Trees & Graphs

Nathalie Henry Riche, Microsoft Research

About

Nathalie Henry Riche

nath@microsoft.com

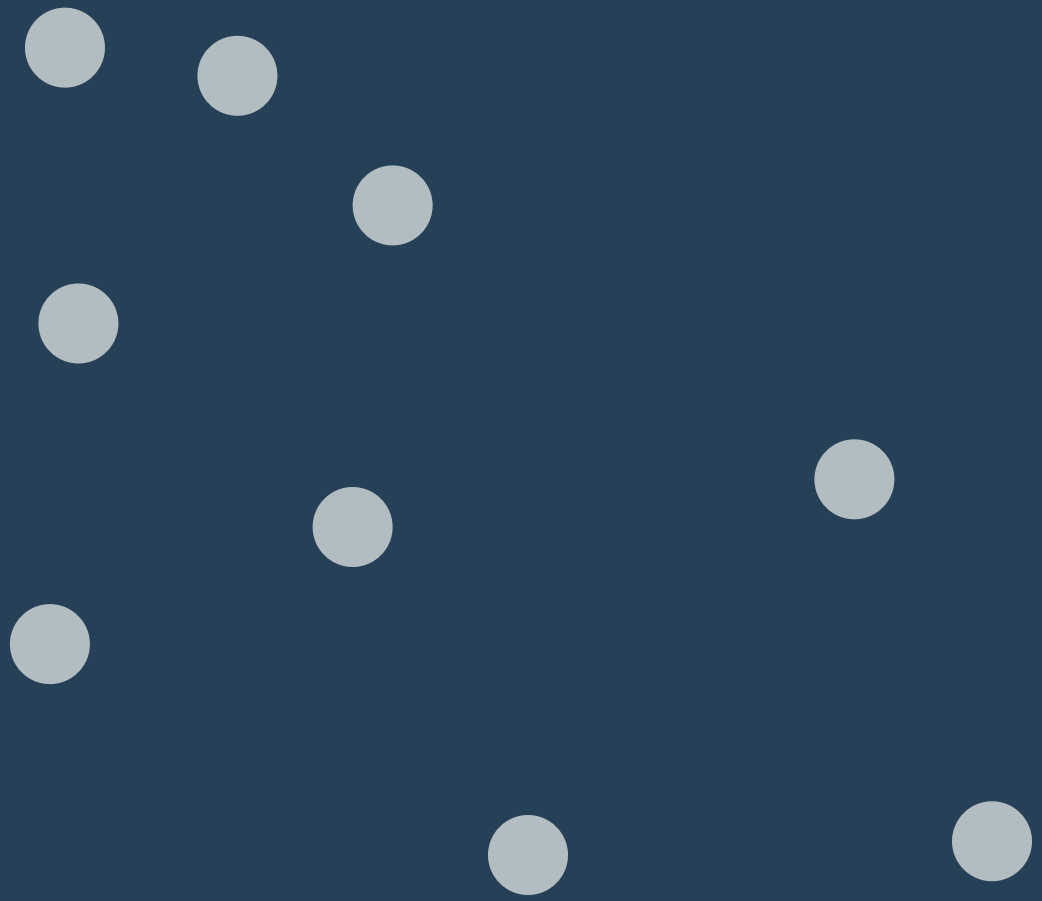
Researcher @ Microsoft Research since 2009

Today:

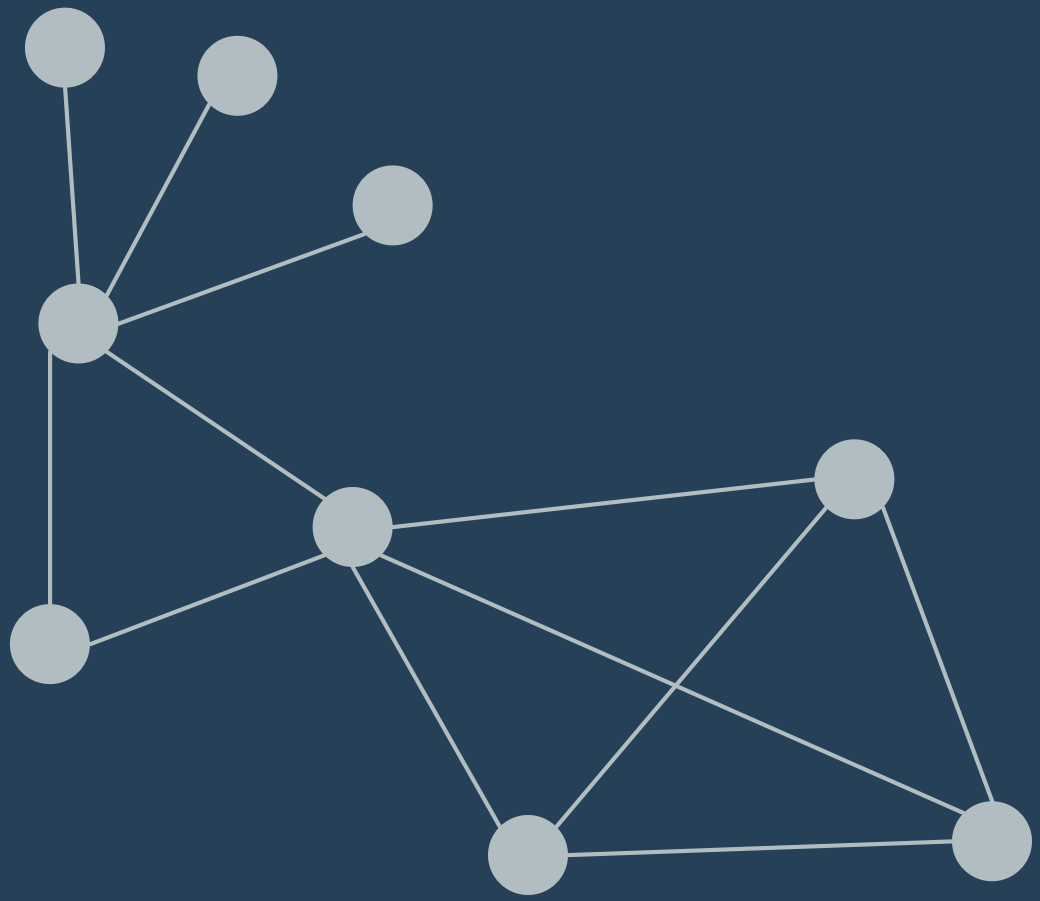
- Overview of techniques to visualize trees & graphs
- Their strengths & weaknesses
- Areas for future research

What’s in a graph?

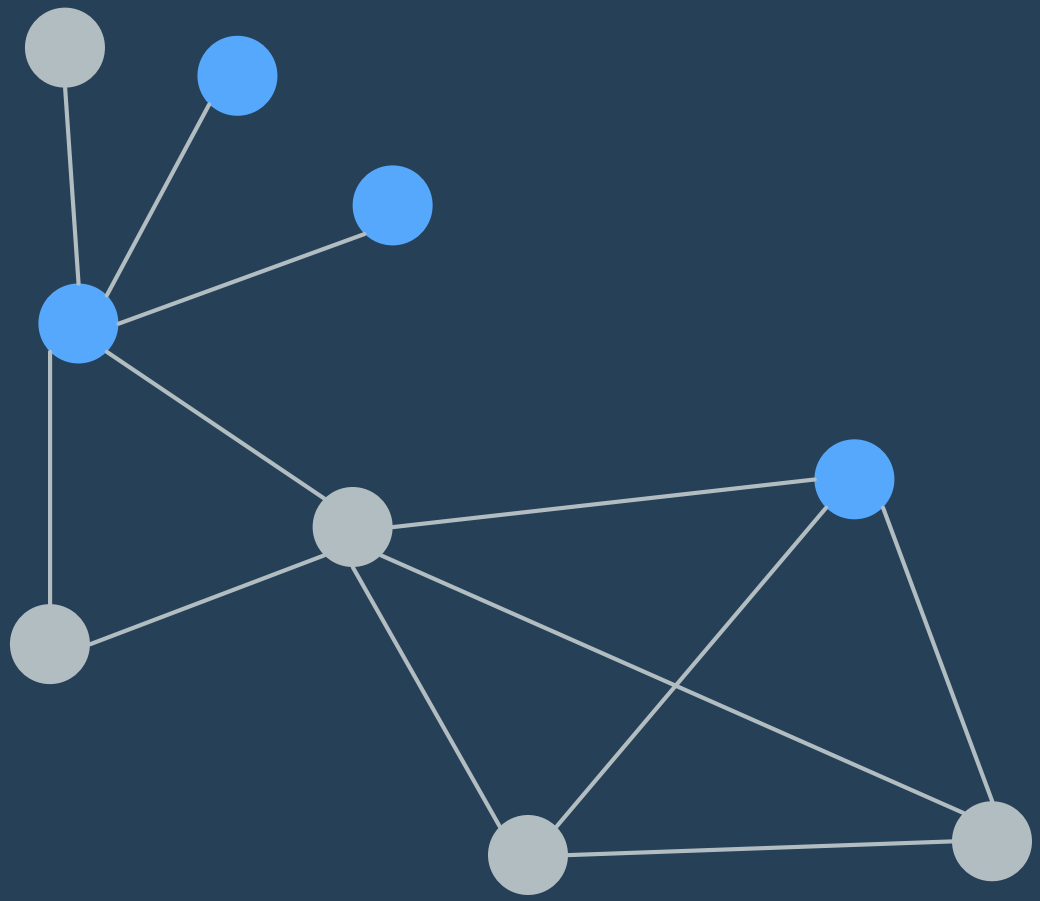
What’s in a graph?



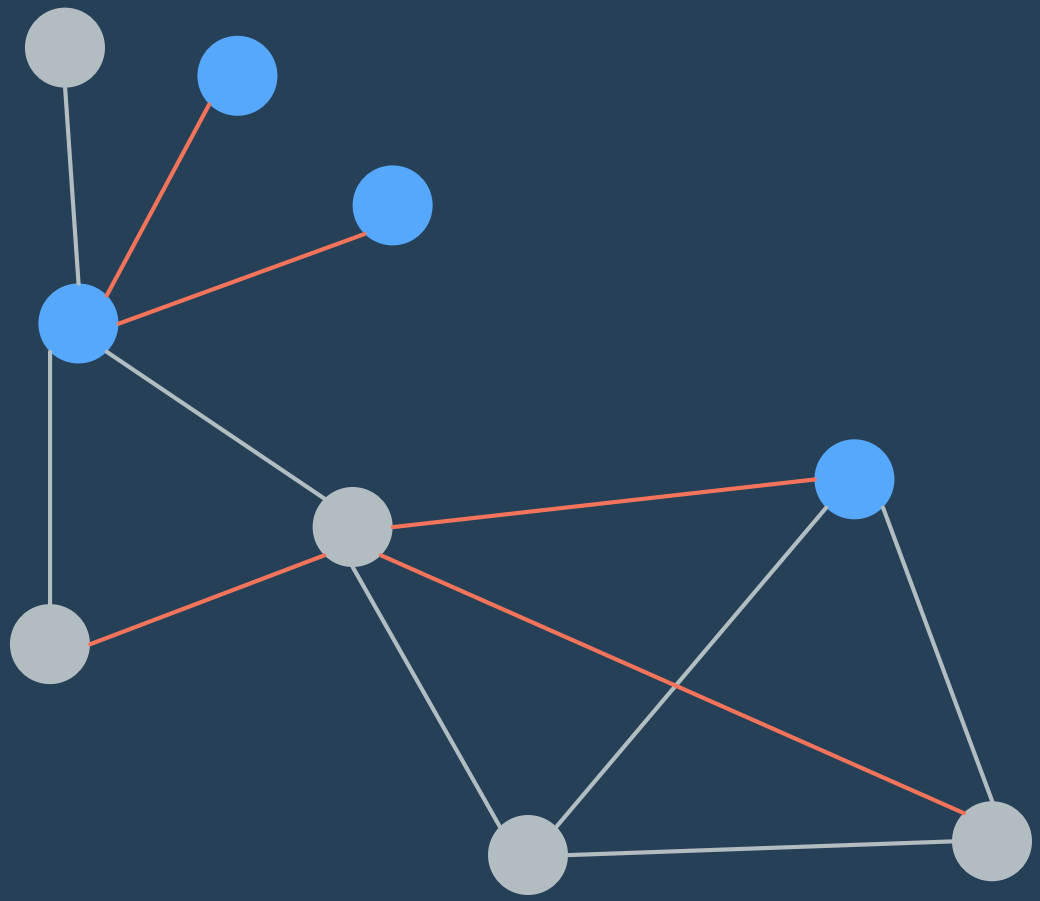
What’s in a graph?



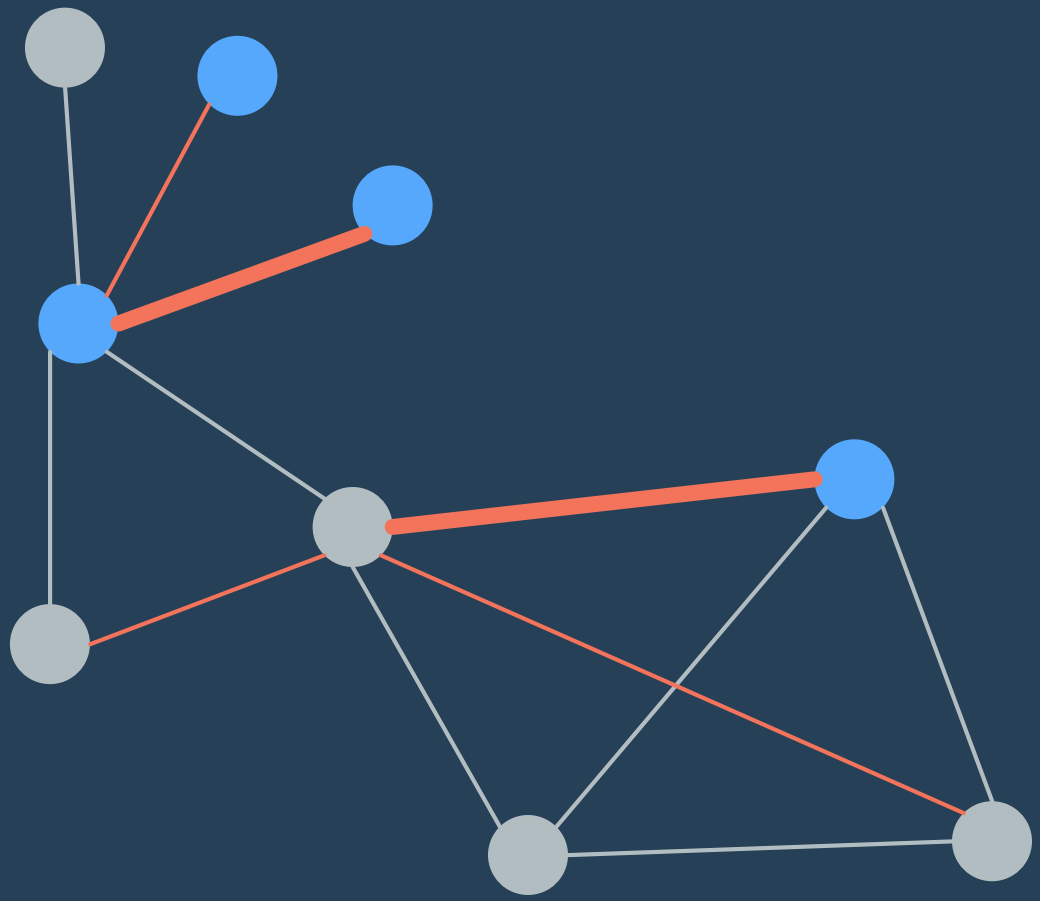
What’s in a graph?



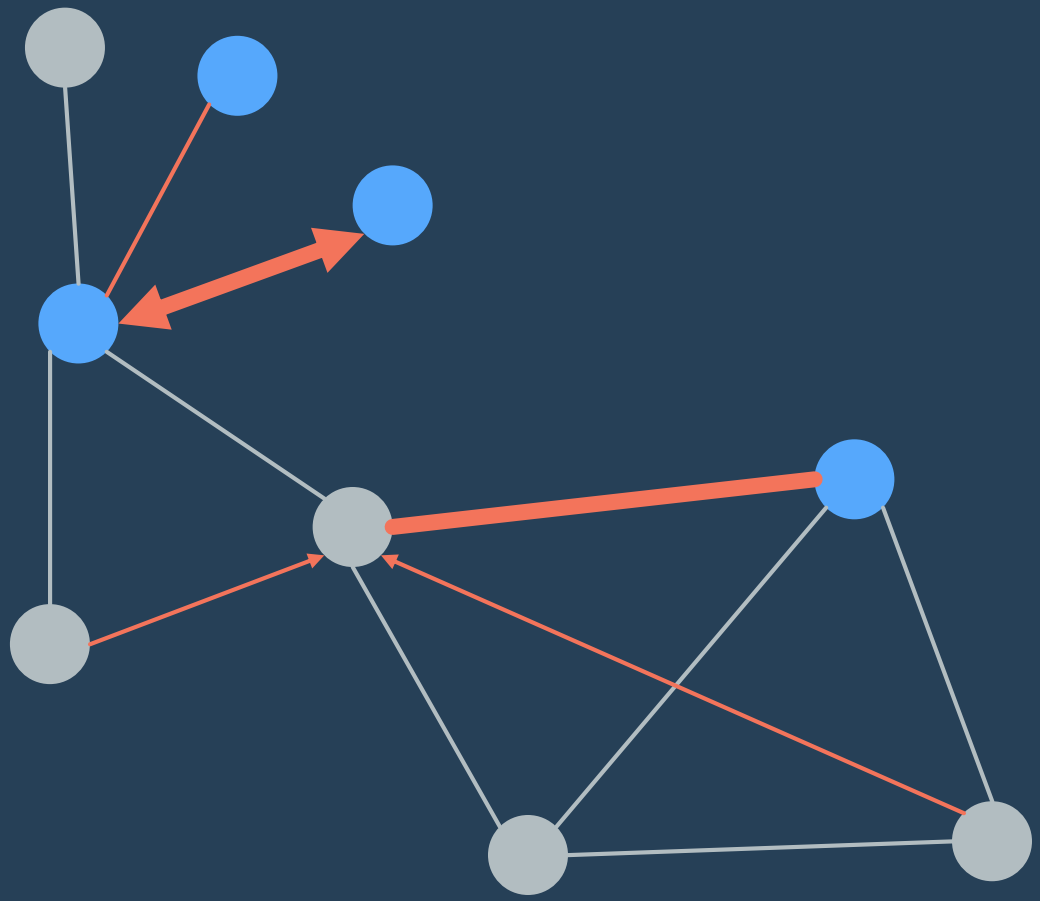
What’s in a graph?



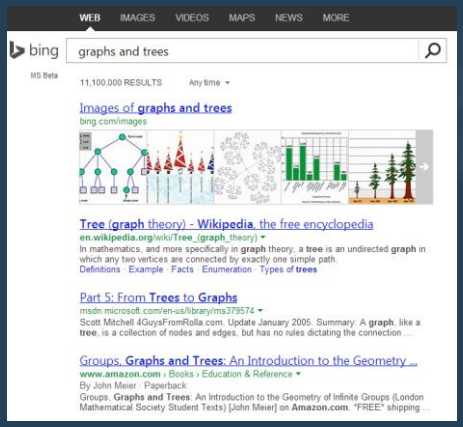
What’s in a graph?

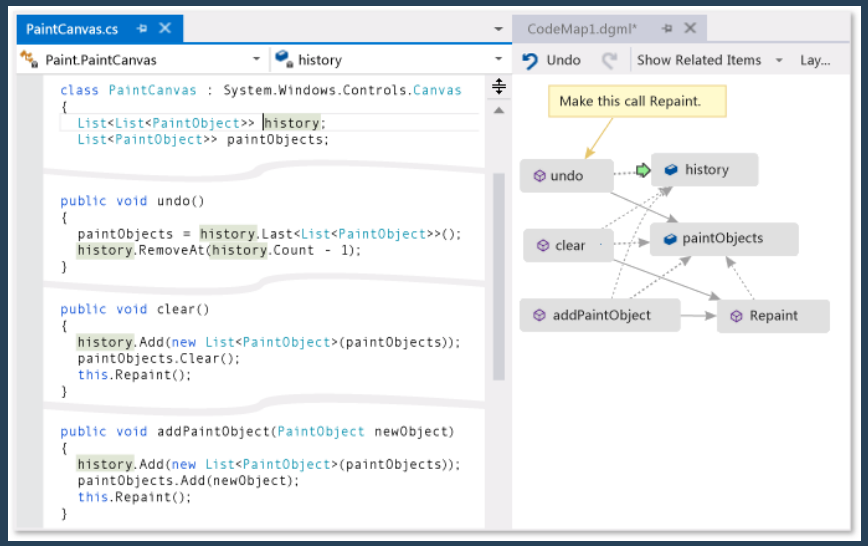


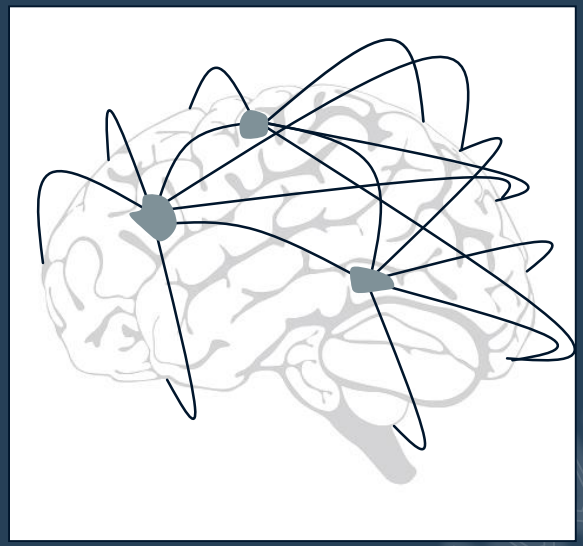
What’s in a graph?



Everything can be a graph!







What questions might we ask?

- How does the brain organize itself to achieve a function?
- How does knowledge disseminate in online communities?
- How are two graphs similar?
- Which entities in a social network might be terrorists?

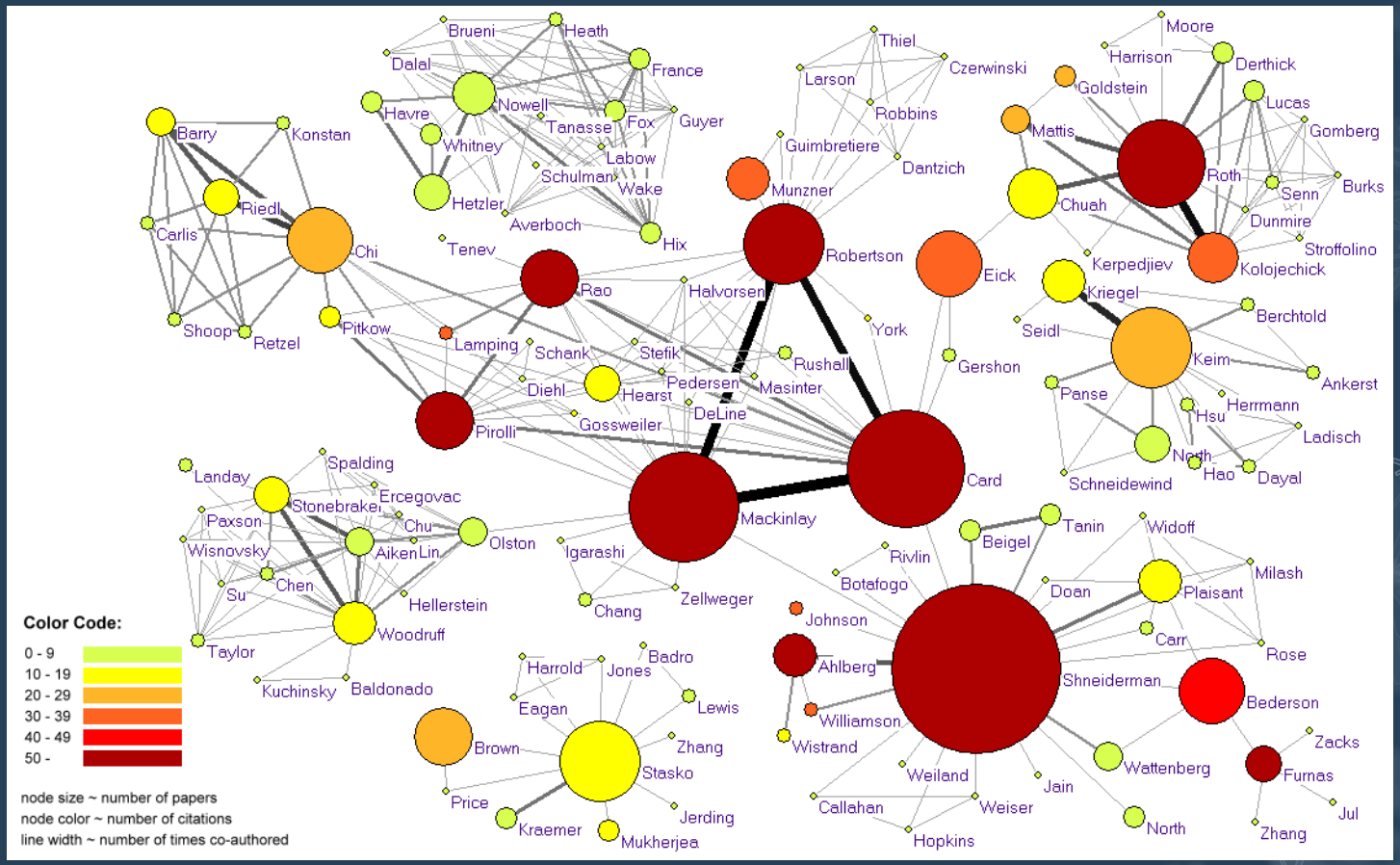
Graph Drawing

The primary concern of graph drawing is
the spatial arrangement of nodes and links

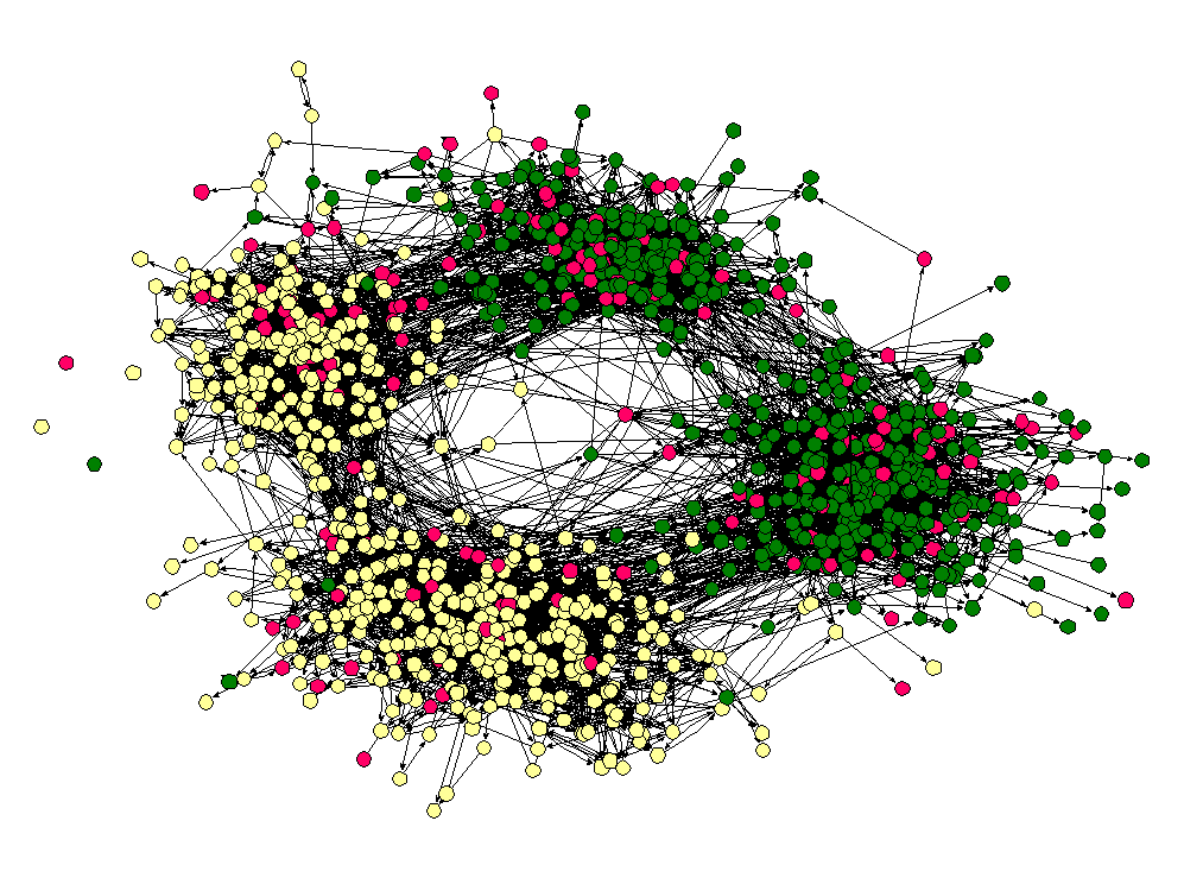
Often (but not always) the goal is
to effectively depict the graph structure:

- Connectivity patterns
- Partitions / Clusters
- Outliers

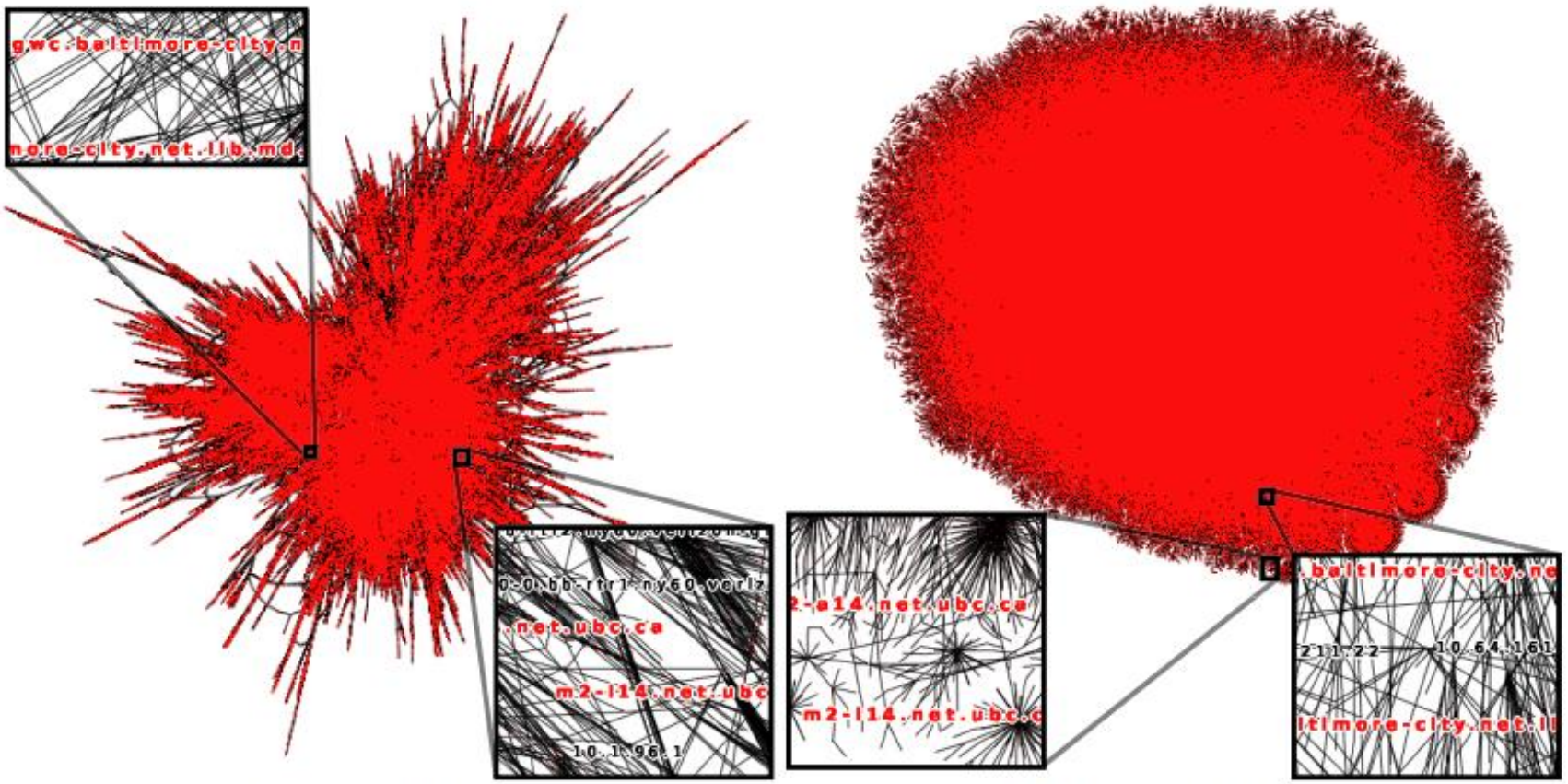
Putting things into perspective



Putting things into perspective



Putting things into perspective



(a) FM^3: 11 Minutes

(b) LGL: 12 Hours

Outline

Tree visualization

Graph visualization

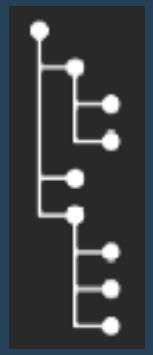
- node-link diagrams
- matrices

Recent research topics

Trees

4 Major tree visualizations

Indented lists

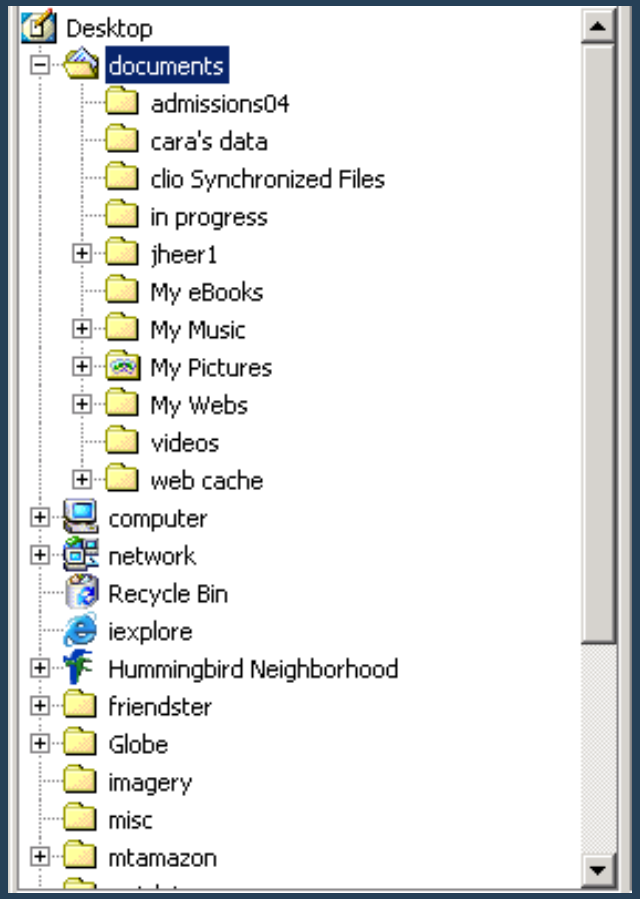


Node-link trees

Layered diagrams

Treemaps

Indented List



Places all items along vertically spaced rows

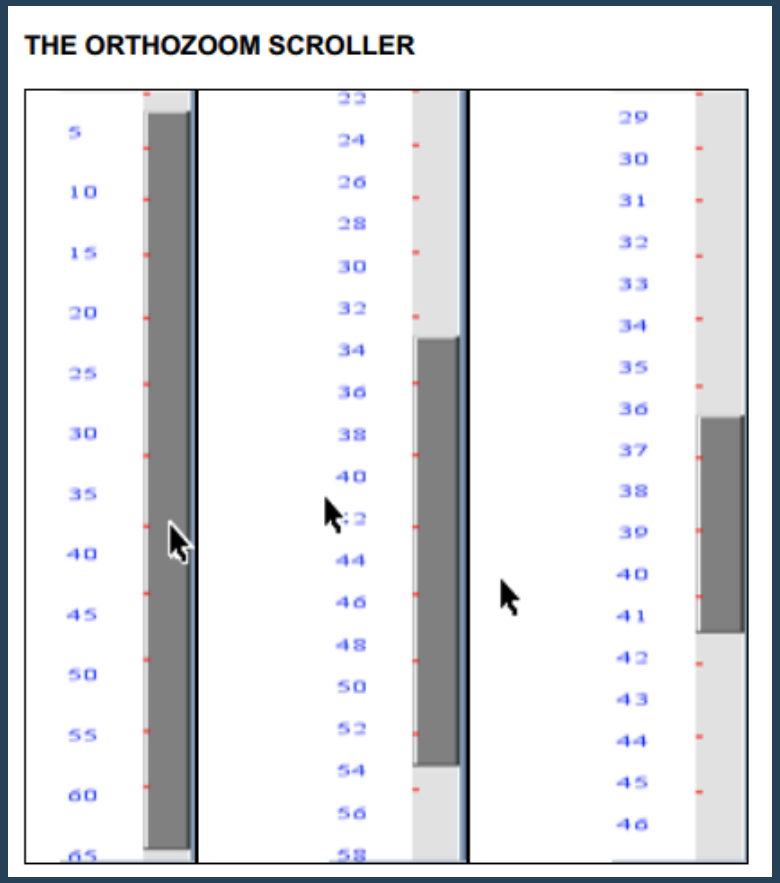
Indentation used to show parent/child relationships

Commonly used as a component in an interface

Breadth and depth contend for space

Often requires a great deal of scrolling

Interaction can help



OrthoZoom, *Appert et al.,* CHI 2006

http://www.lri.fr/~appert/website/orthozoom/orthozoom.html

king, for the queen: and therefore, Peter, have at
thee with a downright blow!
YORK
Dispatch: this knave's tongue begins to double.
Sound, trumpets, alarum to the combatants!
Alarum. They fight, and PETER strikes him down
HORNER
Hold, Peter, hold! I confess, I confess treason.
Dies
YORK
Take away his weapon. Fellow, thank God, and the
good wine in thy master's way.
PETER
O God, have I overcome mine enemy in this presence?
O Peter, thou hast prevailed in right!
KING HENRY VI
Go, take hence that traitor from our sight;
For his death we do perceive his guilt:
And God in justice hath revealed to us
The truth and innocence of this poor fellow,
Which he had thought to have murder'd wrongfully.
Come, fellow, follow us for thy reward.
Sound a flourish. Exeunt

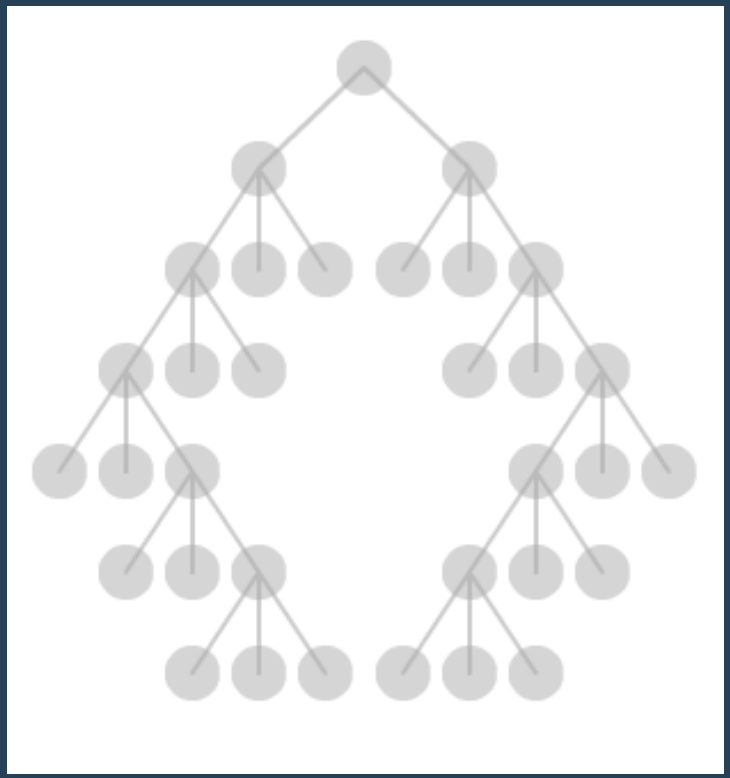
SCENE IV. A street.

Enter GLOUCESTER and his Servingmen, inmourning cloaks
GLOUCESTER
Thus sometimes hath the brightest day a cloud;
And after summer evermore succeeds
Barren winter, with his wrathful nipping cold:
So cares and joys abound, as seasons fleet.
Sirs, what's o'clock?
Servants
Ten, my lord.
GLOUCESTER
Ten is the hour that was appointed me
To watch the coming of my punish'd duchess:
Uneath may she endure the flinty streets,
To tread them with her tender-feeling feet.
Sweet Nell, ill can thy noble mind abrook
The abject people gazing on thy face,
With envious looks, laughing at thy shame,
That erst did follow thy proud chariot-wheels
When thou didst ride in triumph through the streets.
But, soft! I think she comes; and I'll prepare
My tear-stain'd eyes to see her miseries.
Enter the DUCHESS in a white sheet, and a taperburning in her hand; with
STANLEY, the Sheriff,and Officers
Servant
So please your grace, we'll take her from the sheriff.
GLOUCESTER
No, stir not, for your lives; let her pass by.

1108

Zoom: 1.0

Node-Link Trees

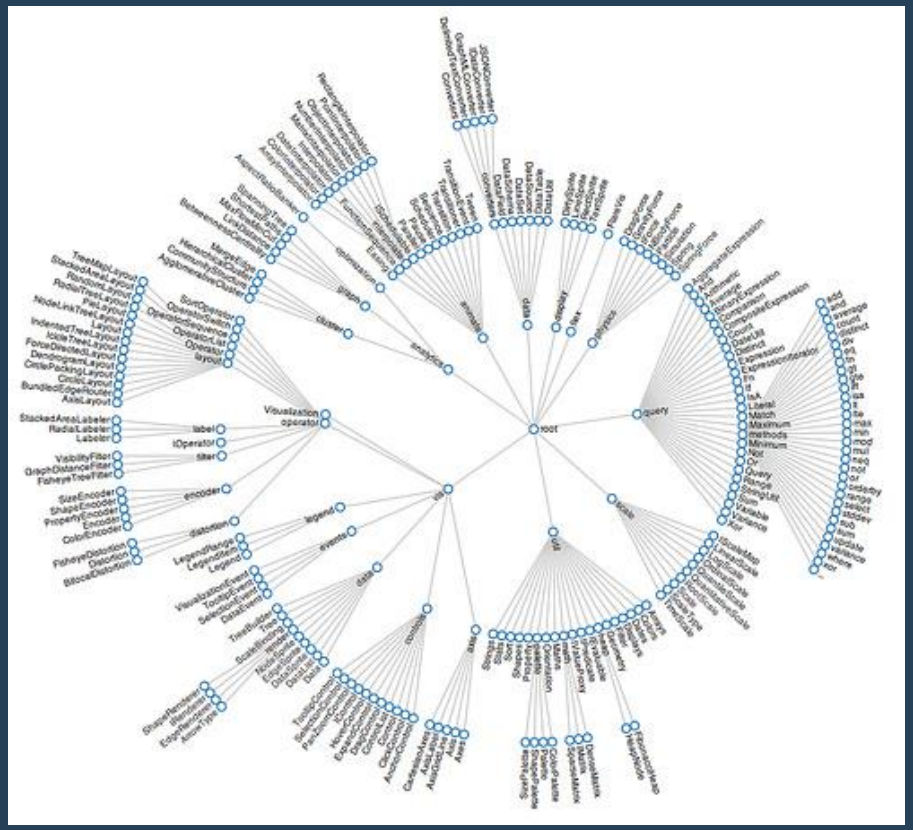


Nodes are distributed in space, connected by straight or curved lines.

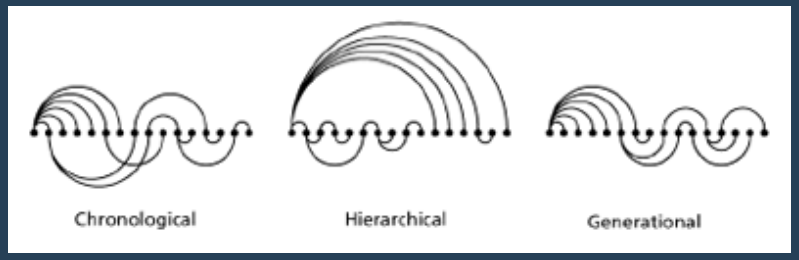
Typical approach is to use 2D space to break apart breadth and depth.

Reingold-Tilford algorithm achieves linear time

Node-Link Trees


root
query
data
scale
operator
visual
Visualization
label
iOperator
filter
encoder
legend
events
DataEvent
SelectionEvent
ToolEvent
VisualizationEvent
Literal
Match
Maximum
Minimum
methods
Not
Or
Query
Range
Variable
Visualization
Expression
Expressionization
Aggregate Expression
Arithmetic
Binary Expression
Composite
Composite Expression
Count
Date Util
Datecd
Expression
Expressionization
In
Is
IsA
Literal
Match
Maximum
methods
Minimum

Other node-Link trees



ThreadArcs,

Kerr,

2003

PhylloTrees,

Neumann et al.,

Eurovis 2006

http://treevis.net

Layered diagrams

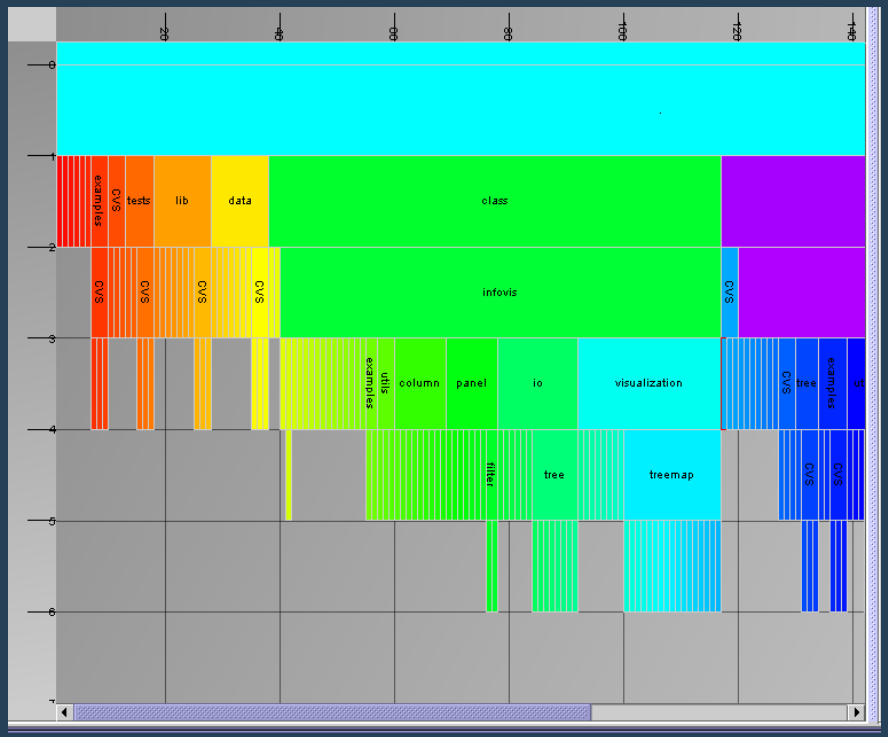
Signify tree structure using

- Layering
- Adjacency
- Alignment

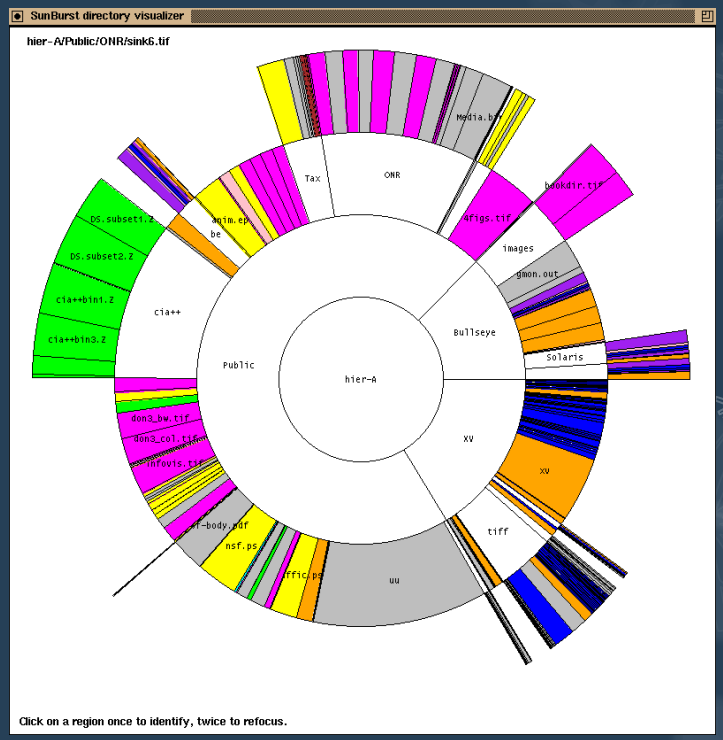
Involves recursive sub-division of space

We can apply the same set of approaches as in node-link layout.

Layered diagrams



Icicle Trees



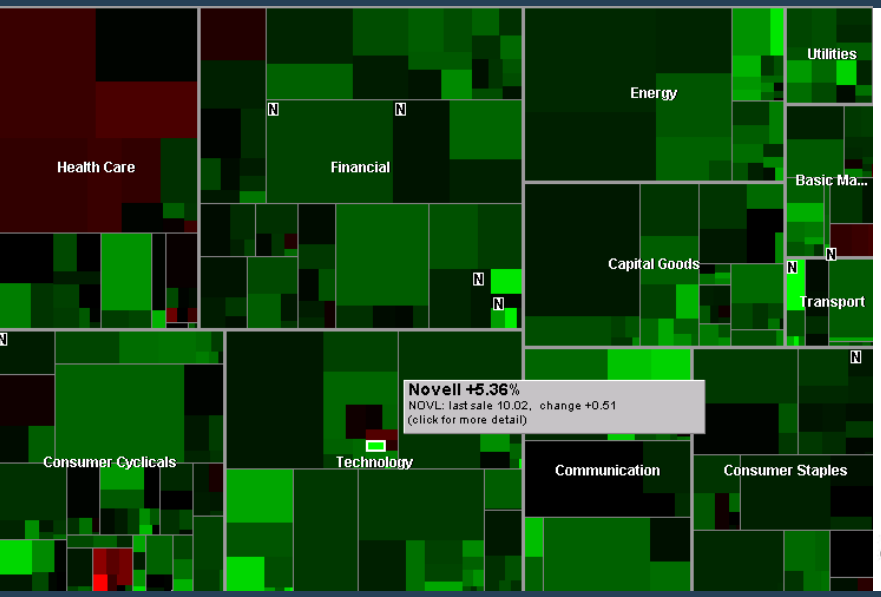
SunBurst, *Stasko et al.*,

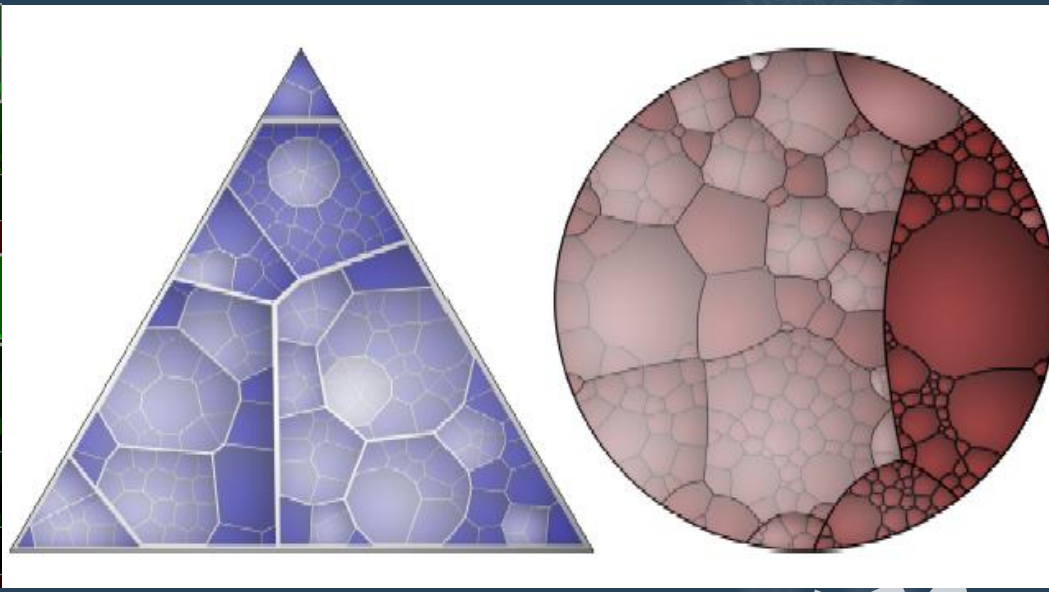
Infovis 2000

Treemaps

Encode hierarchy using spatial enclosure

Space-filling technique





http://www.cs.umd.edu/hcil/treemap-history/

Treemaps

Benefits

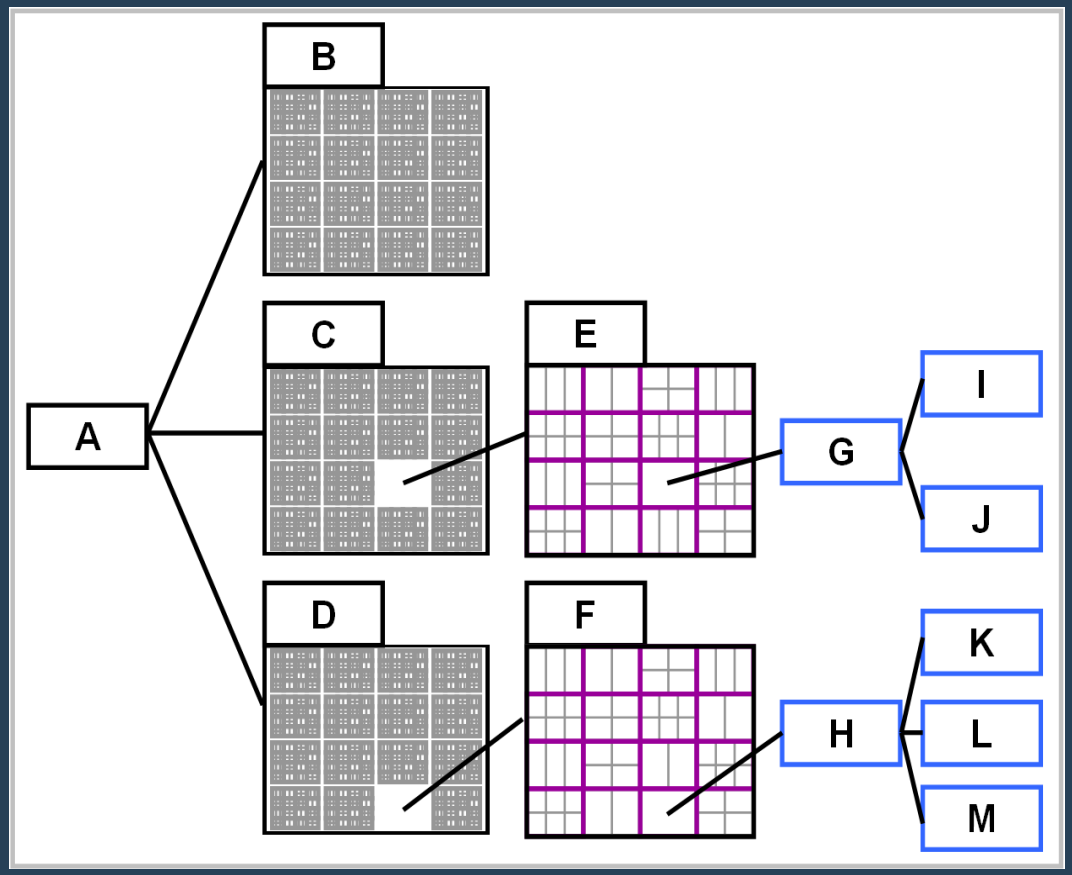
Provides a single view of an entire tree

