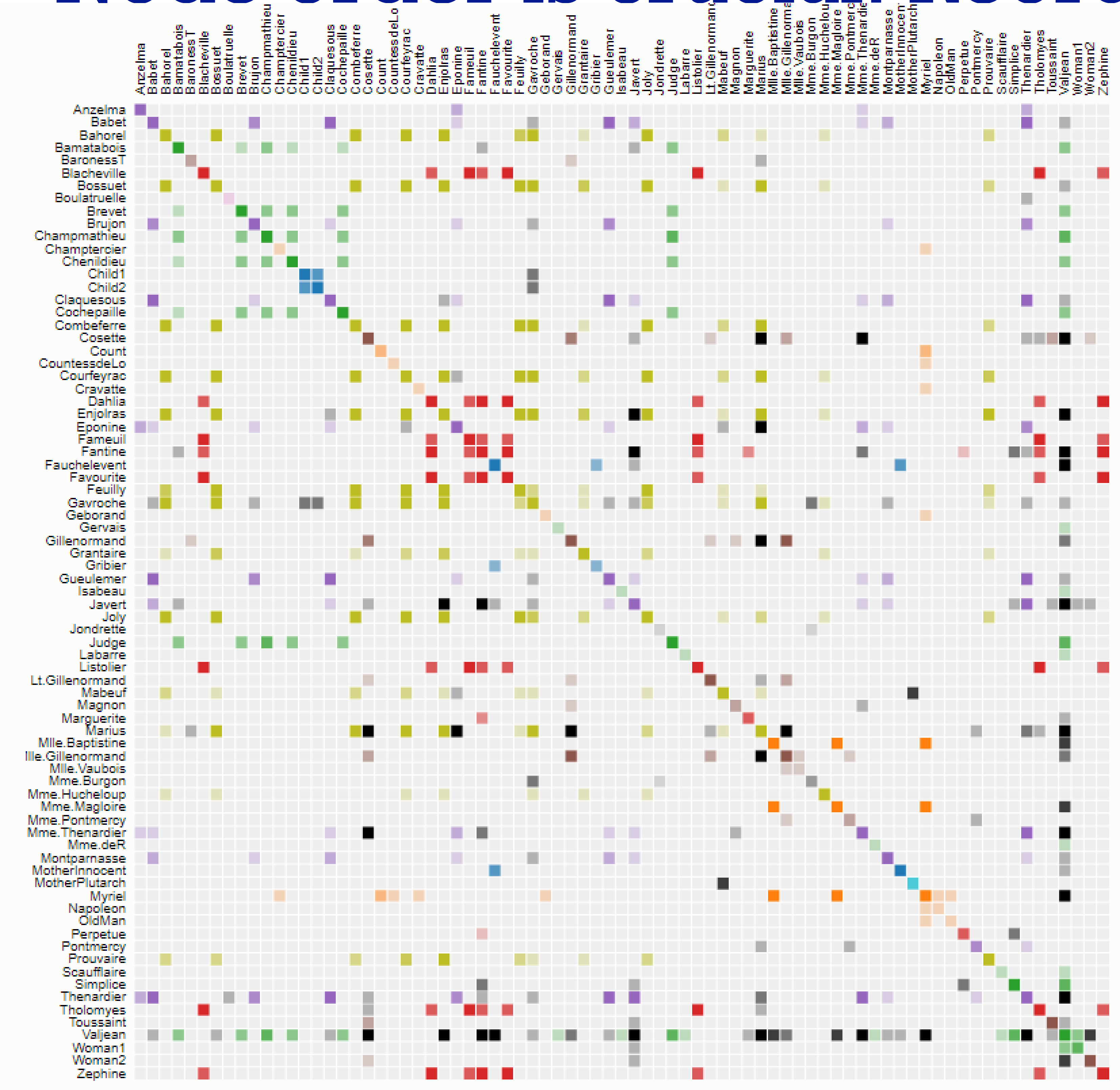


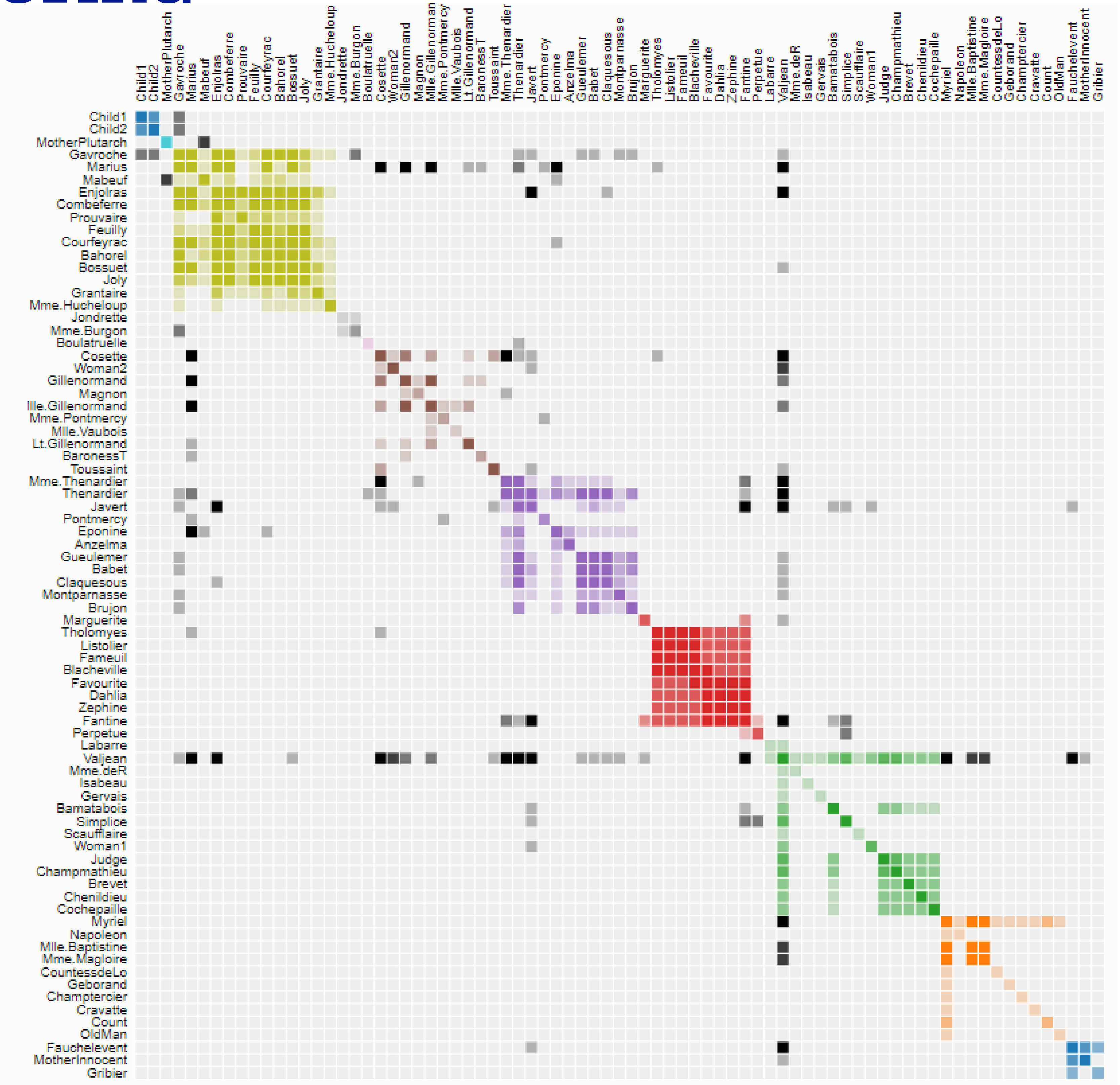






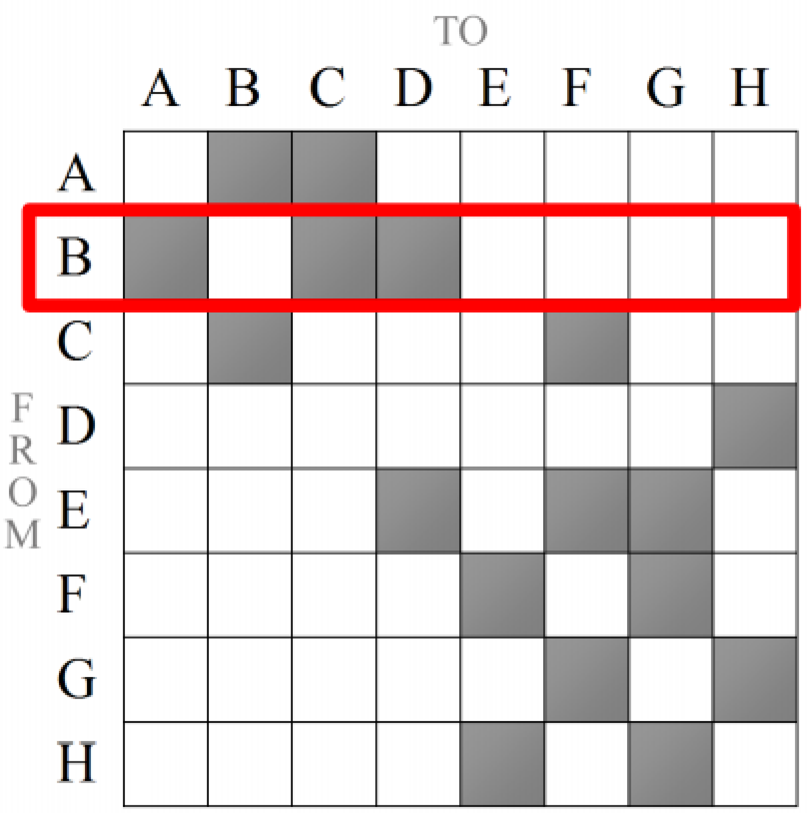
Node order is crucial: Reordering



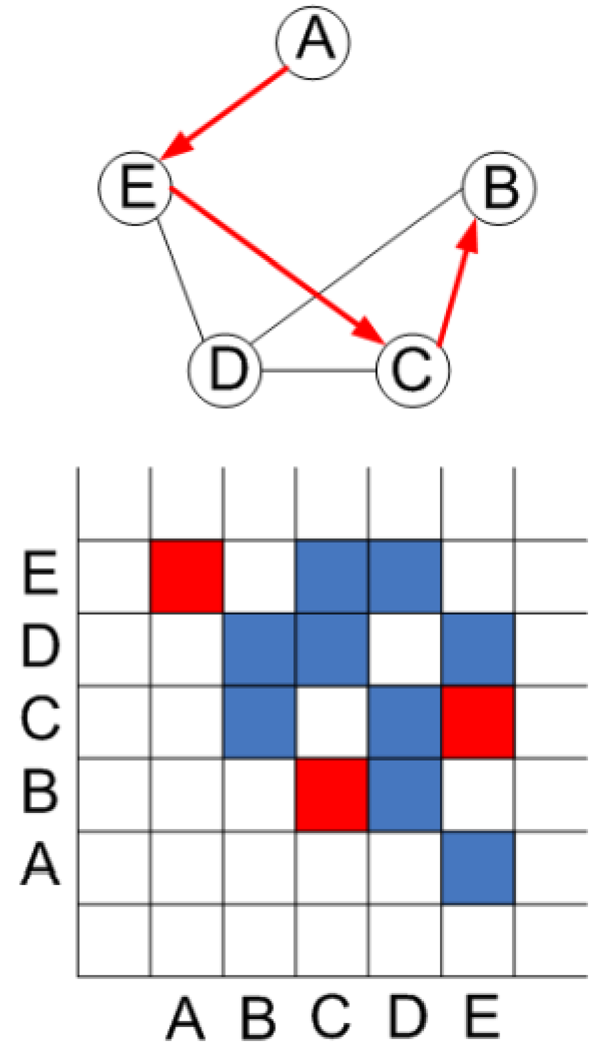


[*Les Misérables* Co-occurrence]

