2D Visualization Techniques: Taxonomy and overview

Lyn Bartram

IAT 814 week 10

10.11.2009

These slides have been largely adapted from D. Keim, B. Zupan and M. Hearst

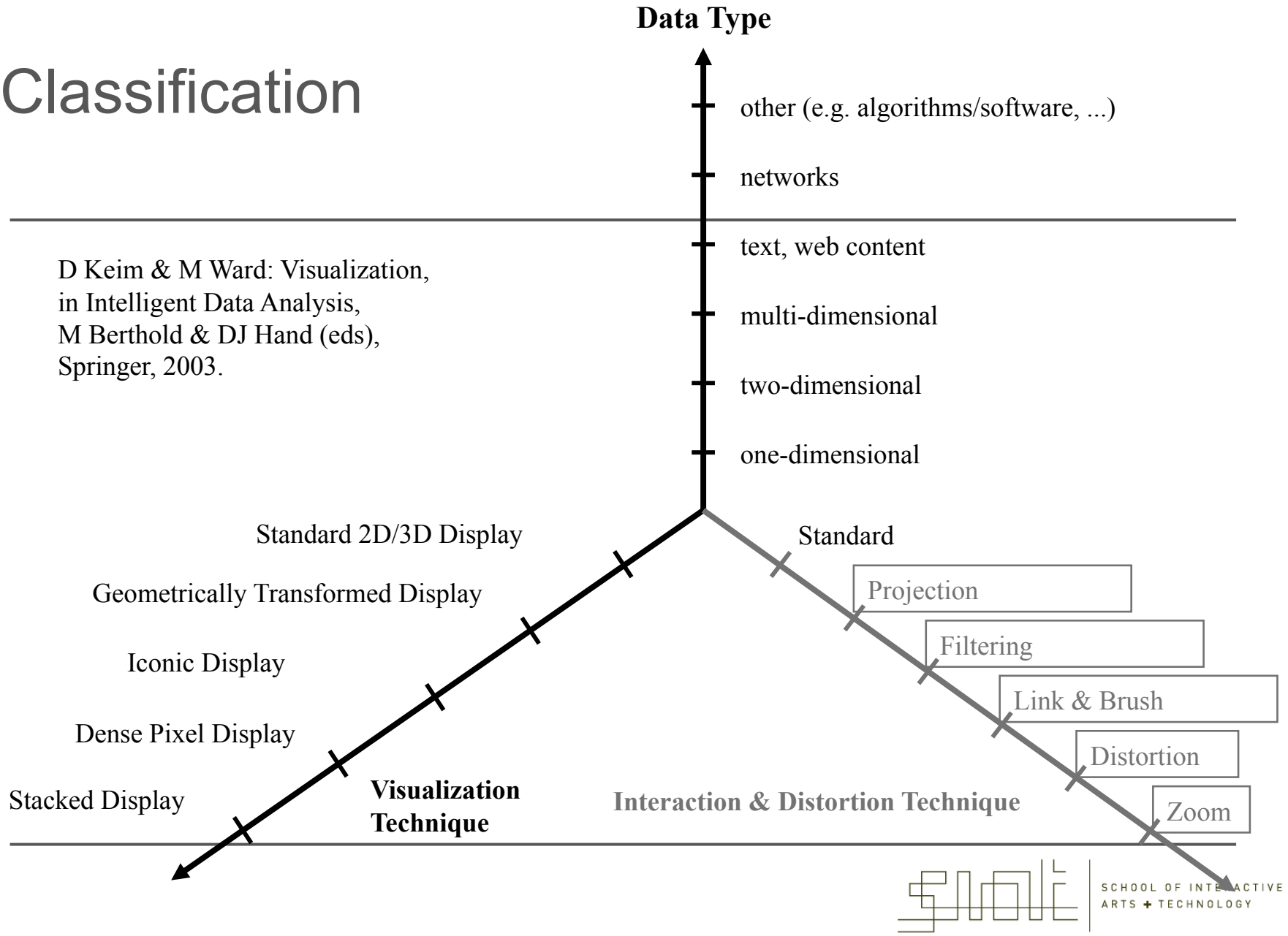
SCHOOL OF INTERACTIVE ARTS + TECHNOLOGY

Today

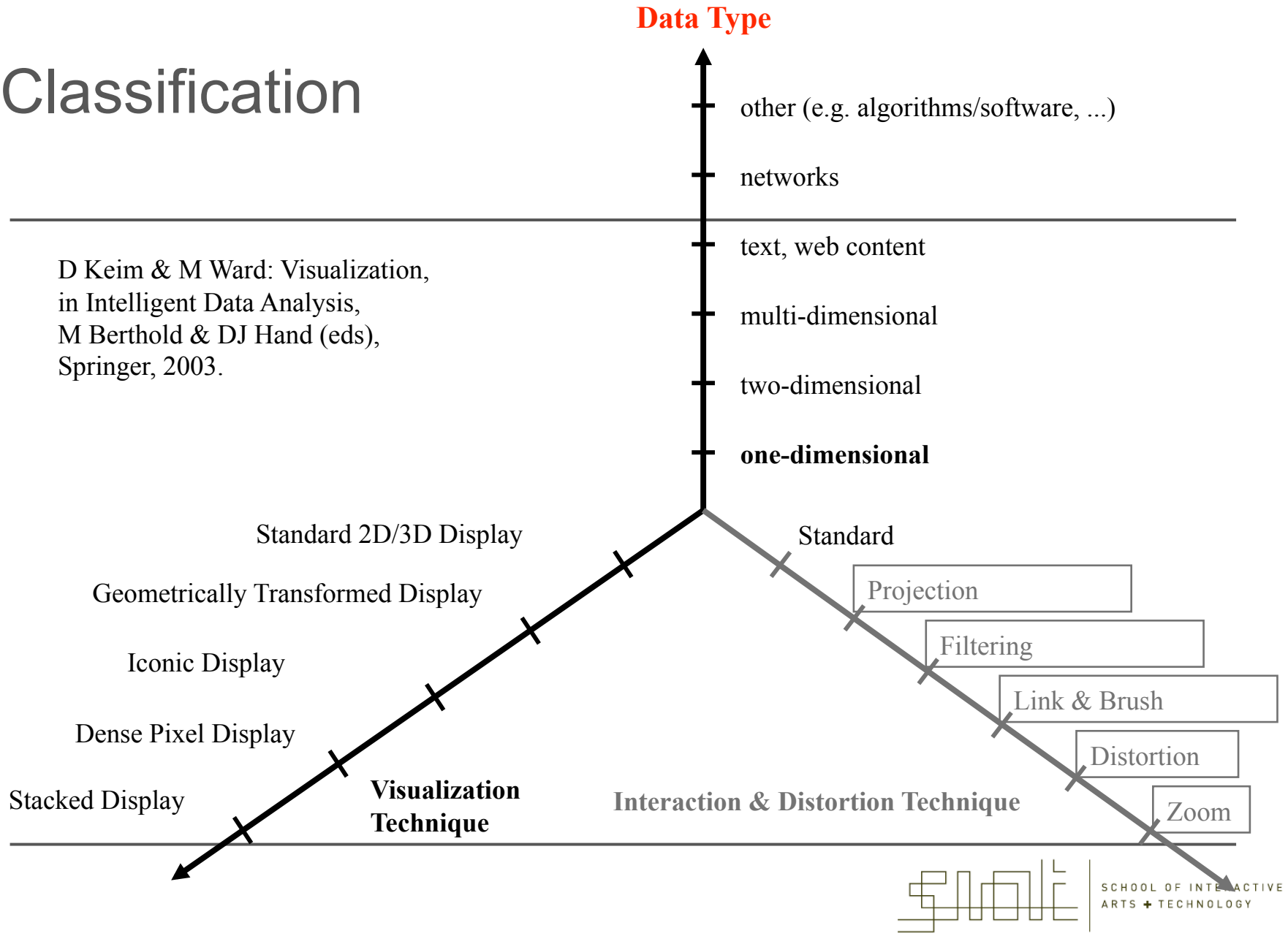
- Assignments and presentations
- Assignment 3 out this week
- Presentations start next week (Thursday)

- Project presentations Week 13 and 14
- Project papers due Week 14.

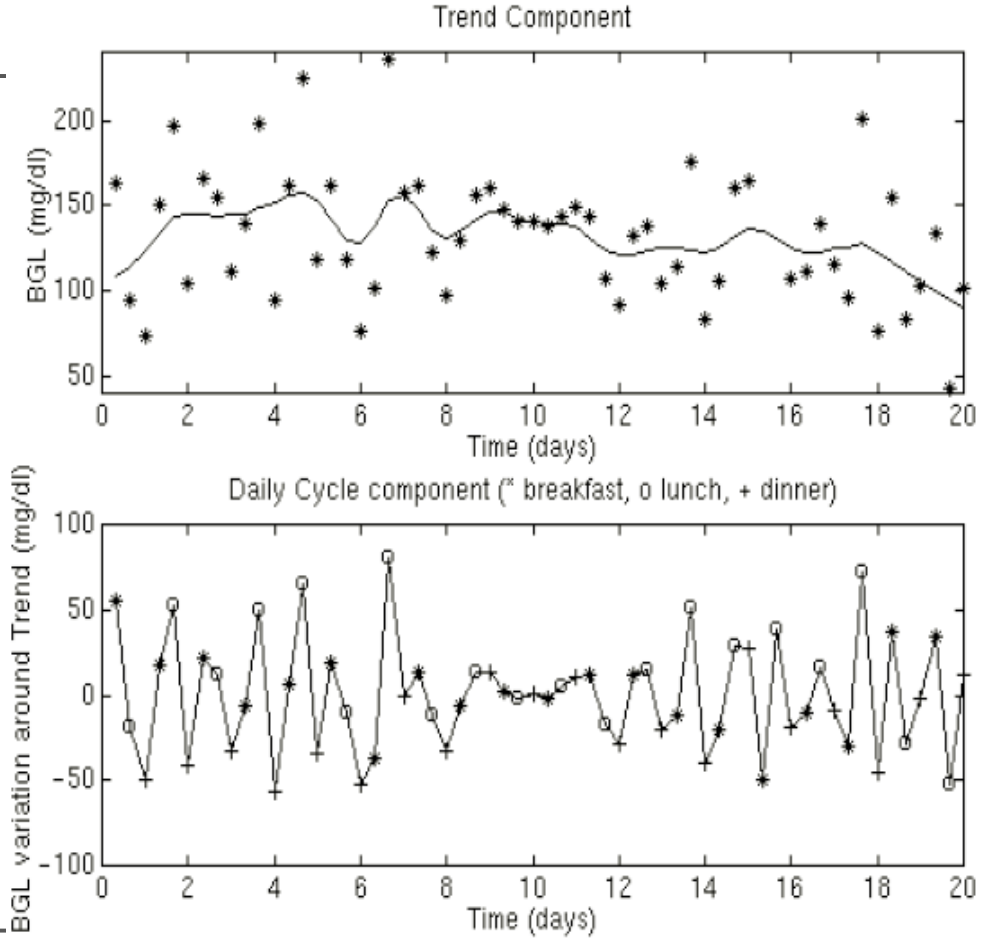
Classification



Classification



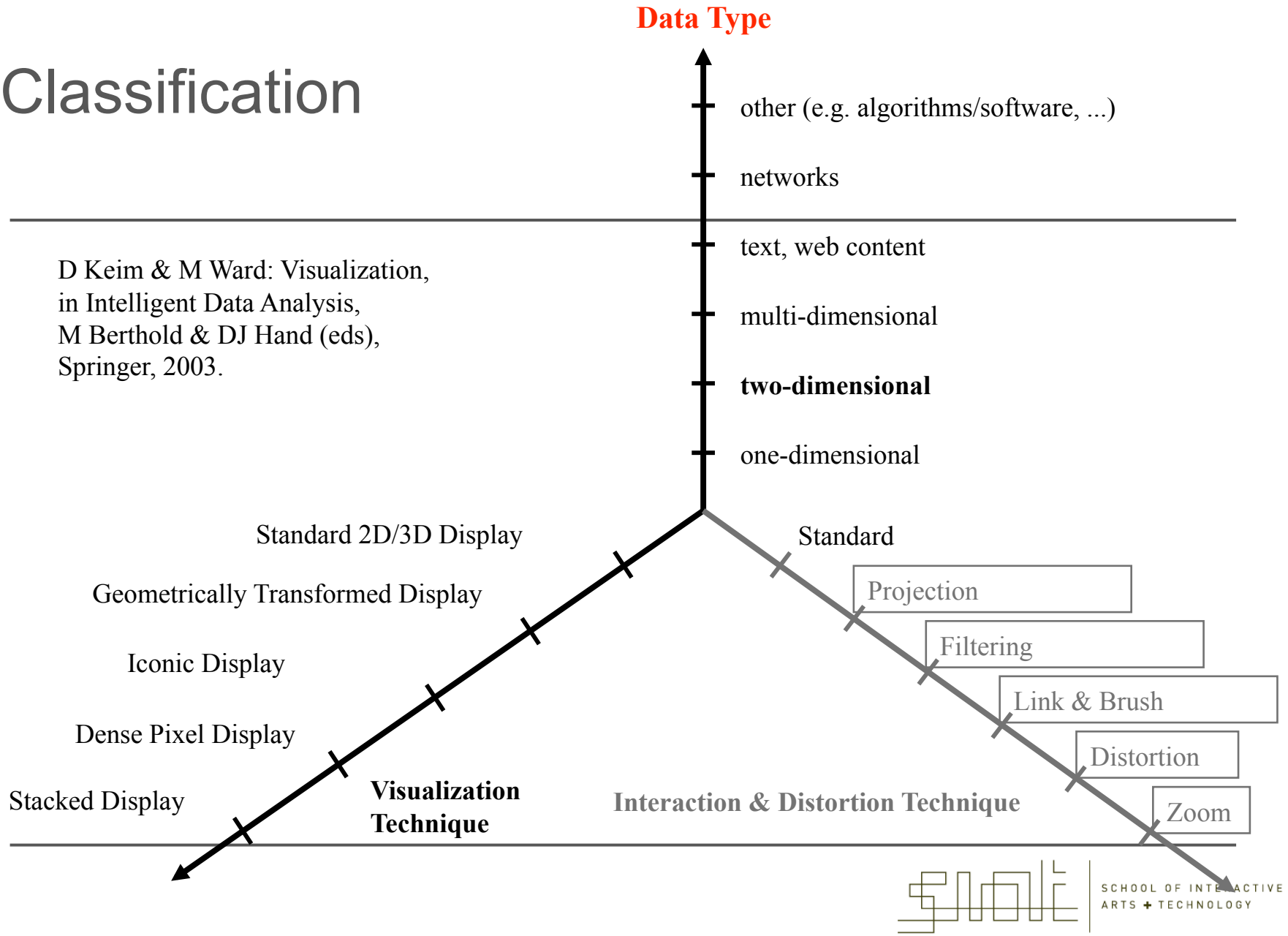
Data: One-Dimensional



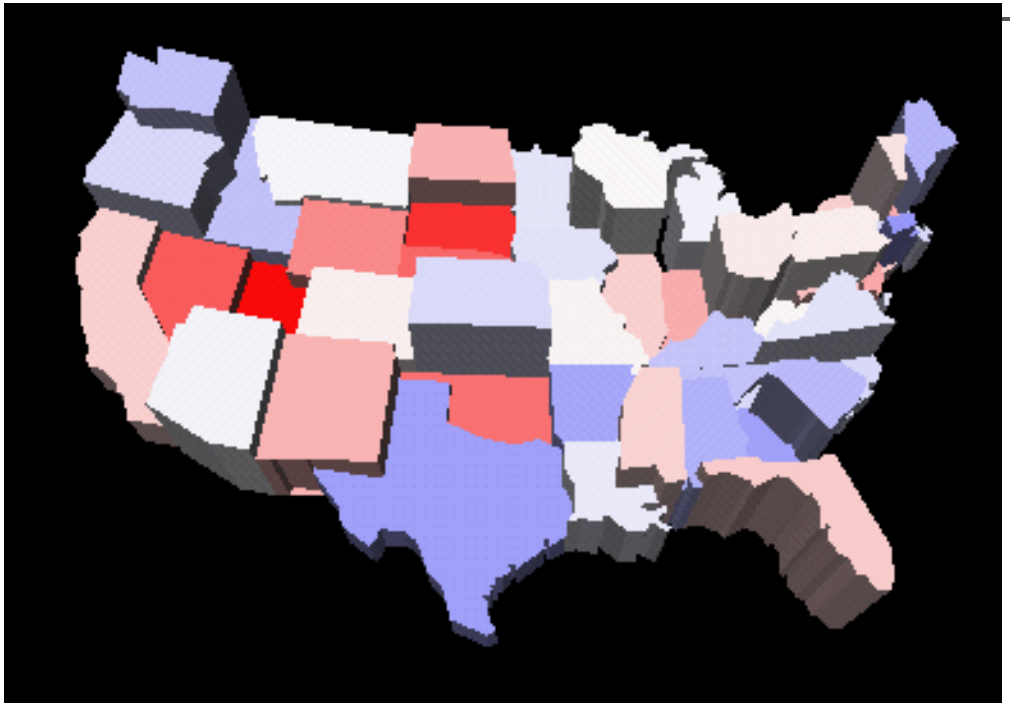
R Bellazzi: Mining Biomedical Time Series by Combining Structural Analysis and Temporal Abstractions, In Proc. of AMIA 1998.