Easier to spot large/small nodes

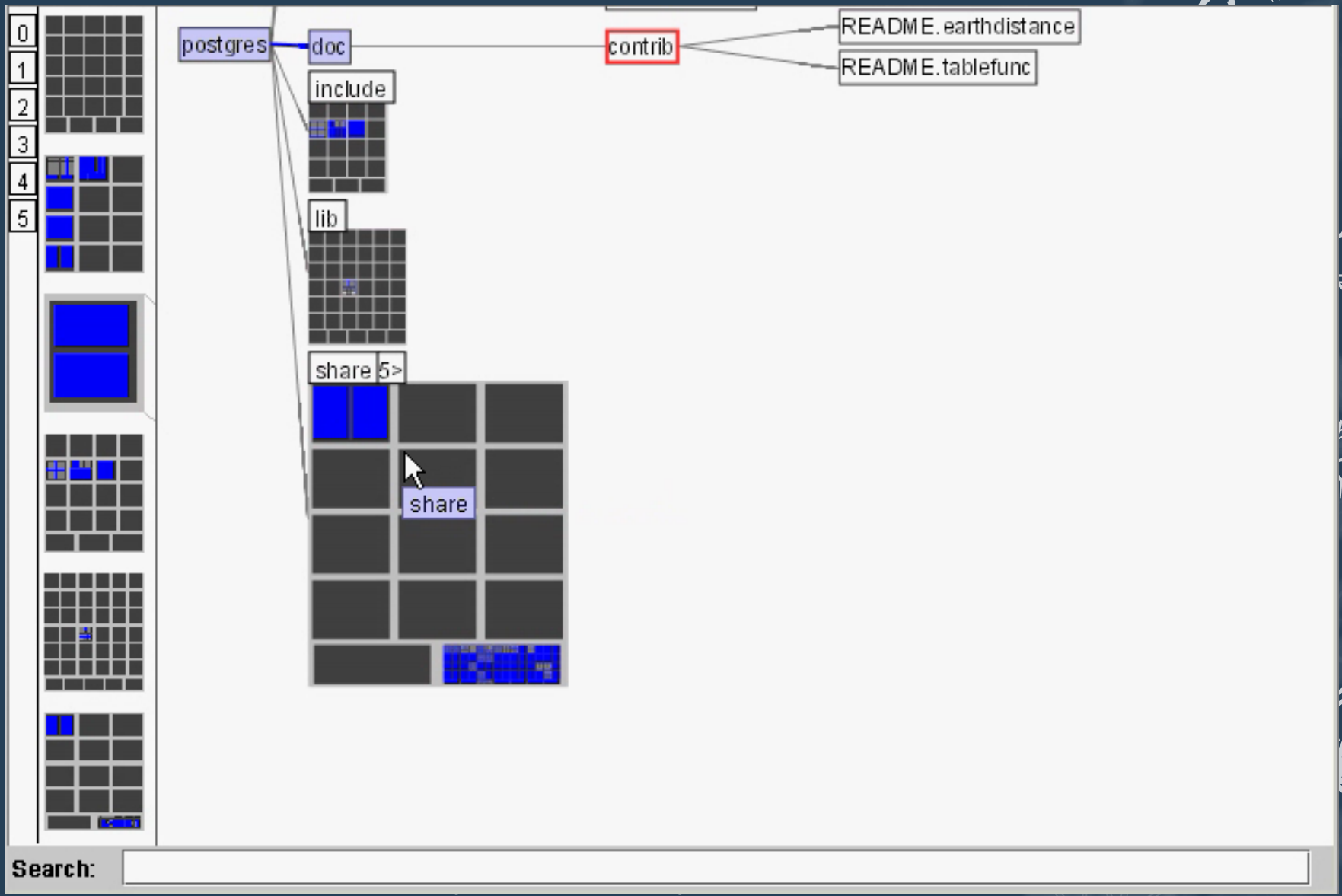
Problems

Difficult to accurately read depth

Hybrids

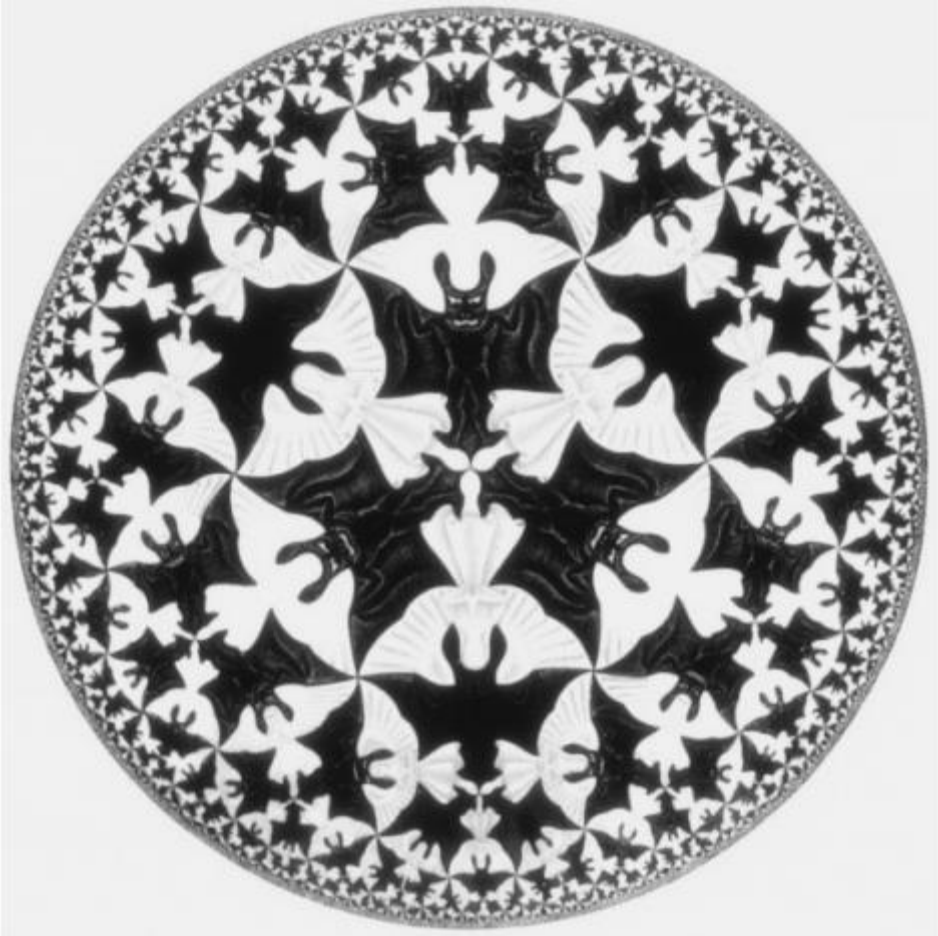


Elastic Hierarchies, *Zhao et al*., Infovis 2005

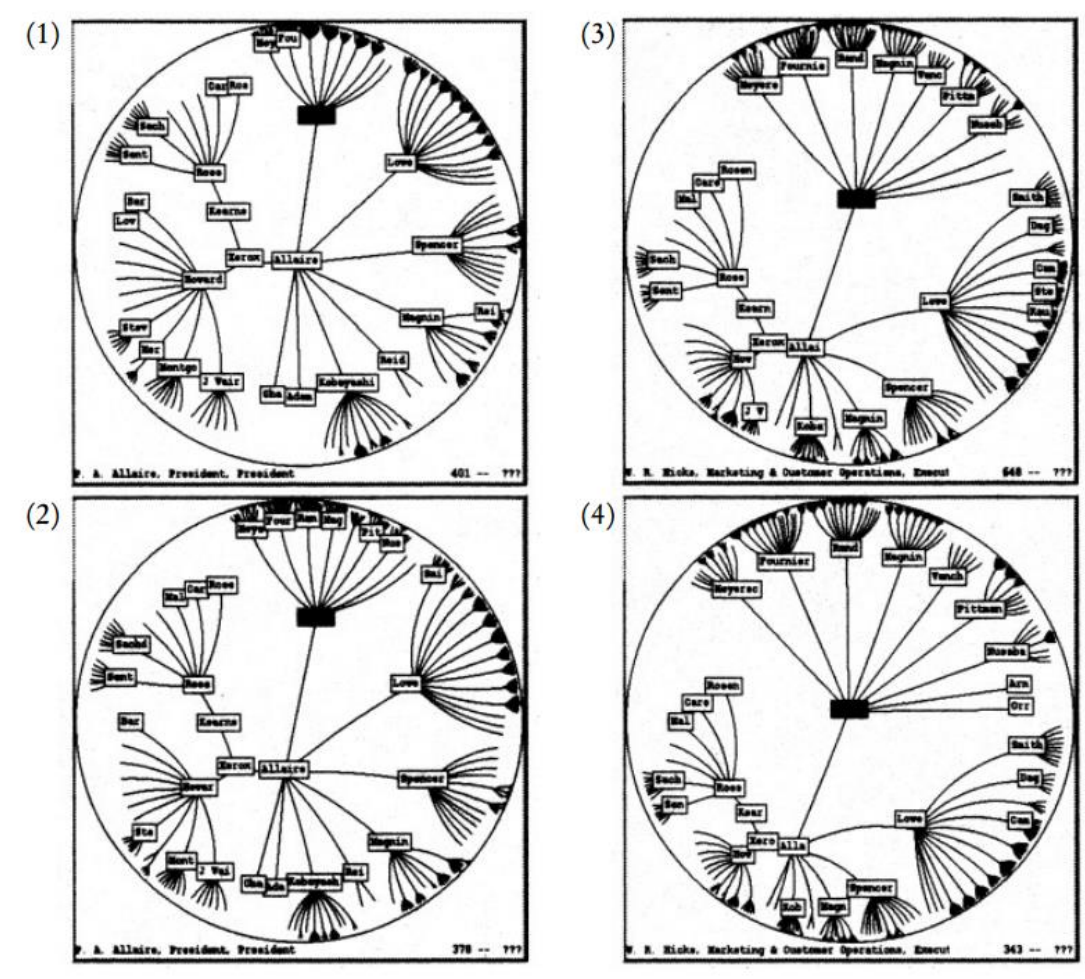

0
1
2
3
4
5
postgres
doc
include
lib
share
5>
share
contrib
README.earthdistance
README.tablefunc
Search:

The issue of scale

Hyperbolic Space

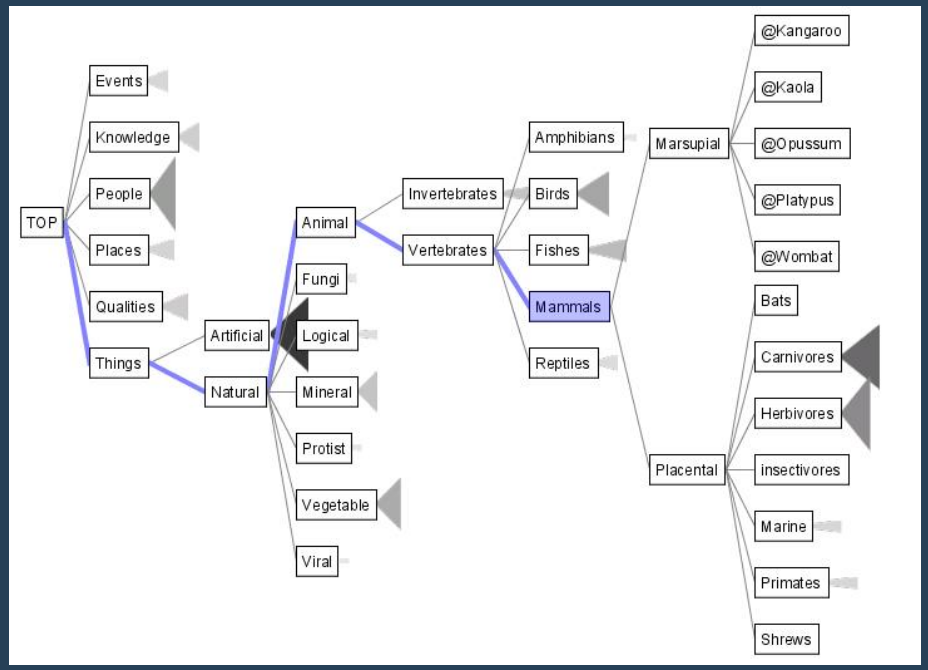


Hyperbolic Tree Browser

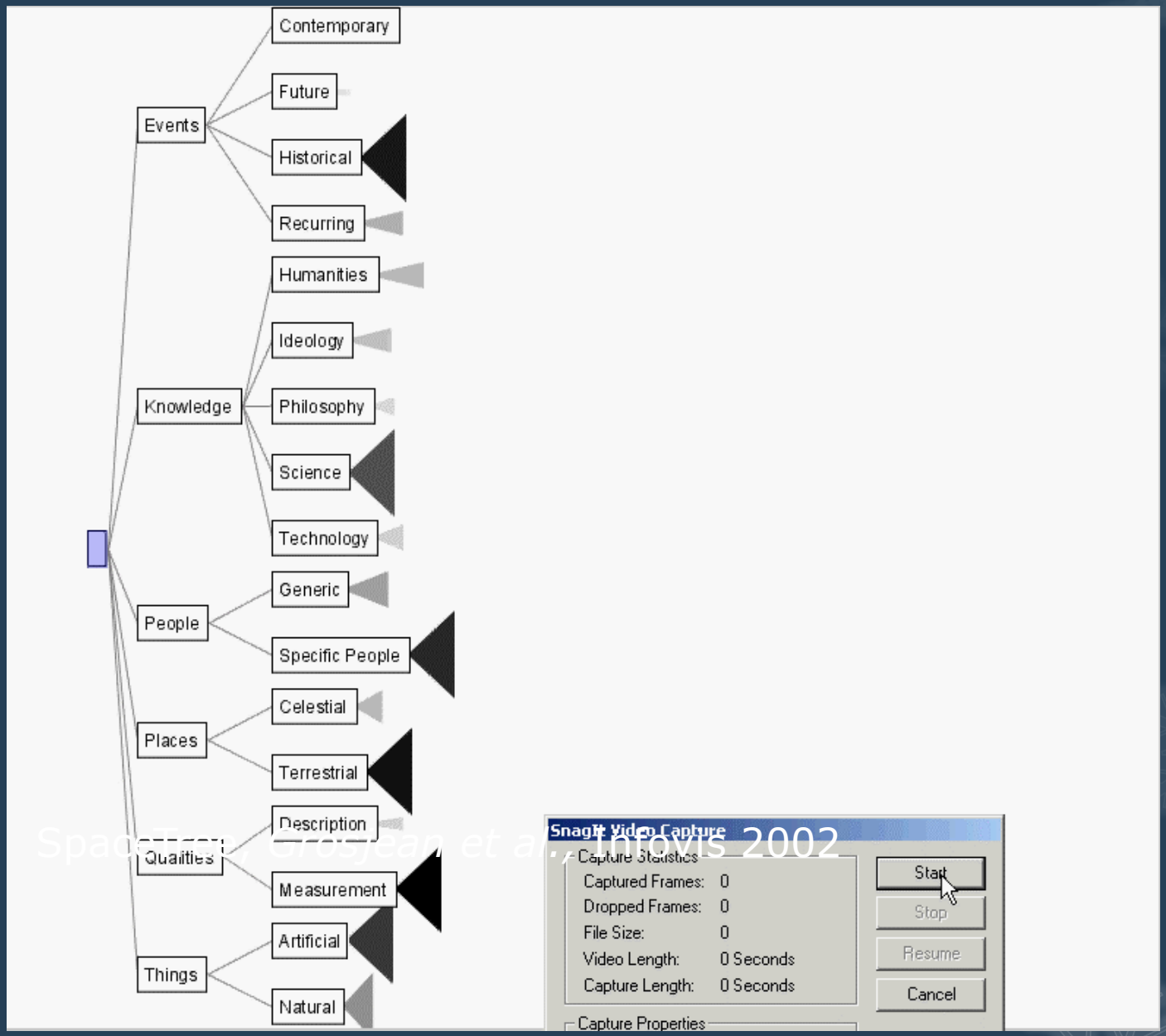


Hyberbowlic tree browser, *Lamping et al*., CHI 1995

Aggregation



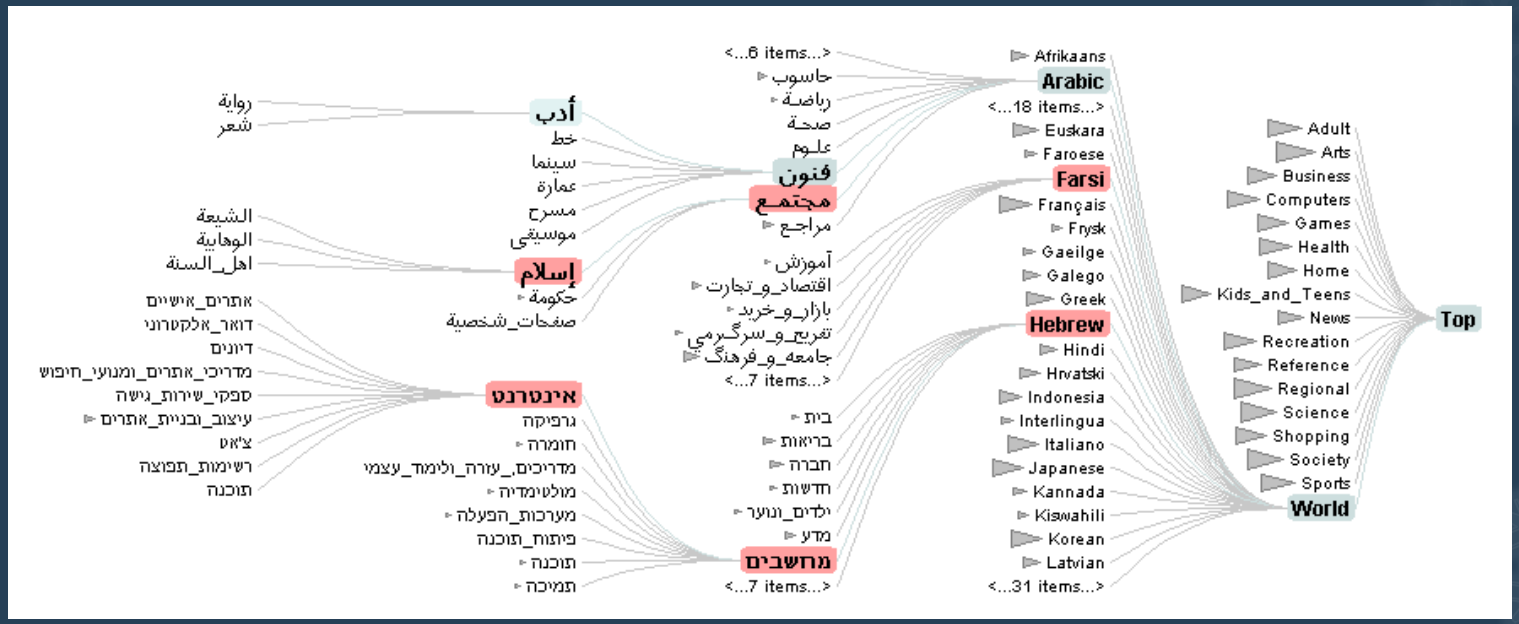
SpaceTree, *Grosjean et al.,* Infovis 2002



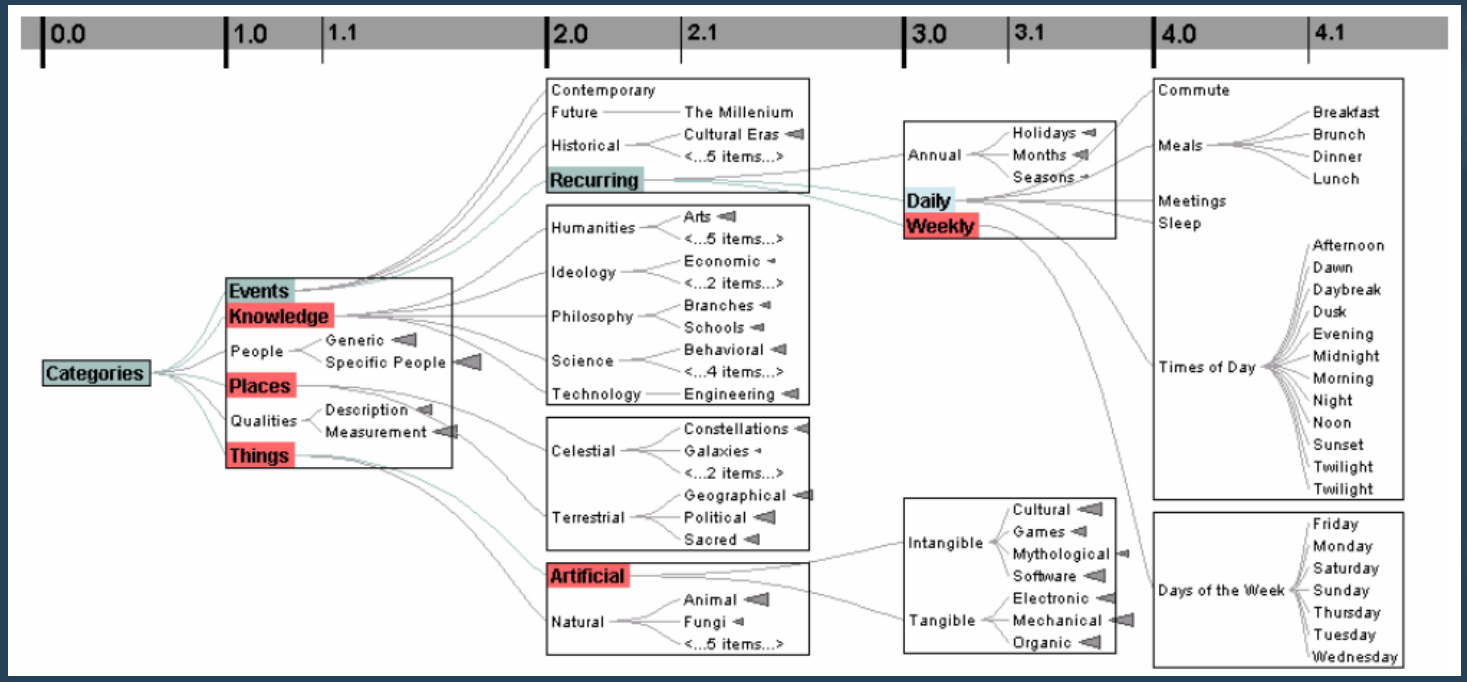
Spaceball, Grosjean et al., Infovis 2002

Degree-of-interest trees

Degree-of-interest trees



Degree-of-interest trees



Cull "un-interesting" nodes on a per block basis until all blocks on a level fit within bounds.

Attempt to center child blocks beneath parents.

Graphs

Graph Visualization

Two representations:

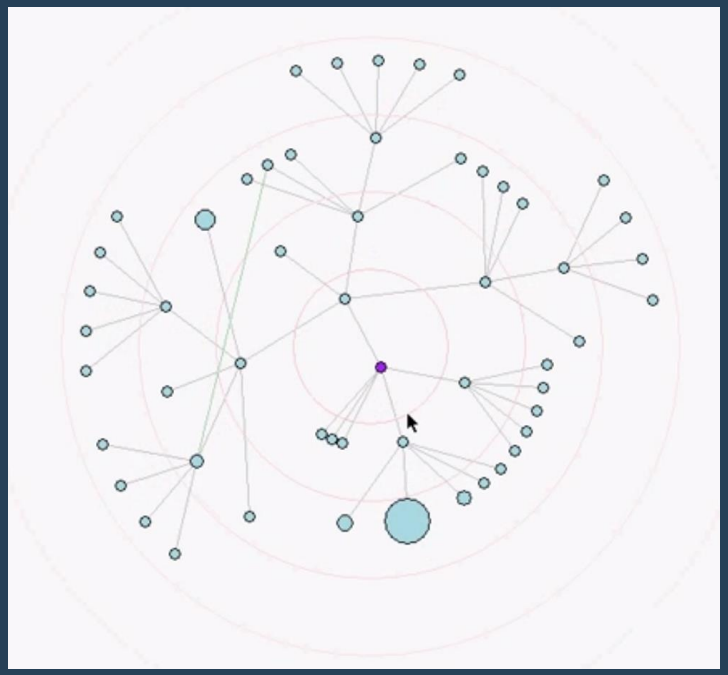
- Node-link diagrams
- Matrices

Major Node-Link Layouts

Scalability issues and solutions

Matrix-based representations

See the tree in this graph?

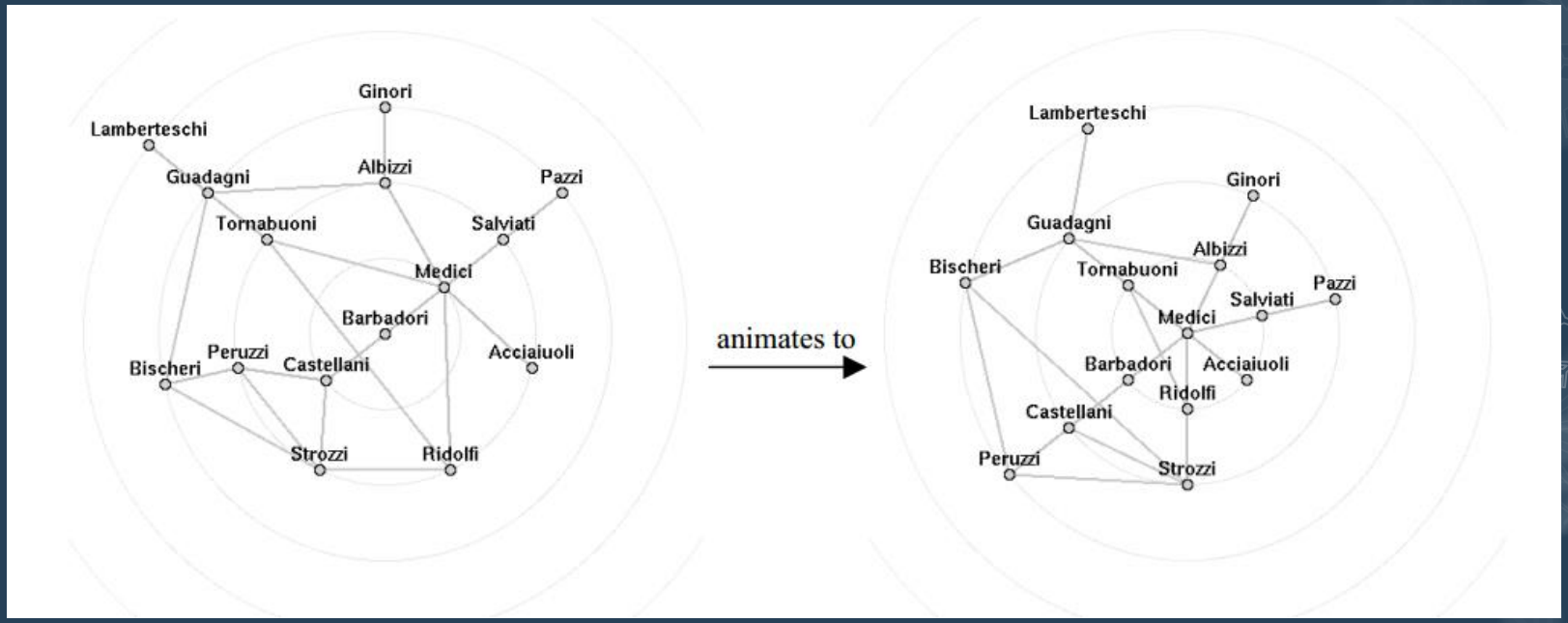


Many graphs are tree-like or have useful spanning trees

Spanning trees lead to arbitrary roots

Fast tree layouts allow graph layouts to be recalculated at interactive rates

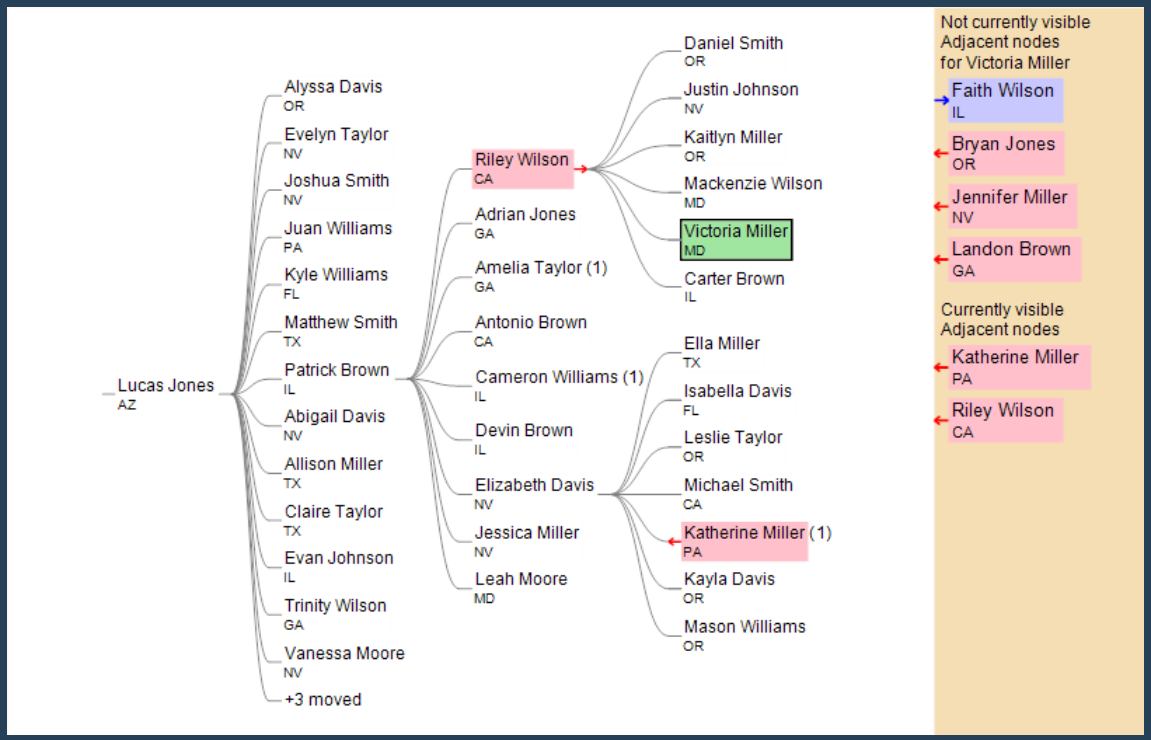
See the tree in this graph?



Animated Graphs with Radial Layout, *Yee et al.,* Infovis 2001

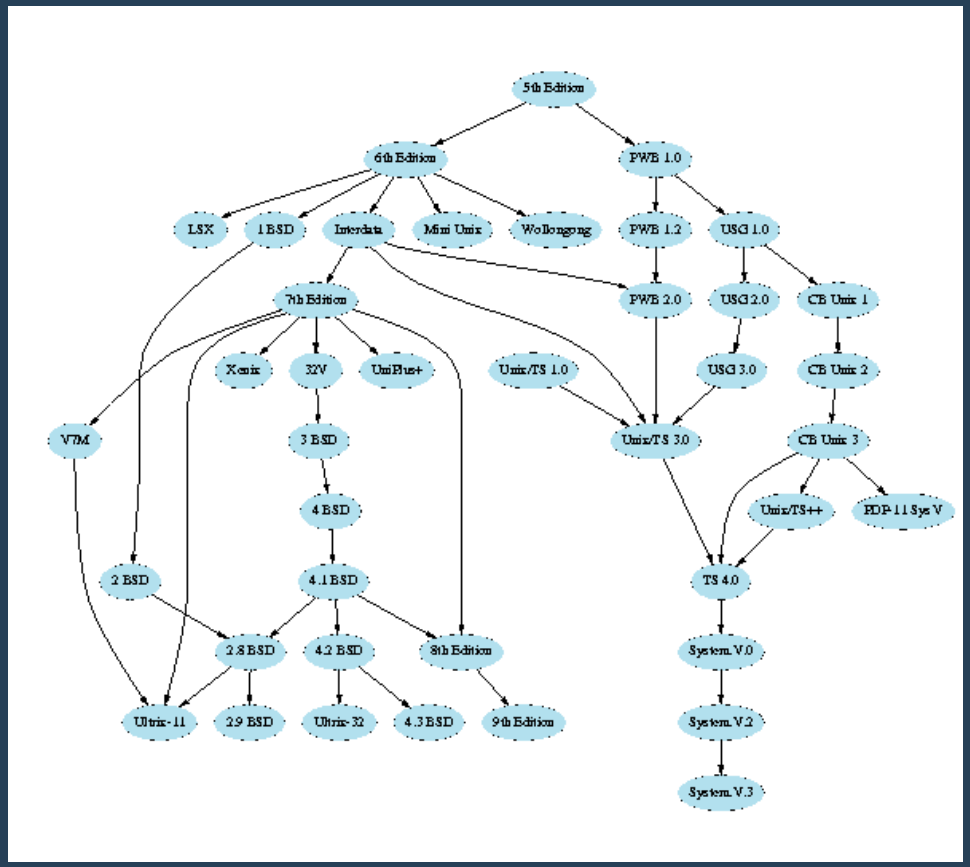
http://www.youtube.com/watch?v=OPX5iGro_lA

See the tree in this graph?