Adjacency matrix



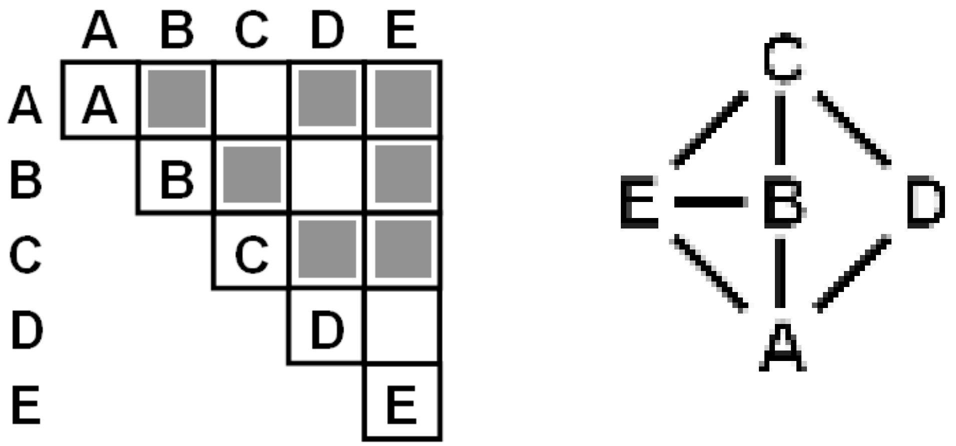
good for topology tasks related to neighborhoods (node 1-hop neighbors)



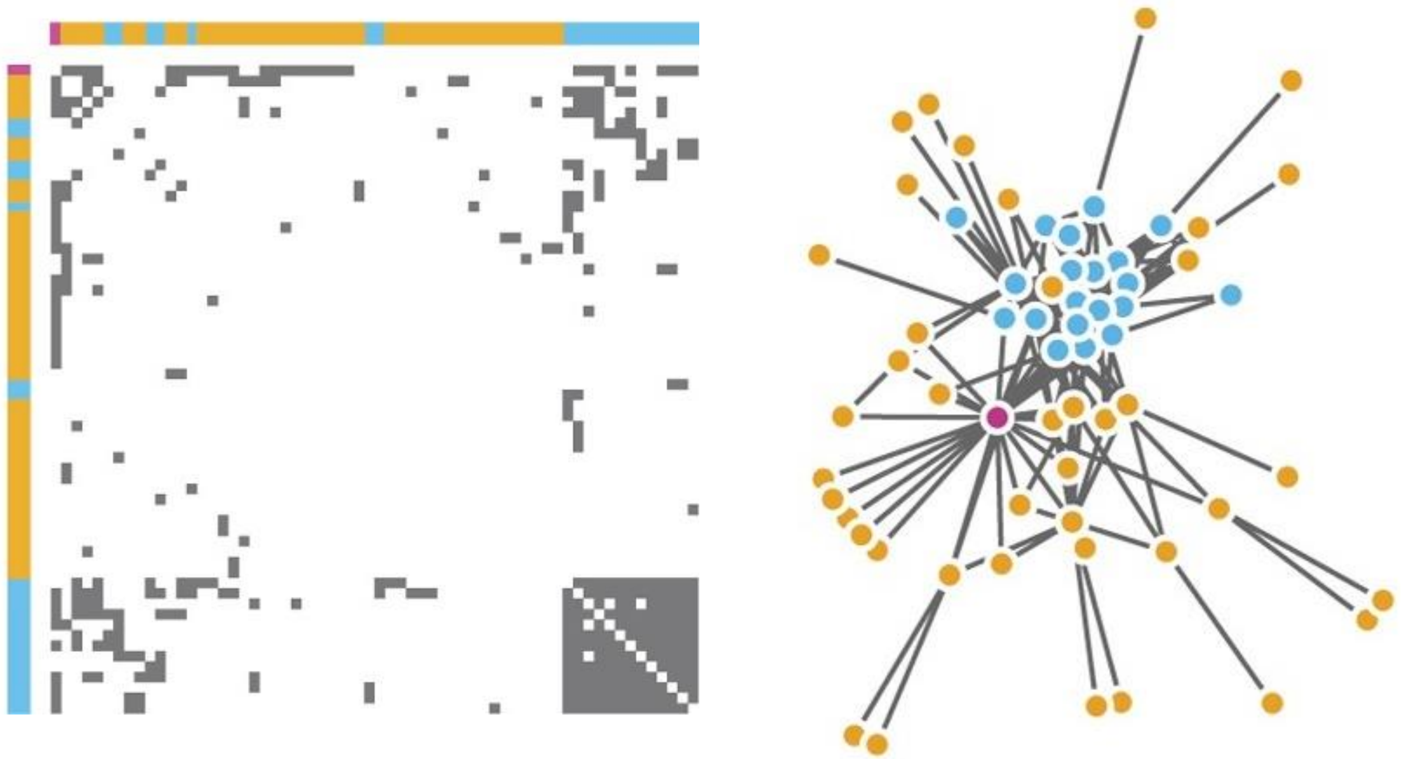
Bad for topology task related to paths

Idiom: **adjacency matrix view**

- data: network
  - transform into same data/encoding as heatmap
- derived data: table from network
  - 1 quant attrib
    - weighted edge between nodes
  - 2 categ attribs: node list x 2
- visual encoding
  - cell shows presence/absence of edge
- scalability
  - 1K nodes, 1M edges



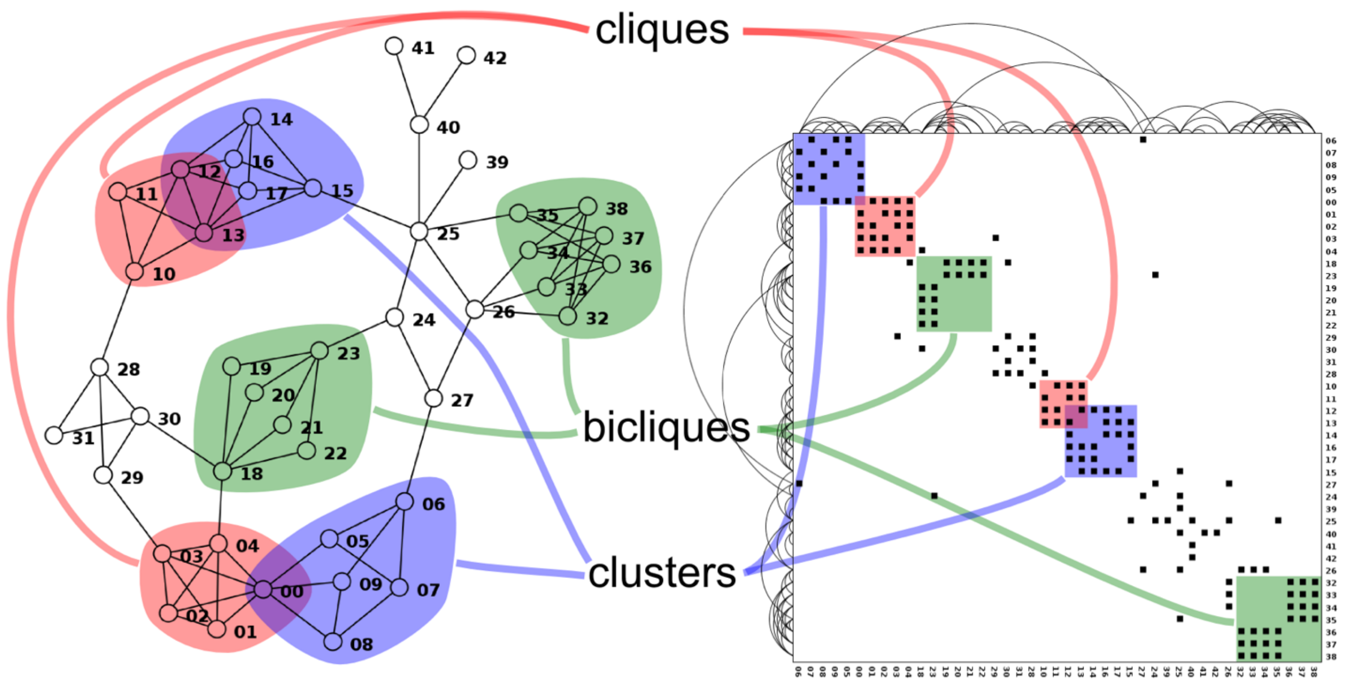
[NodeTrix: a Hybrid Visualization of Social Networks. Henry, Fekete, and McGuffin. IEEE TVCG (Proc. InfoVis) 13(6):1302-1309, 2007.]



[Points of view: Networks. Gehlenborg and Wong. Nature Methods 9:115.]

Connection vs. adjacency comparison

- adjacency matrix strengths
  - predictability, scalability, supports reordering
  - some topology tasks trainable
- node-link diagram strengths
  - topology understanding, path tracing
  - intuitive, no training needed
- empirical study
  - node-link best for small networks
  - matrix best for large networks
    - if tasks don't involve topological structure!