SCHOOL OF INTERACTIVE ARTS + TECHNOLOGY

Classification

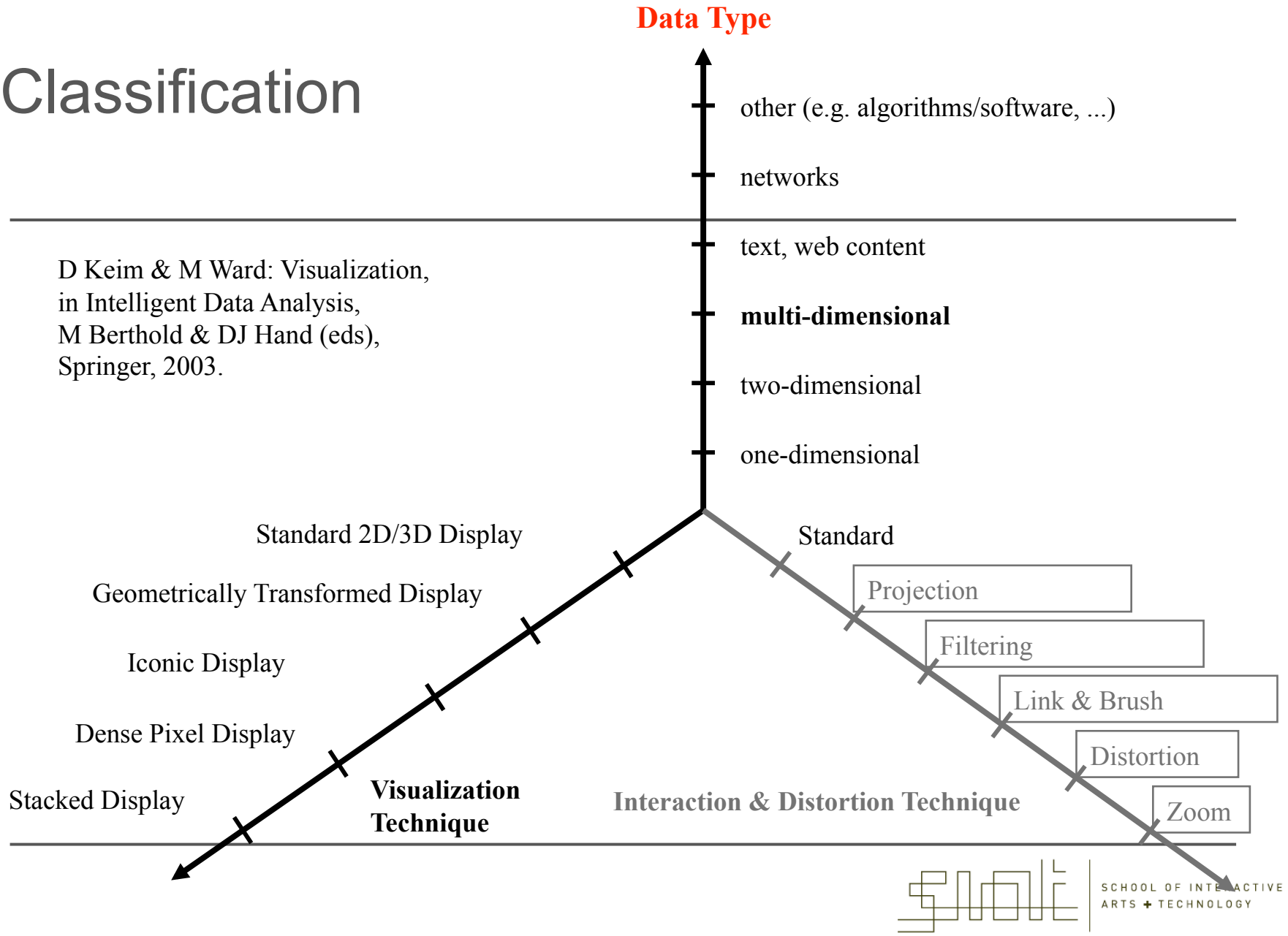


Data: Two-Dimensional

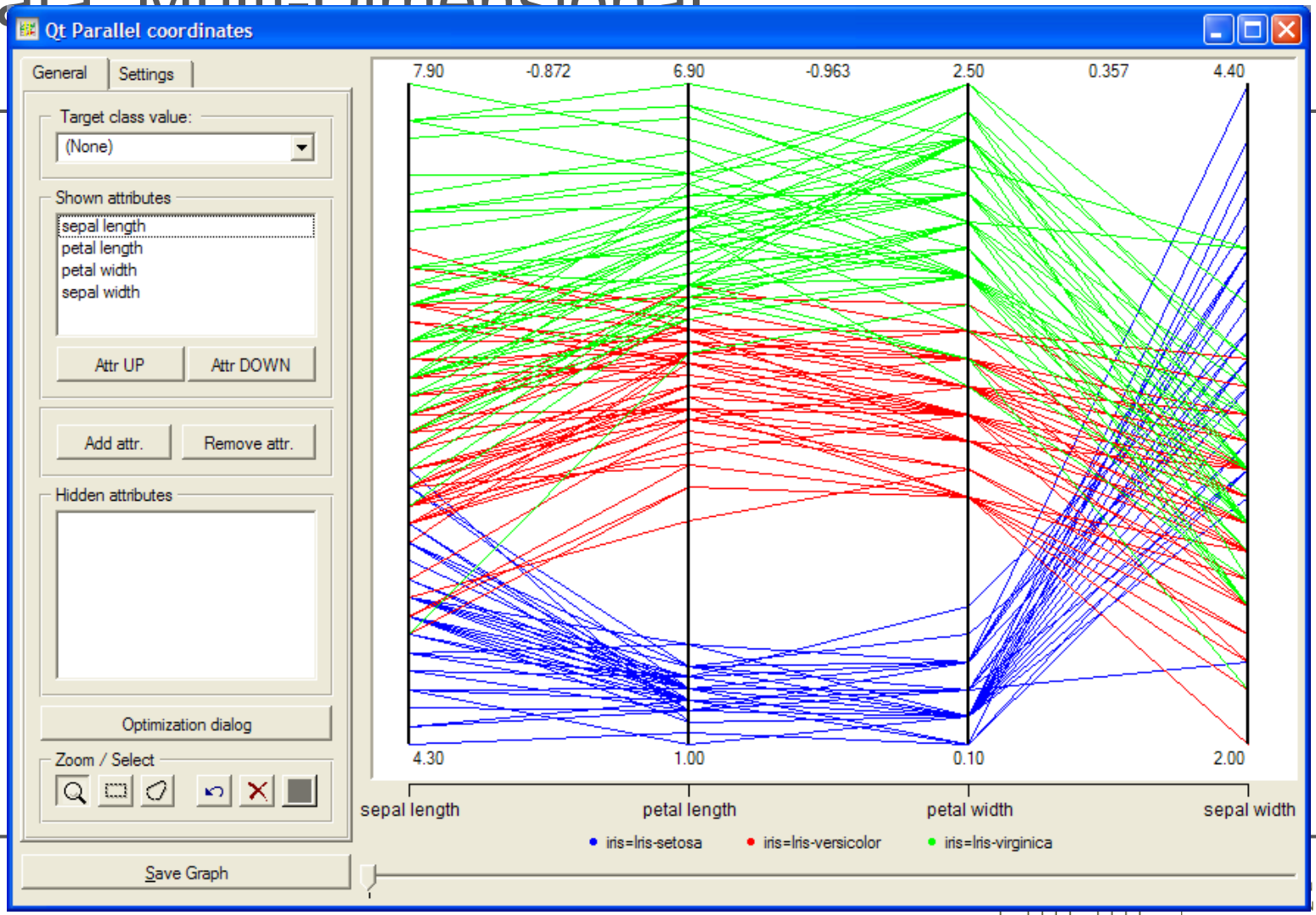


MineSet's Map Visualizer.