TreePlus, *Lee et al.*, VAST 2006

Hierarchical graph layout

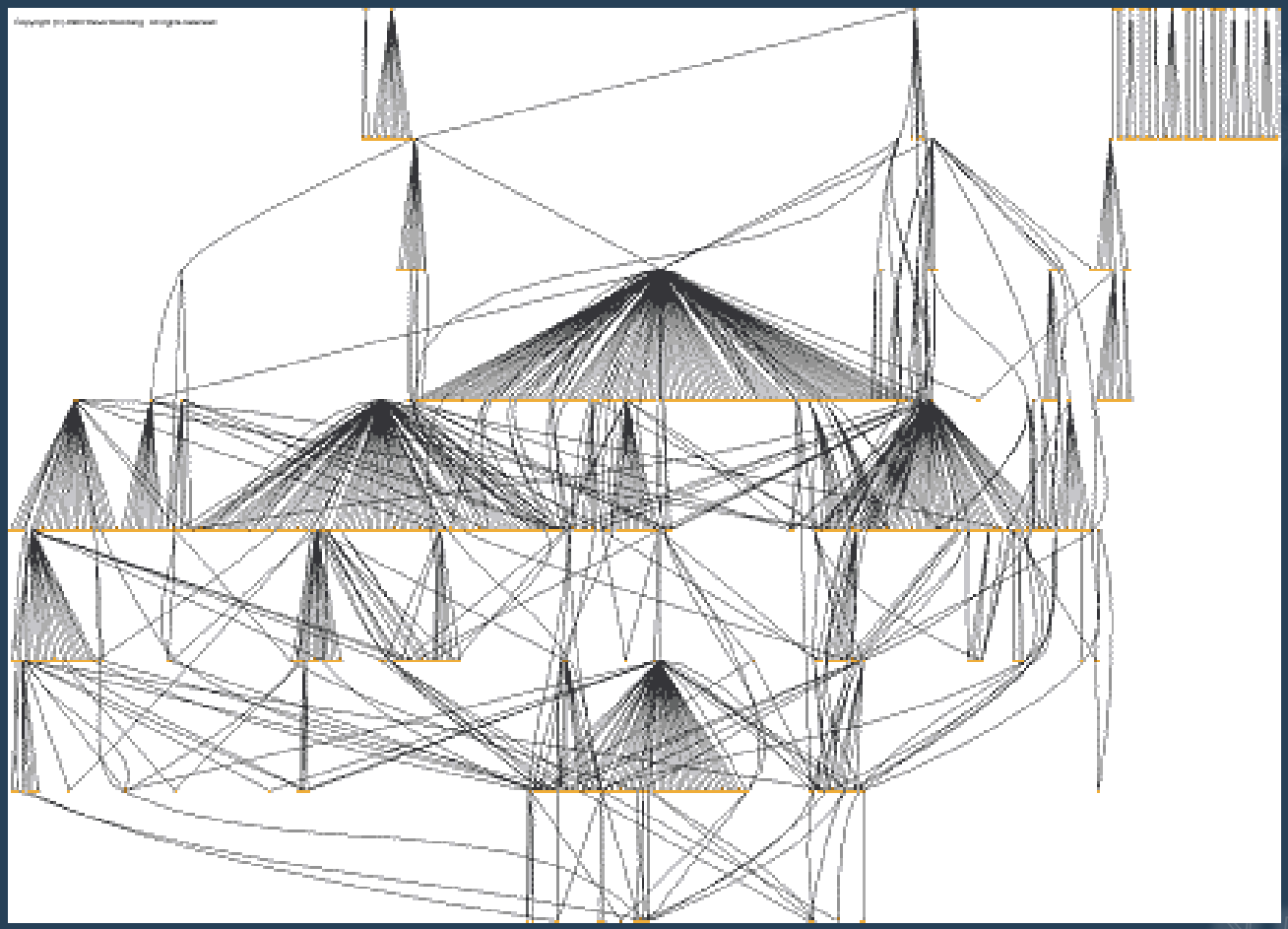


Sugiyama-style or layered graph drawing

Layout of a Direct Acyclic Graph

Hierarchical layering based on descent

Hierarchical graph layout



Optimization techniques

Treat layout as an *optimization problem*

- Define layout using an *energy model* and/or a set of *constraints*: equations the layout should try to obey
- Use optimization algorithms to solve

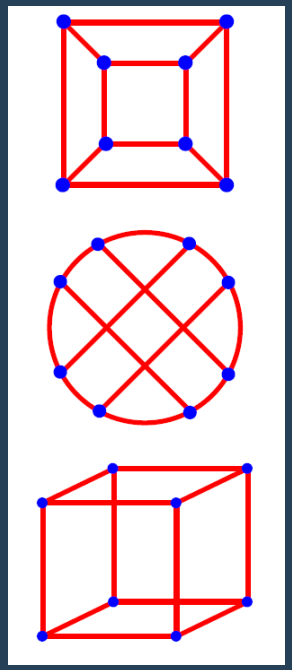
Regularly used for undirected graphs

- *Force-Directed Layout* most common

We can introduce directional constraints

- *DiG-CoLa* (Di-Graph Constrained Optimization Layout) [Dwyer 05]
- Iterative constraint relaxation

“Aesthetic” constraints



Minimize edge crossings

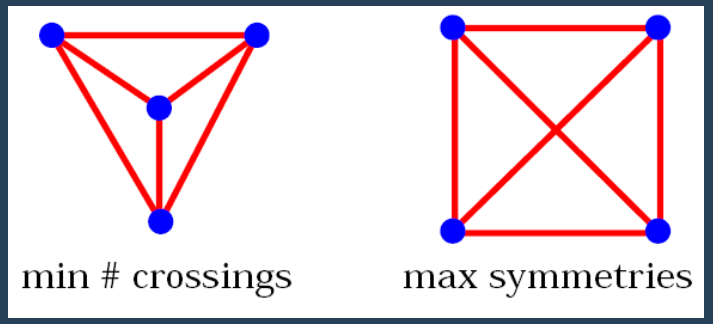
Minimize area

Minimize line bends

Minimize line slopes

Maximize smallest angle between edges

Maximize symmetry



but, can’t do it all.

Force-directed layout

Nodes = charged particles $F = G*m_1*m_2 / (x_i - x_j)^2$

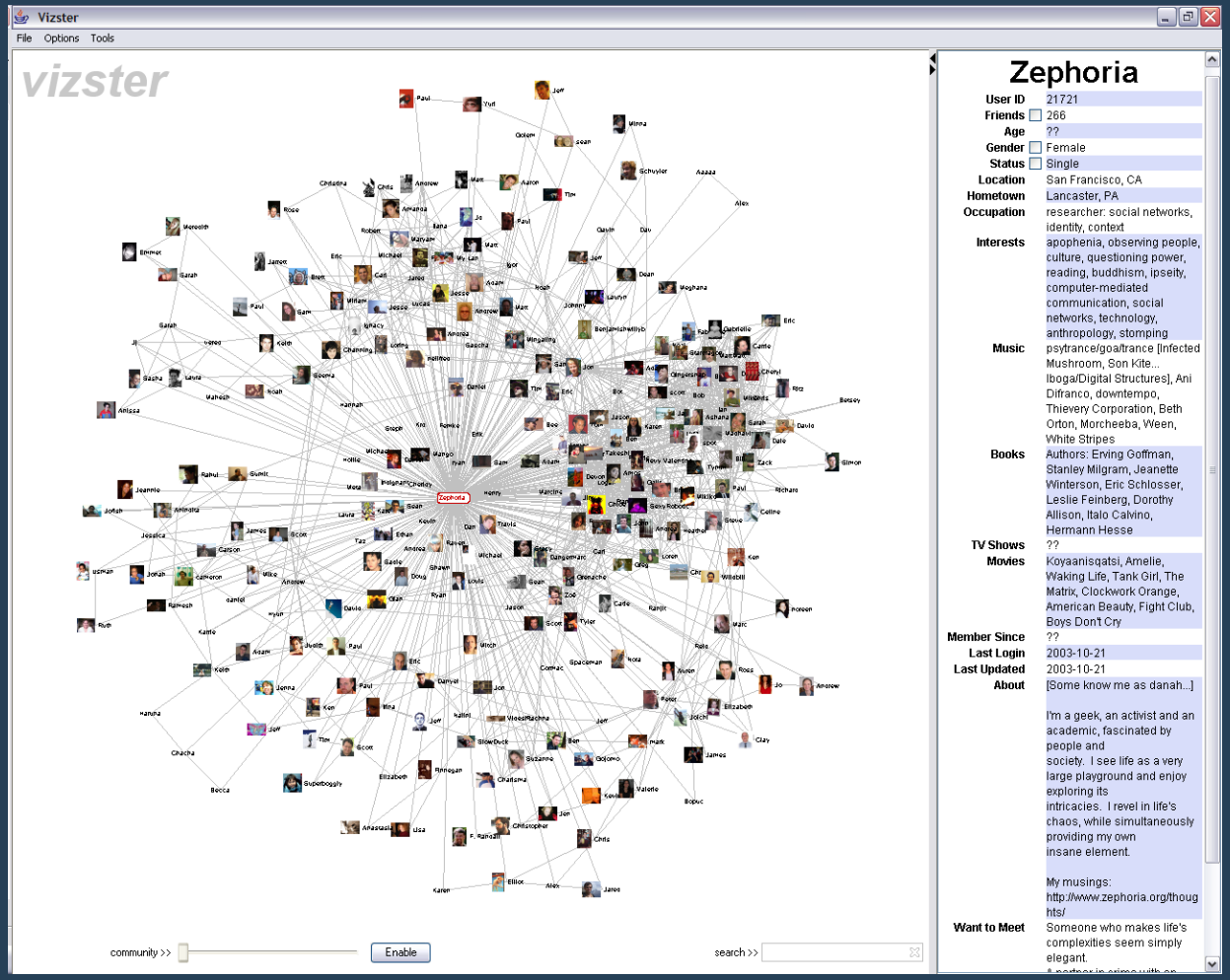
with air resistance $F = -b * v_i$

Edges = springs $F = -k * (x_i - x_j - L)$

Repeatedly calculate forces, update node positions

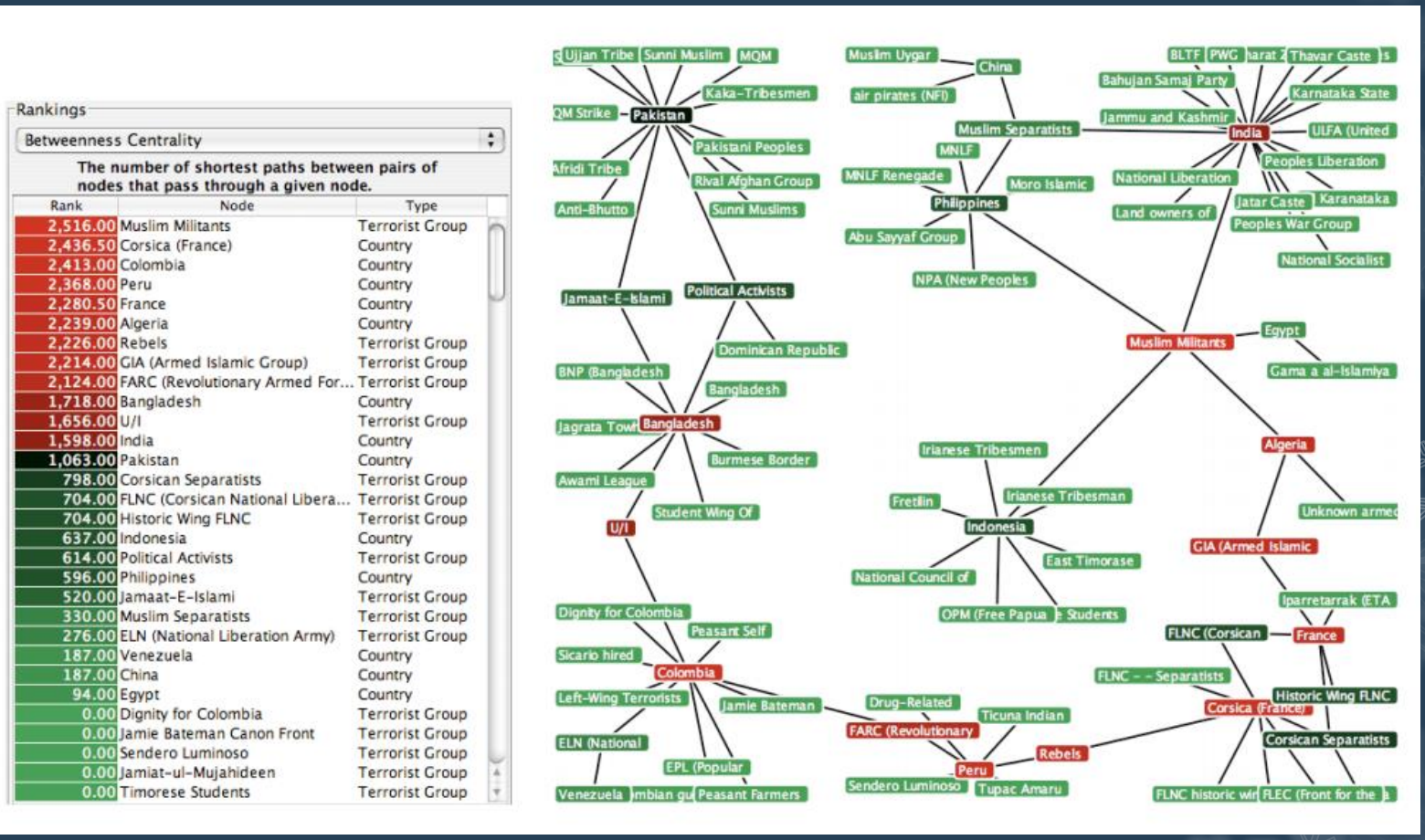
- Naïve approach $O(N^2)$
- Speed up to O(N log N) using quadtree or k-d tree
- Numerical integration of forces at each time step

Ego-Centered Networks



Vizster, *Heer et al.*, Infovis 2005

Filtered Networks



Social Action, *Perer et al.*, Infovis 2006

Constraint Optimization layout

Minimize stress function

$$\text{stress}(X) = \Sigma_{i<j}\, w_{ij}\, (\, \|X_i - X_j\| - d_{ij}\,)^2$$

- X: node positions, d: optimal edge length,
- w: normalization constants
- Use global (*majorization*) or localized (*gradient descent*) optimization

→ Says: Try to place nodes d_{ij} apart

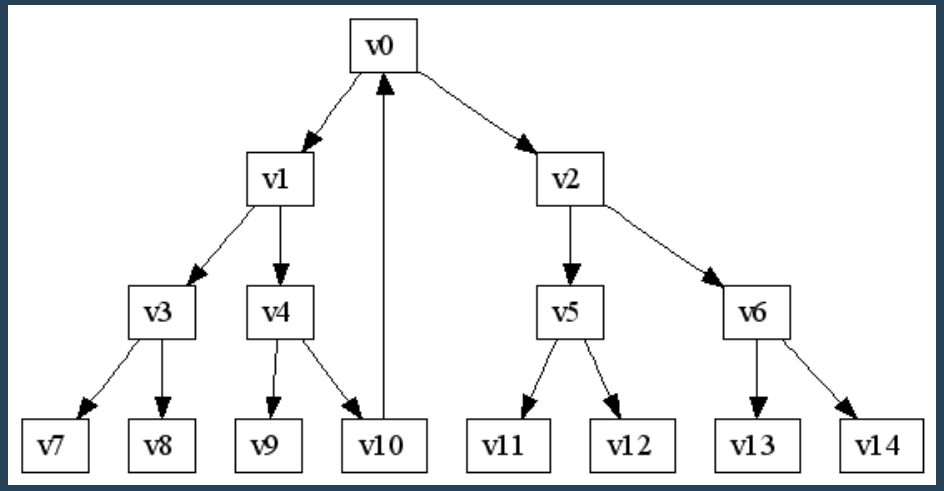
Add hierarchy ordering constraints

$$E_H(y) = \Sigma_{(i,j)\in E}\, (\, y_i - y_j - \delta_{ij}\,)^2$$

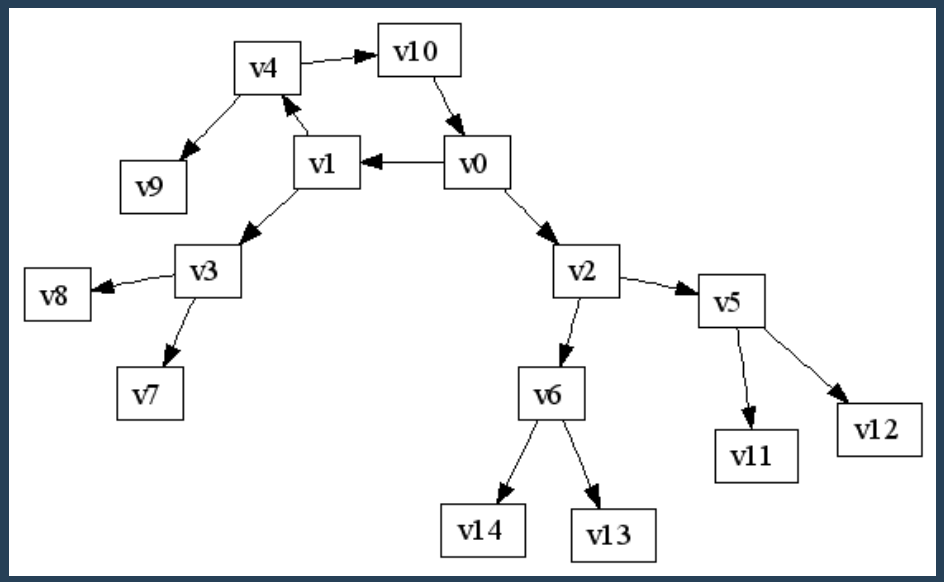
- y: node y-coordinates
- δ : edge direction (e.g., 1 for i→j, 0 for undirected)

→ Says: If *i* points to *j*, it should have a lower y-value

Constraint Optimization layout

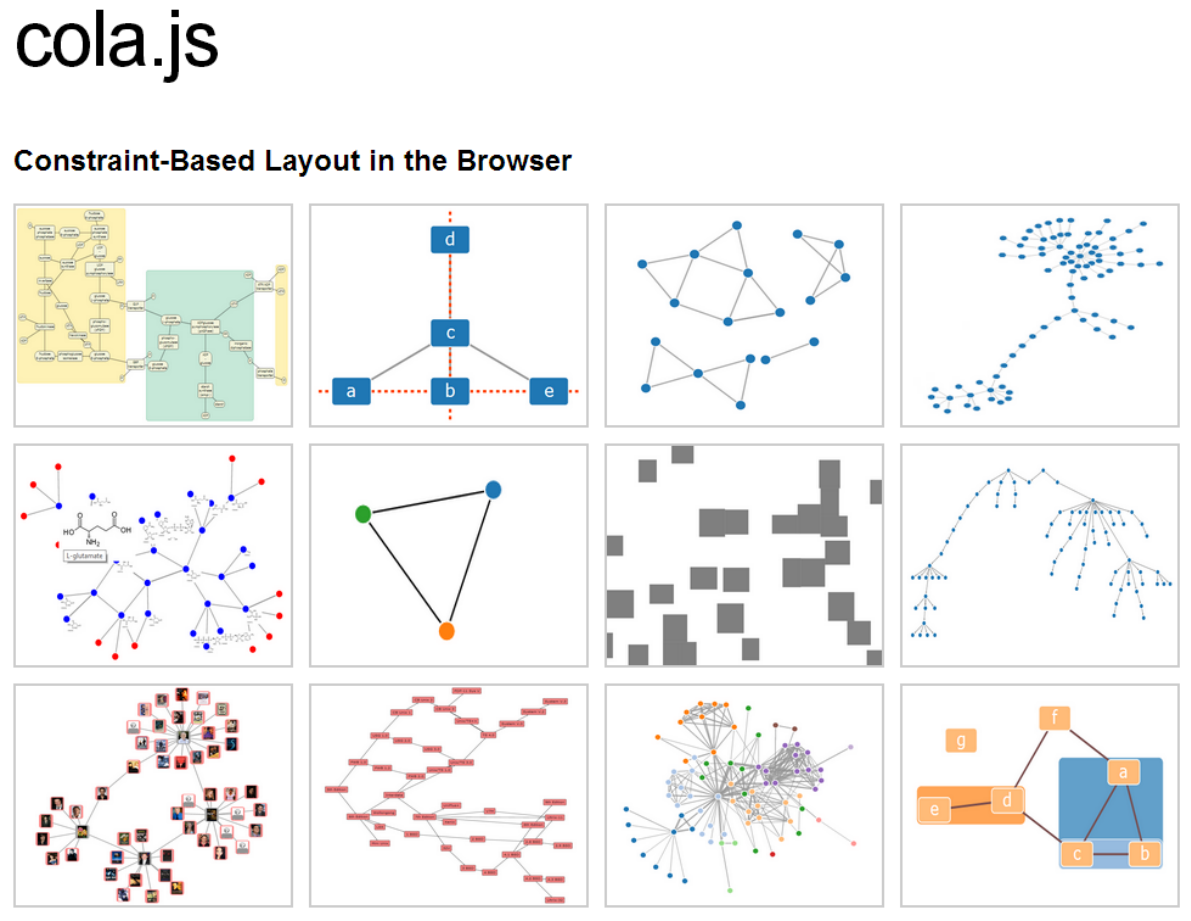


Sugiyama layout (dot)
Preserve tree structure



DiG-CoLa method
Preserve edge lengths

Constraint-based layout



http://marvl.infotech.monash.edu/webcola/

Coping with messiness

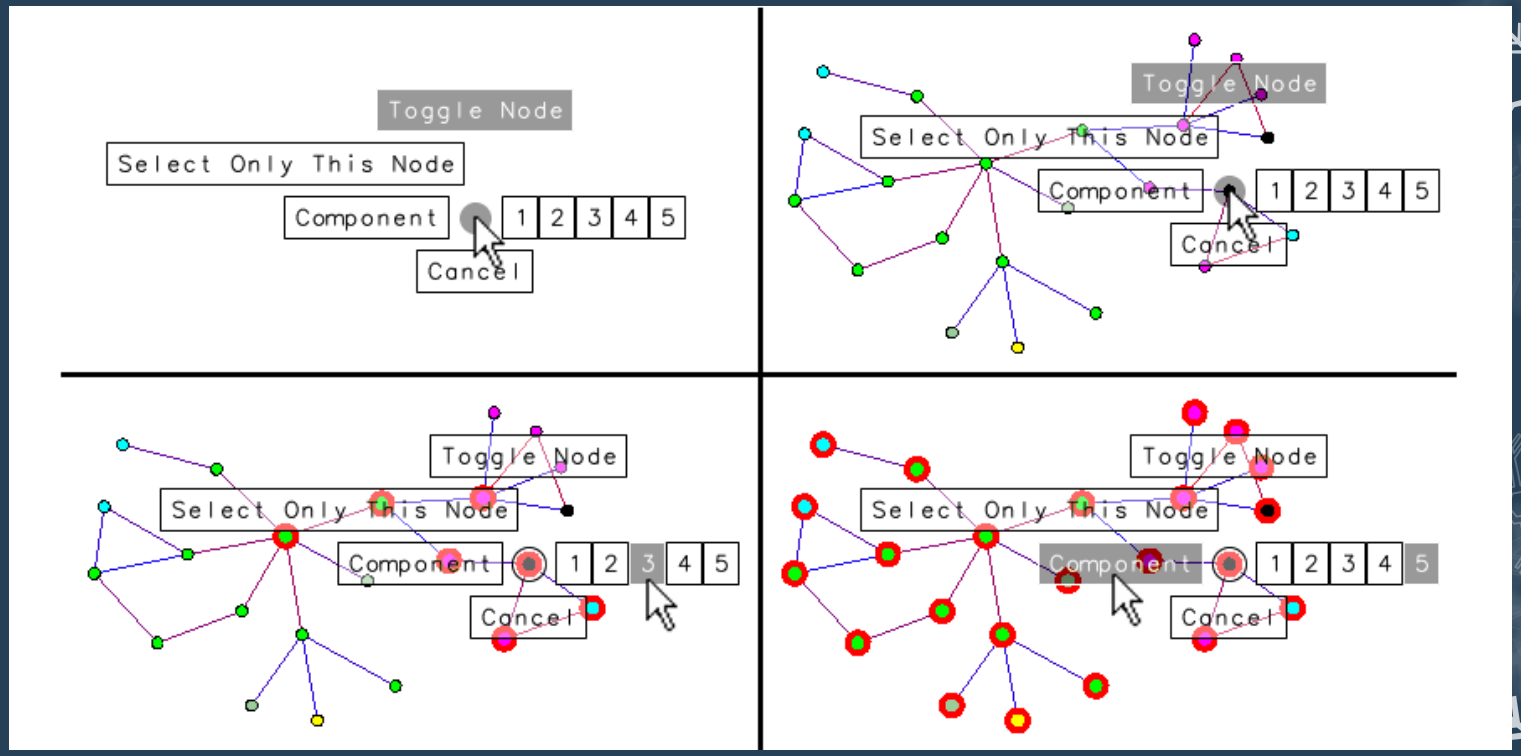
Interaction Techniques for Selecting and Manipulating Subgraphs in Network Visualizations

Presented at IEEE InfoVis 2009

Michael J. McGuffin
École de technologie supérieure
Montréal, Canada, http://profs.logti.etsmtl.ca/mmcguffin/

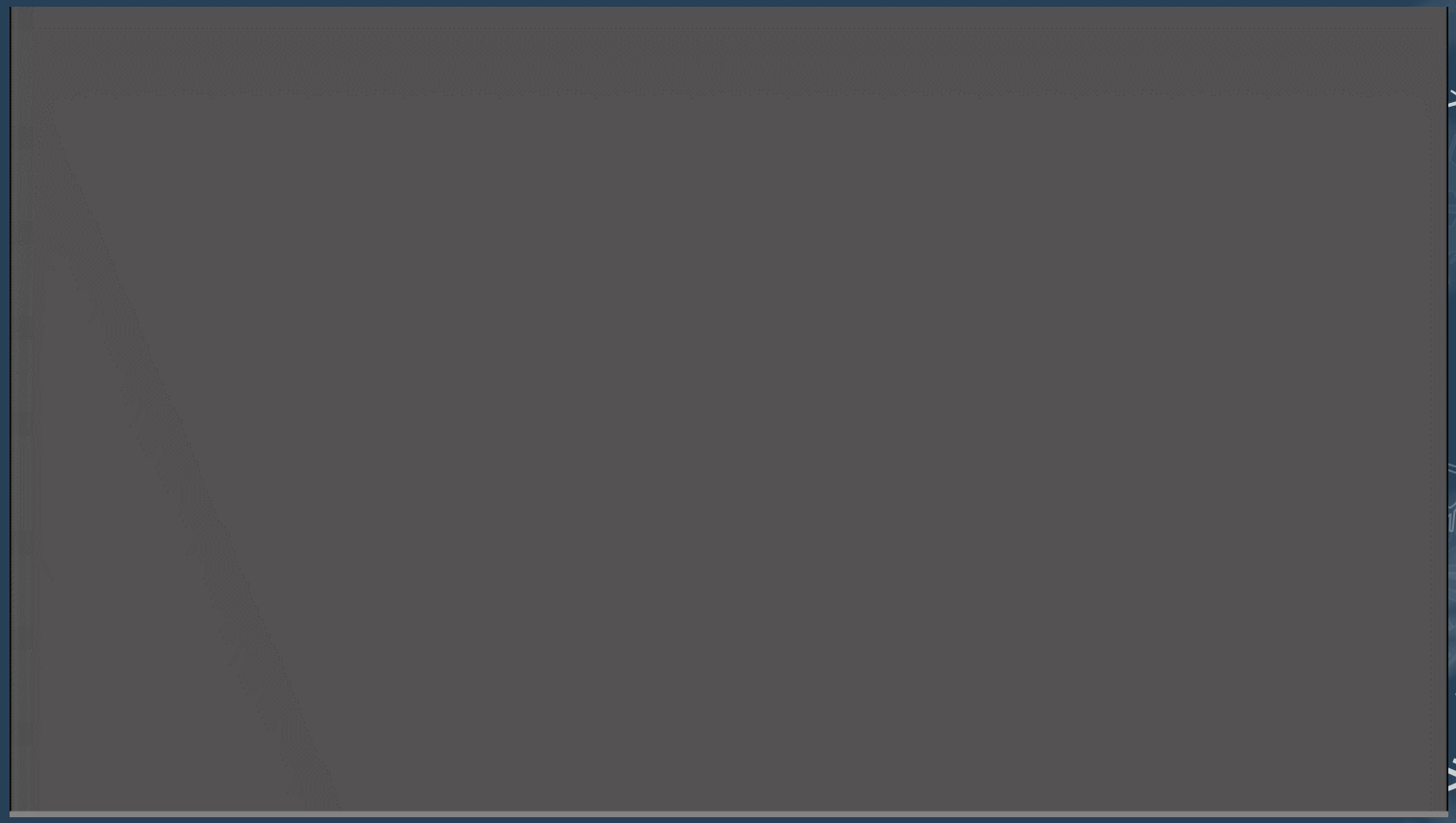
Igor Jurisica
Ontario Cancer Institute, PMH/UHN
Toronto, Canada, http://www.cs.toronto.edu/~juris/