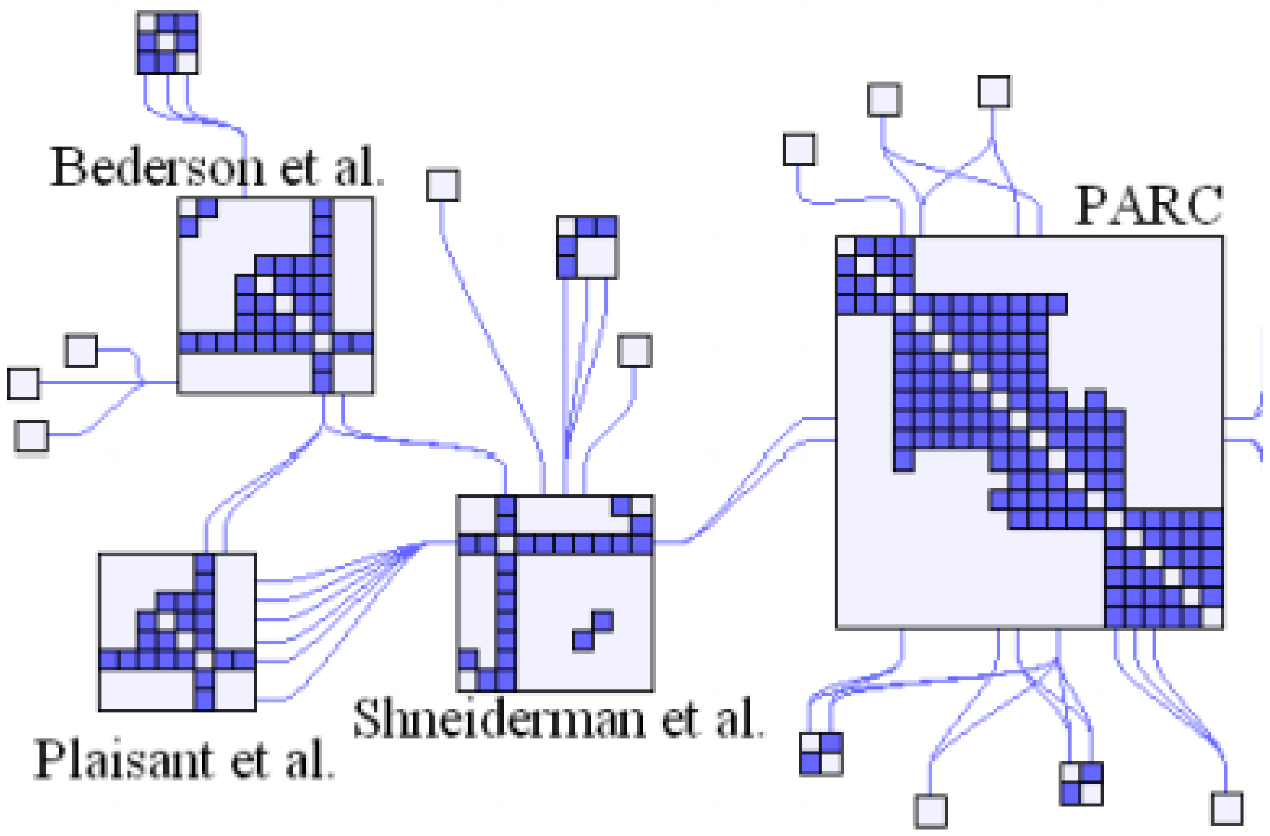


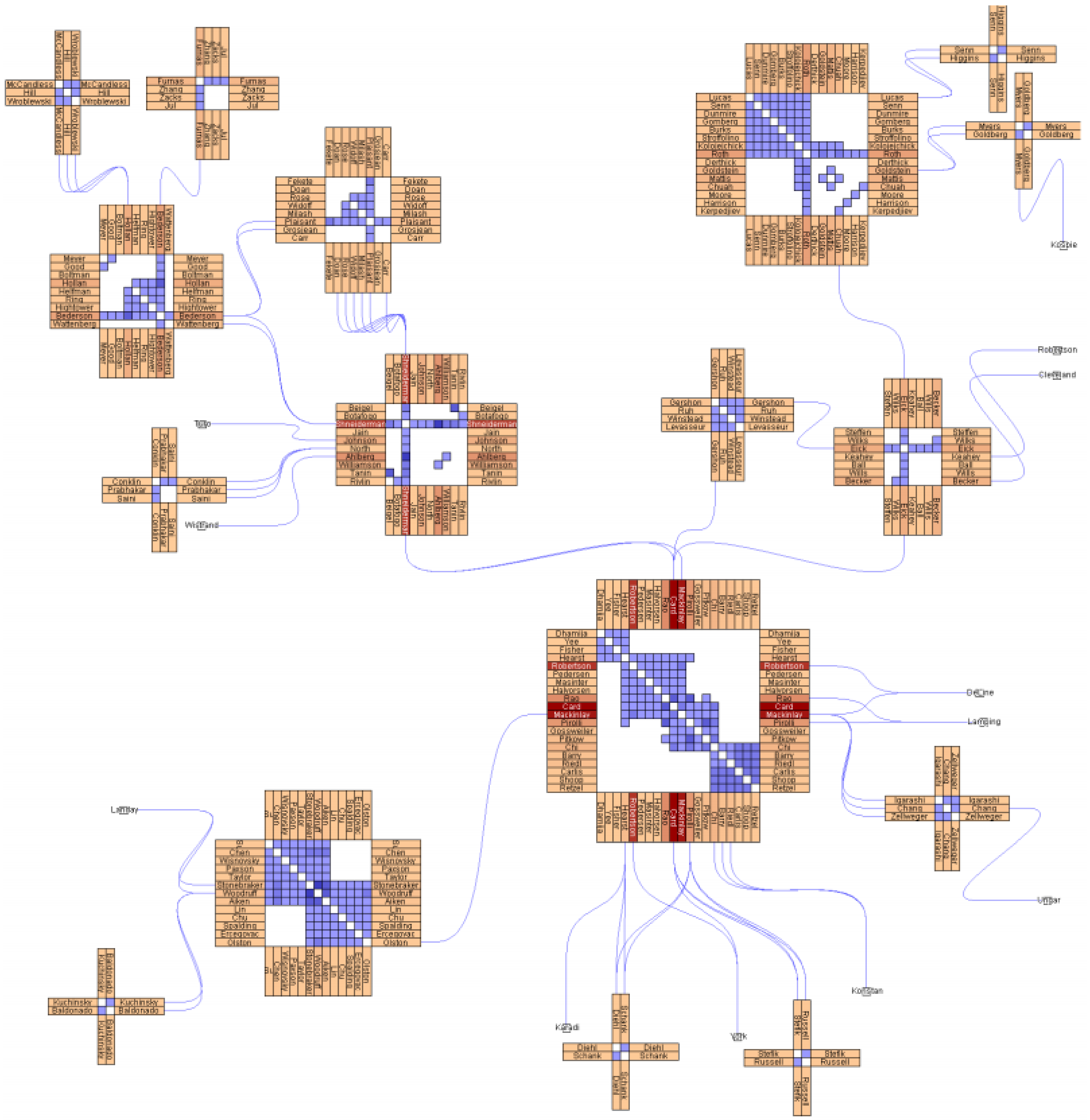
http://www.michaelmcguffin.com/courses/vis/patternsInAdjacencyMatrix.png

[On the readability of graphs using node-link and matrix-based representations: a controlled experiment and statistical analysis. Ghoniem, Fekete, and Castagliola. Information Visualization 4:2 (2005), 114–135.]

Idiom: **NodeTrix**

- hybrid nodelink/matrix
- capture strengths of both



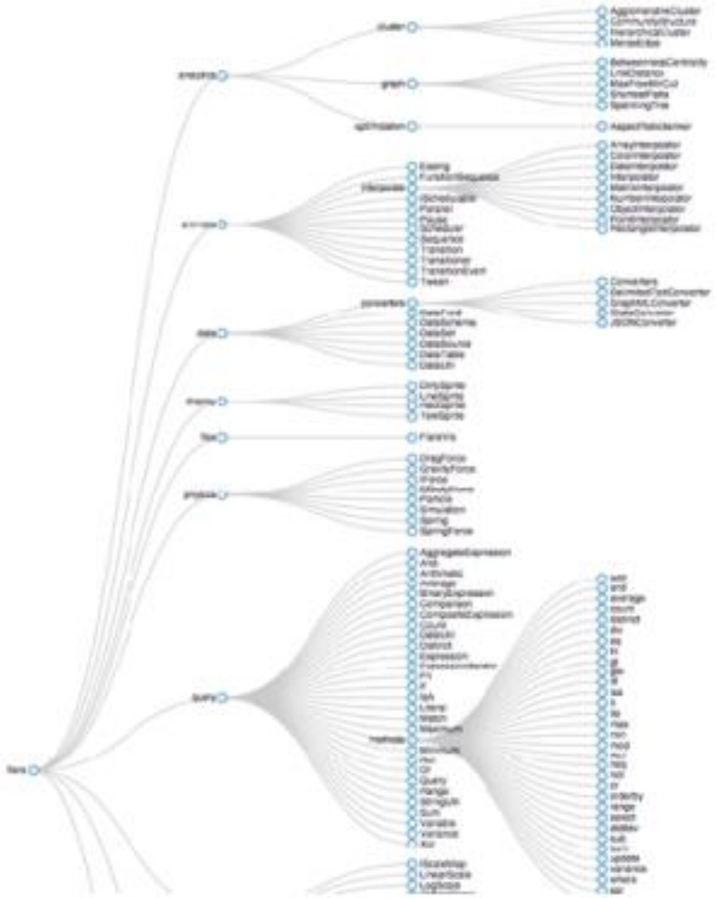


Tree

Idiom: **radial node-link tree**

- Reingold-Tilford
  - tidy drawings of trees
    - exploit parent/child structure
  - allocate space: compact w/o overlap
    - rectilinear and radial variants
  - nice algorithm writeup

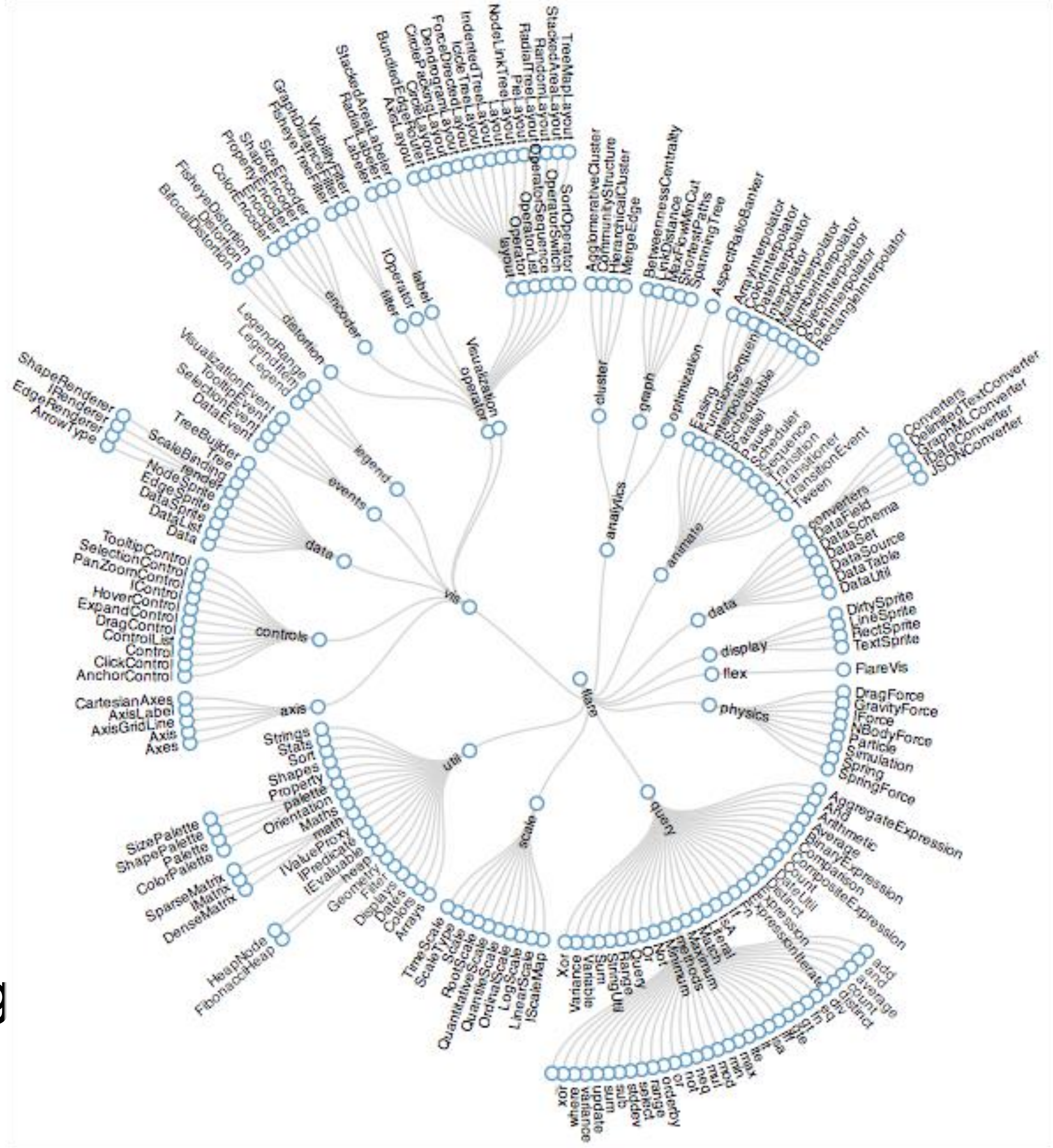
http://billmill.org/pymag-trees/



http://mbostock.github.com/d3/ex/tree.html

Idiom: **radial node-link tree**

- data
  - –tree
- encoding
  - –link connection marks
  - –point node marks
  - –radial axis orientation
    - angular proximity: siblings
    - distance from center: depth in tree
- tasks
  - –understanding topology, following
- scalability
  - –1K - 10K nodes