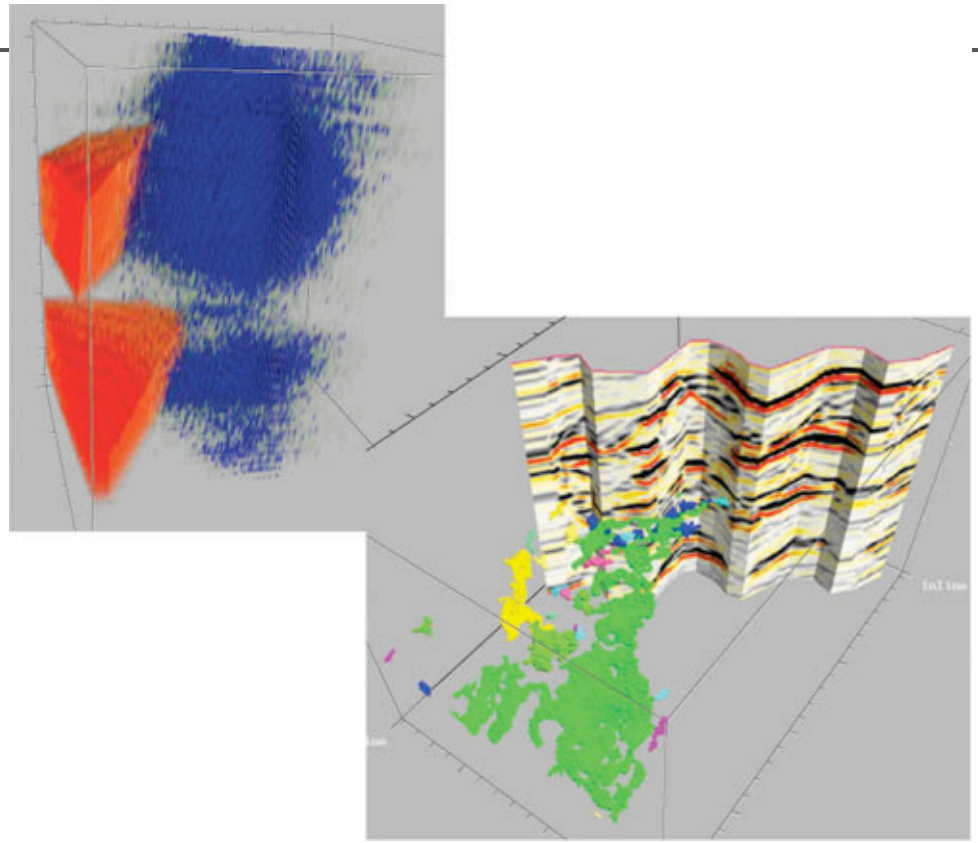
Classification



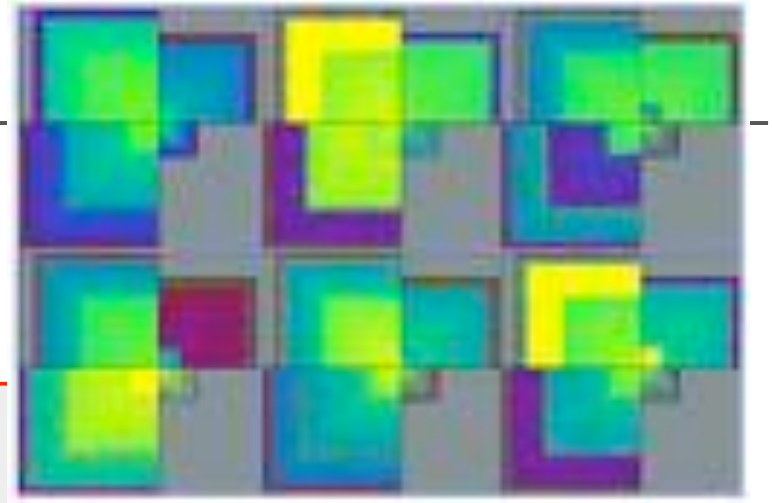
Data: Multi-Dimensional

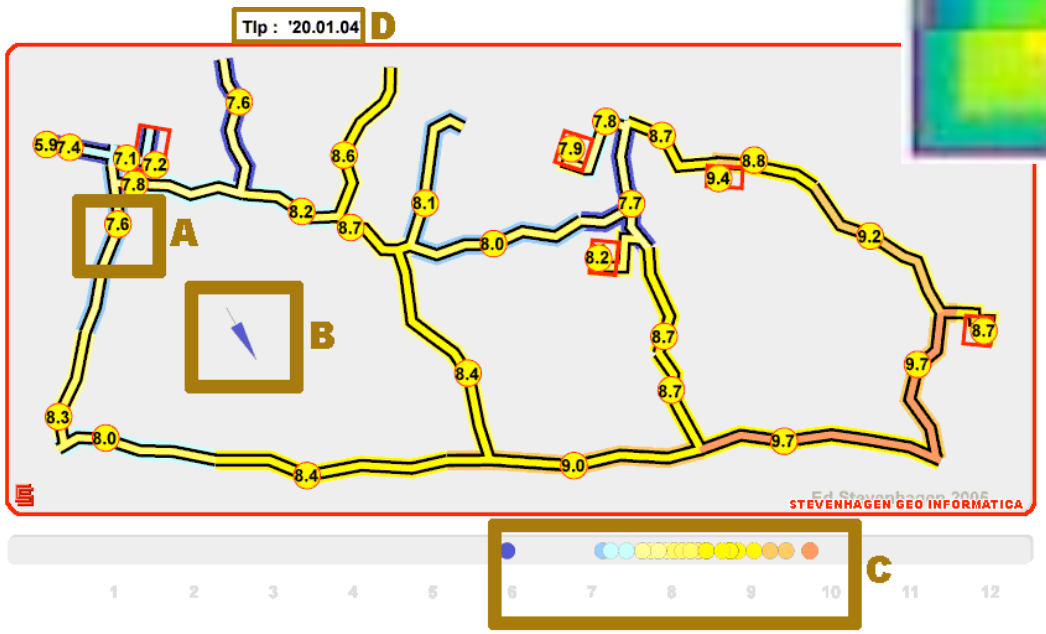


Data: Multi-Dimensional

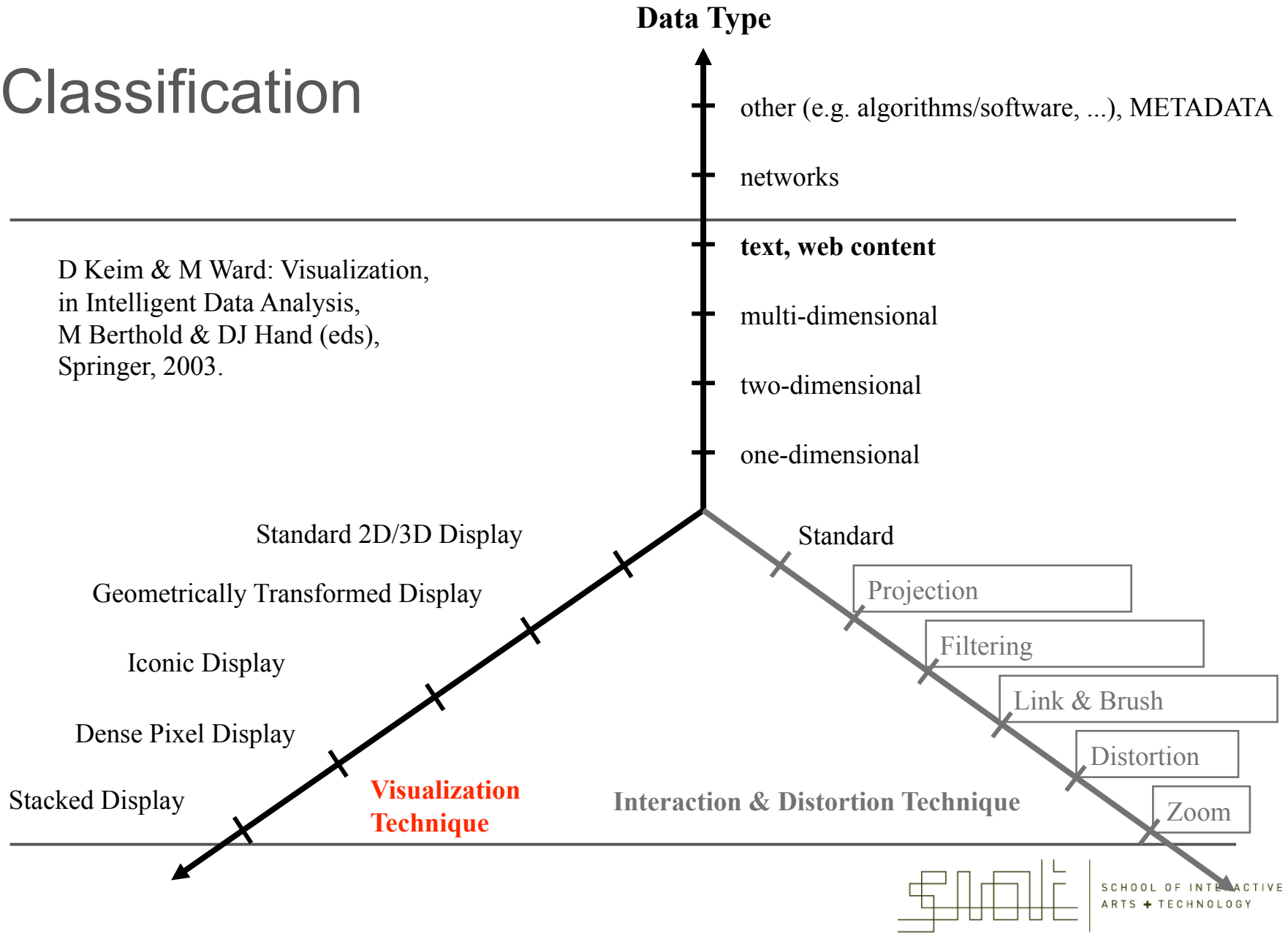


Data: Multi-Dimensional

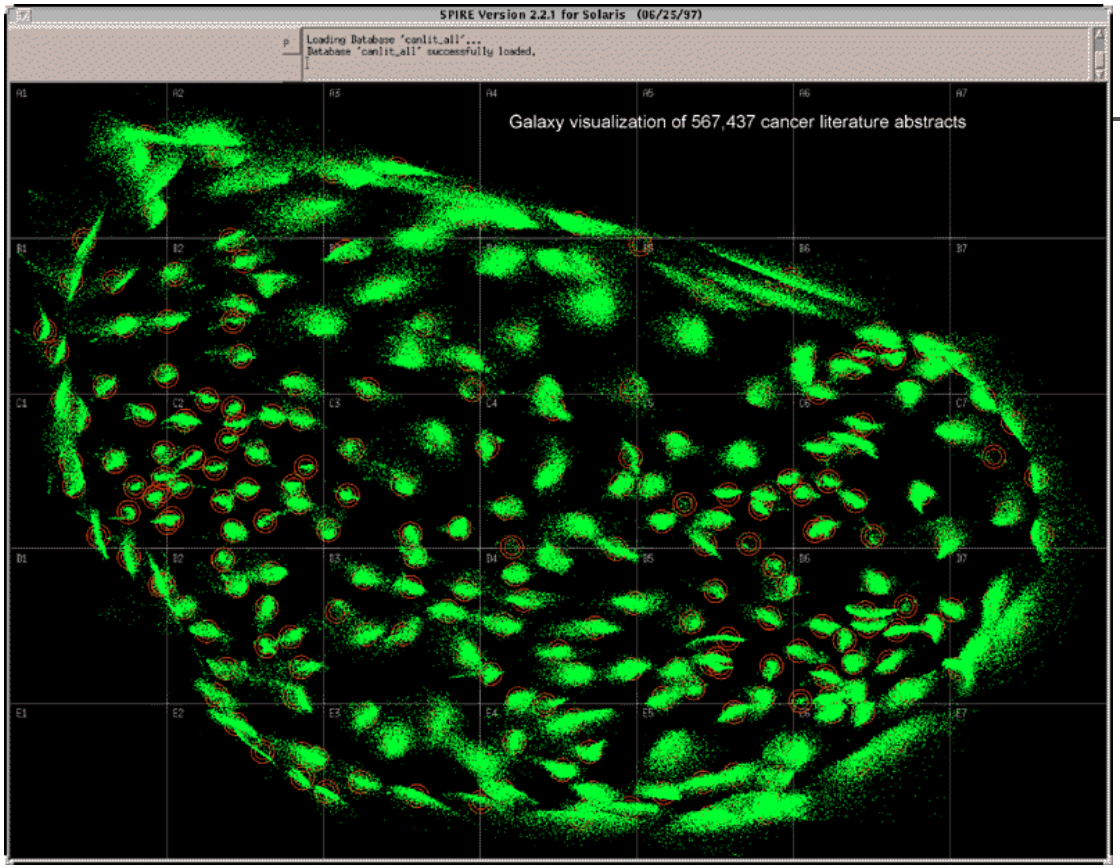




Classification



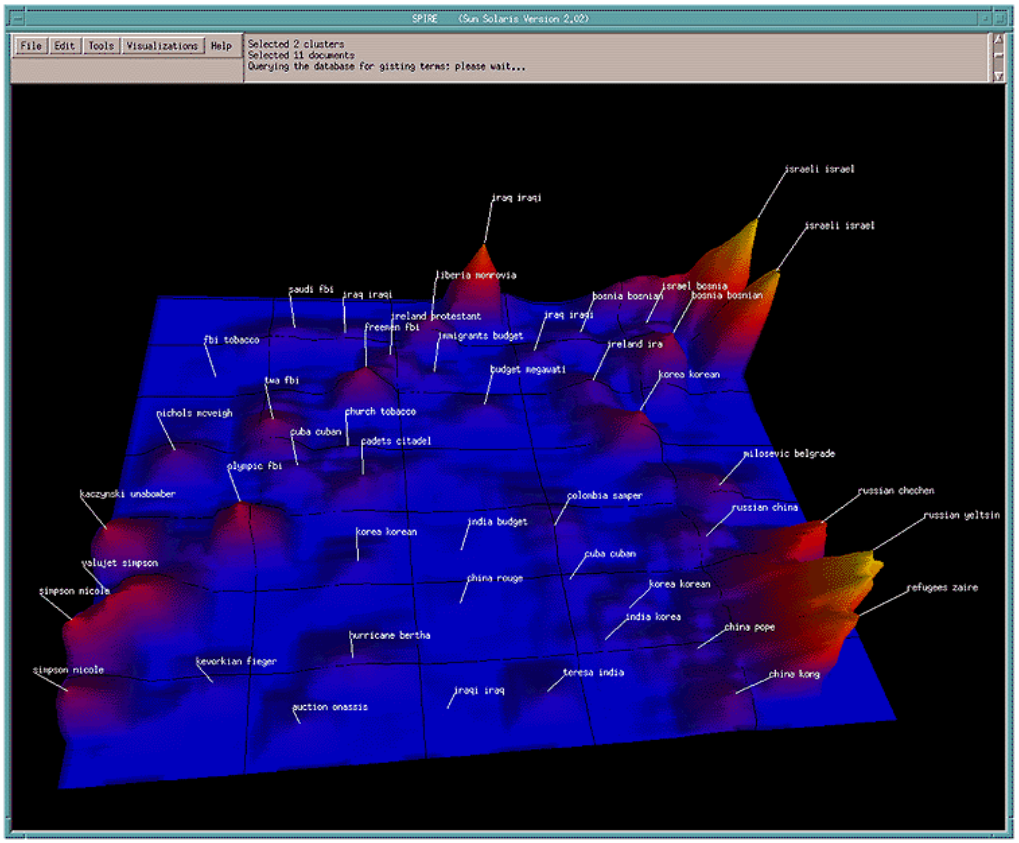
Data: Text



From Inspire (TM) Software, see www.pnl.gov/infoviz/technologies.html

- Galaxies visualization
- Uses the "night sky" visualization to represent a set of documents
- One document – one star
- Stars clustered together represent related documents
- Includes analytical tools to investigate groups and time-based trends, query contents

Data: Text

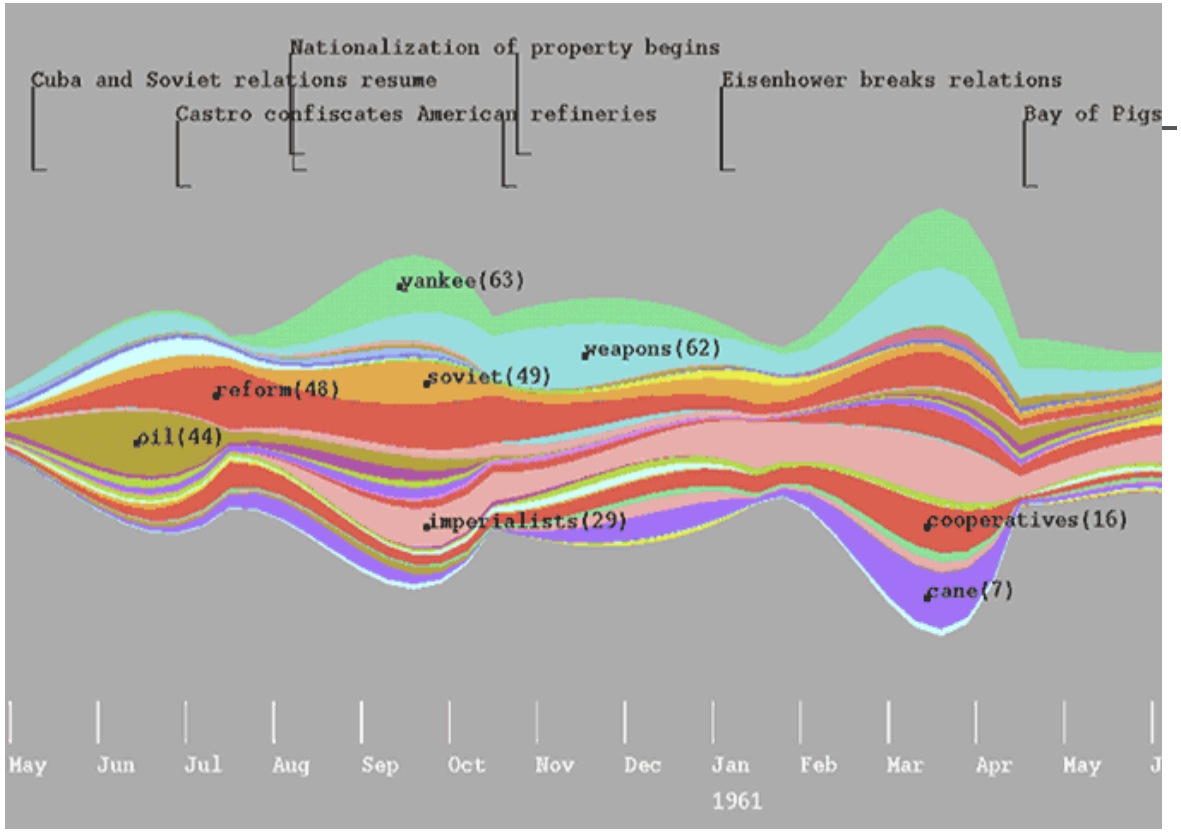


ThemeView (TM)

- Topics or themes of text documents shown in relief map of a natural terrain
- The height of a peek relates to the strength of the topic

From Inspire (TM) Software, see www.pnl.gov/infoviz/technologies.html

Data: Text

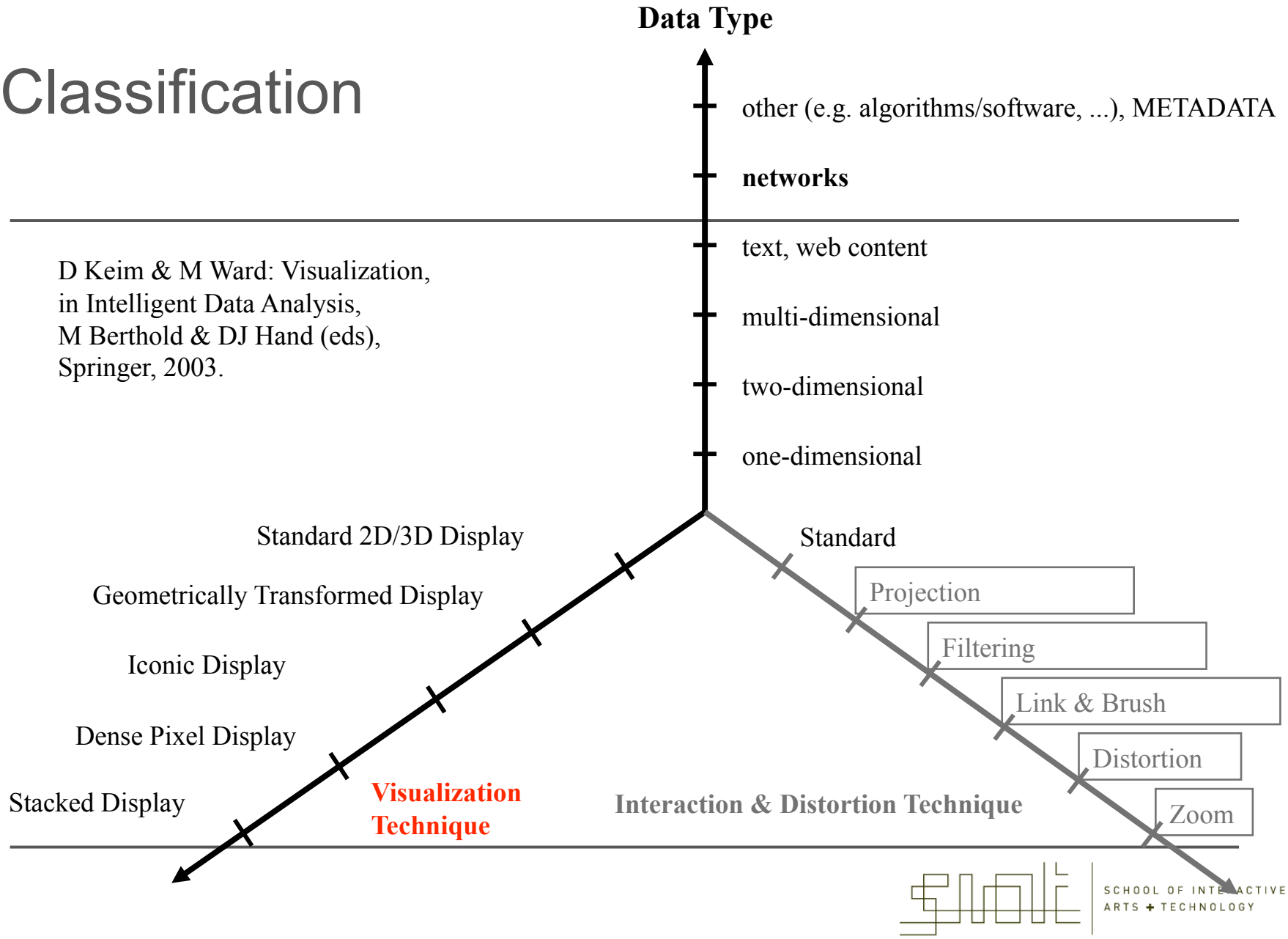


Theme River (TM)