Layout Interaction Techniques



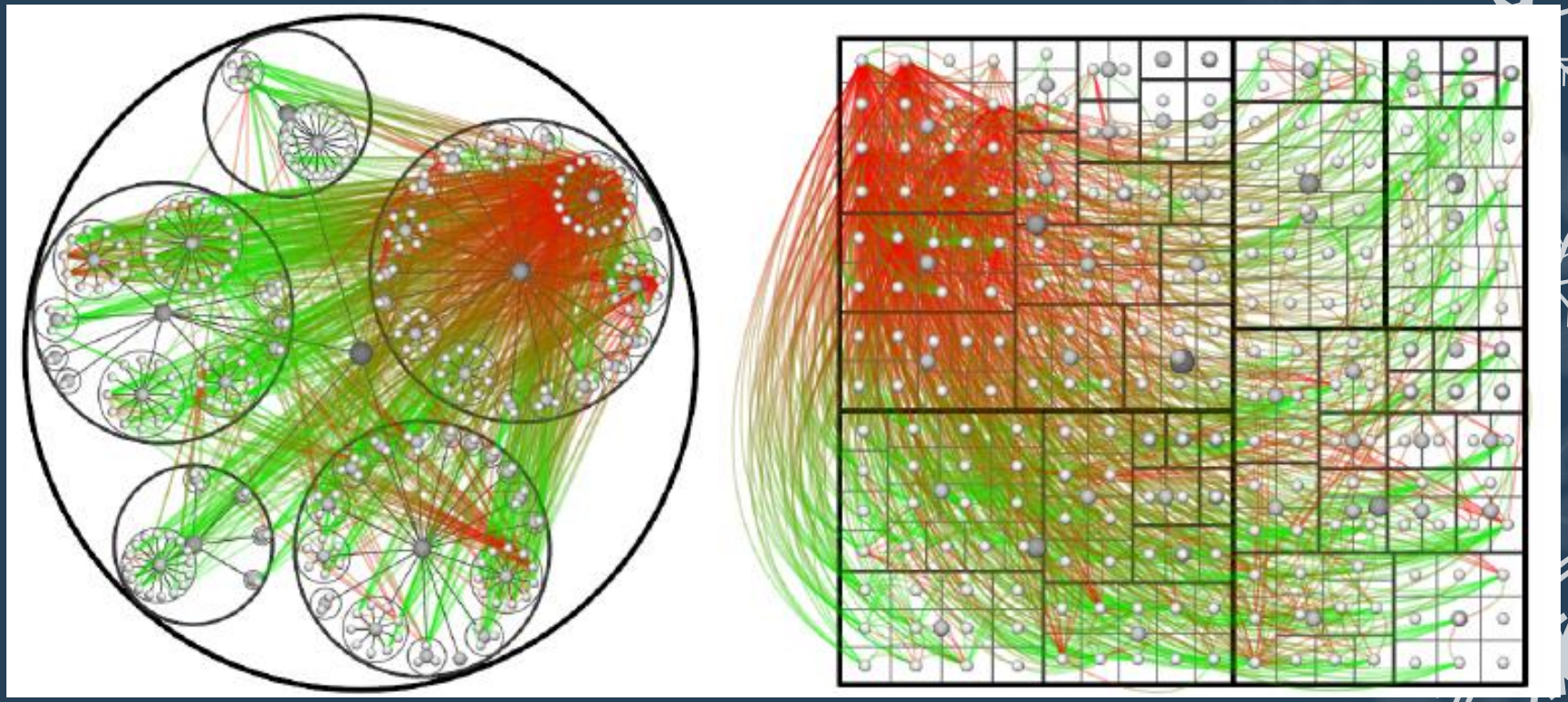
HotBox, *McGuffin et al.*, Infovis 2009

Edge Interaction Techniques

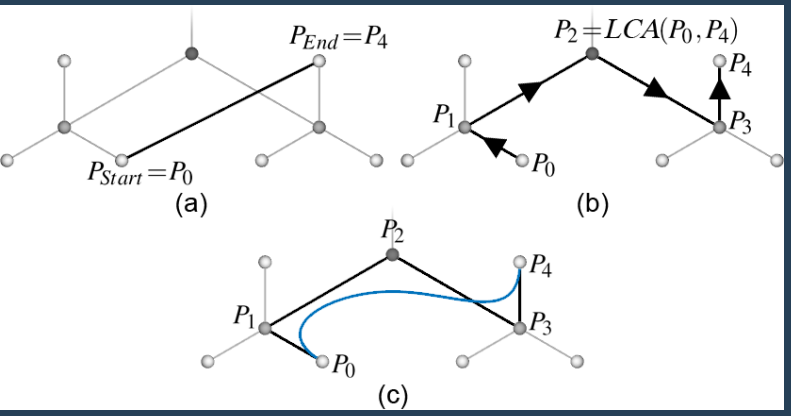


Multitouch Edge Interaction, *Schmidt et al.,* ITS 2010

Edge Bundling

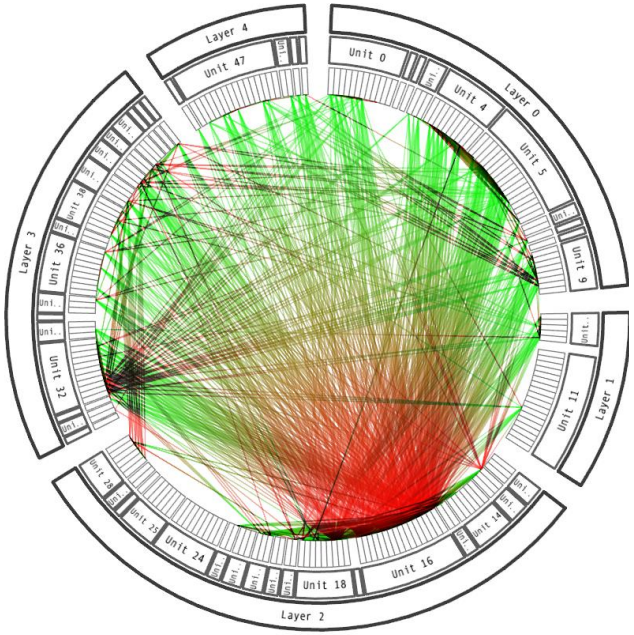


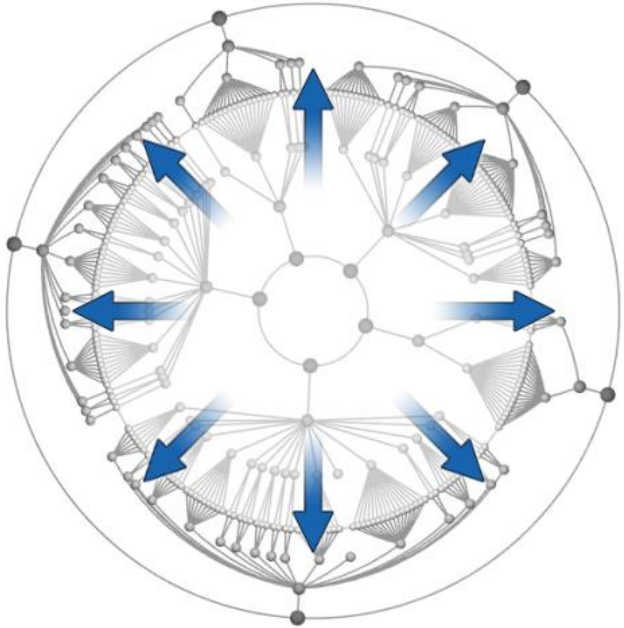
Hierarchical Edge Bundling

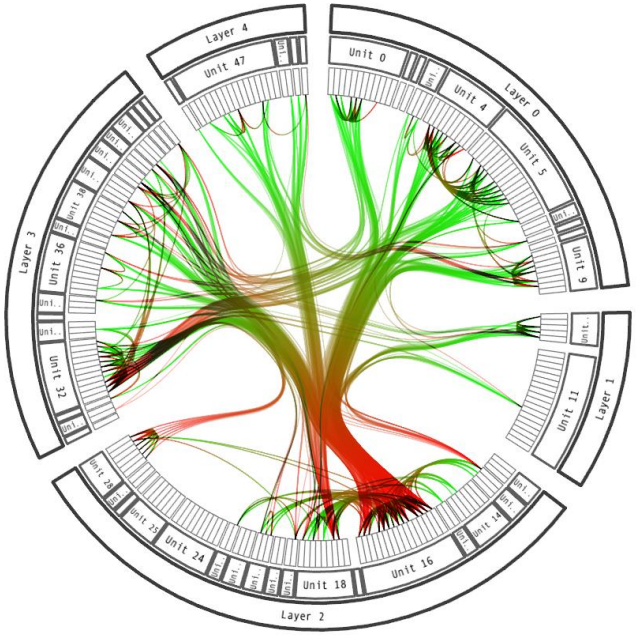


Use the hierarchy to bundle edges

Holten, Infovis 2006





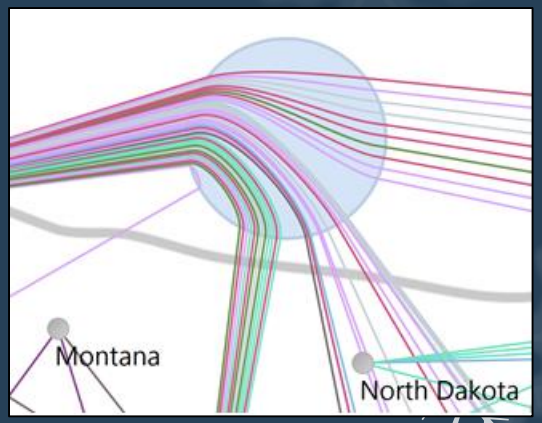


Bundle-based exploration

Recorded trajectories
Bundling and unbundling exploration

MoleView, *Hurter et al.,* Infovis 2011

Interactive Bundling



Curvature in Networks, *Henry Riche et al.,* AVI 2012

The issue of scale

Solutions

Extracting network motifs

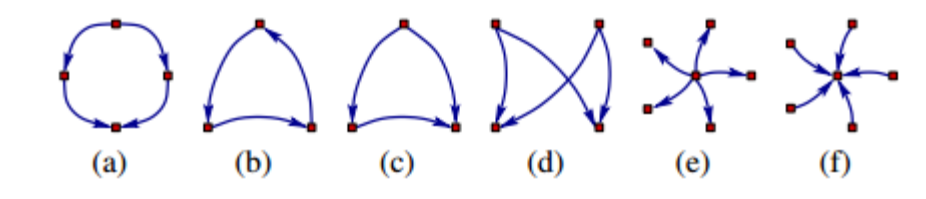
Taking advantage of node attributes

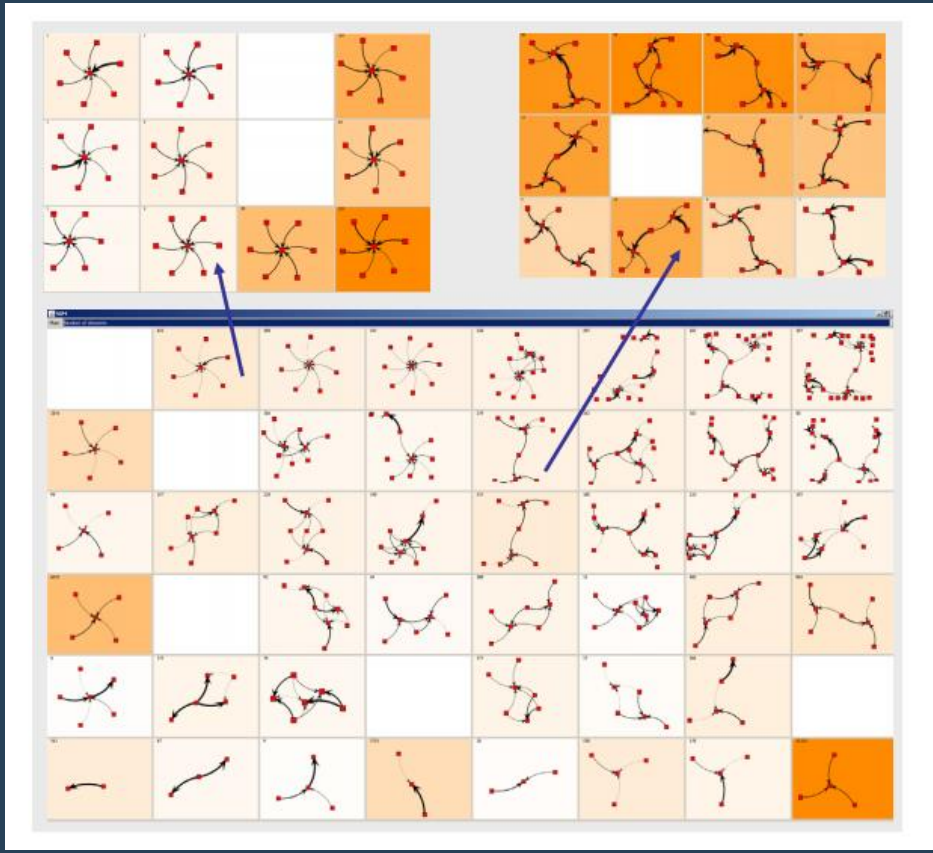
- to layout/filter
- to aggregate

Degree-of-Interest graphs

Use the alternative representation

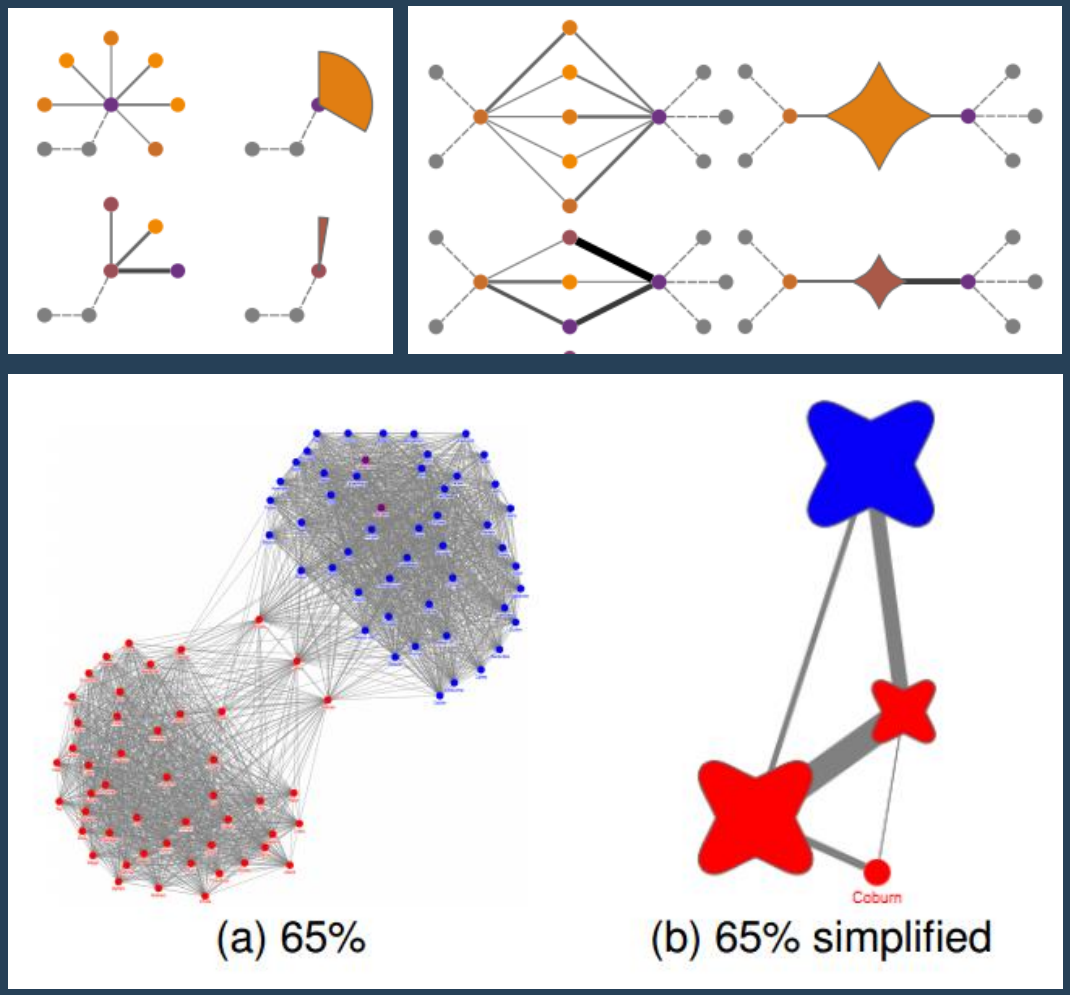
Motifs





Motifs, *von Landsberger et al.,* VAST 2009

Motifs of higher order



Motifs, *Dunne et al.,* CHI 2013

Attribute-driven layout

Large node-link diagrams get messy!

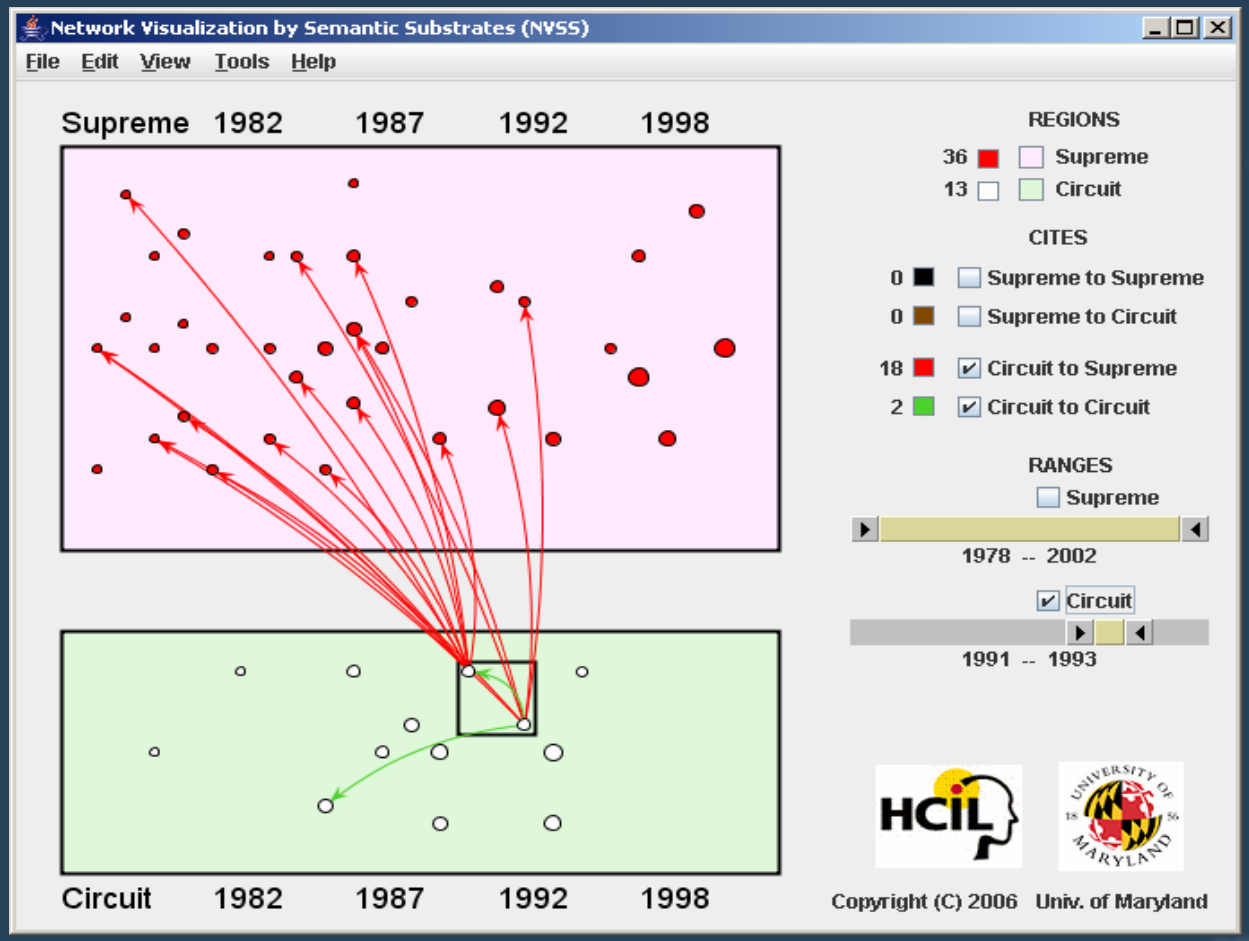
Is there additional structure we can exploit?

Idea: Use data attributes to perform layout

- e.g., scatter plot based on node values

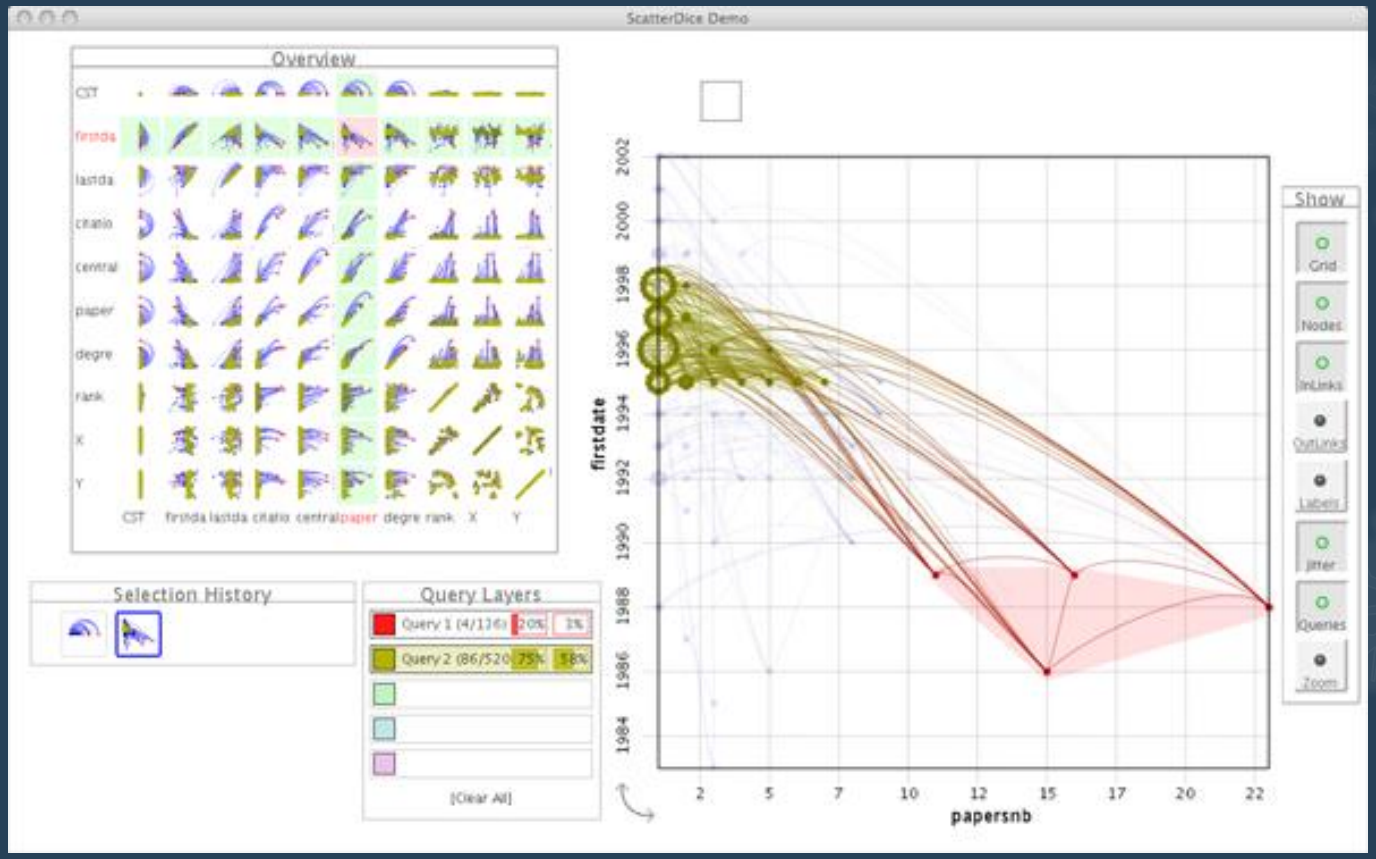
Dynamic queries and/or brushing can be used to explore connectivity

Attribute-driven layout

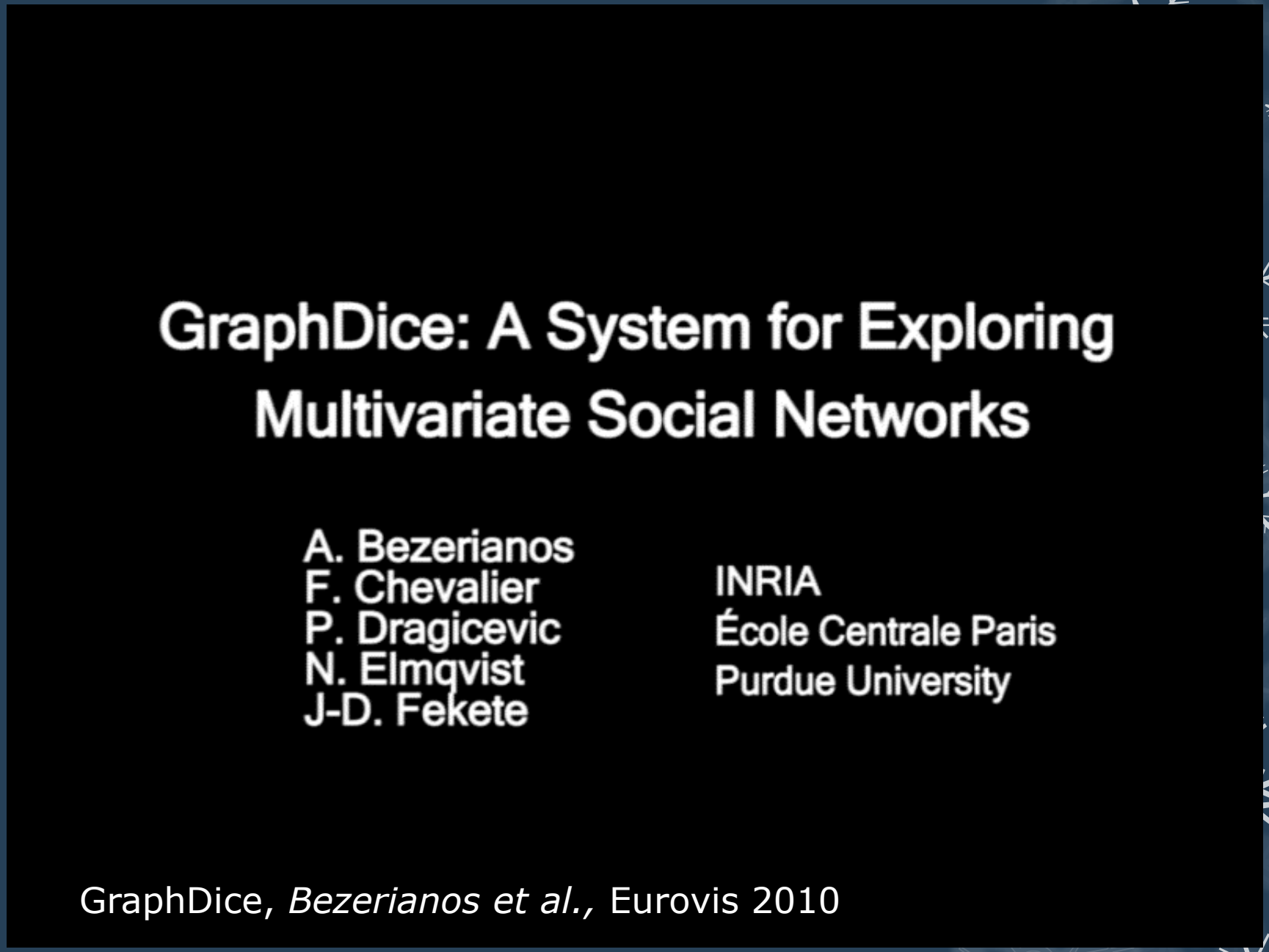


Semantic Substrates, *Shneiderman et al.*, Infovis 2006

Attribute-driven layout

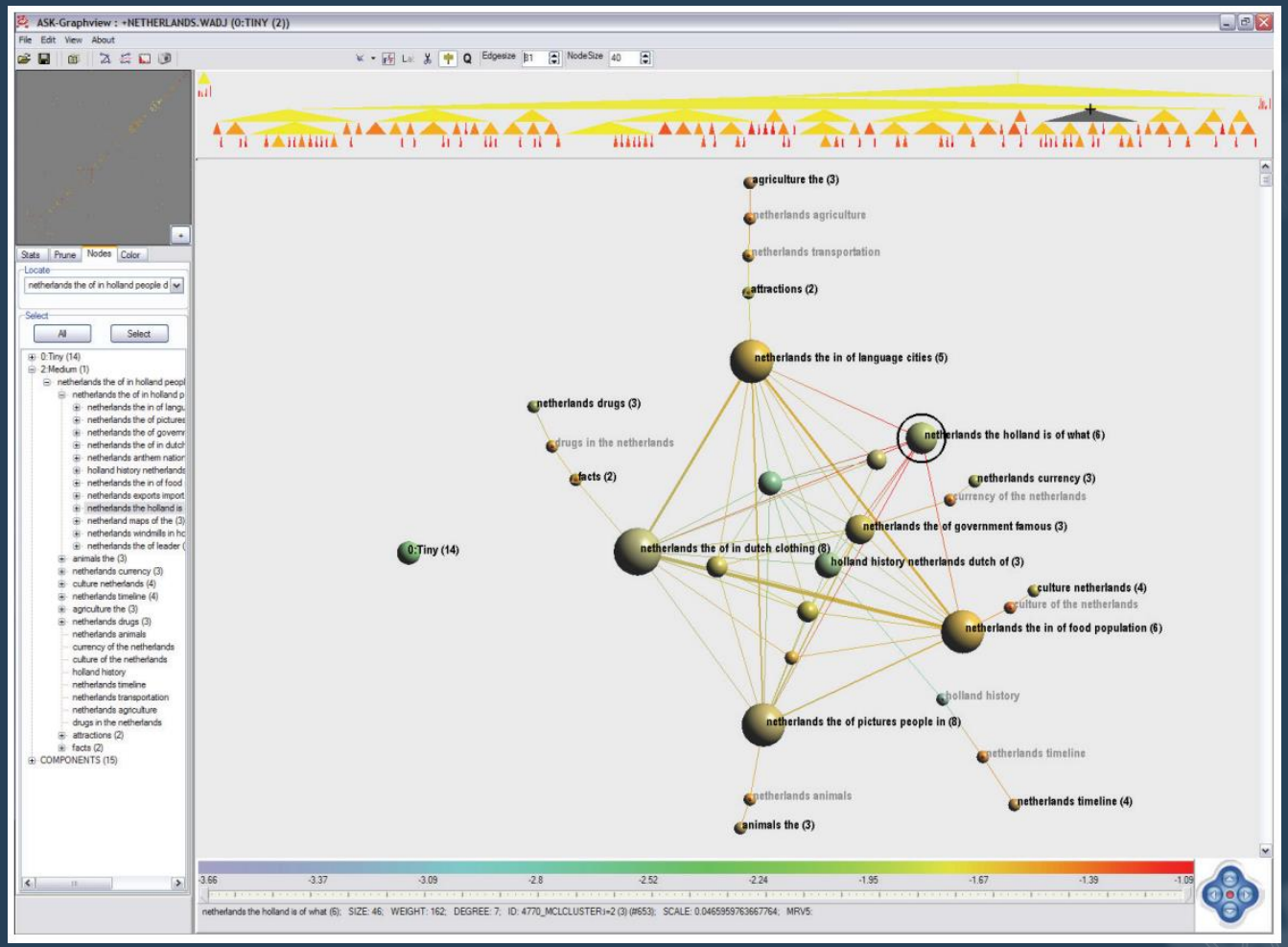


GraphDice, *Bezerianos et al.*, Eurovis 2010



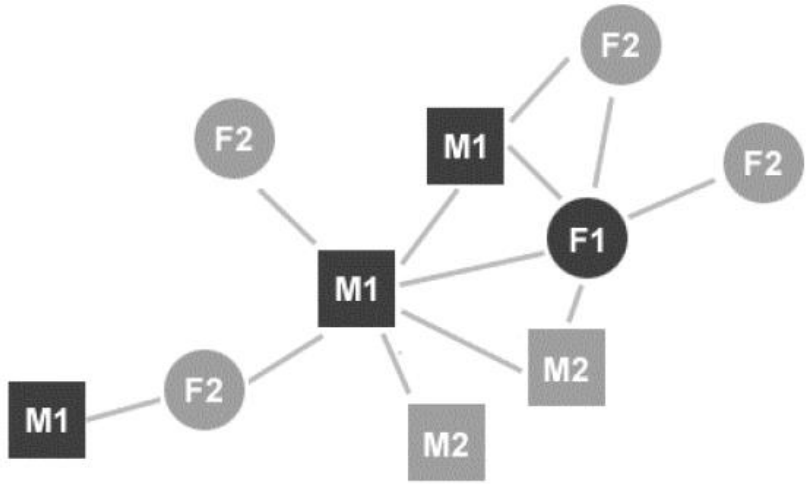
GraphDice, *Bezerianos et al.,* Eurovis 2010