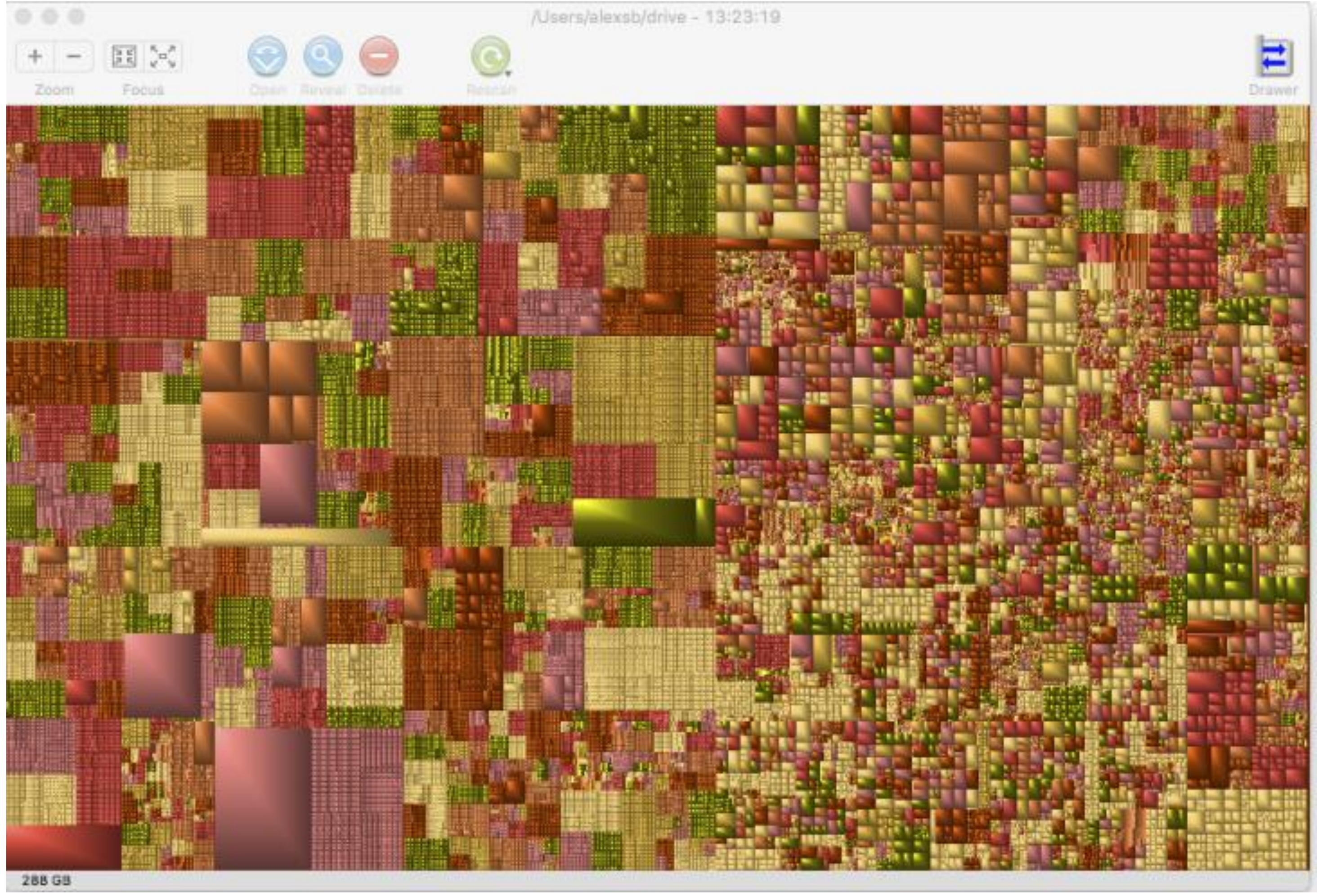
http://mbostock.github.io/d3/ex/tree.html

Idiom: **treemap**

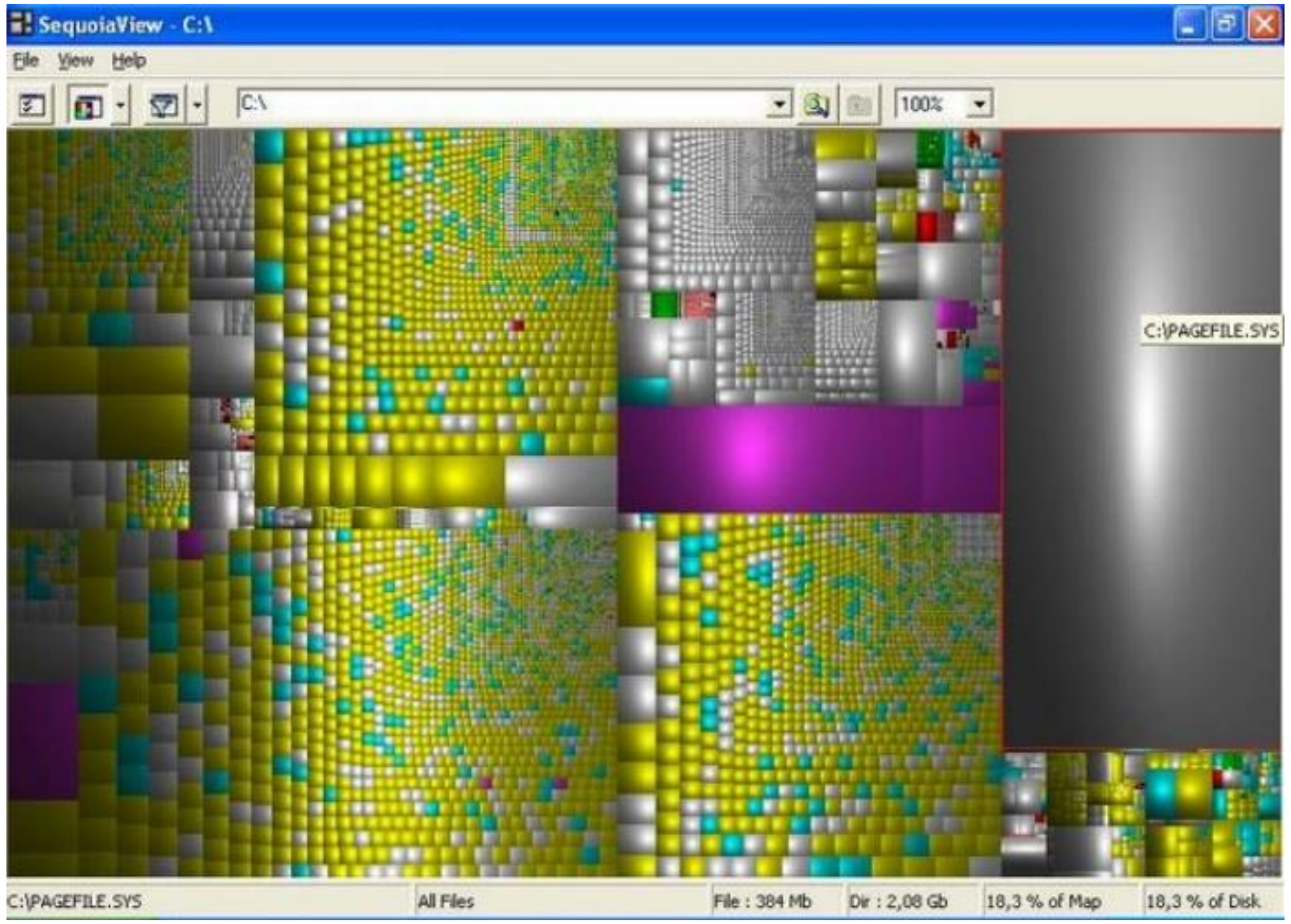
- data
  - tree
  - 1 quant attrib at leaf nodes
- encoding
  - area containment marks for hierarchical stru
  - rectilinear orientation
  - size encodes quant attrib
- tasks
  - query attribute at leaf nodes
- scalability
  - 1M leaf nodes