- Identification of time related trends and patterns
- Themes represented as colored streams
- The width of the stream relates to the collective strength of a theme

From Inspire (TM) Software, see www.pnl.gov/infoviz/technologies.html

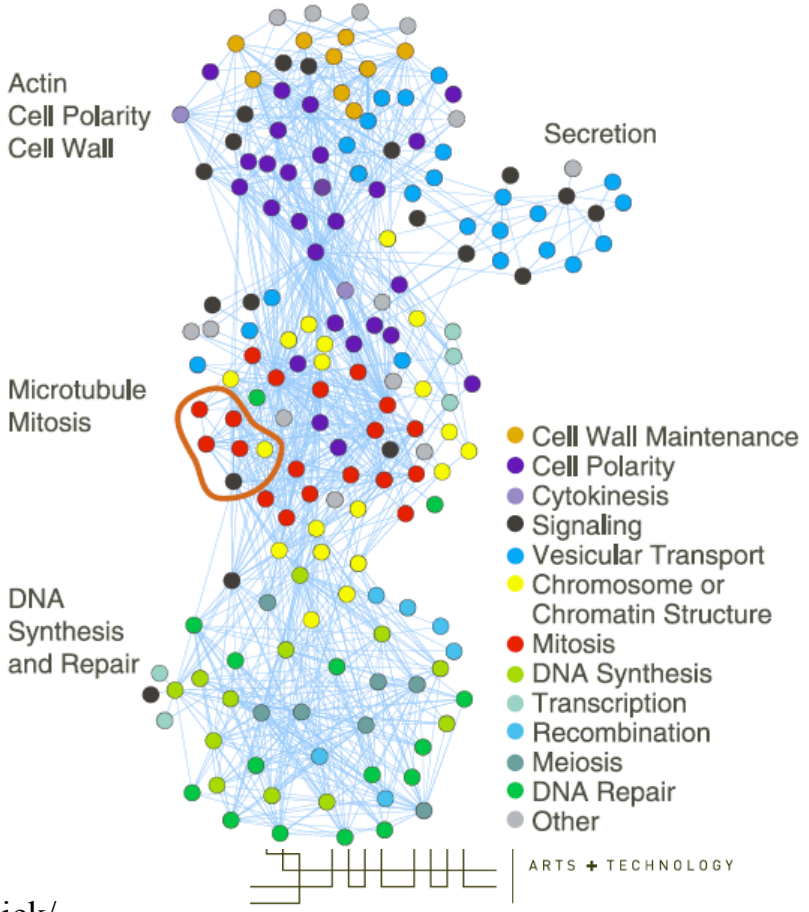
Classification



Data: Networks

E. coli metabolic network (colors denote predominant biochemical class of metabolites)
Ravasz et al., Science 297, 30 Aug 2002.

S. cerevisiae gene interaction network
Tong et al., Science 303, 6 Feb 2004.

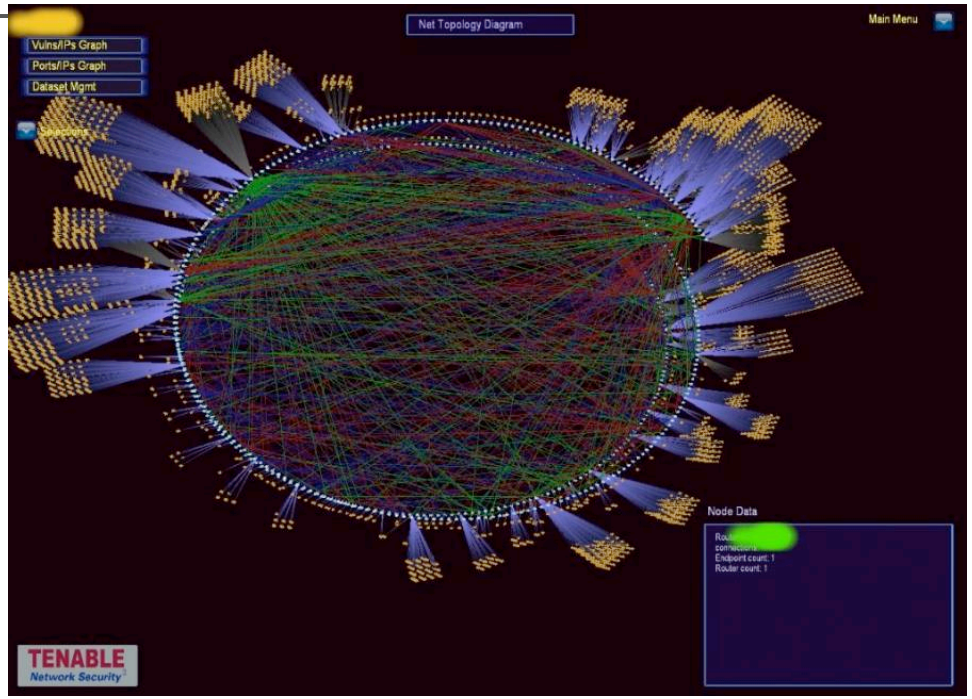


V Batagelj, A Mrvar: Pajek @ vlado.fmf.uni-lj.si/pub/networks/pajek/

Data: Networks

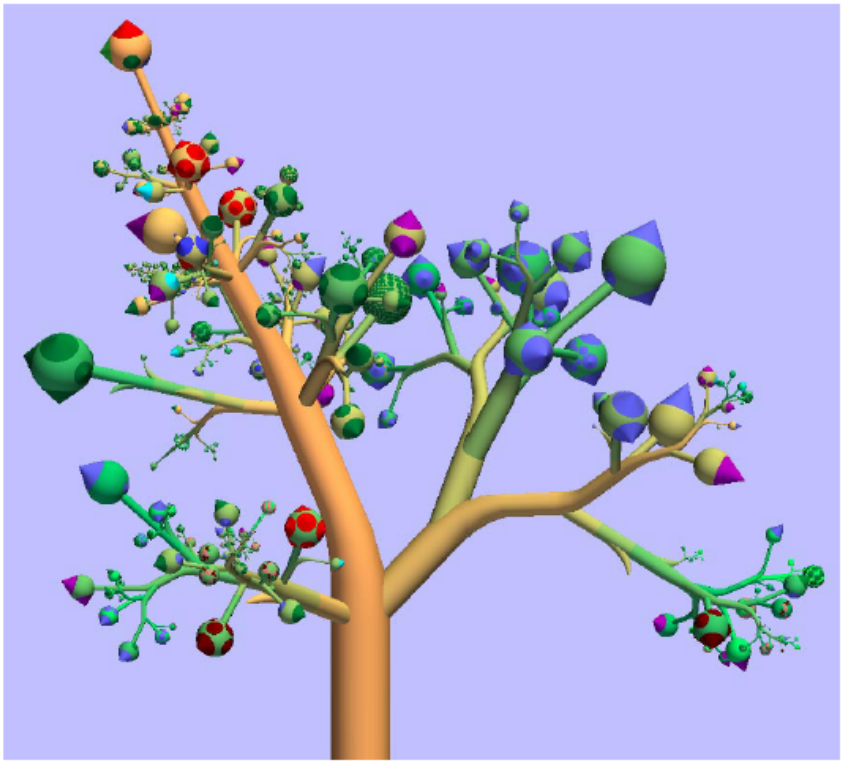
Topology

- What's connected to what
- Strength/nature of relationship
- Balance

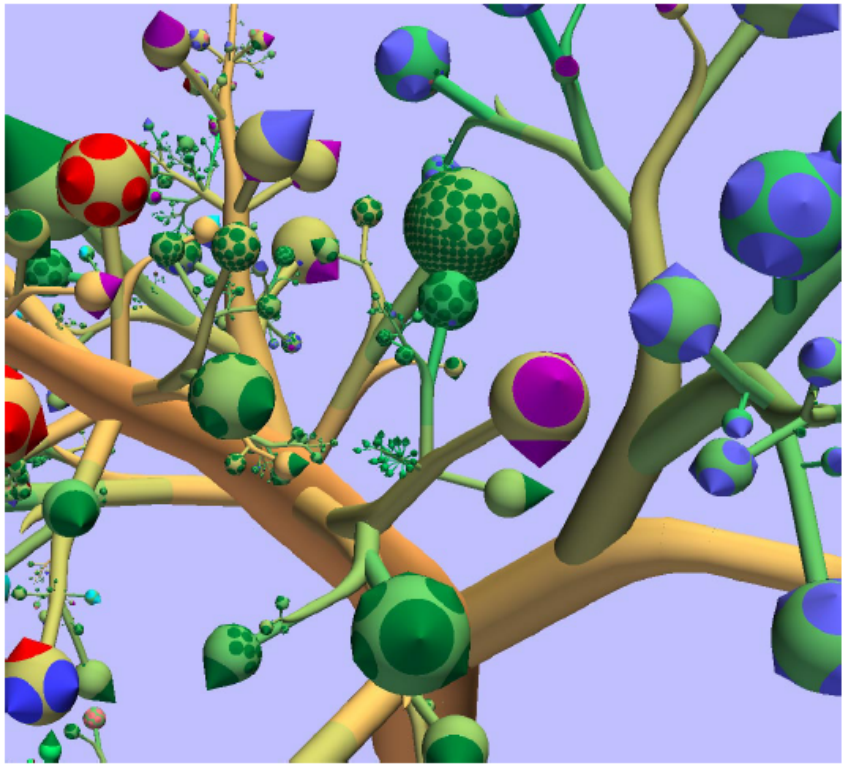


Tenable Network Security's Security Center

Data: Tree Hierarchies



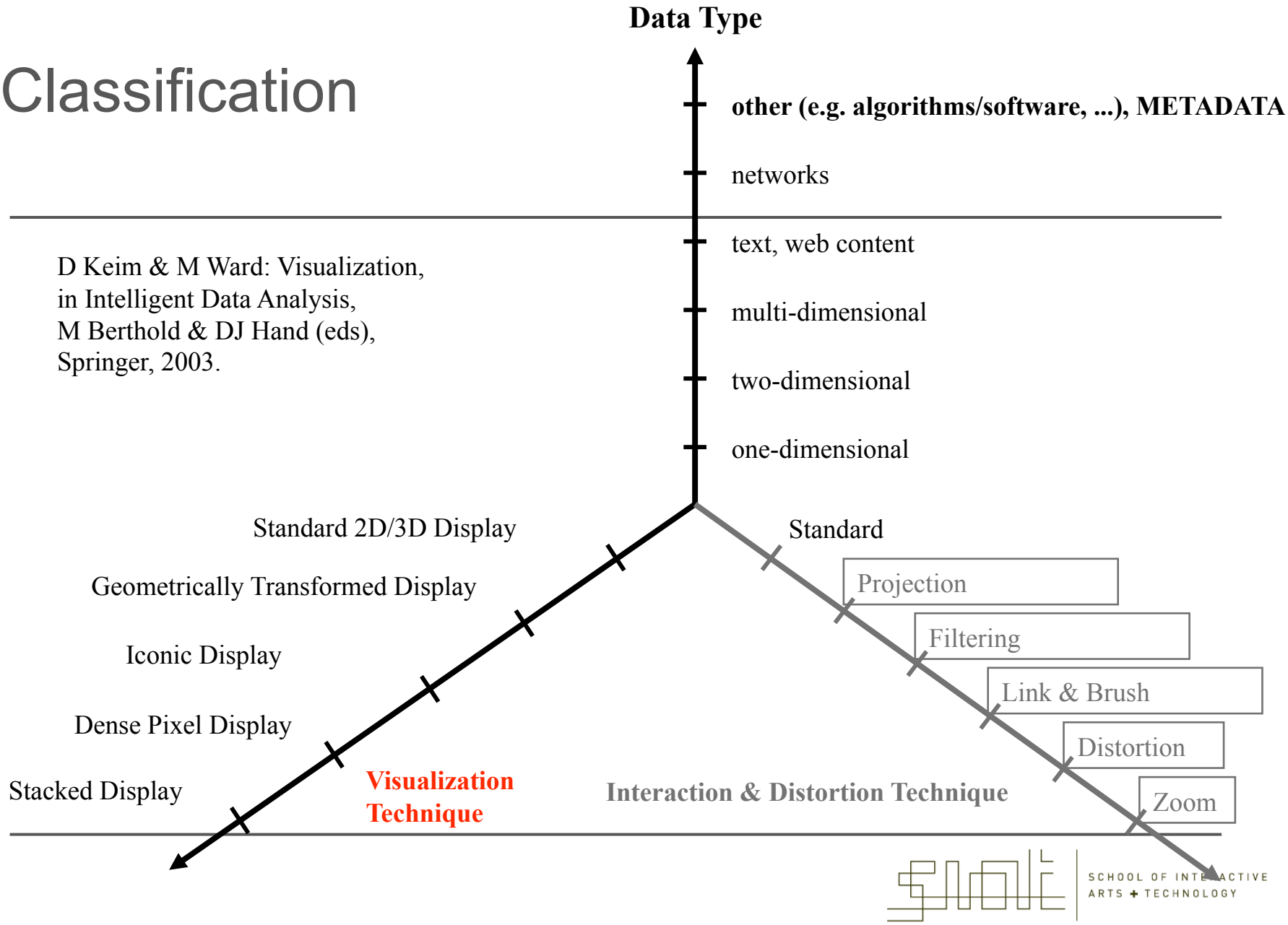
Unix home directory



Selected detail

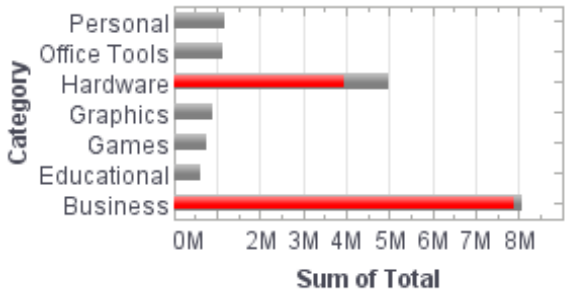
Kleiberg *et al.*: Botanic Visualization of Huge Hierarchies, In InfoVis, 2001.