Hierarchical Aggregation

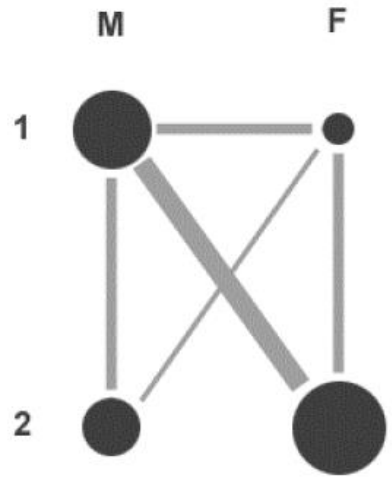


ASK-GraphView, *Abello et al.*, Infovis 2006

Attribute-driven aggregation

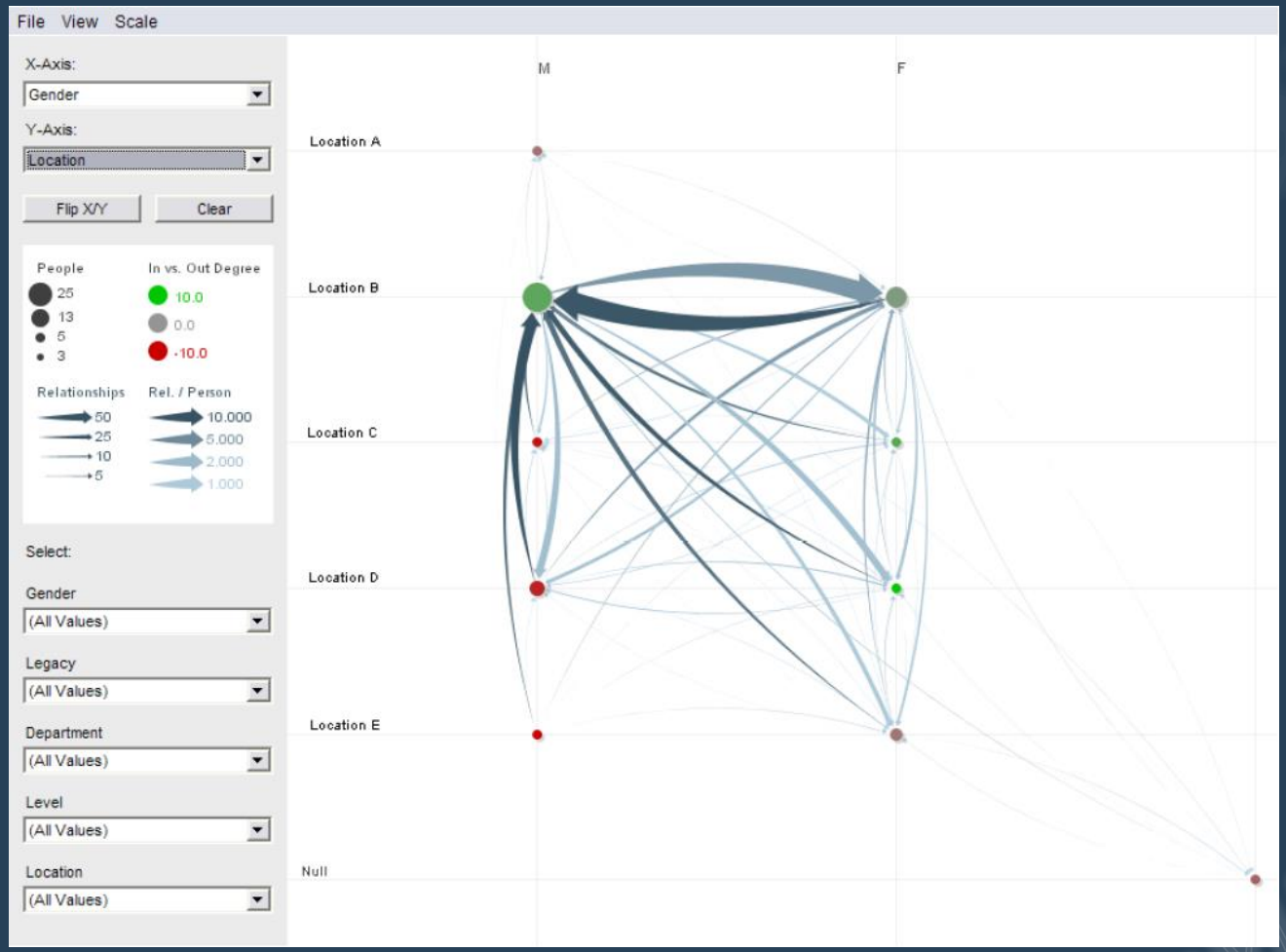


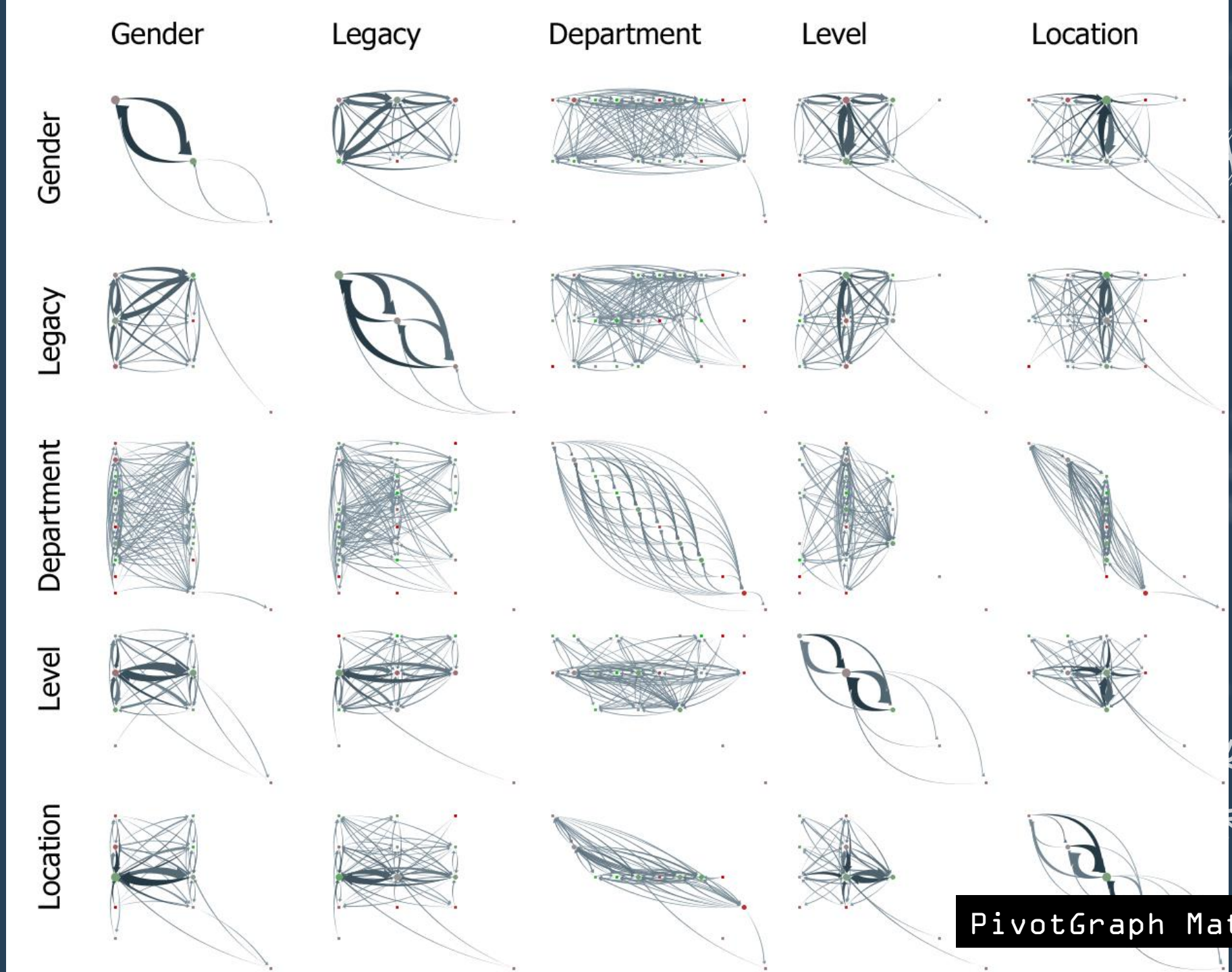
Node and Link Diagram



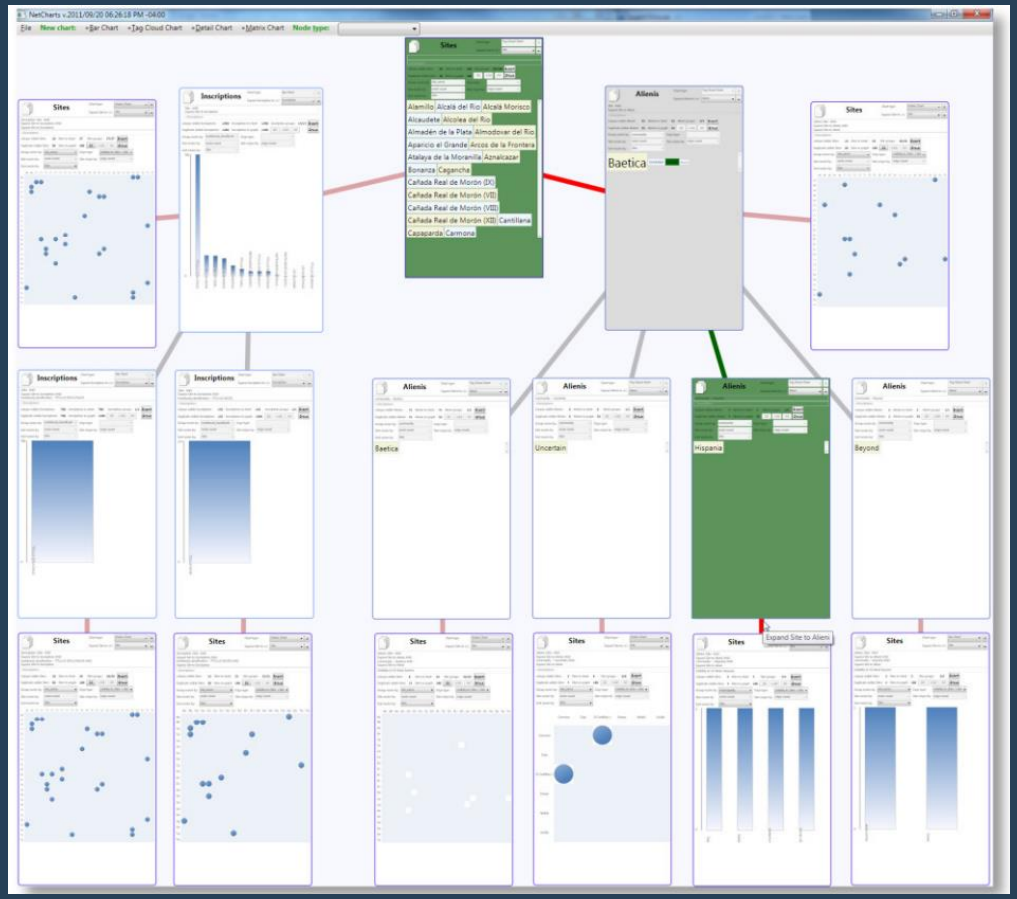
PivotGraph Roll-up

PivotGraph


File View Scale
X-Axis:
Gender
Y-Axis:
Location
Flip X/Y
Clear
People
25
13
5
3
In vs. Out Degree
10.0
0.0
-10.0
Relationships
50
25
10
5
Rel. / Person
10.000
5.000
2.000
1.000
Select:
Gender
(All Values)
Legacy
(All Values)
Department
(All Values)
Level
(All Values)
Location
(All Values)
M
F
Location A
Location B
Location C
Location D
Location E
Null

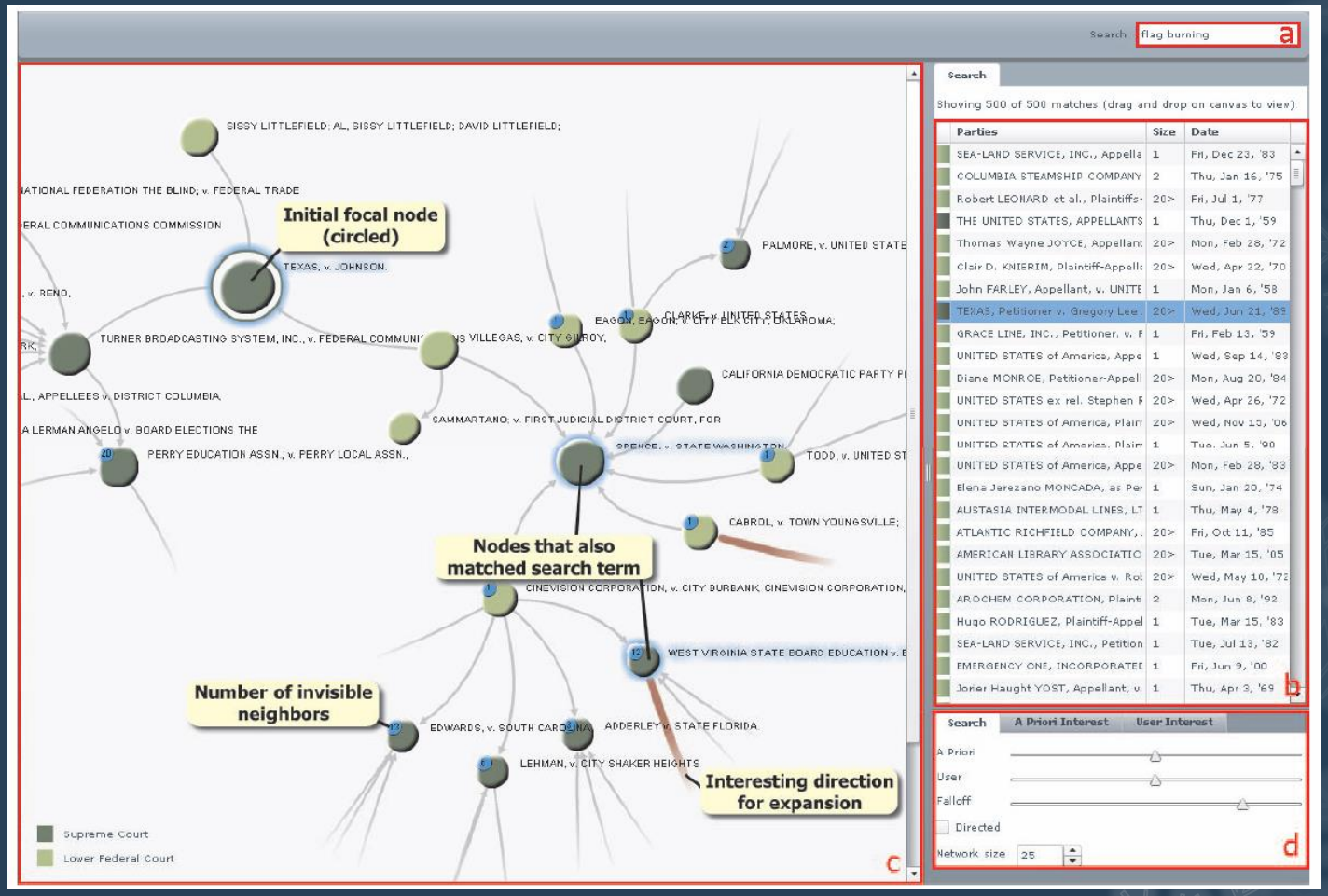

Gender
Legacy
Department
Level
Location
Gender
Legacy
Department
Level
Location
PivotGraph Matr