http://tulip.labri.fr/Documentation/3_7/userHandbook/html/ch06.ht

Treemap software: disk space

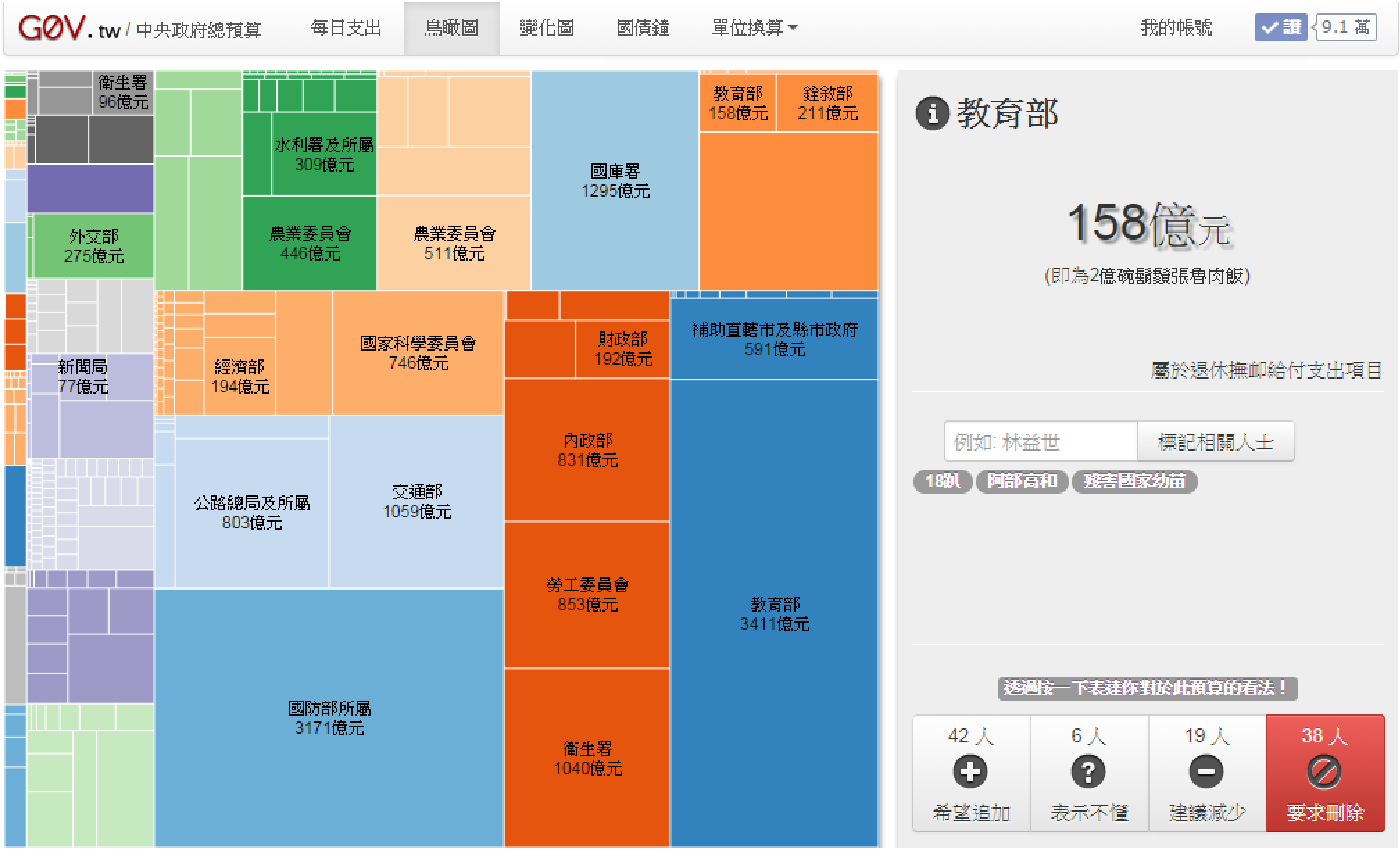


Mac: GrandPerspective



Windows: Sequoia View

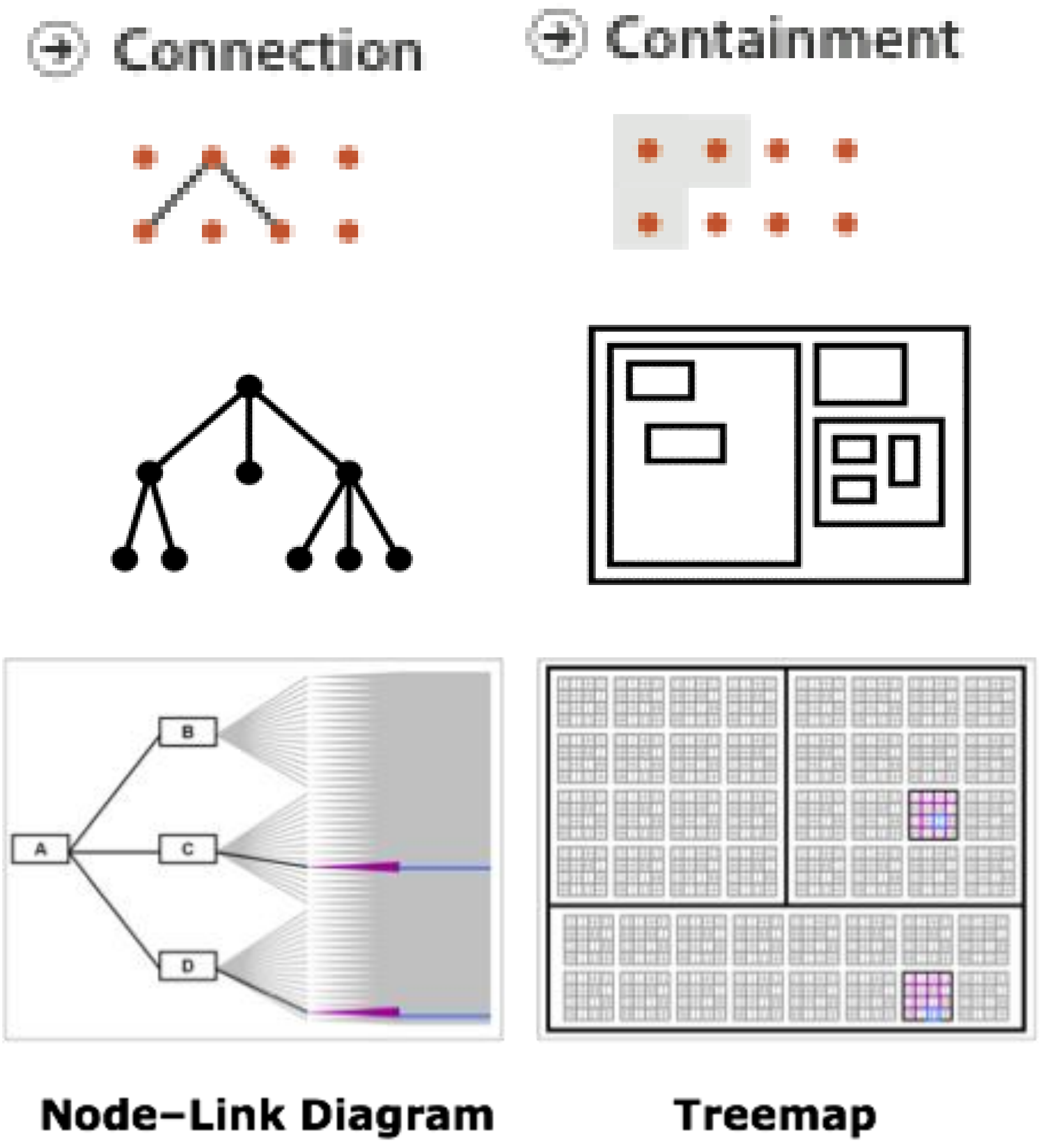
g0v 中央政府總預算



G0V.tw / 中央政府總預算
每日支出
鳥瞰圖
變化圖
國債鐘
單位換算
我的帳號
讚
9.1 萬
衛生署 96億元
水利署及所屬 309億元
外交部 275億元
農業委員會 446億元
農業委員會 511億元
國庫署 1295億元
教育部 158億元
銓敘部 211億元
新聞局 77億元
經濟部 194億元
國家科學委員會 746億元
財政部 192億元
補助直轄市及縣市政府 591億元
內政部 831億元
公路總局及所屬 803億元
交通部 1059億元
勞工委員會 853億元
教育部 3411億元
國防部所屬 3171億元
衛生署 1040億元
教育部
158億元
(即為2億碗鬍鬚張魯肉飯)
屬於退休撫卹給付支出項目
例如: 林益世
標記相關人士
18趴
阿部高和
殘害國家幼苗
透過按一下表達你對於此預算的看法！
42 人
希望追加
6 人
表示不懂
19 人
建議減少
38 人
要求刪除

Link marks: Connection and containment

- marks as links (vs. nodes)
  - common case in network drawing
  - 1D case: connection
    - ex: all node-link diagrams
    - emphasizes topology, path tracing
    - networks and trees
  - 2D case: containment
    - ex: all treemap variants
    - emphasizes attribute values at leaves (size coding)
    - only trees

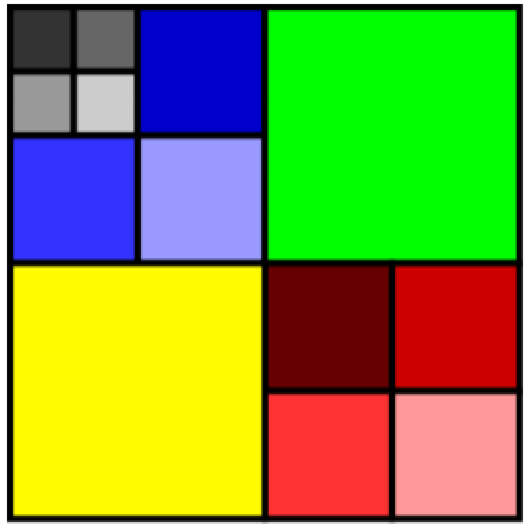


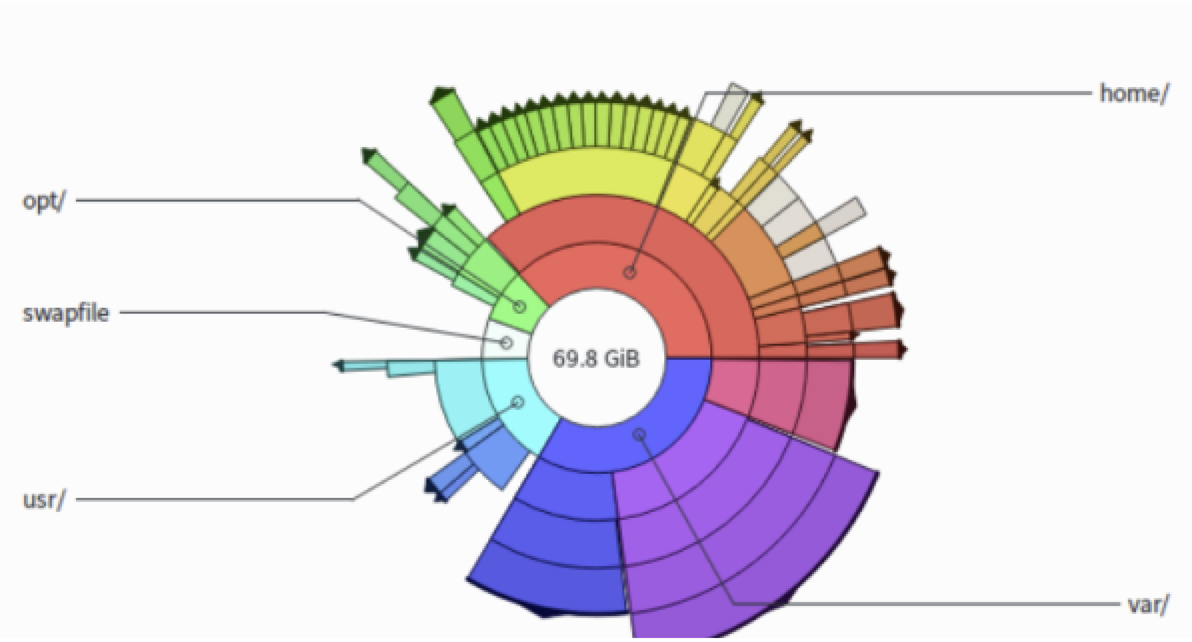
[Elastic Hierarchies: Combining Treemaps and Node-Link Diagrams. Dong, McGuffin, and Chignell. Proc. InfoVis 2005, p. 57-64.]