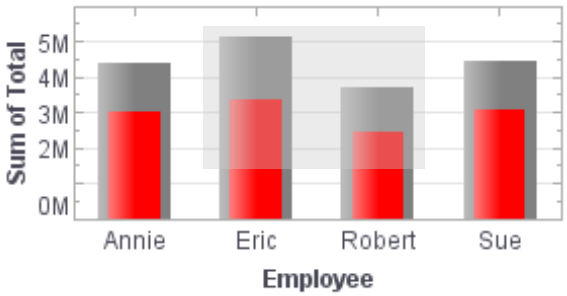
Classification

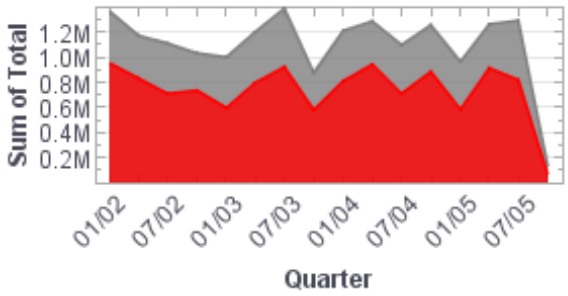


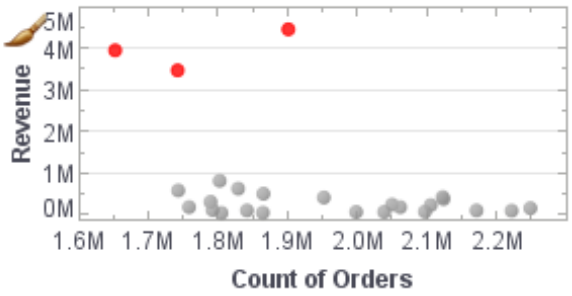
Data: Metadata

- Brushing
- Uncertainty
- History
- Navigation

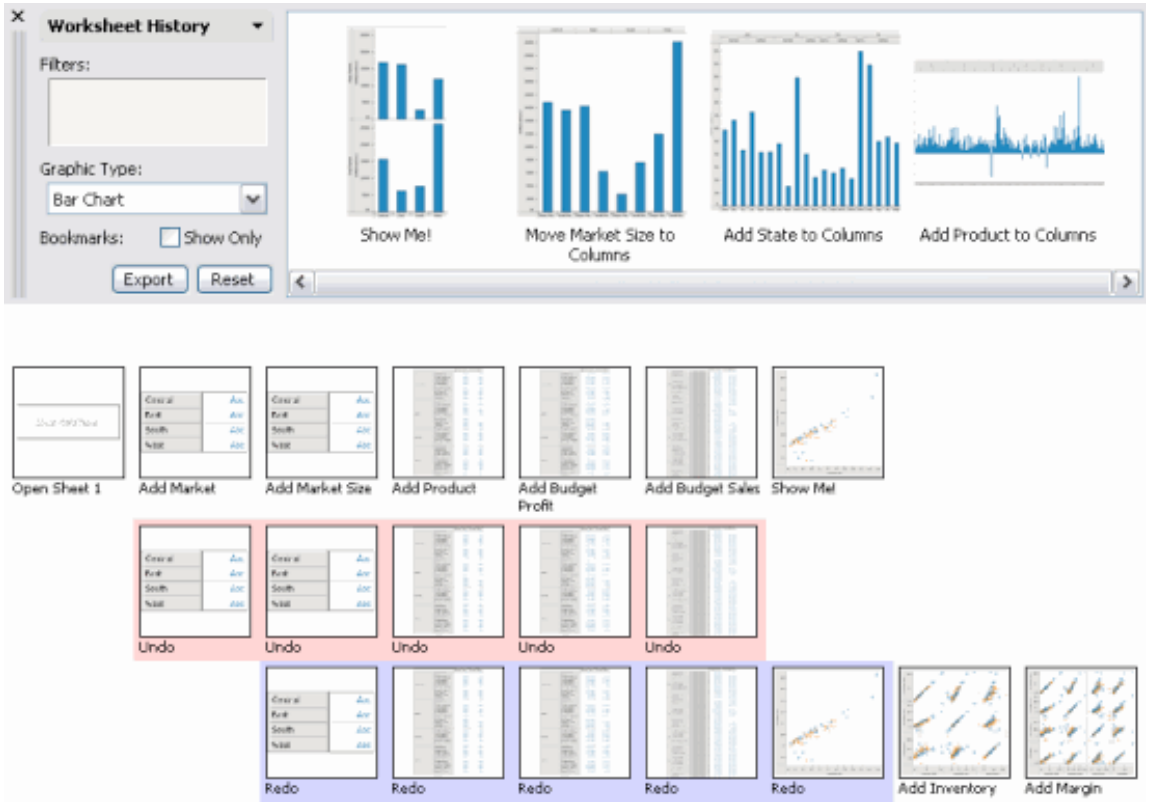








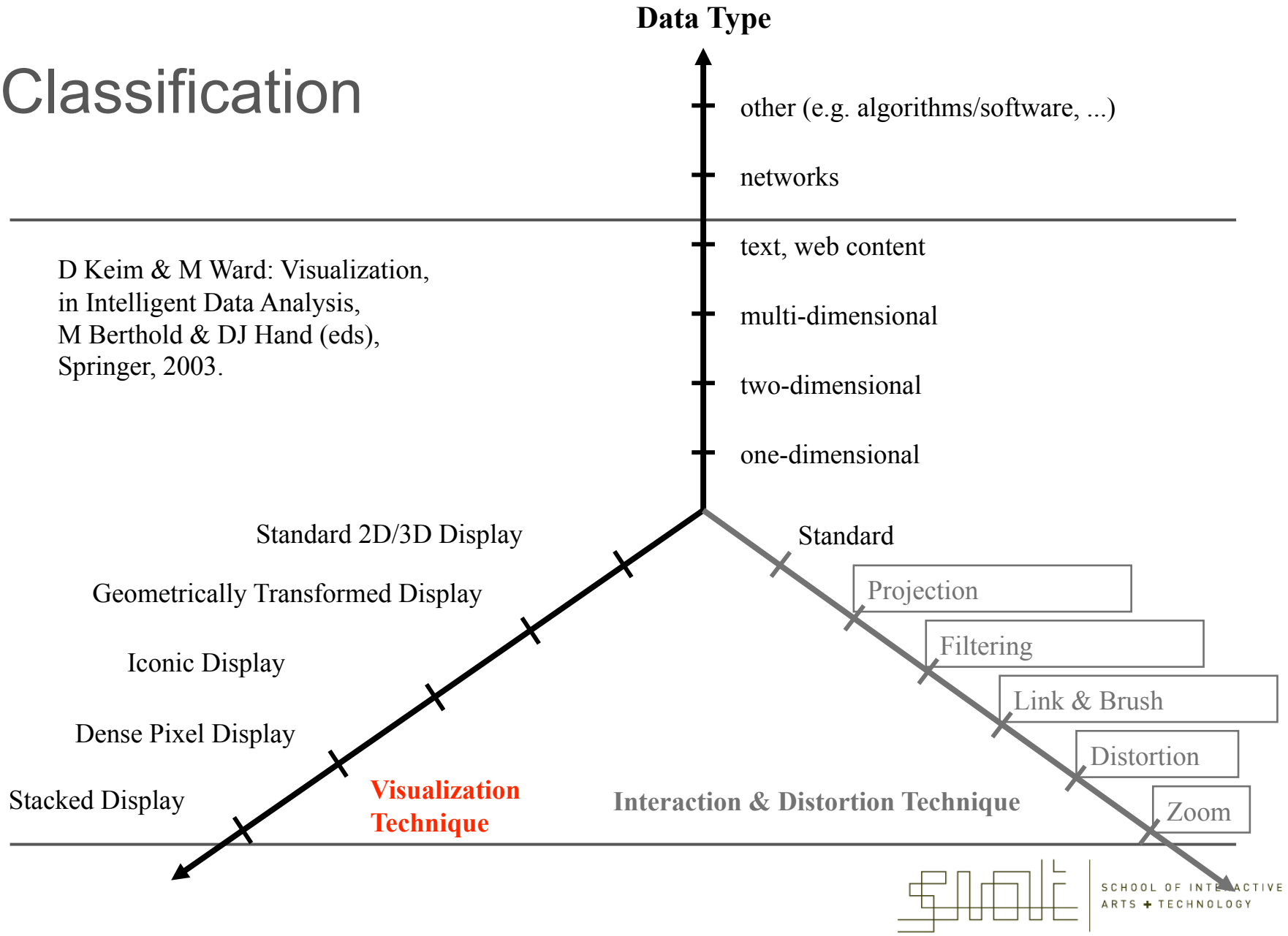
Data: Metadata



- History
- Navigation

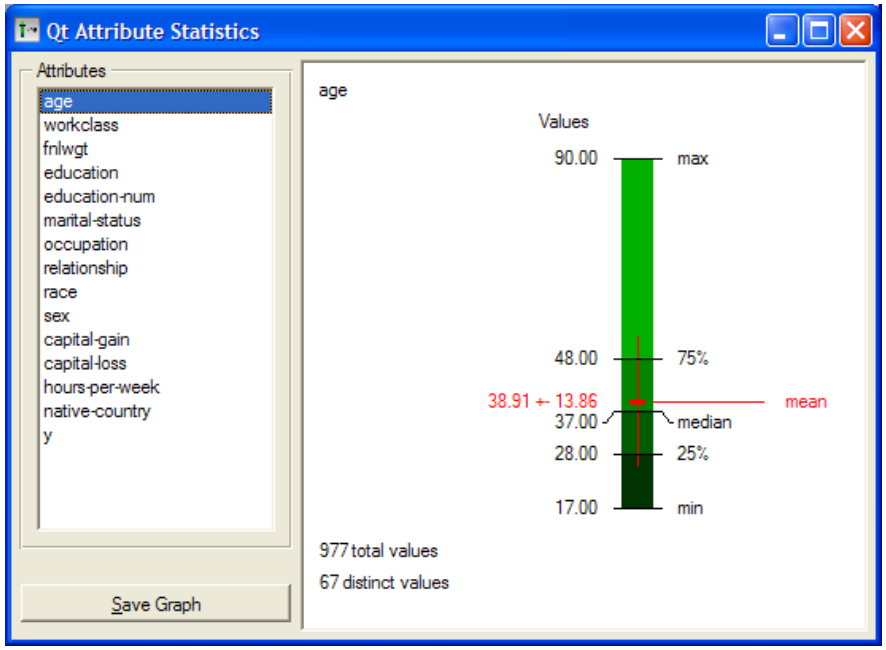
SCHOOL OF INTERACTIVE ARTS + TECHNOLOGY

Classification



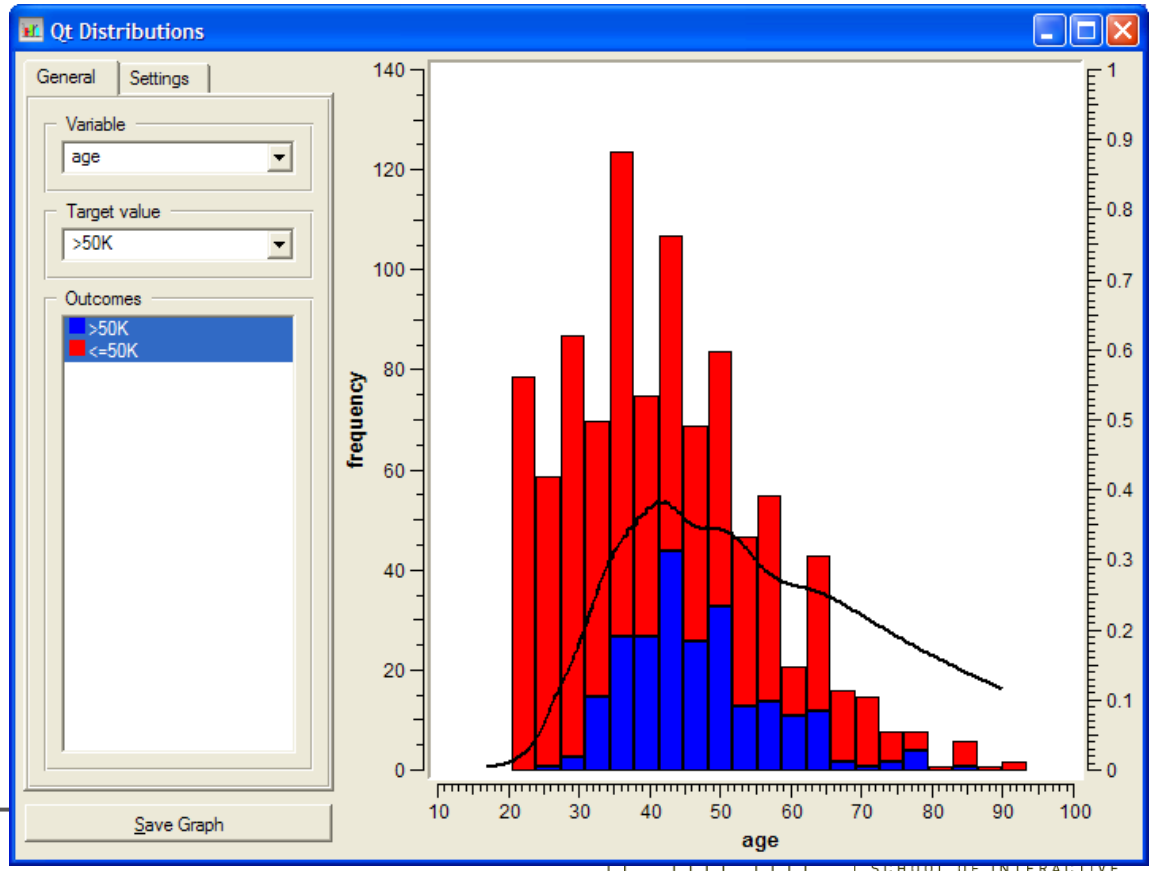
Standard 2D/3D

- x-y (x-y-z) plots
- bar charts
- line graphs
- histograms
- maps



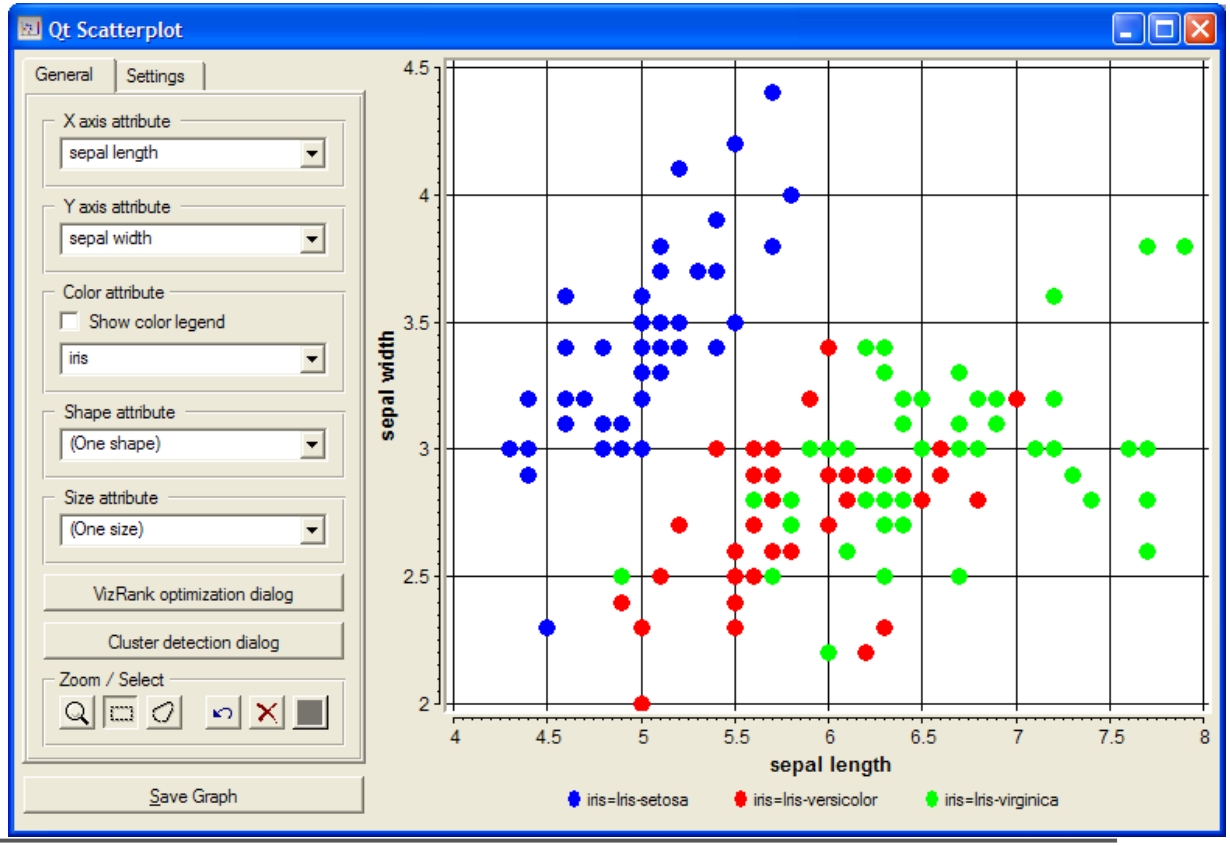
Standard 2D/3D

- x-y (x-y-z) plots
- bar charts
- line graphs
- histograms
- maps



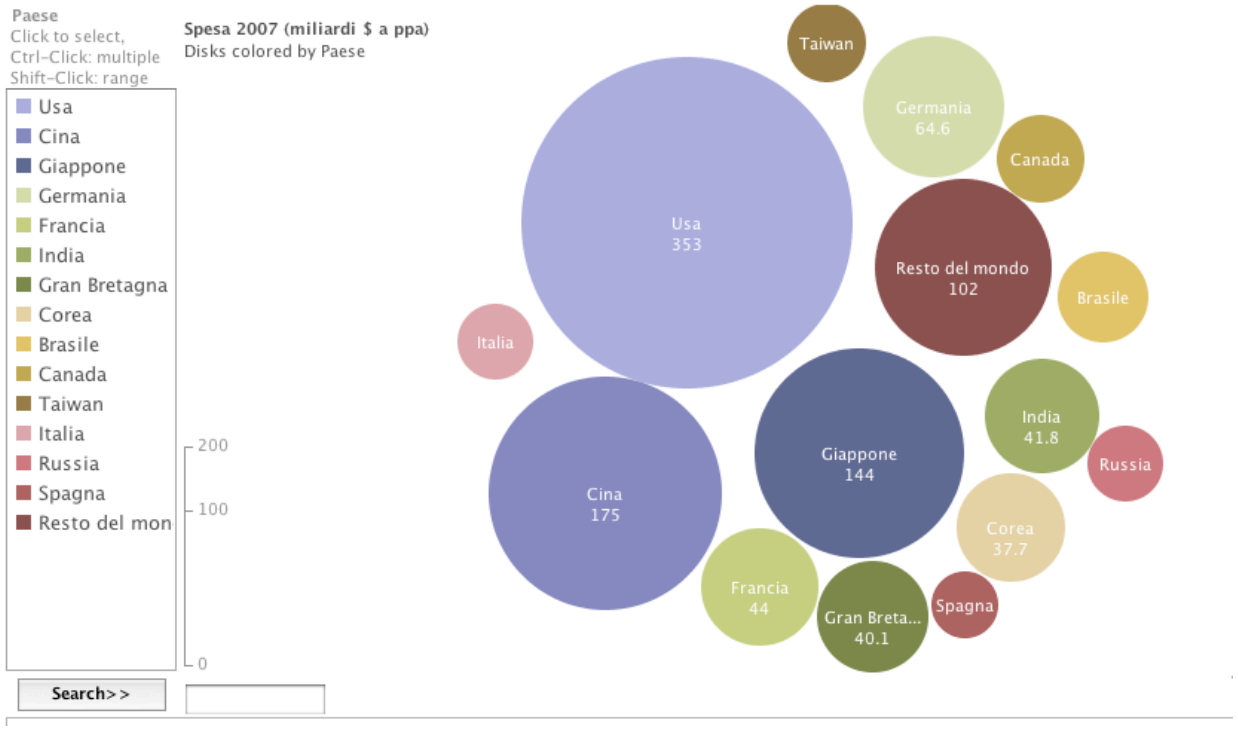
Standard 2D/3D

- x-y (x-y-z) plots
- bar charts
- line graphs
- histograms
- maps



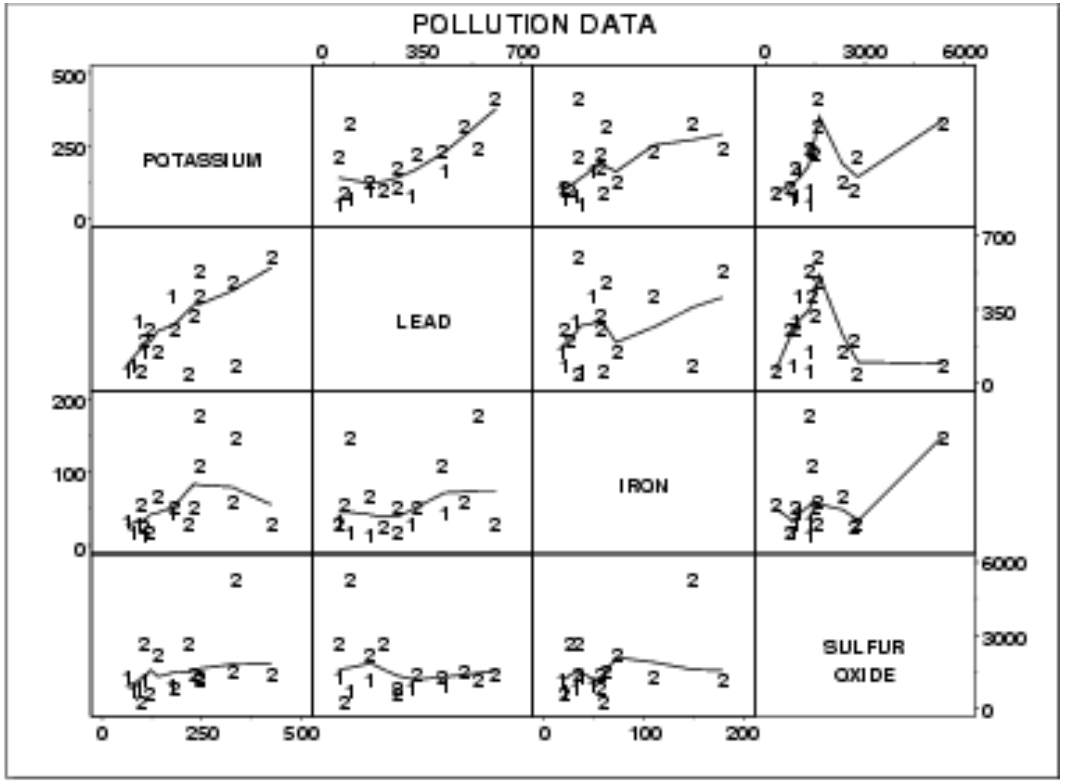