GraphTrail



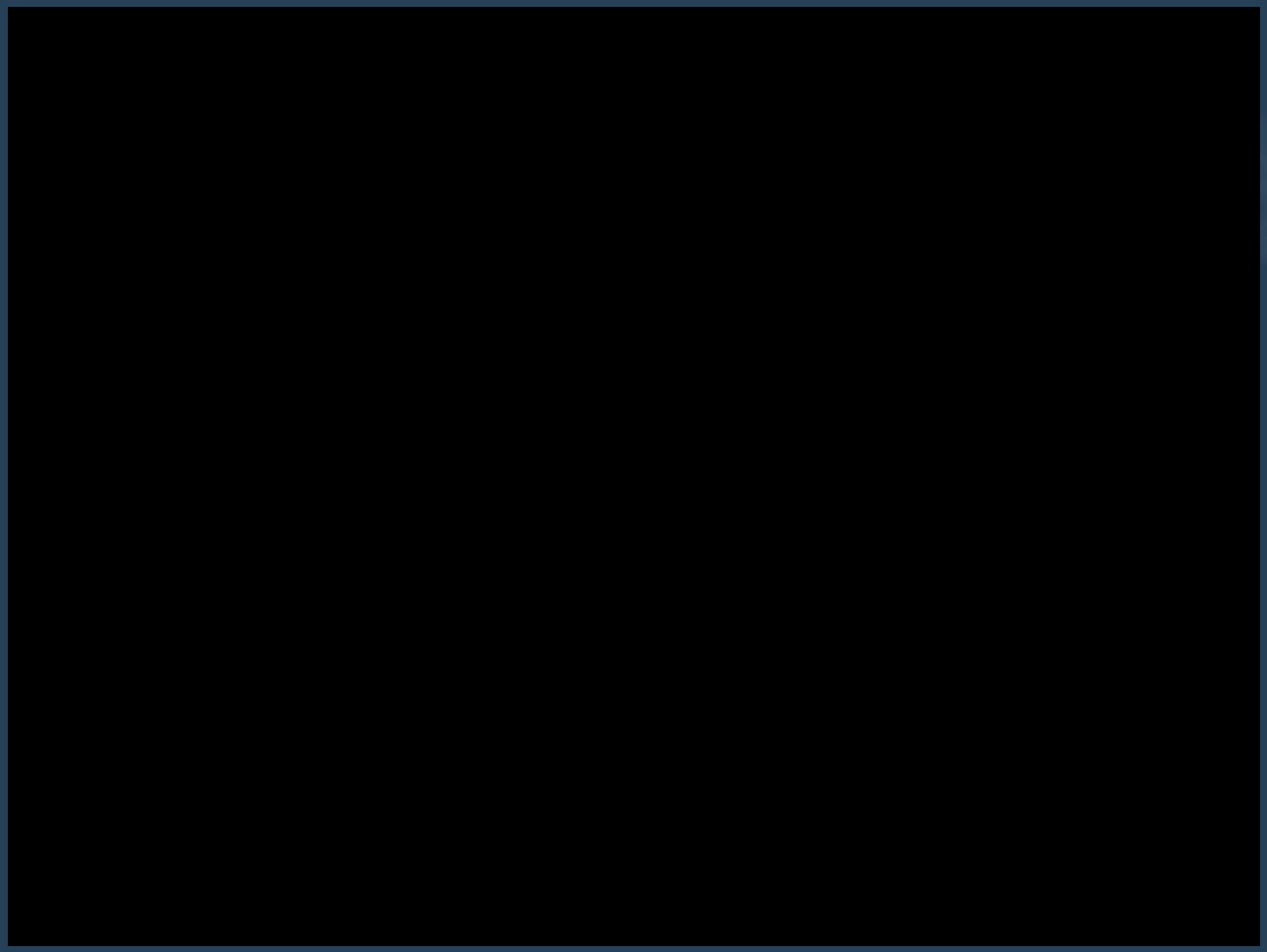
GraphTrail, *Dunne et al.*, CHI 2012

Degree-of-Interest Graphs



Search, Show Context, Expand, *Perer et al.*, Infovis 2009

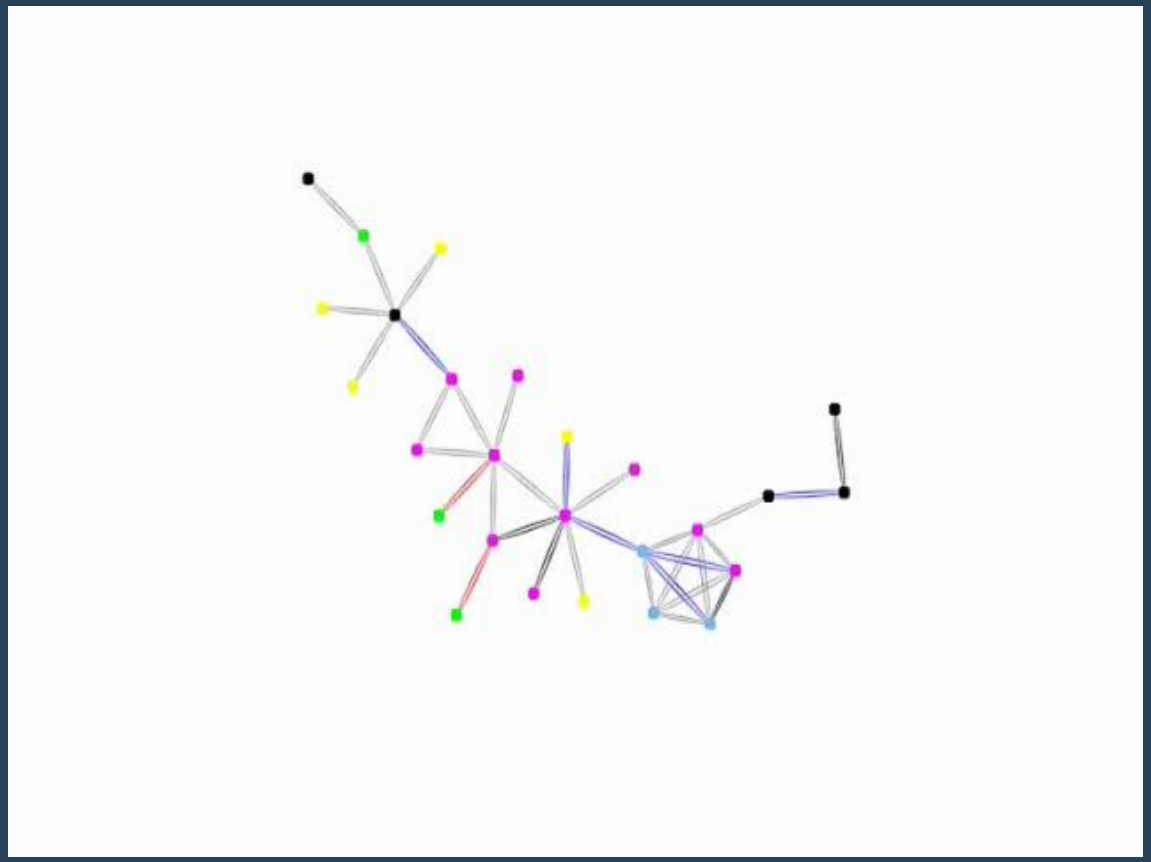
Search & Browse



PivotPaths, *Doerk et al.,* Infovis 2012

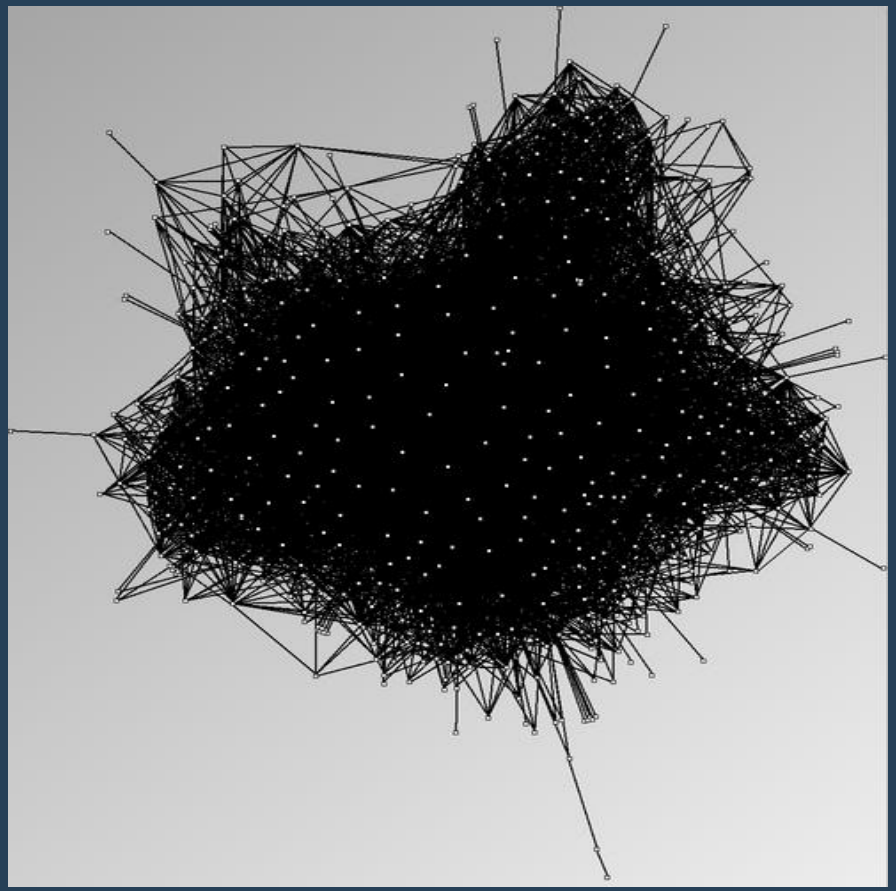
Use the Alternative

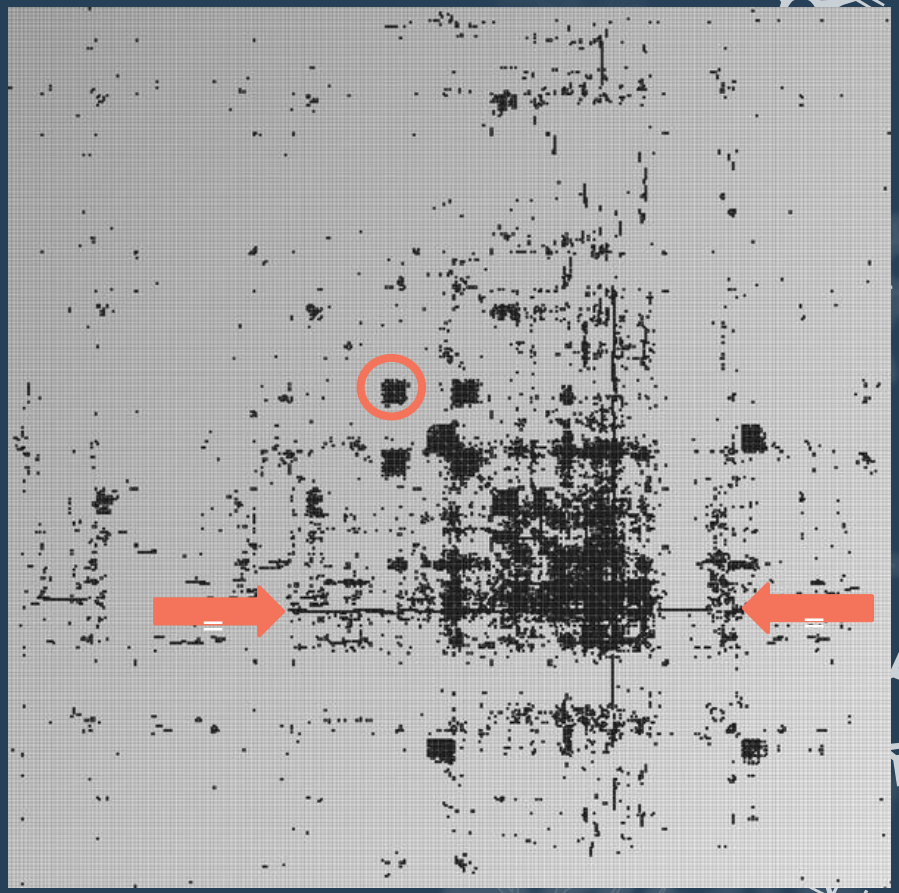
Matrices



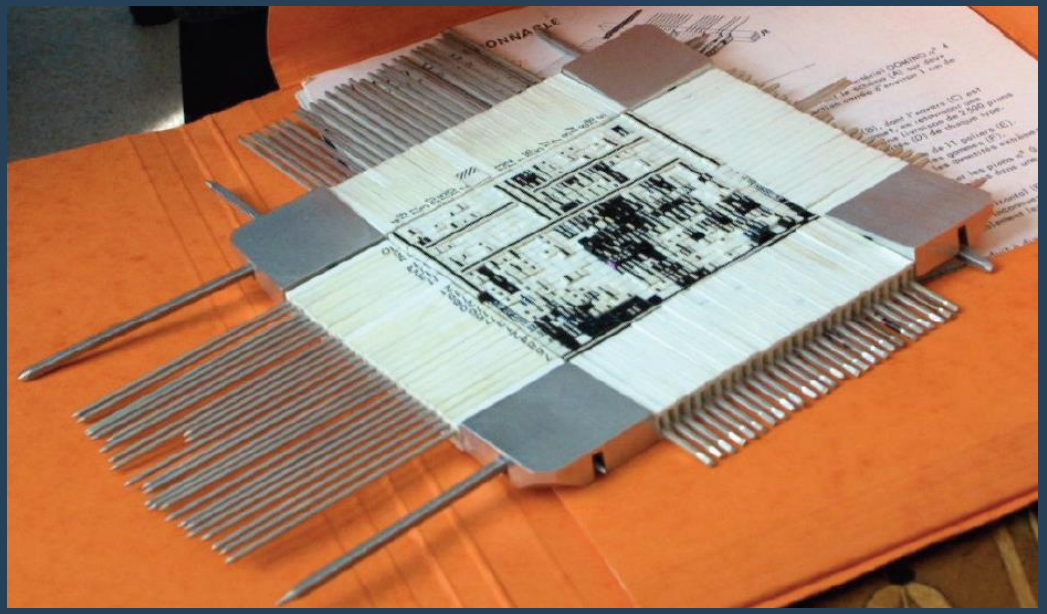
Matrices

one year of email between ~500 researchers

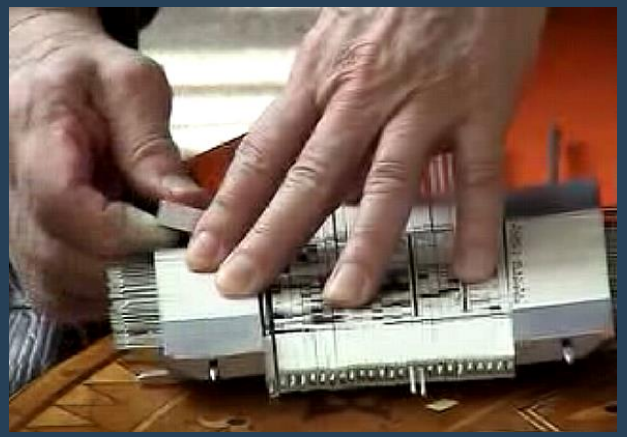




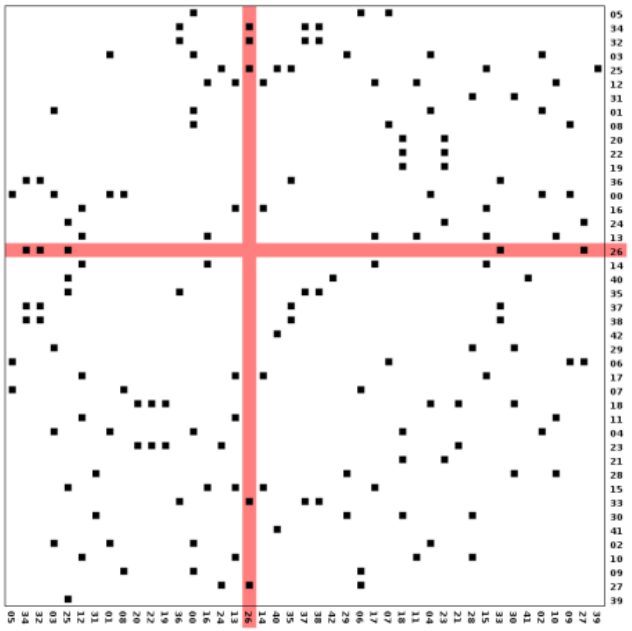
The Reorderable Matrix

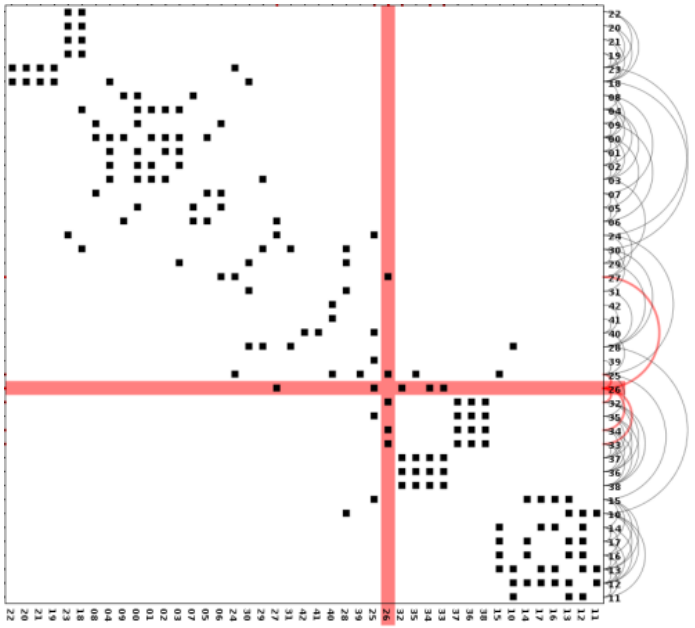


Jacques Bertin 1967



Revealing patterns



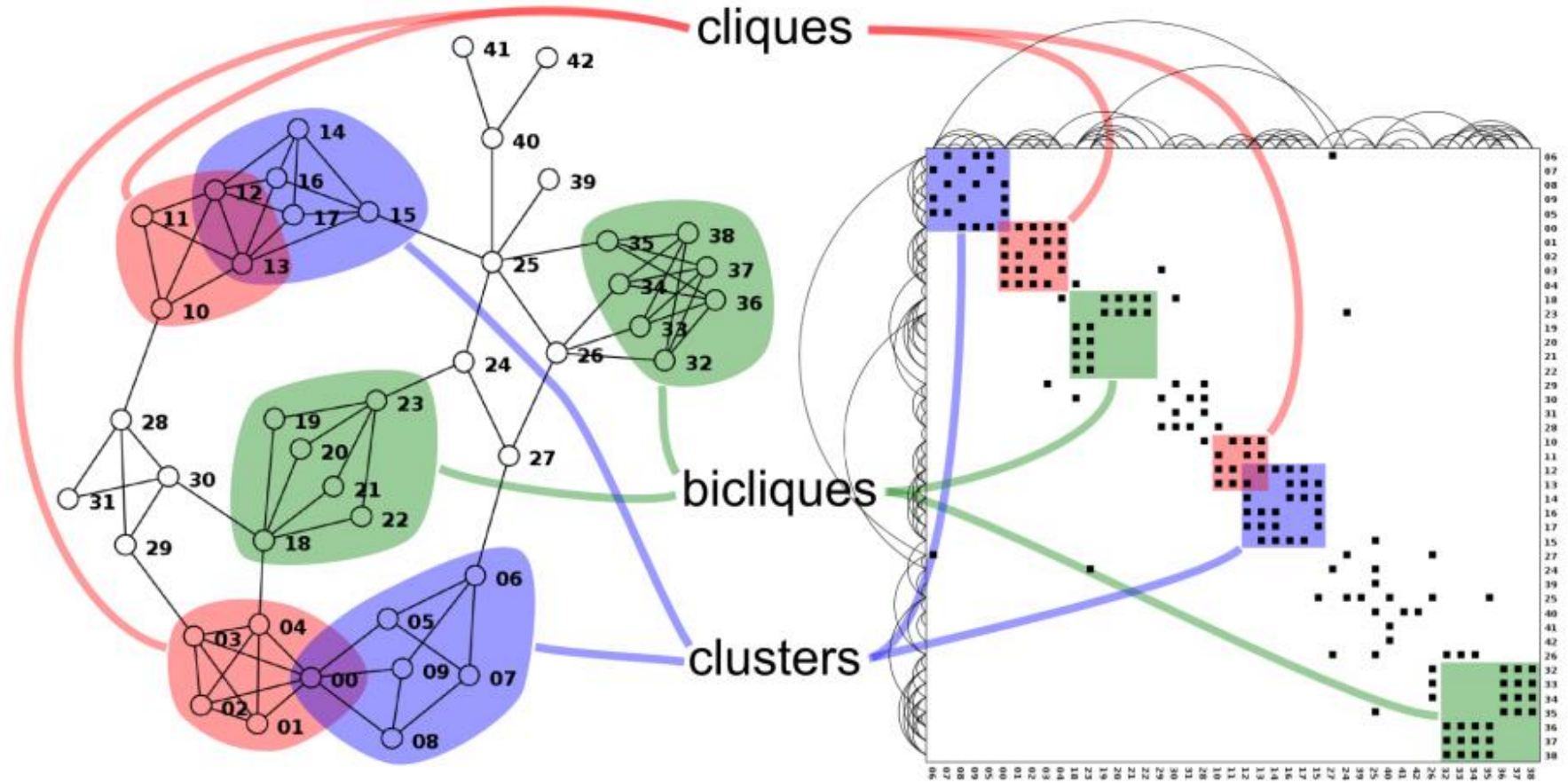


Matrix vs Node-Link

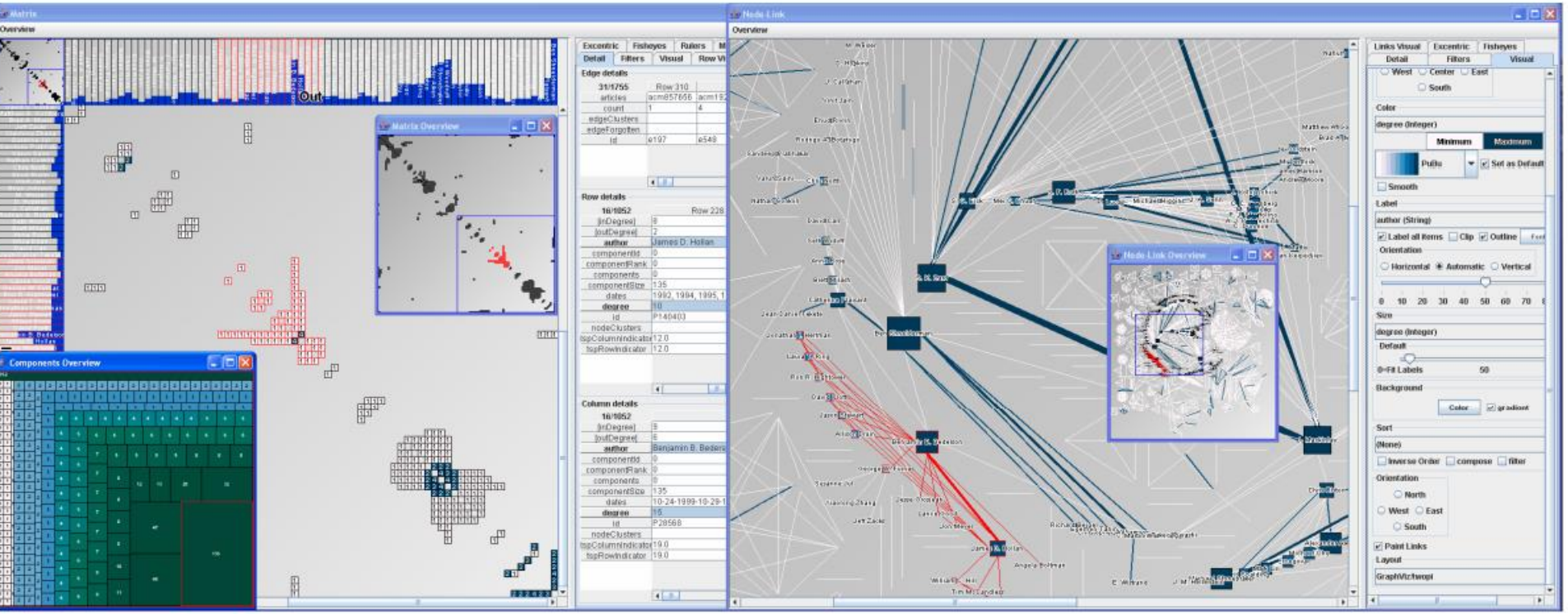
Matrix	Node-Link
× Require learning	✓ Familiar
✓ No overlap	× Node overlap
✓ No crossings	× Link crossing
× Use a lot of space	✓ More compact
✓ Dense graphs	× Dense graphs
× Sparse graphs	✓ Sparse graphs