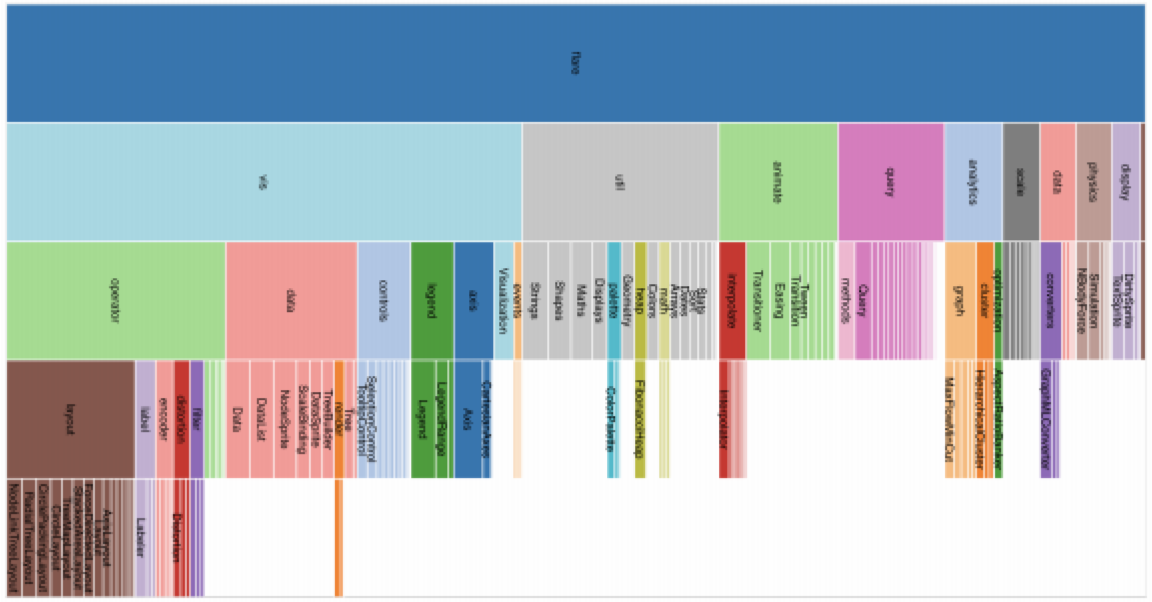
Implicit layouts

- alternative to connection and containment: position
  - show parent-child relationships only through relative positions

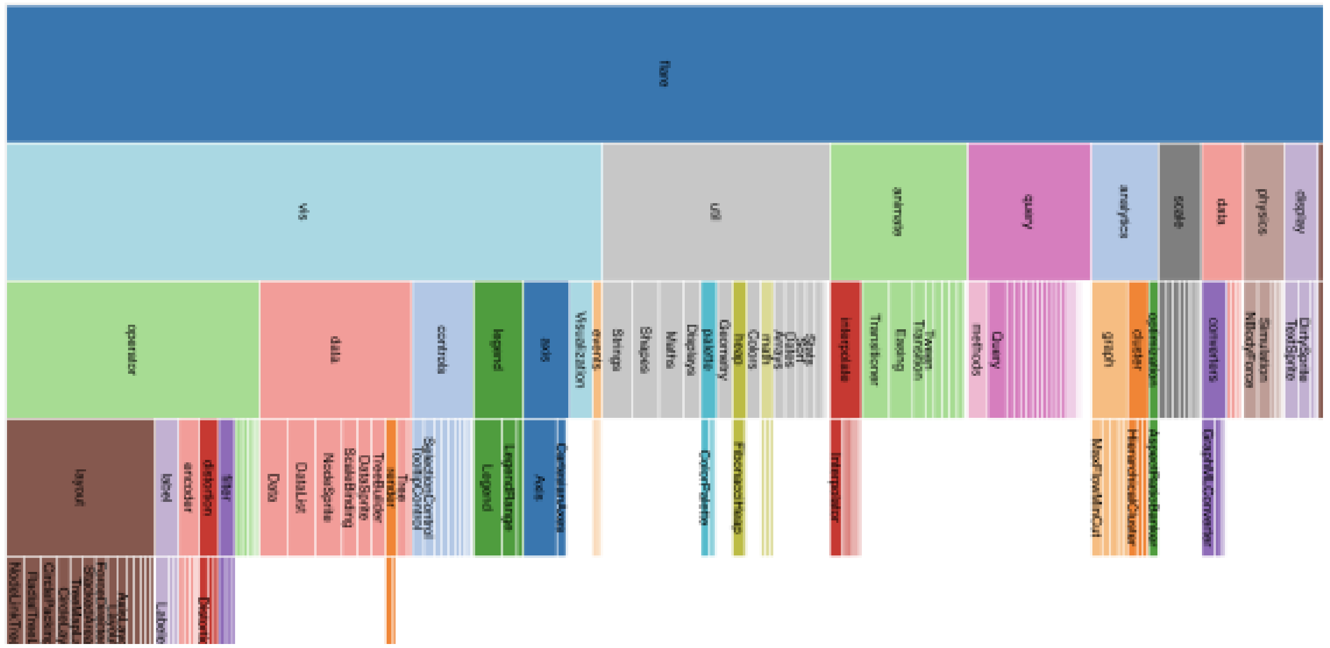
Treemap	Sunburst	Icicle Plot
containment	position	position