Standard 2D/3D

- x-y (x-y-z) plots
- bar charts
- line graphs
- Bubble charts
- histograms
- maps



Geometrically Transformed Displays

- includes several classes of visualizations
- projection pursuit, finding “interesting transformations” of multi-dim data set
- scatterplot matrix
- parallel coordinates



Geometrically Transformed I

- Parallel coordinates

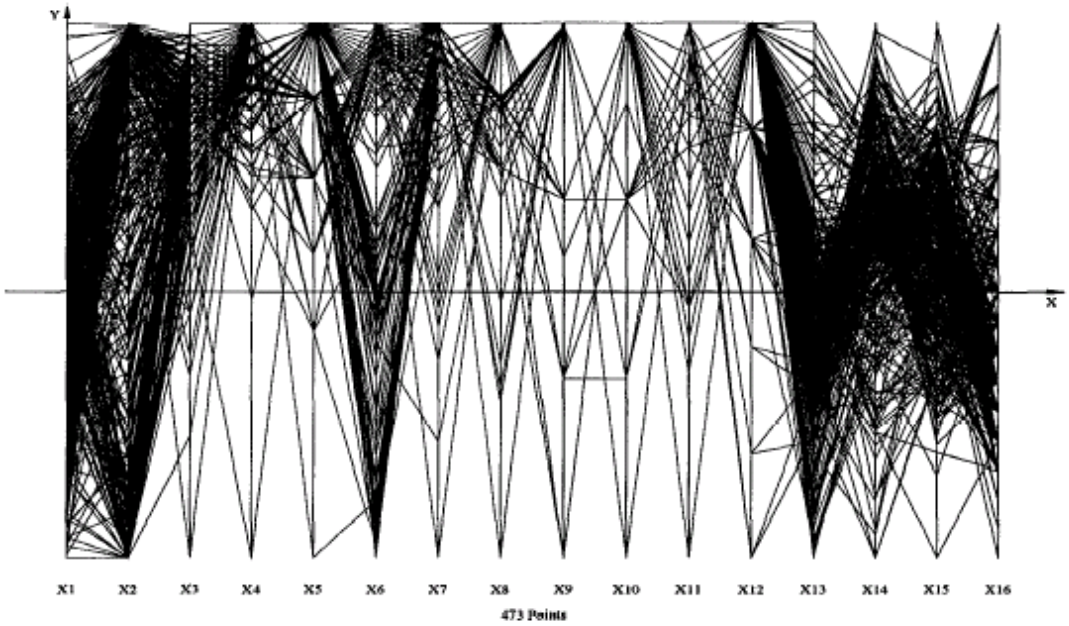


Figure 1: The full dataset consisting of 473 batches

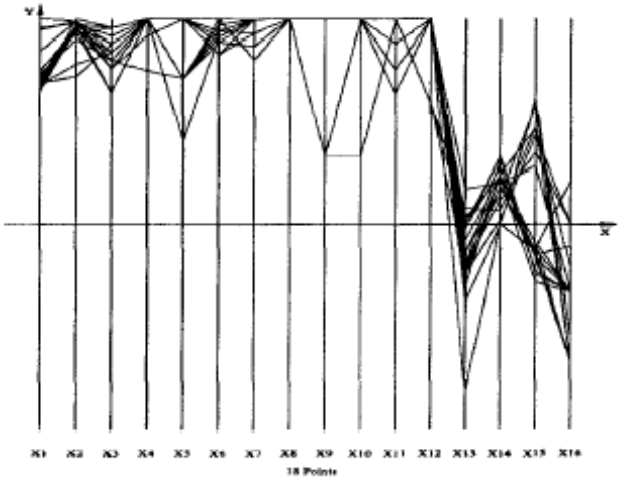


Figure 2: The batches high in Yield, $X1$, and Quality, $X2$.

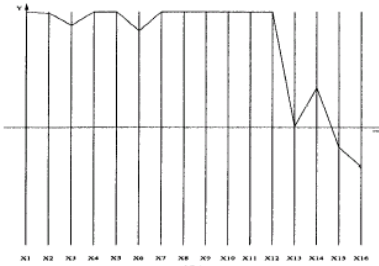


Figure 5: The best batch. Highest in Yield, $X1$, and very high in Quality, $X2$.

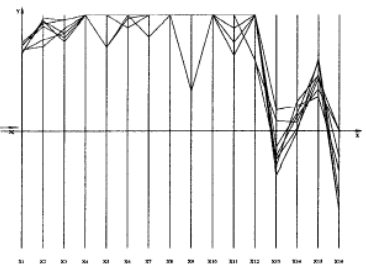
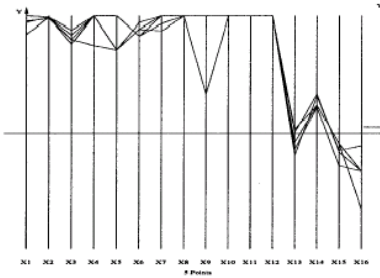
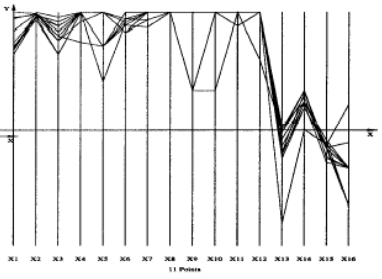
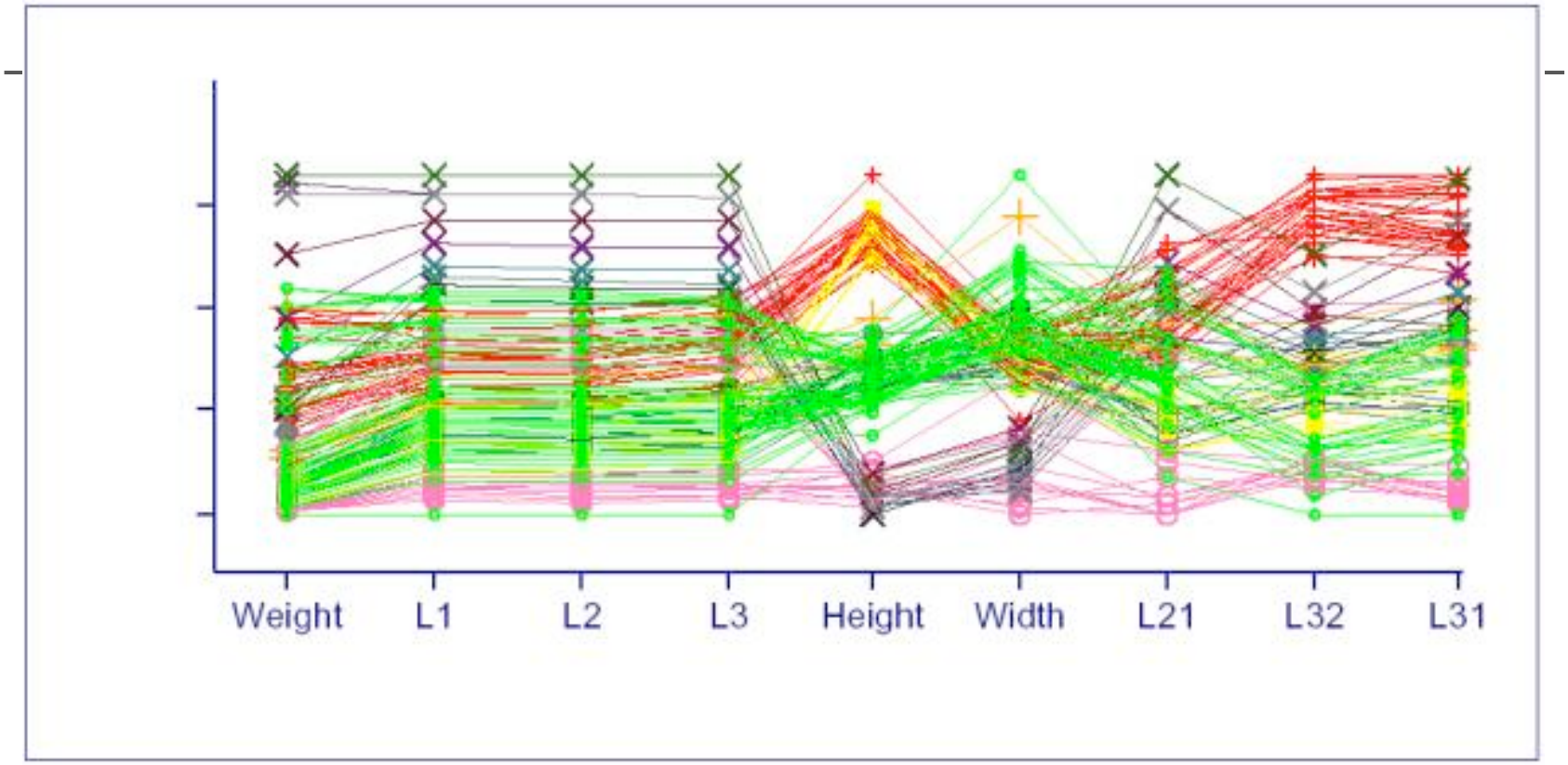