Comparison Study, *Ghoniem et al.,* Information Visualization Journal 2005

Learning phase

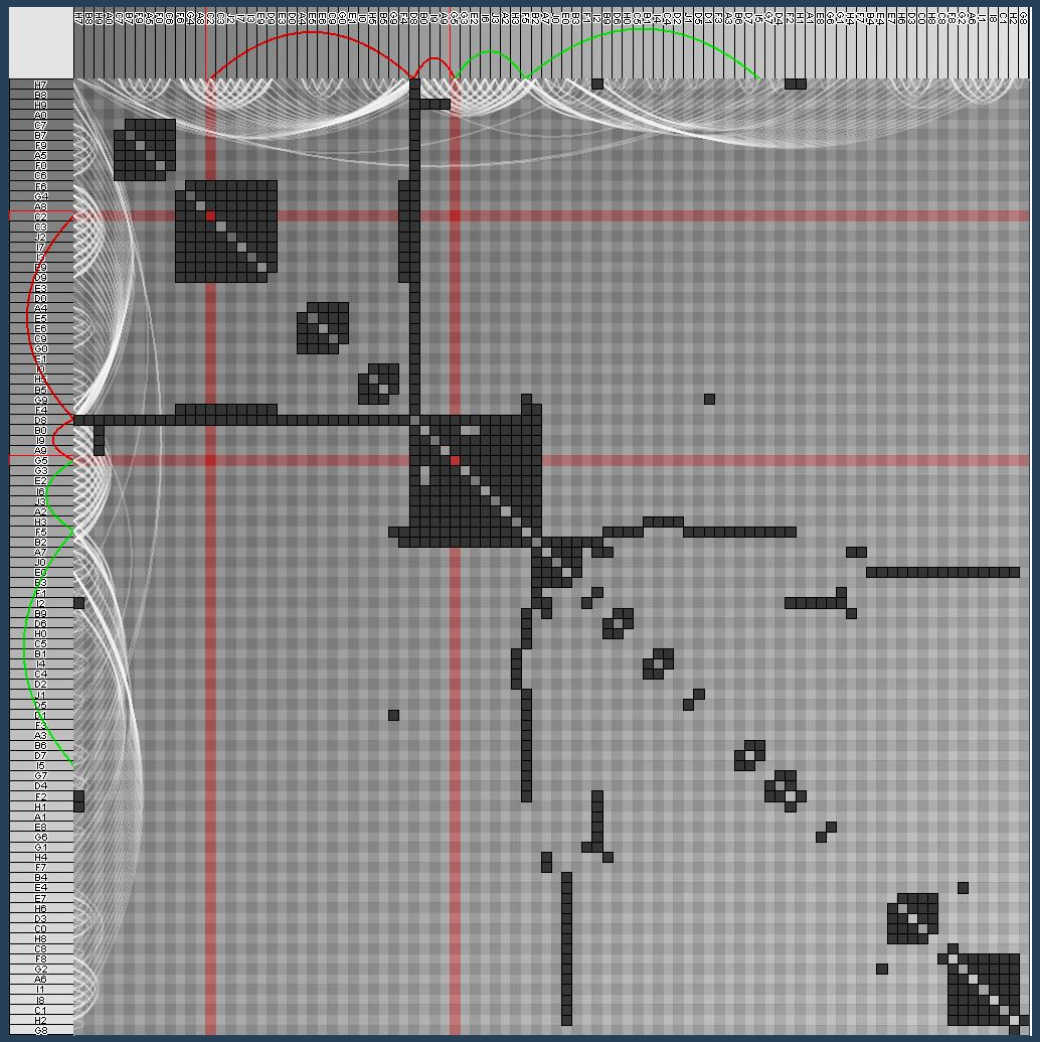


Matrix + Node-Link

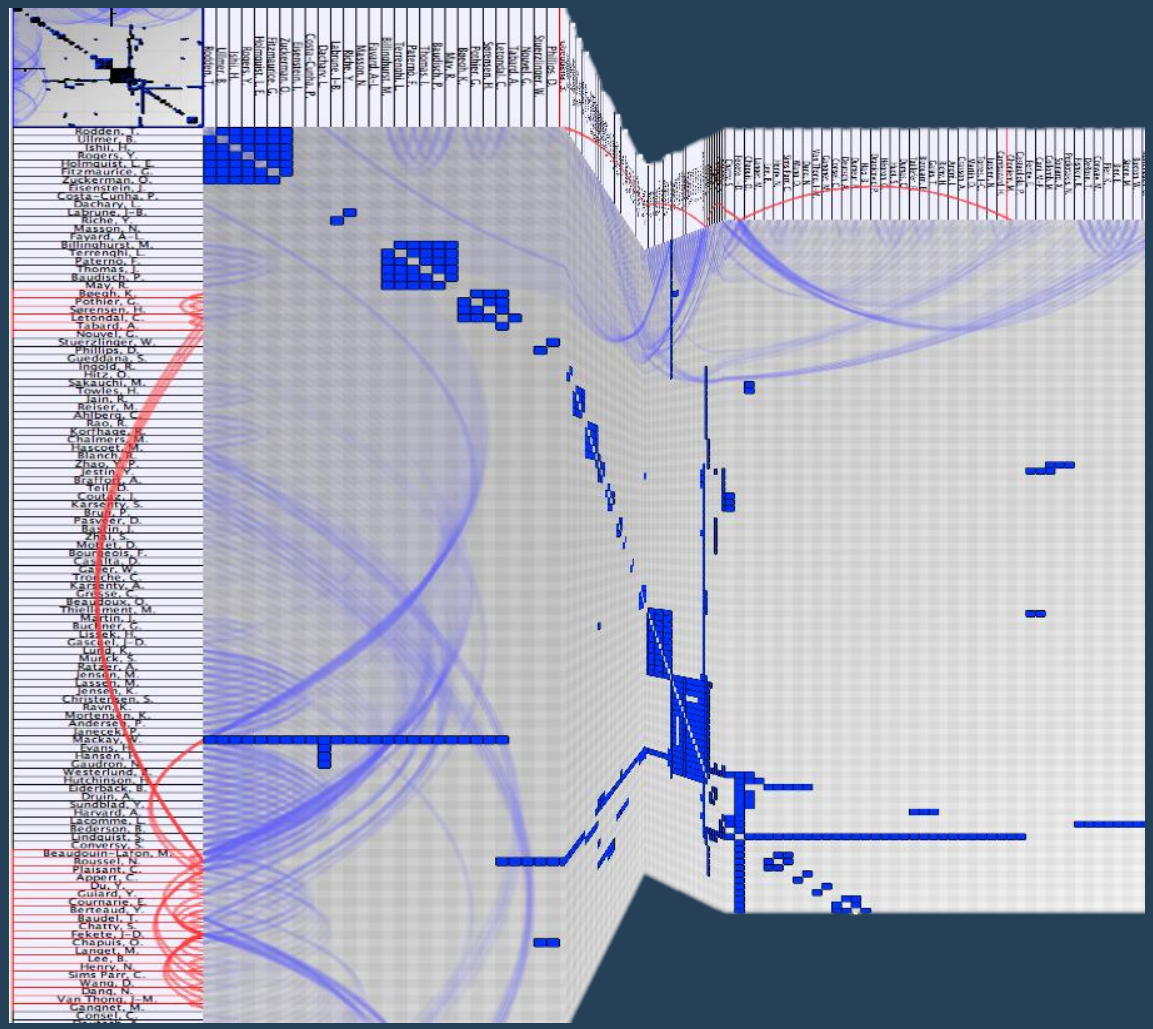


MatrixExplorer, *Henry et al.*, Infovis 2006

Following paths in Matrices

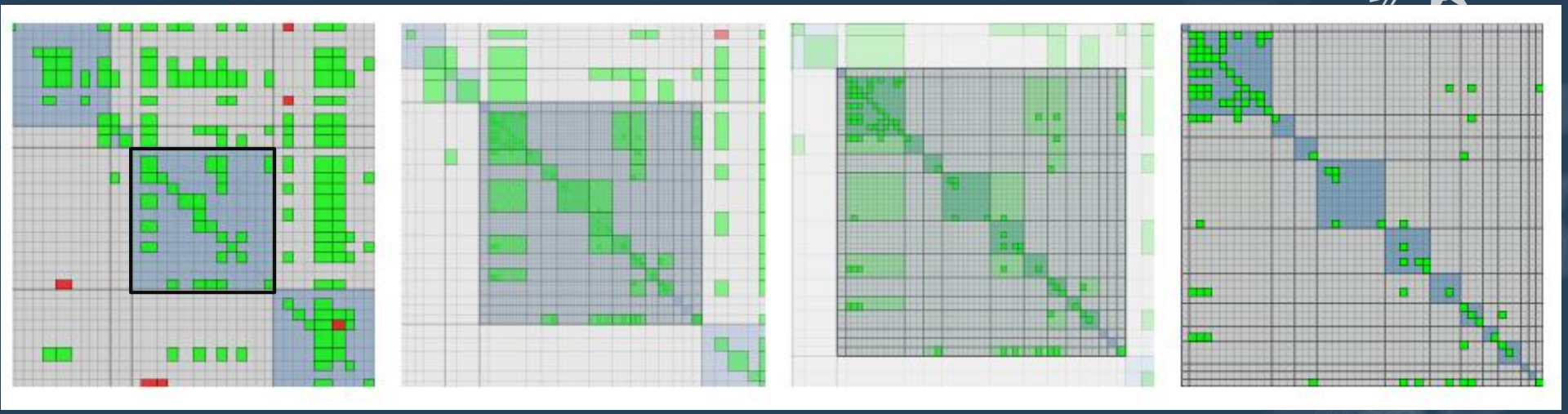


Navigation Techniques



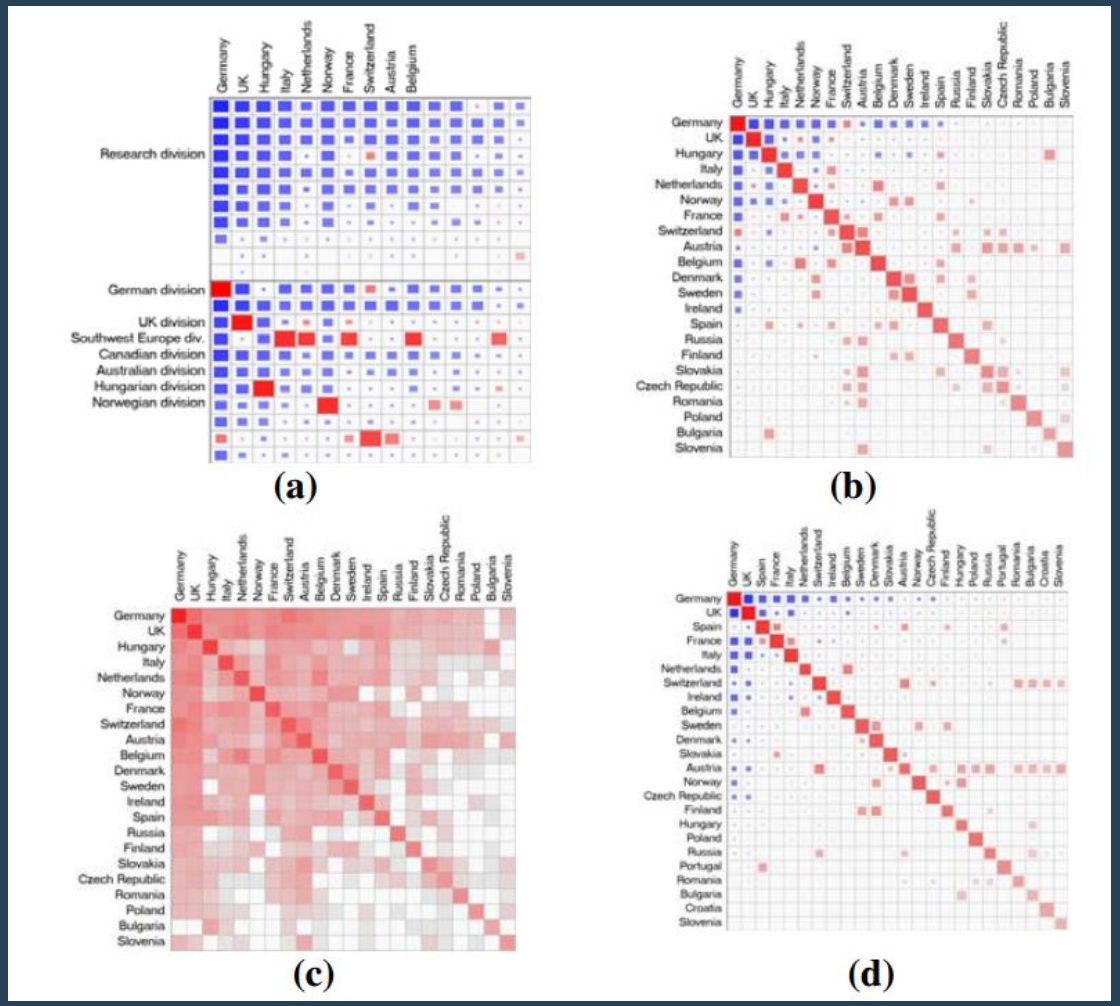
Melange, *Elmqvist et al.*, CHI 2008

Hierarchical Aggregation



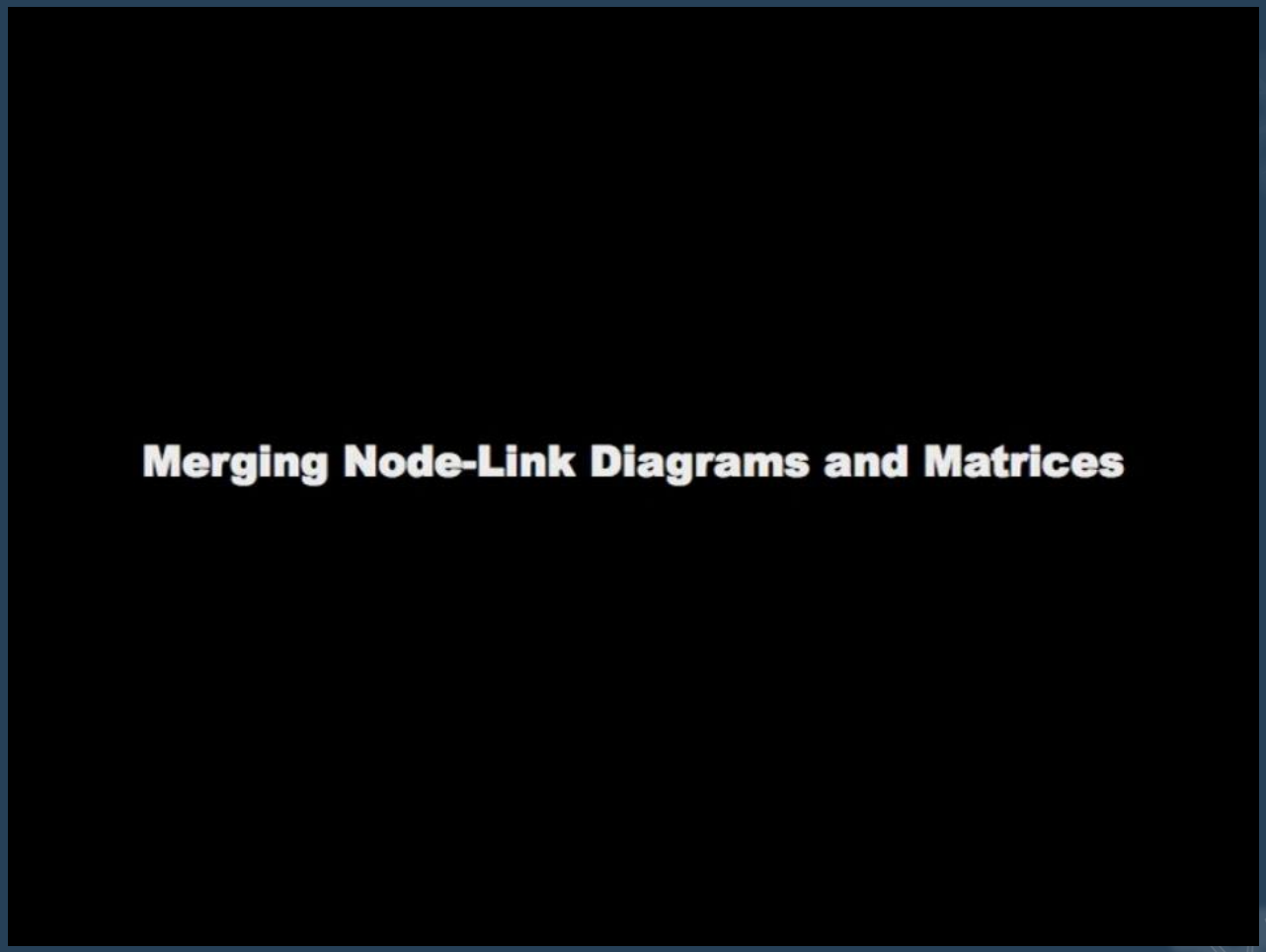
MatrixZoom, *van Ham,* Infovis 2003

Attribute-driven Aggregation



Honeycomb, *van Ham et al.,* Interact 2009

Hybrid Graph Representation



NodeTrix, *Henry et al.,* Infovis 2008

Hybrid Graph Representation

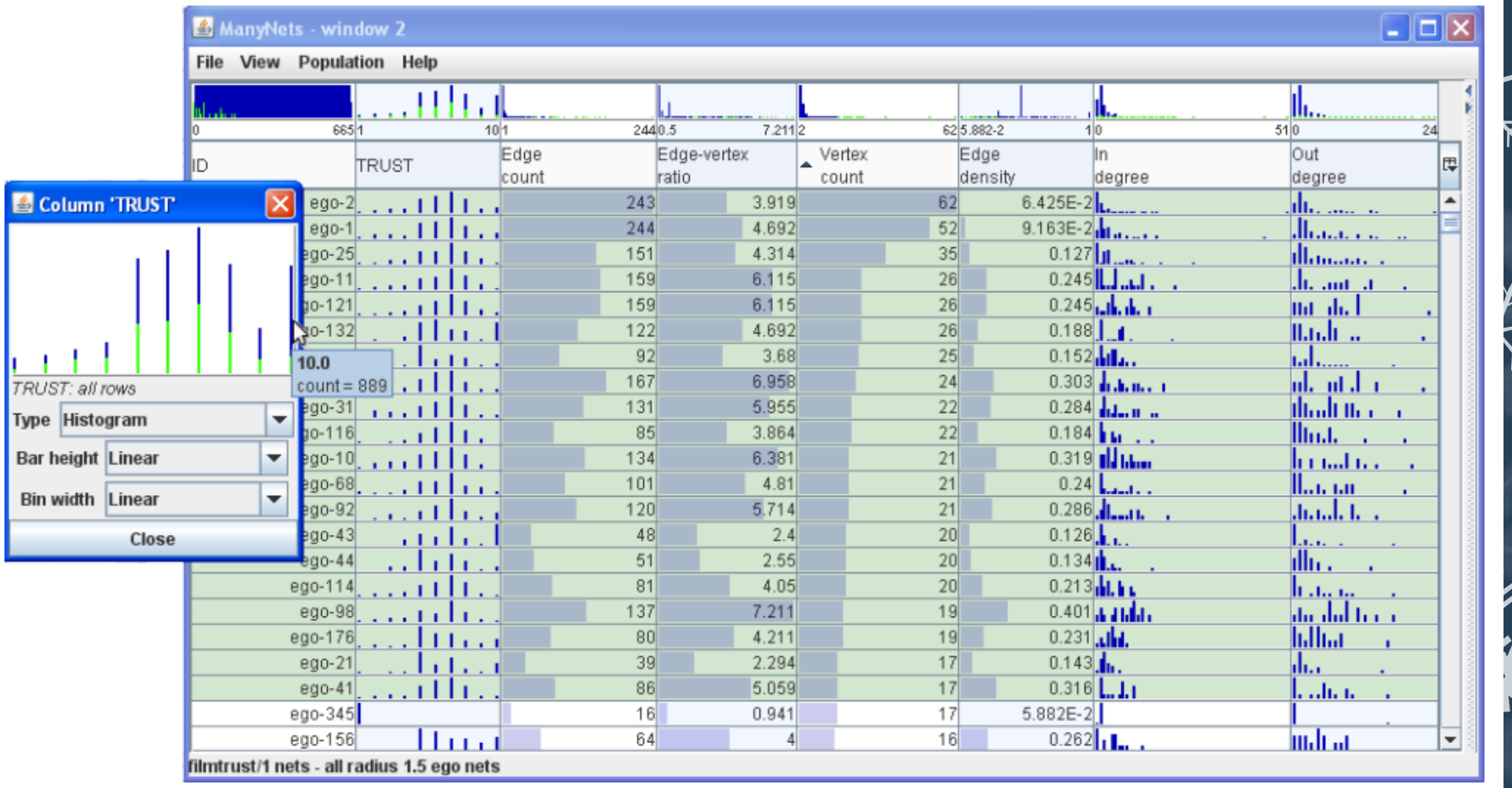
Active research topics

Network Comparison

Heterogeneous Networks

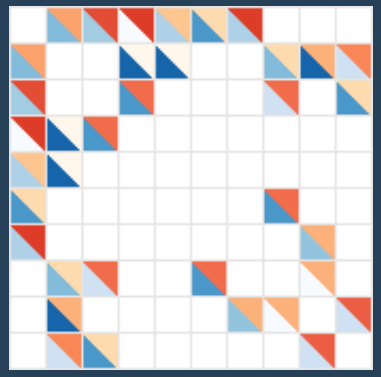
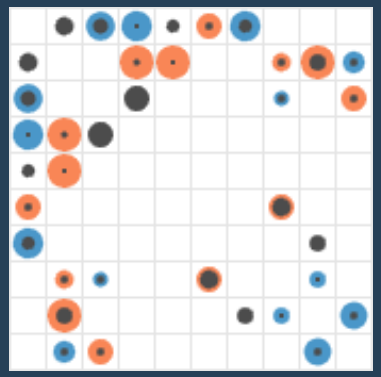
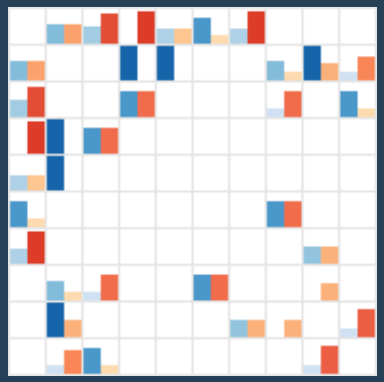
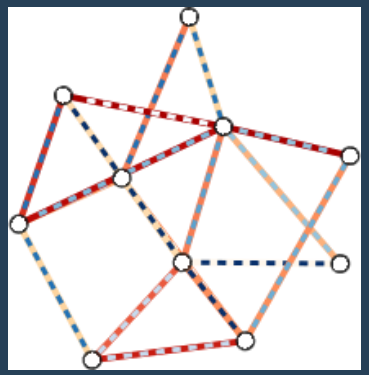
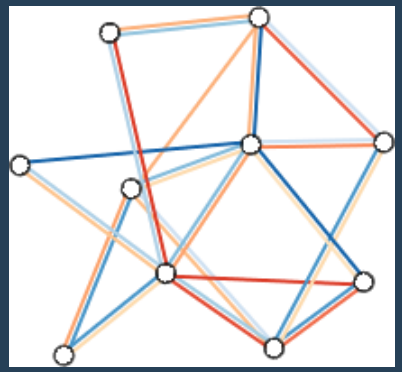
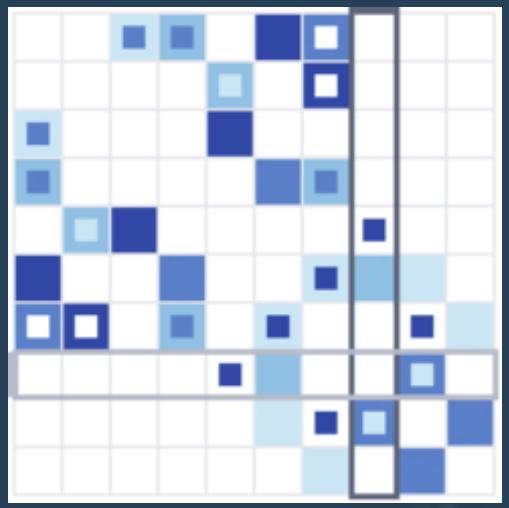
Dynamic Networks

Comparing Networks



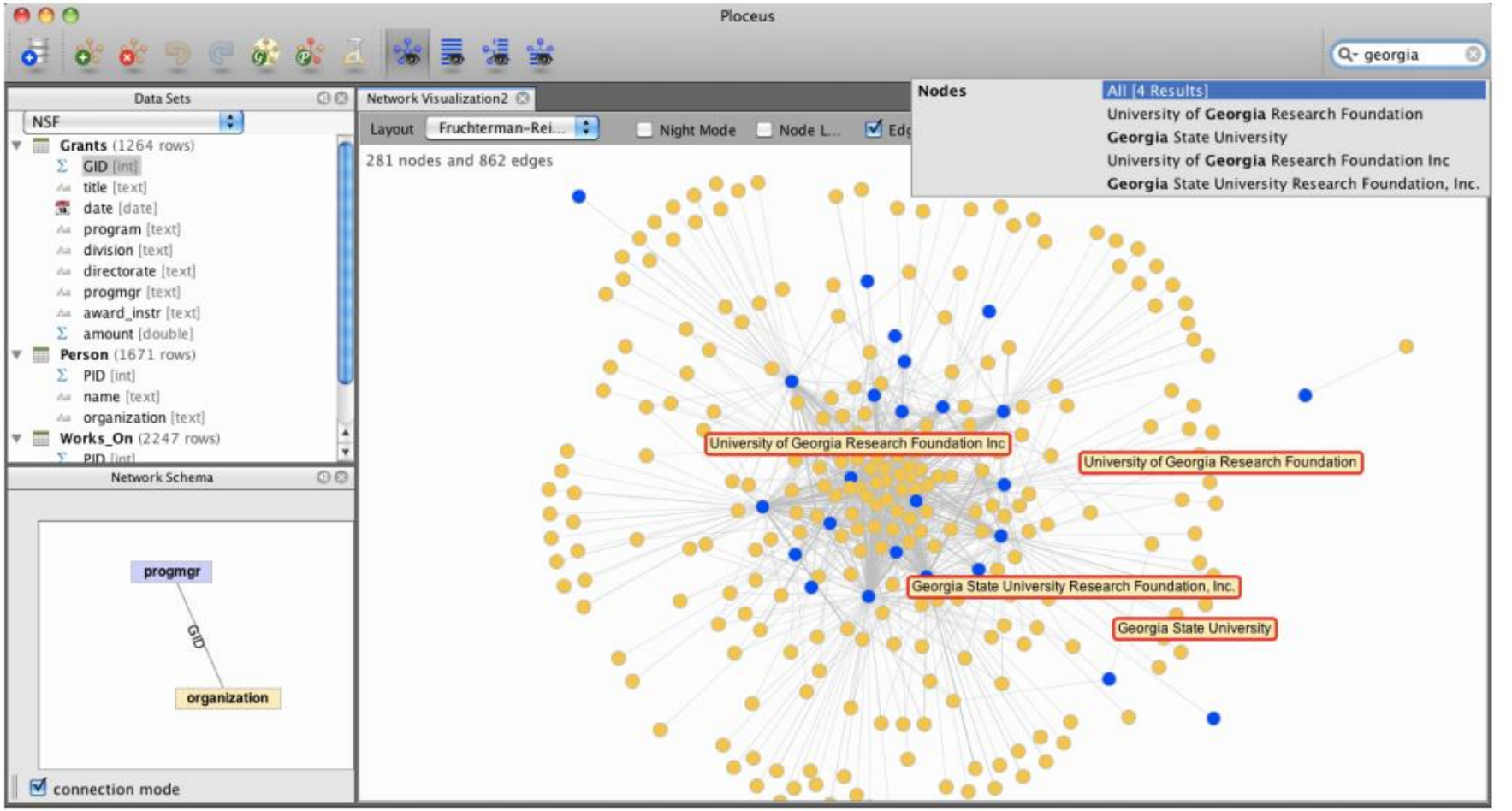
ManyNets, *Freire et al.*, CHI 2010

Comparing Weigthed Networks

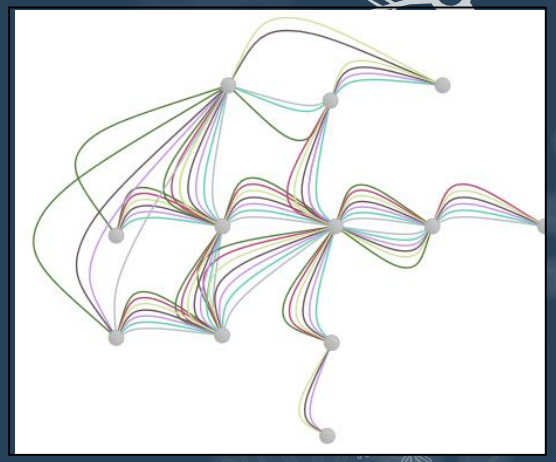
Comparison Matrix, *Alper et al.*, CHI 2013

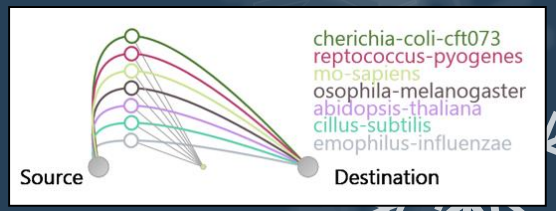
Heterogeneous Networks



Heterogeneous Networks

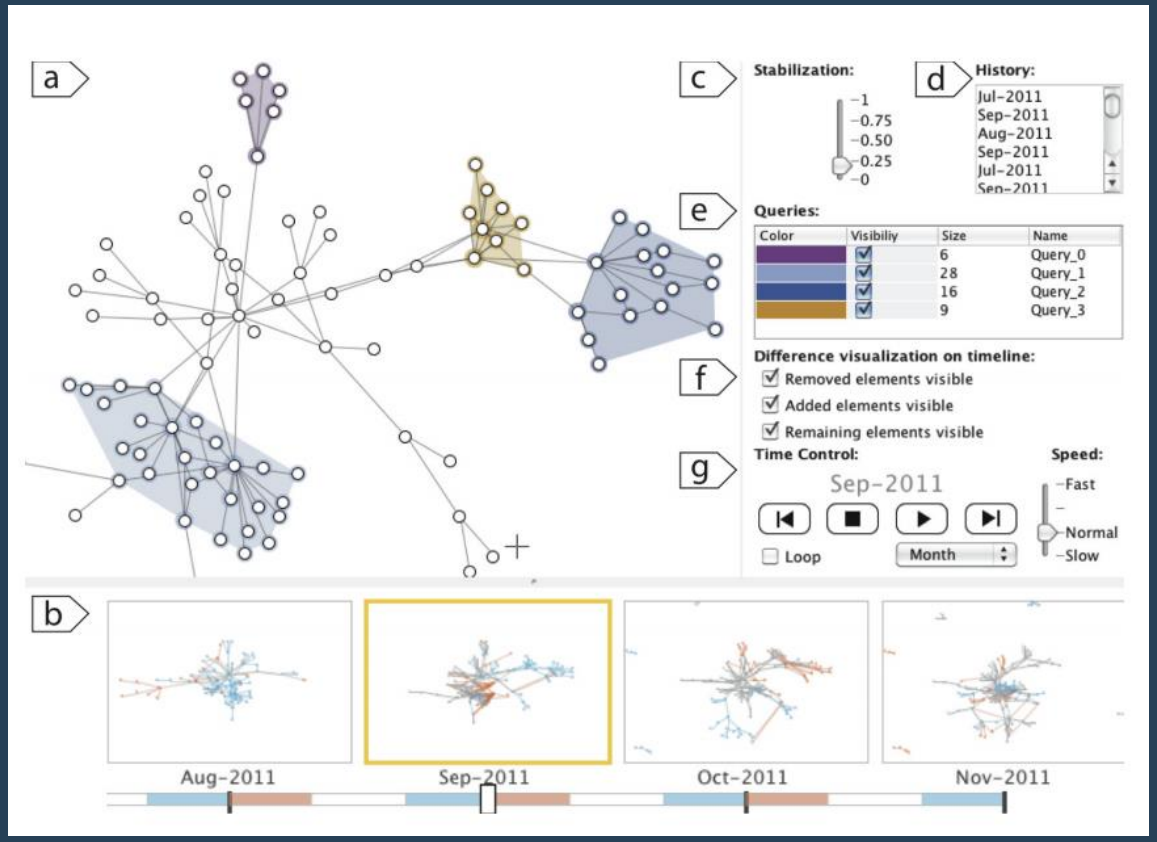
Interactive Link Legends





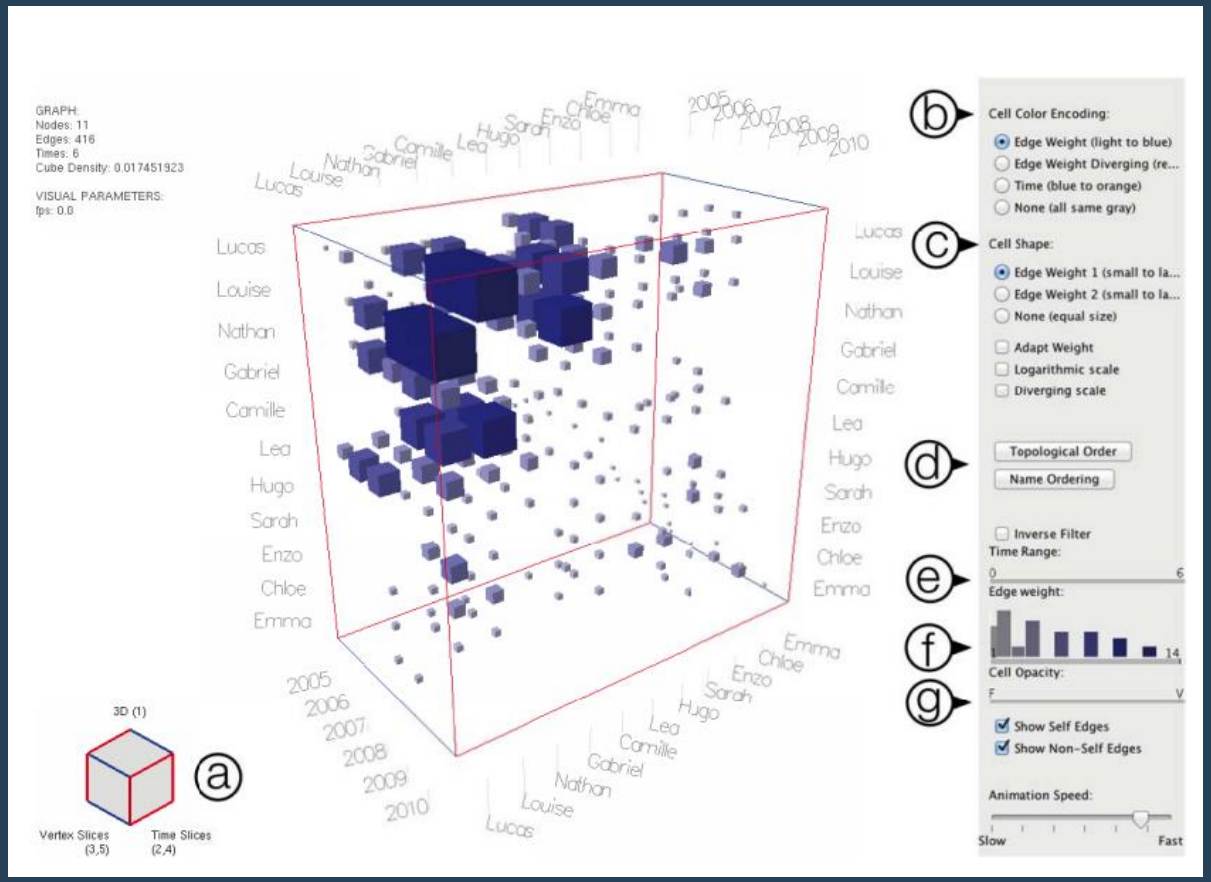
Curvature in Networks, *Henry Riche et al.,* AVI 2012

Dynamic graphs



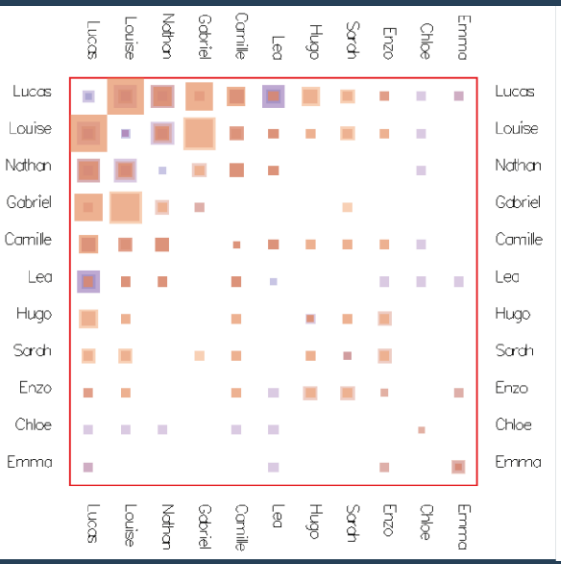
Graph Diaries, *Bach et al.*, TVCG 2013

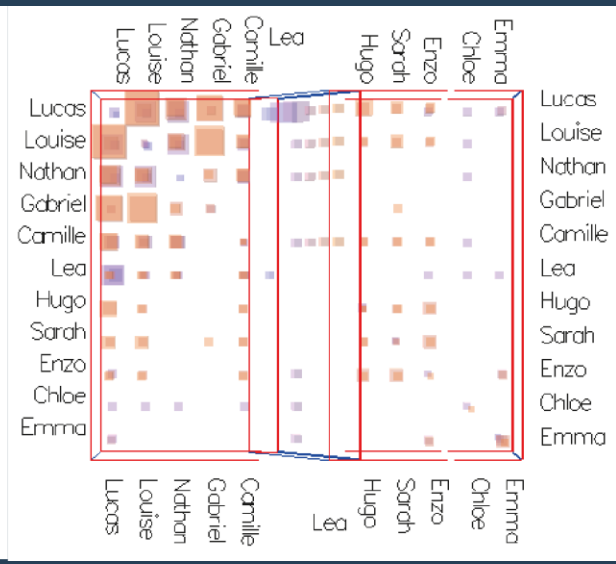
Dynamic graphs

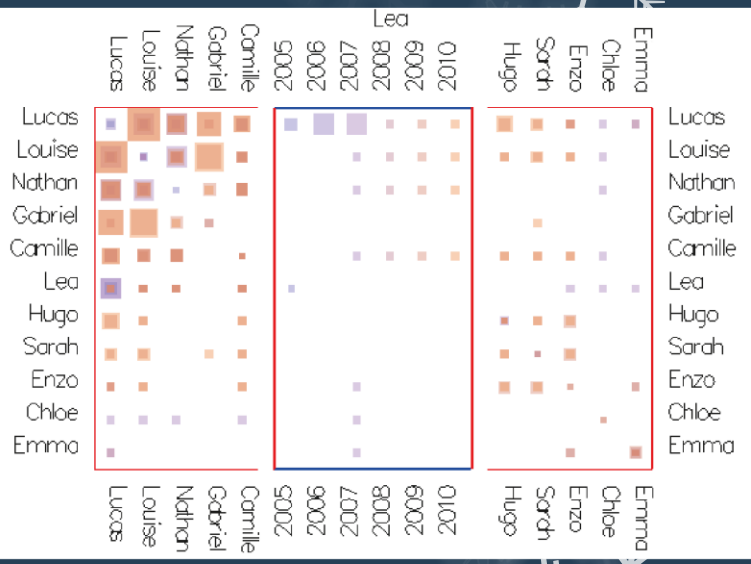


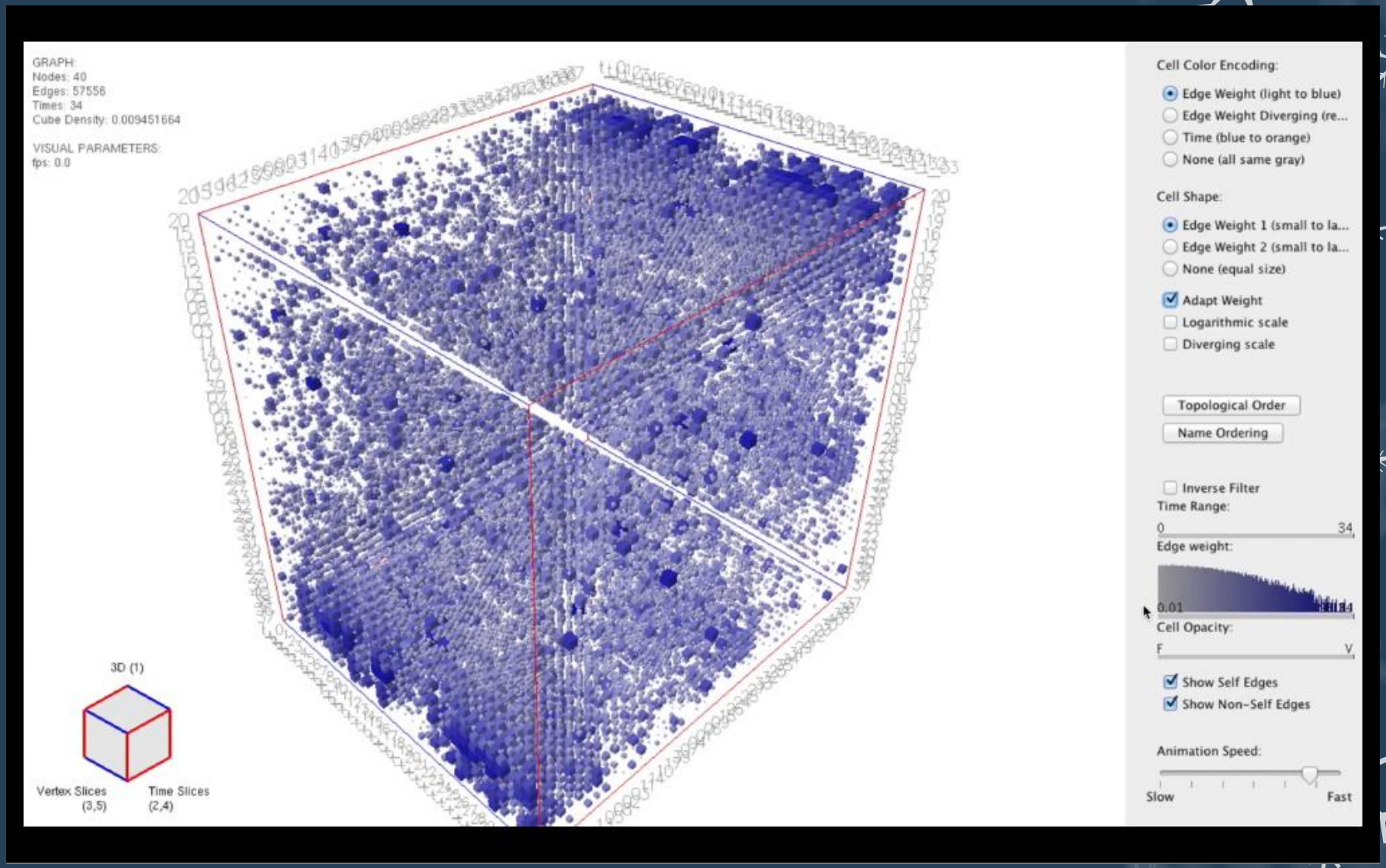
Cubix, *Bach et al.*, CHI 2014

Dynamic graphs








GRAPH:
Nodes: 40
Edges: 57556
Times: 34
Cube Density: 0.009451664
VISUAL PARAMETERS:
fps: 0.0
3D (1)
Vertex Slices
(3,5)
Time Slices
(2,4)
Cell Color Encoding:
Edge Weight (light to blue)
Edge Weight Diverging (re...
Time (blue to orange)
None (all same gray)
Cell Shape:
Edge Weight 1 (small to la...
Edge Weight 2 (small to la...
None (equal size)
Adapt Weight
Logarithmic scale
Diverging scale
Topological Order
Name Ordering
Inverse Filter
Time Range:
0
34
Edge weight:
0.01
Cell Opacity:
F
V
Show Self Edges
Show Non-Self Edges
Animation Speed:
Slow
Fast

Summary

Fast algorithms exists for tree visualizations

While most familiar representations, node-link diagrams have many issues

- Several can be fixed by interaction techniques

- Others require using different visualization paradigms, such as matrices

Graph visualization is still an active research topic!!