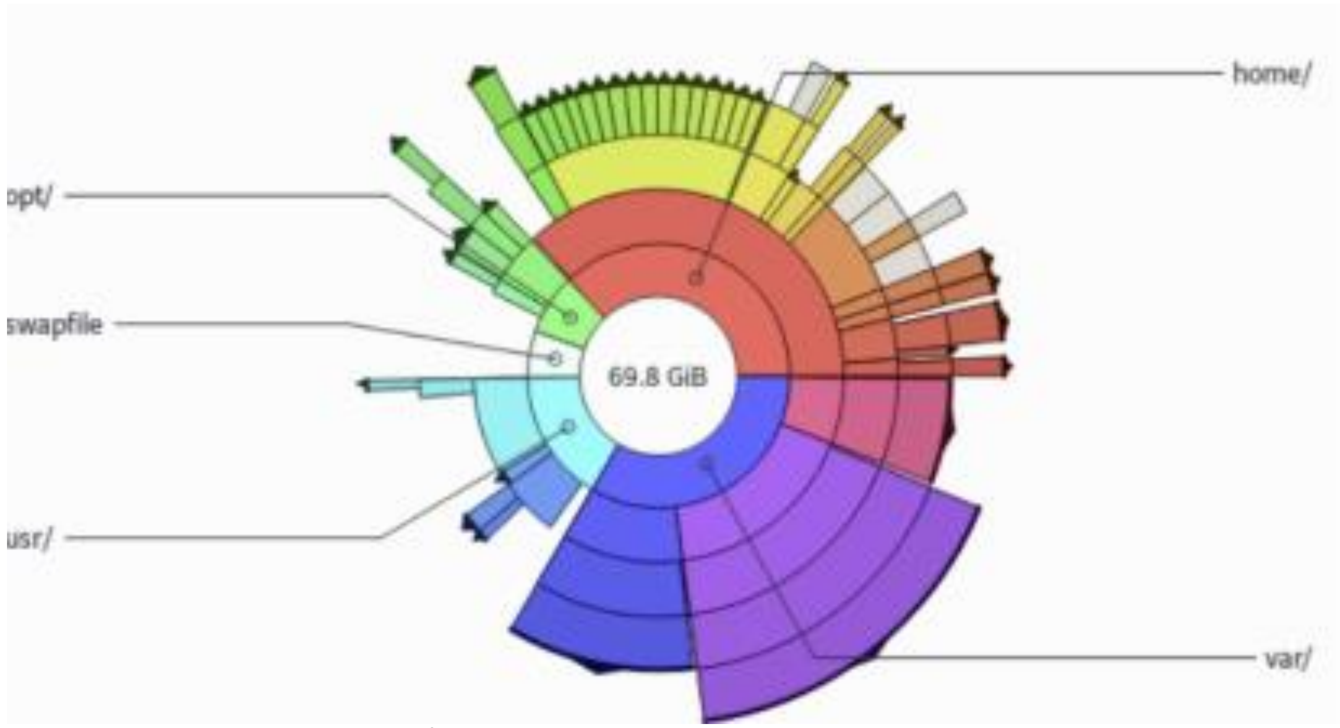




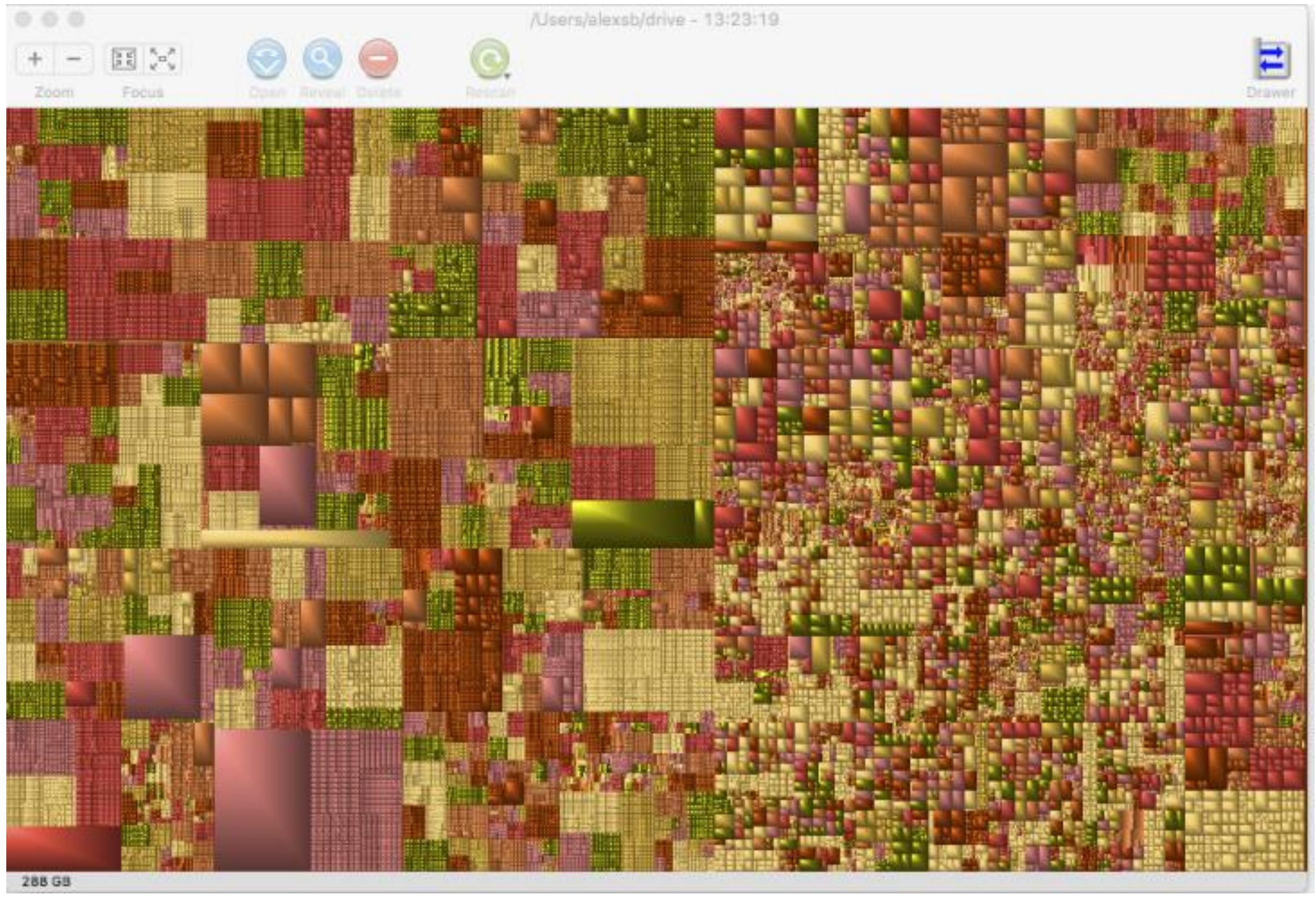


Implicit: Approaches compared





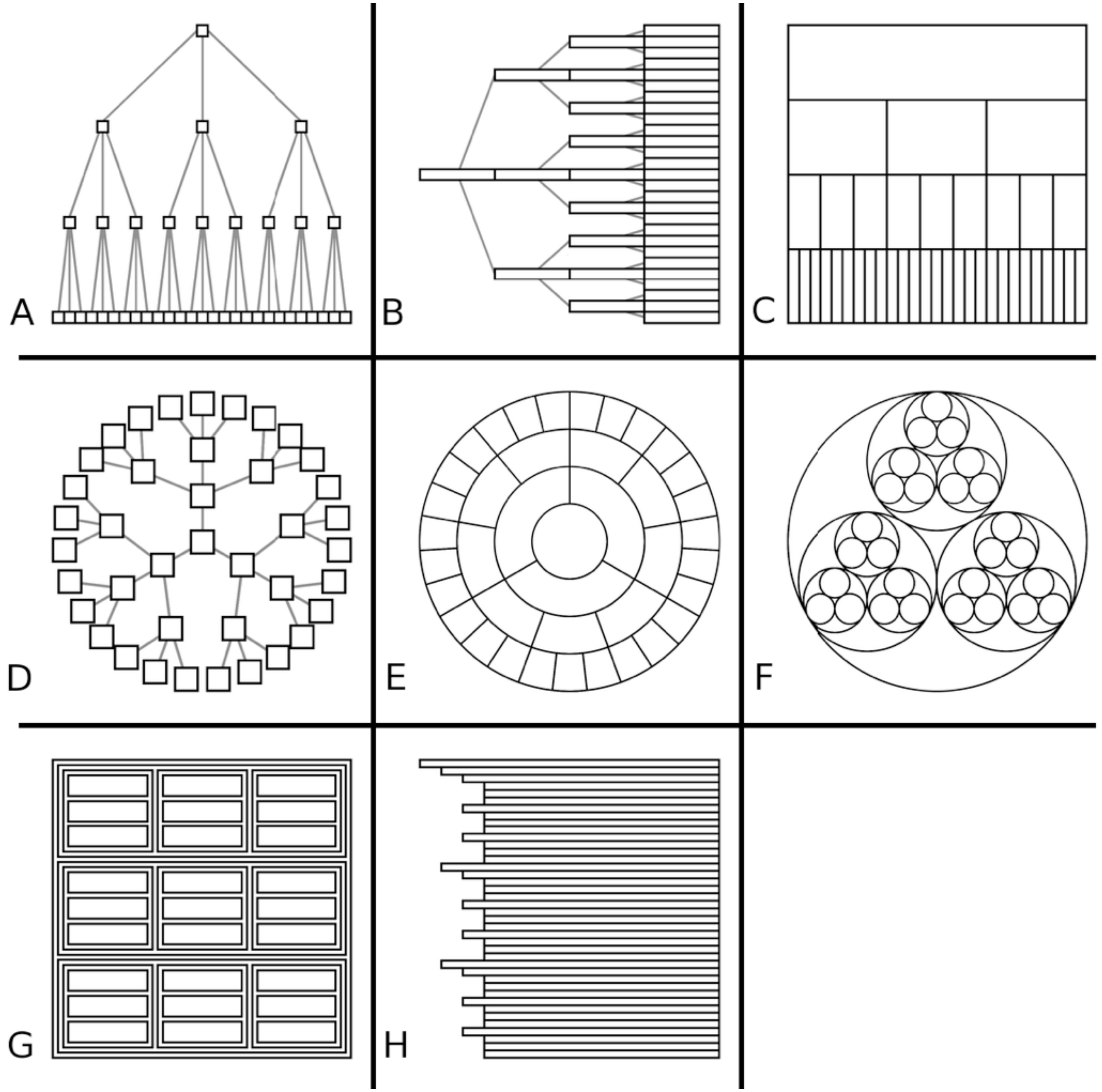
Both inner nodes and leaves visible



Only leaves visible

Tree drawing idioms comparison

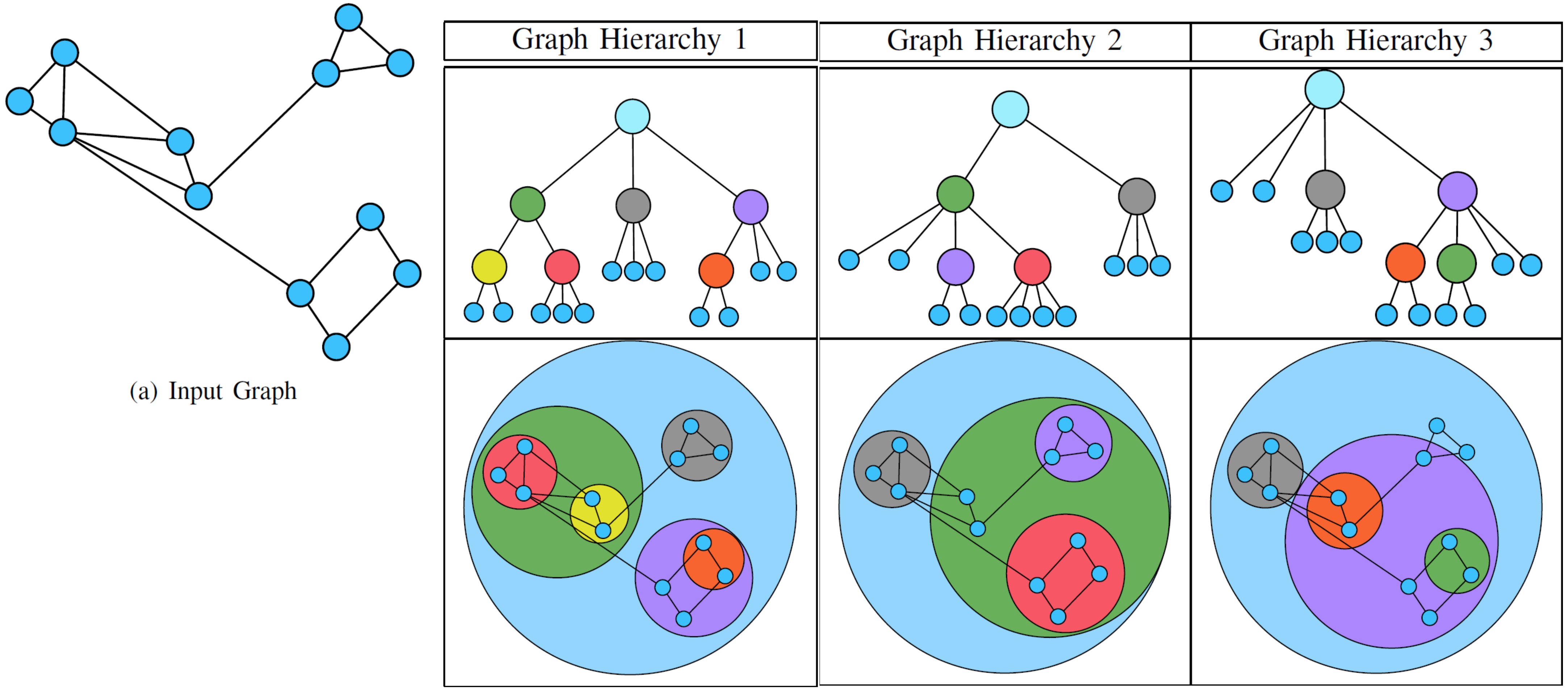
- data shown
  - link relationships
  - tree depth
  - sibling order
- design choices
  - connection vs containment link marks
  - rectilinear vs radial layout
  - spatial position channels
- considerations
  - redundant? arbitrary?
  - information density?
    - avoid wasting space



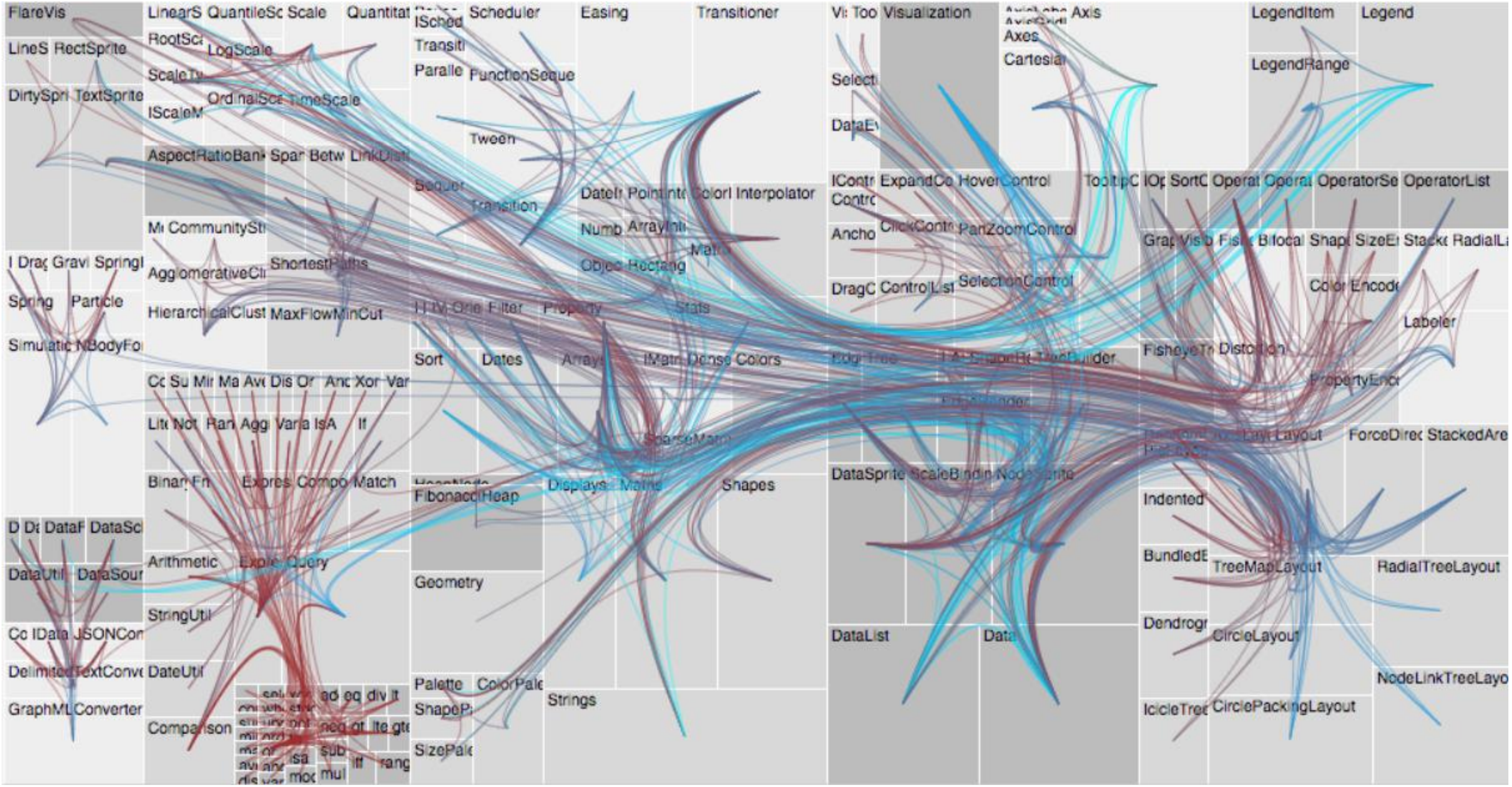
[Quantifying the Space-Efficiency of 2D Graphical Representations of Trees. McGuffin and Robert. Information Visualization 9:2 (2010), 115–140.]