Figure 7: Upper range of split in $X15$



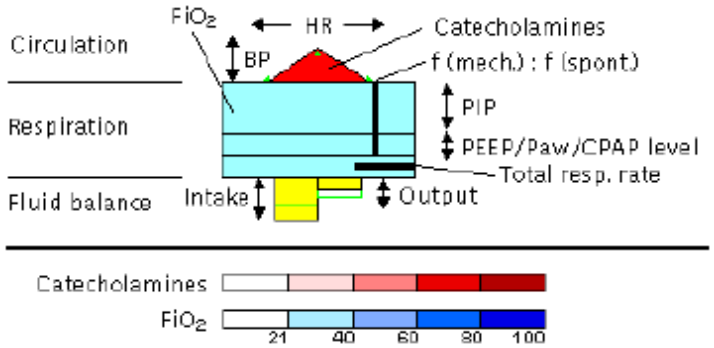


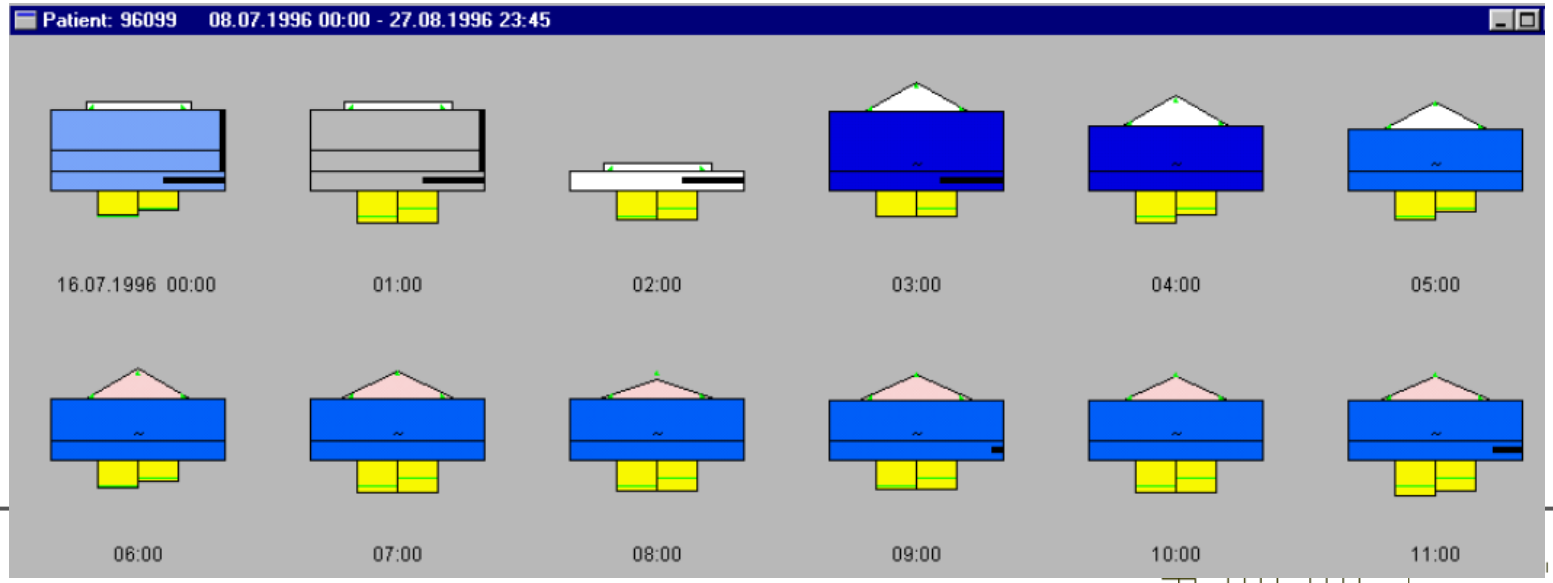
Parallel Coordinates



http://www.evl.uic.edu/aej/526/kyoung/Training-parallelcoordinate.html

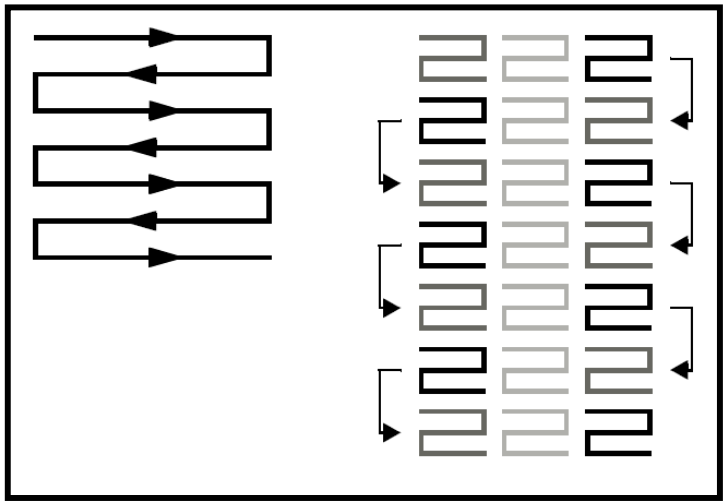
Iconic Displays

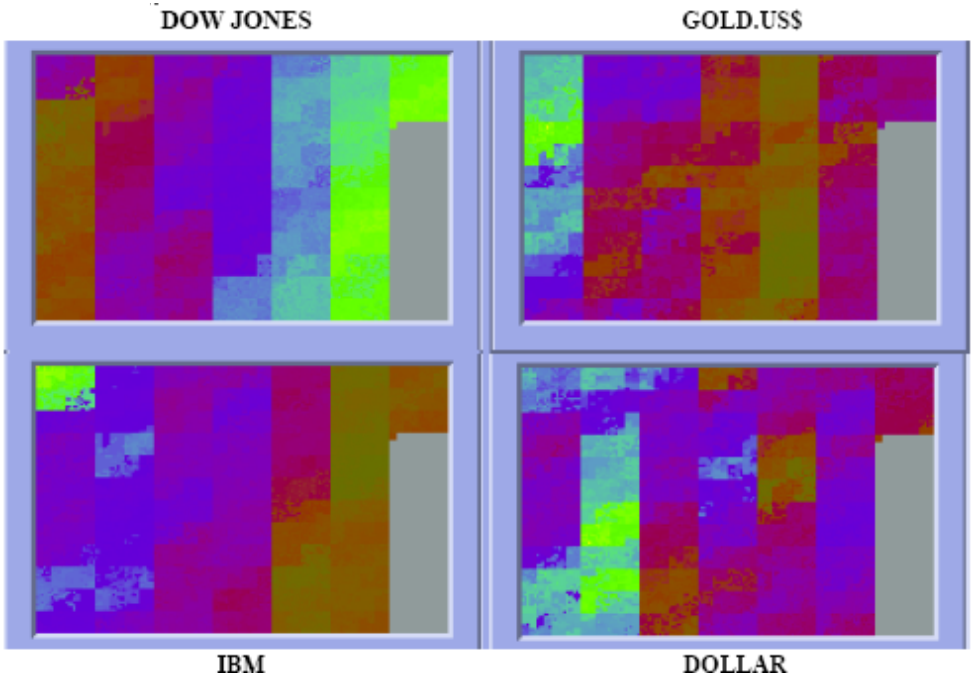




W Horn *et al.*: Metaphor graphics to visualize ICU data over time, In IDAMAP 1998.

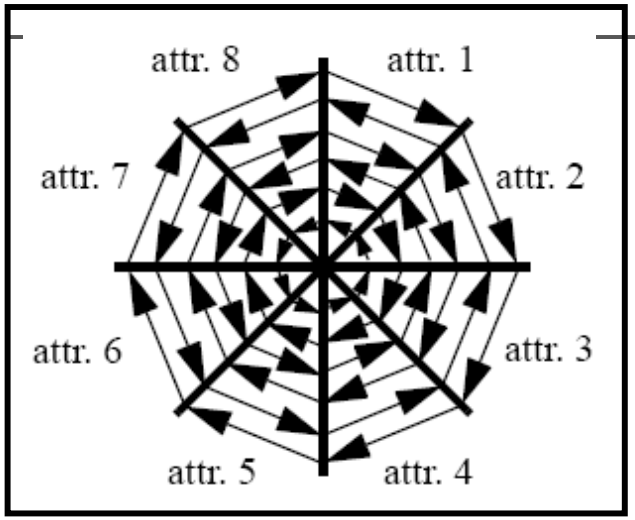
Dense Pixel Displays

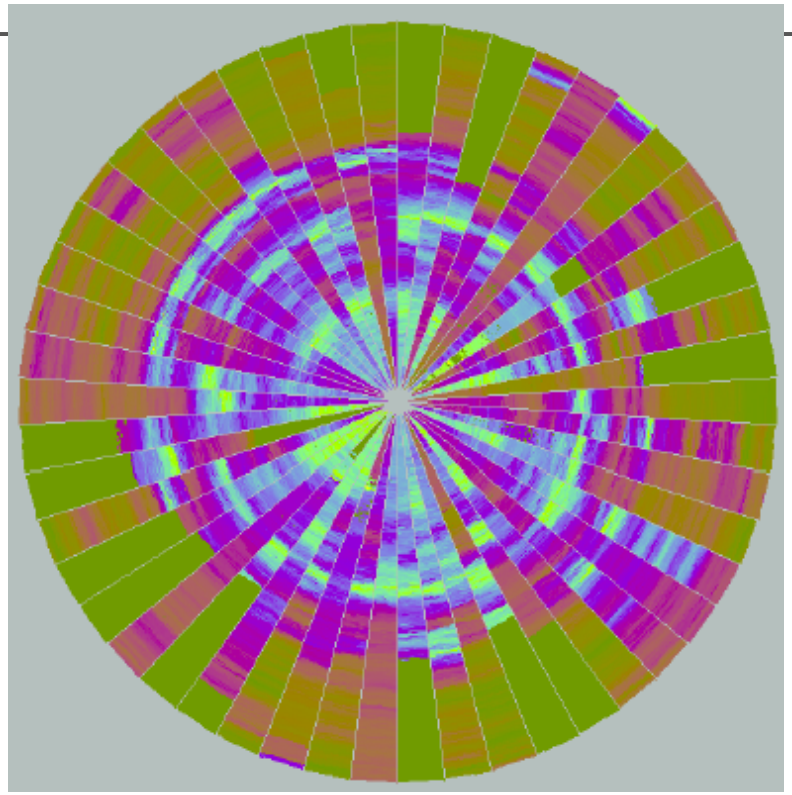




DA Keim *et al.*: Recursive Pattern: A technique for visualizing very large amounts of data Proc. Visualization 95, pages 279-286, 1995.

Dense Pixel Displays

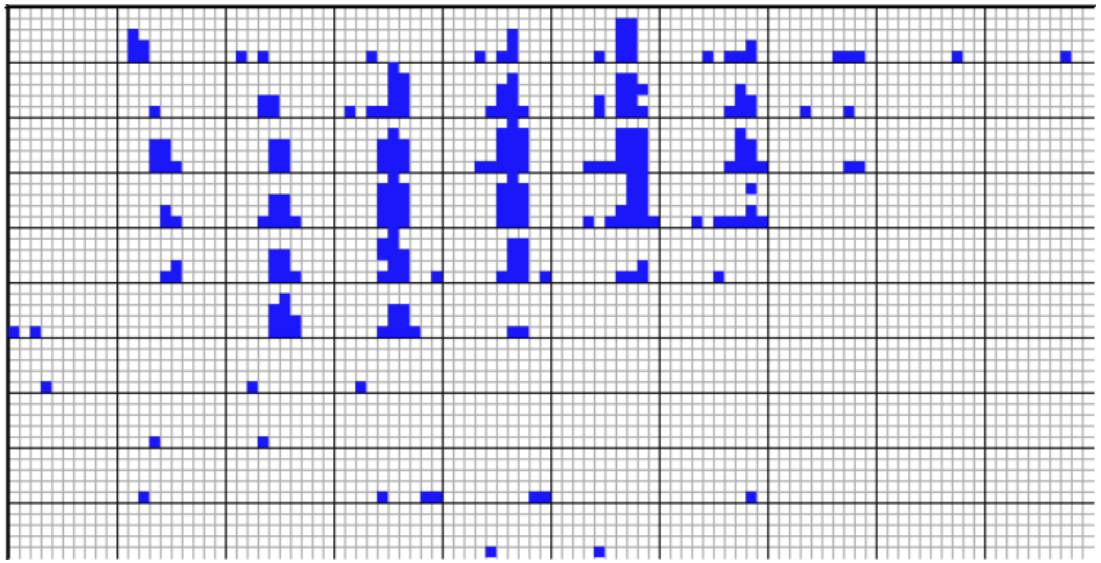




Ankerst *et al.:* Circle Segments: A technique for visually exploring large multidimensional data sets. In Proc. Visualization 96, Hot Topic Session, 1996.