GrouseFlocks: Steerable Exploration of Graph Hierarchy Space[2008]

- A Network + A tree
- Interactively explore and modify graph hierarchies.



Hierarchical edge bundling: treemap vs radial



Hierarchies in D3


Treemap

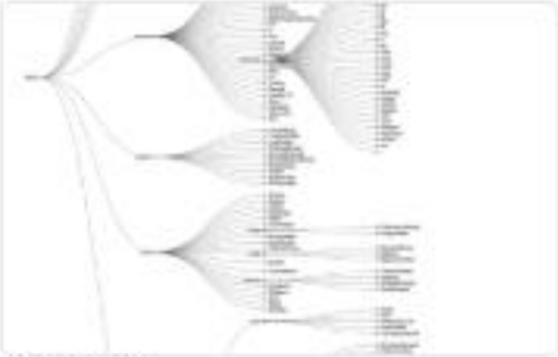

Cascaded treemap


Nested treemap

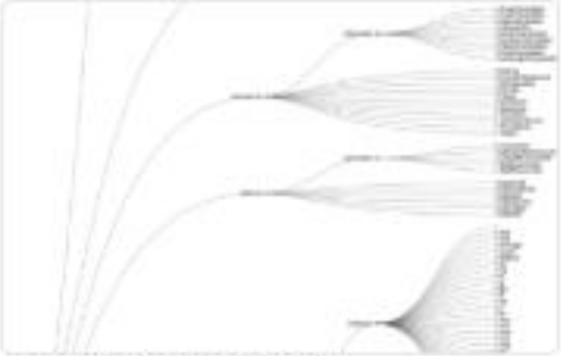

Circle packing


Circle packing (monochrome)


Indented tree


Tidy tree


Radial tidy tree


Cluster dendrogram

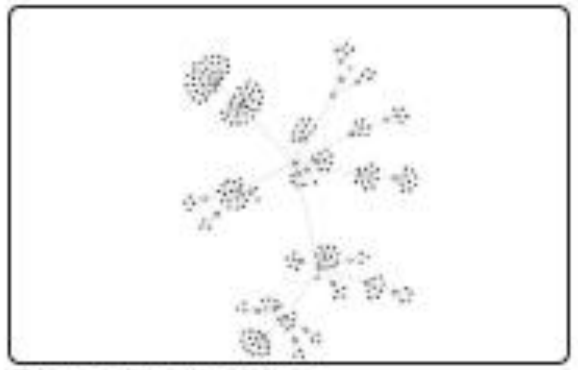

Radial dendrogram


Sunburst


Icicle

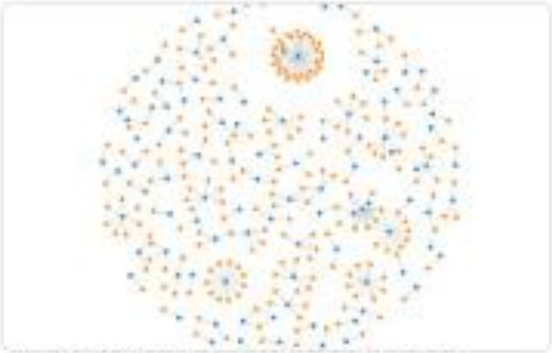

Tangled tree visualization


Phylogenetic tree


Force-directed tree

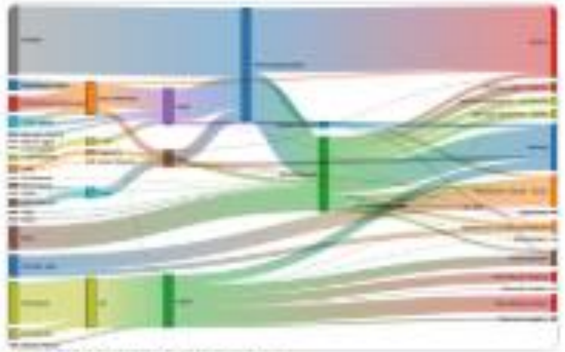
Networks in D3

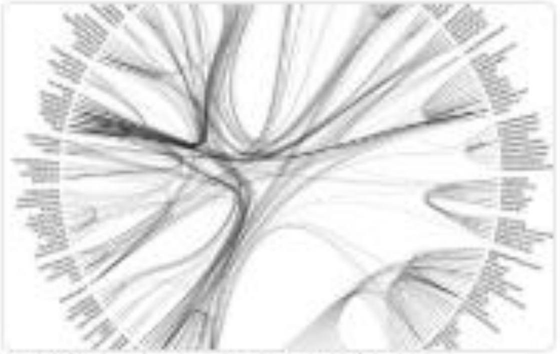

Force-directed graph


Disjoint force-directed graph


Mobile patent suits

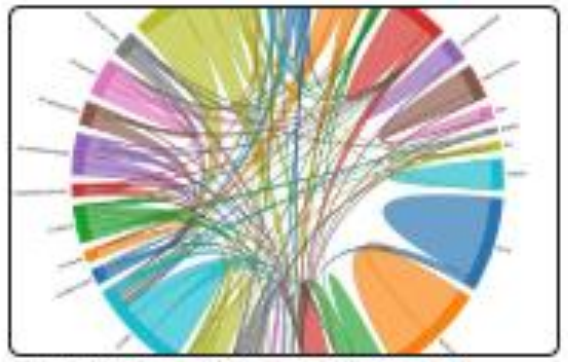

Arc diagram


Sankey diagram

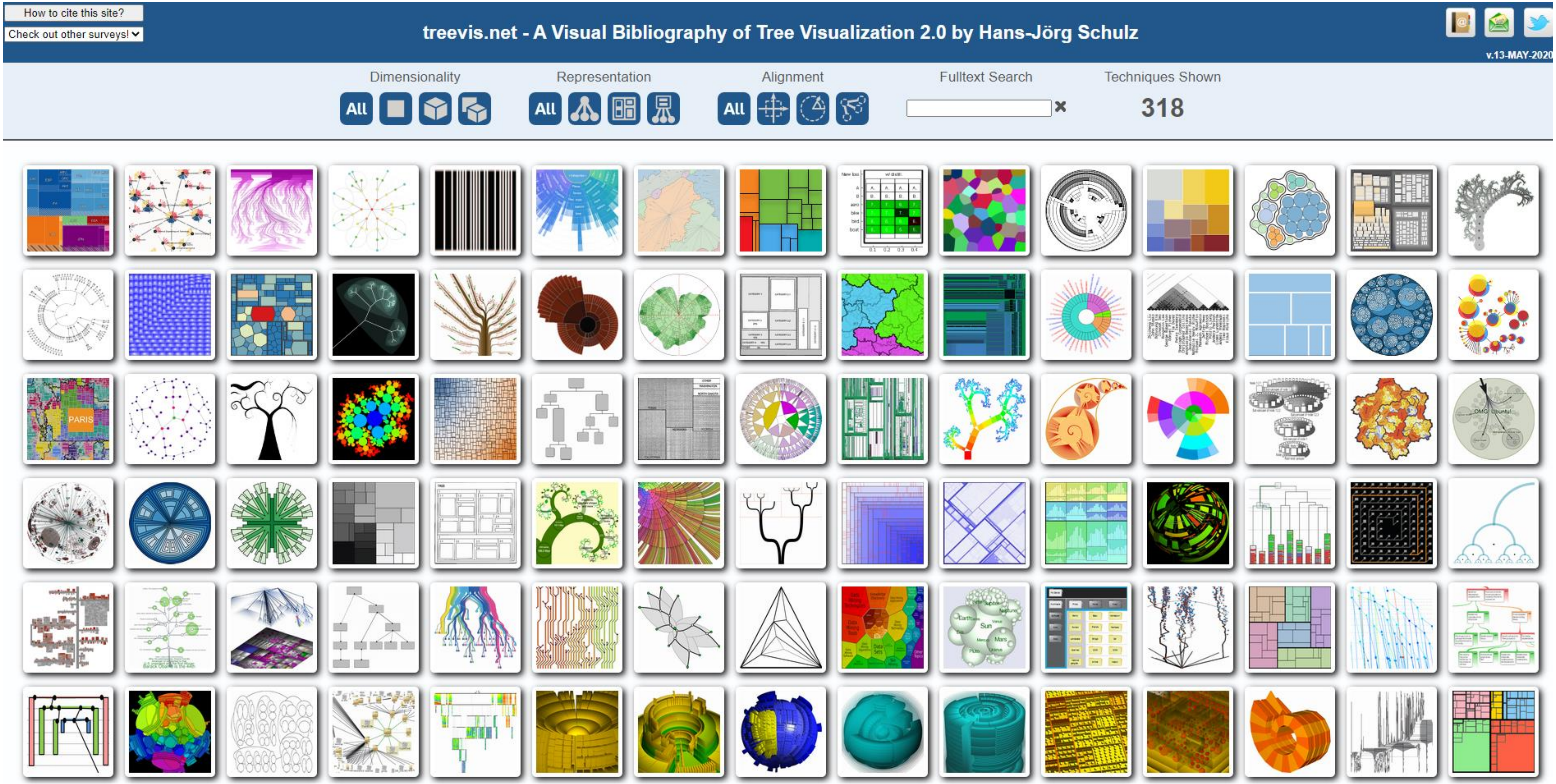

Hierarchical edge bundling


Hierarchical edge bundling


Chord diagram


Chord dependency diagram

treevis.net: many, many options



Further reading

- Visualization Analysis and Design. Munzner. AK Peters Visualization Series, CRC Press, 2014.
  - *Chap 9: Arrange Networks and Trees*
- Visual Analysis of Large Graphs: State-of-the-Art and Future Research Challenges. von Landesberger et al. Computer Graphics Forum 30:6 (2011), 1719–1749.
- Simple Algorithms for Network Visualization: A Tutorial. McGuffin. Tsinghua Science and Technology (Special Issue on Visualization and Computer Graphics) 17:4 (2012), 383–398.
- Drawing on Physical Analogies. Brandes. In Drawing Graphs: Methods and Models, LNCS Tutorial, 2025, edited by M. Kaufmann and D. Wagner, LNCS Tutorial, 2025, pp. 71–86. Springer-Verlag, 2001.
- http://www.treevis.net Treevis.net: A Tree Visualization Reference. Schulz. IEEE Computer Graphics and Applications 31:6 (2011), 11–15.
- Perceptual Guidelines for Creating Rectangular Treemaps. Kong, Heer, and Agrawala. IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis) 16:6 (2010), 990–998.