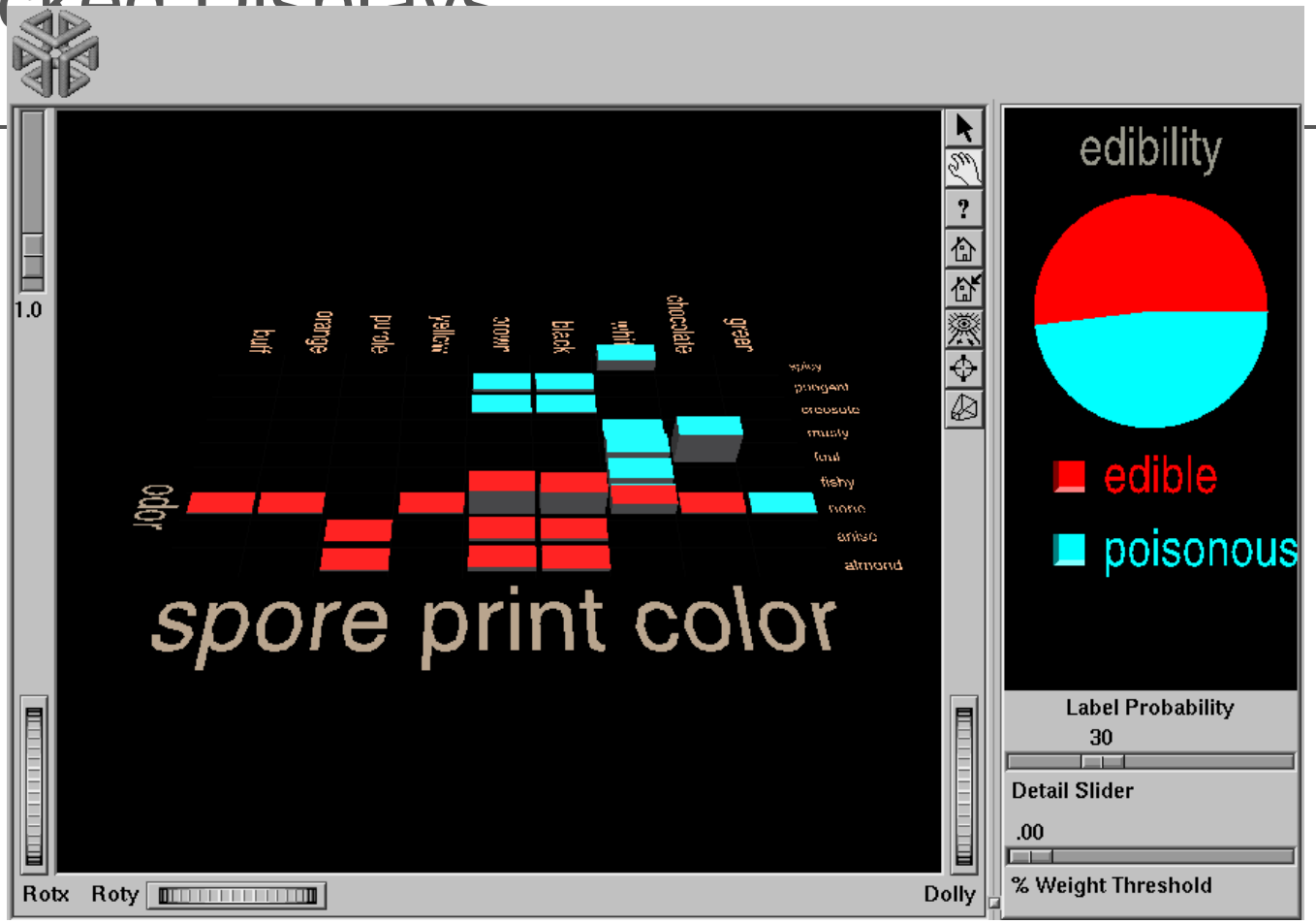
Stacked Displays

- an example is dimensional stacking
- embed one coordinate system within the other
- e.g. two attributes in one system, then another two when drilling down



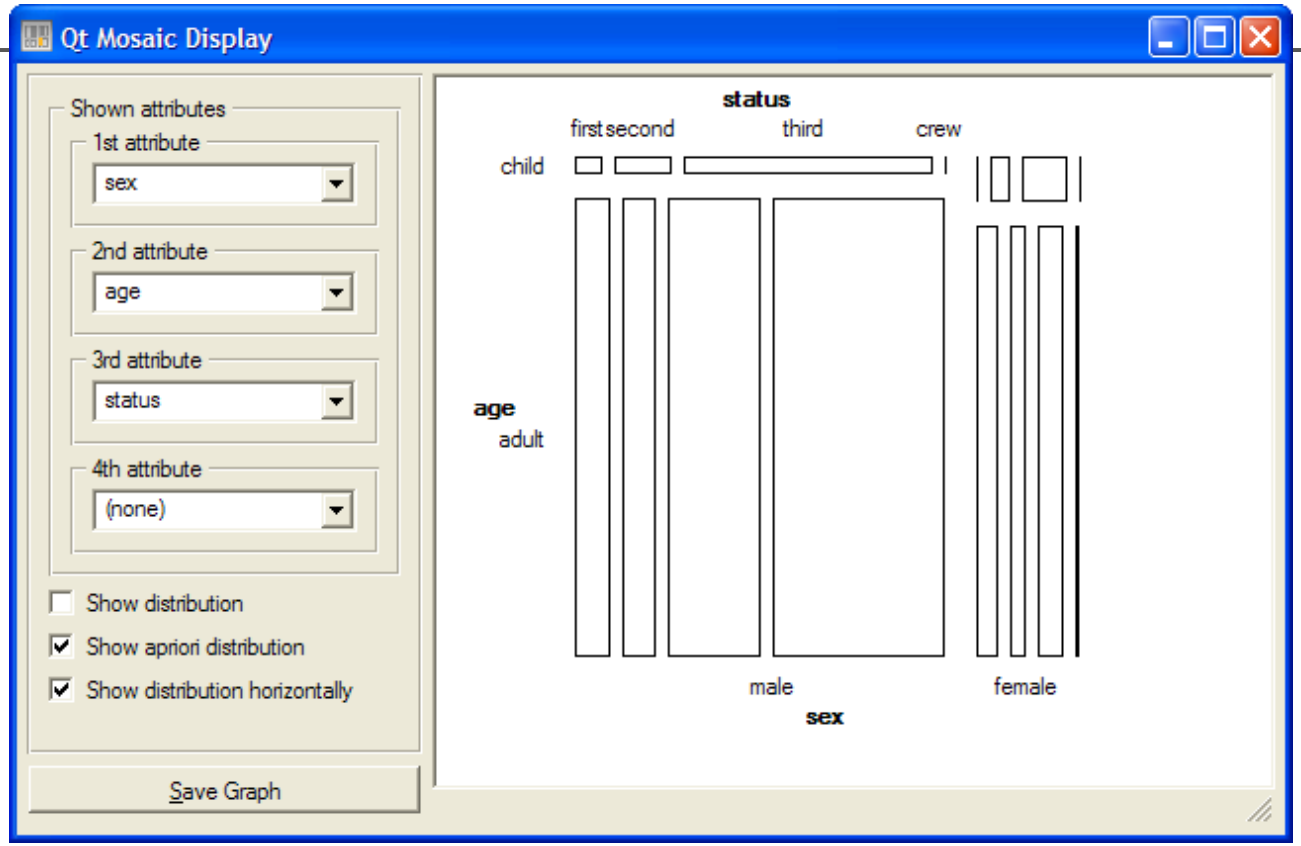
J LeBlanc *et al.*: Exploring n-dimensional databases. In Proc. Visualization 90, pages 230-239, 1990.

Stacked Displays



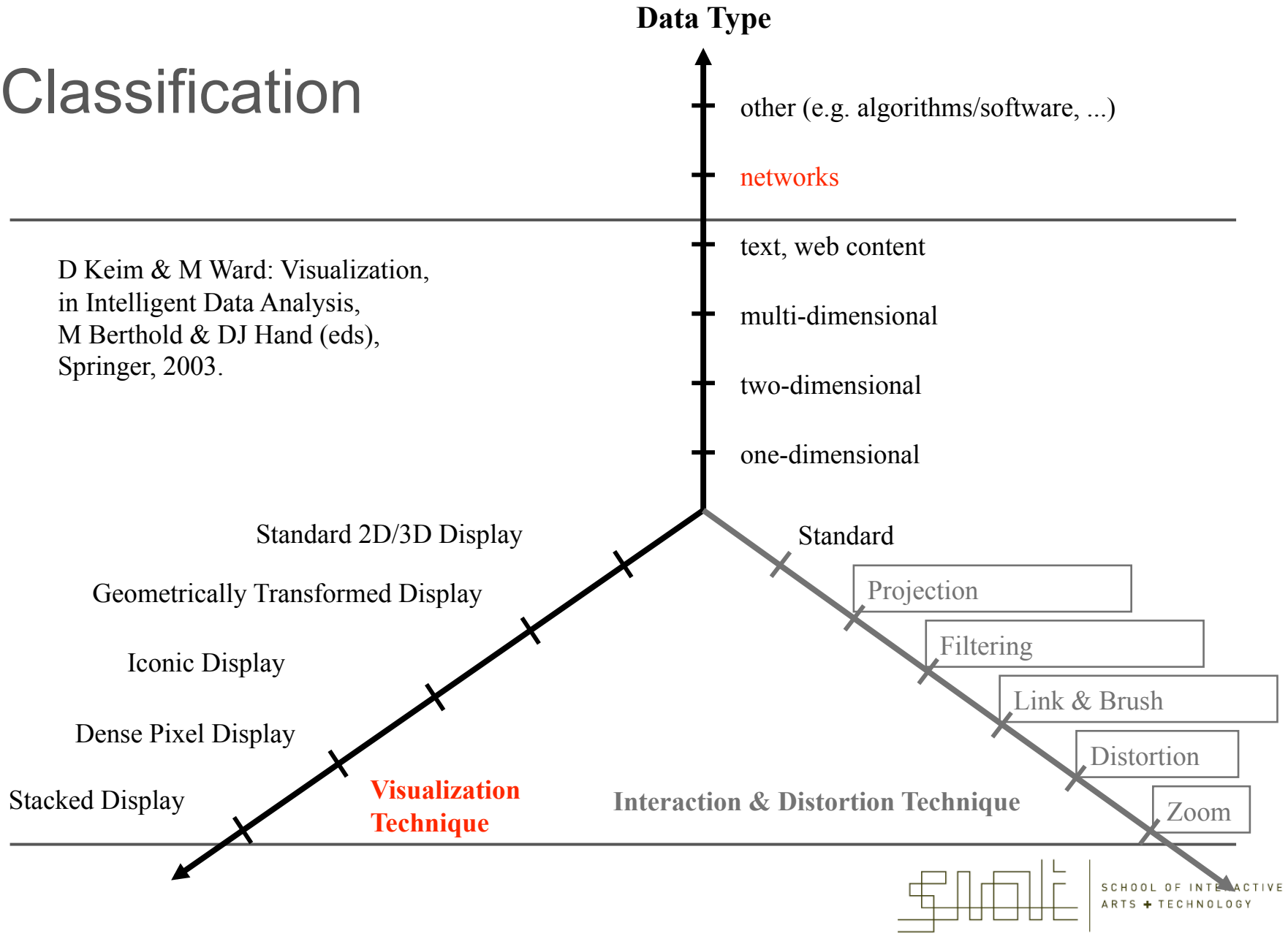
Decision table visualization from SGI's MineSet

Stacked Displays



Mosaic display in Orange.

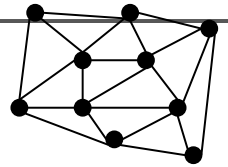
Classification



Visualizing Networks

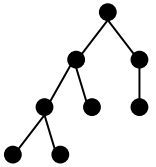
- Networks/graphs vs trees
- Algorithms for network layout
- Multidimensional networks
- New toolkits

Graph and Tree Data Structures



- **Graphs**:
  - Representations of structured, connected data
  - Consist of a set of ***nodes*** (data) and a set of ***edges*** (relations)
  - Edges can be **directed** or **undirected**

- **Trees**:
  - Graphs with a specific structure
    - connected graph with n-1 edges
  - Representations of data with natural hierarchy
  - Nodes are either ***parents*** or ***children***



When is Graph Visualization Applicable?

- Ask the question: *is there an inherent relation among the data elements to be visualized?*
  - If YES – then the data can be represented by nodes of a graph, with edges representing the relations.
  - If NO – then the data elements are “unstructured” and goal is to use visualization to analyze and discover relationships among data.

    Source: Herman, Graph Visualization and Navigation in Information Visualization: a Survey

Common Applications

- Process Visualization (e.g., Visio)
- Dependency Graphs
- Biological Interactions (Genes, Proteins)
- Computer Networks
- Social Networks
- Simulation and Modeling

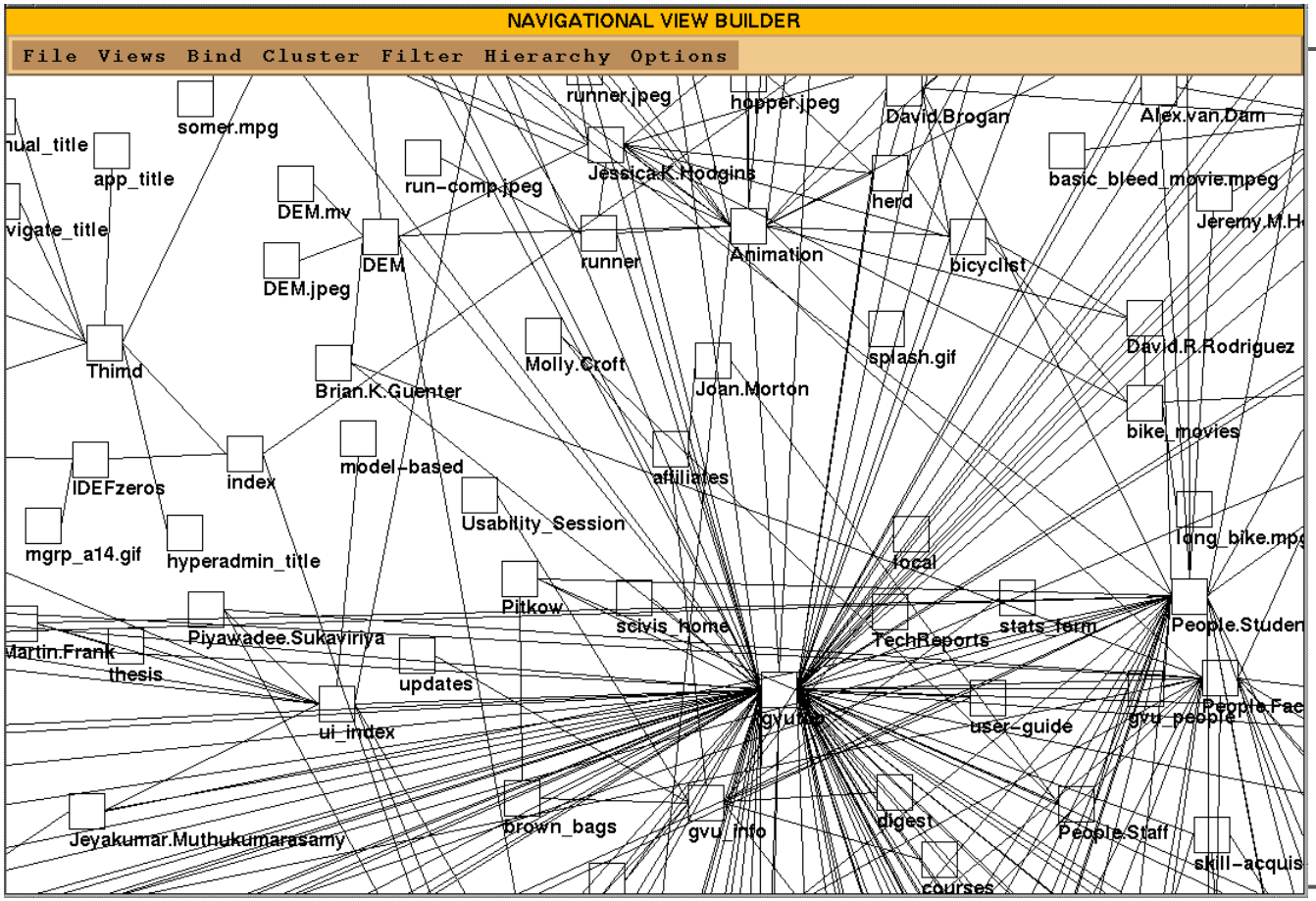
Slide adapted from Jeff Heer

Graph Layout

- How to position the nodes and edges?
- The **primary** concern with networks
  - while inheriting other issues such as color, size, etc
- The topic of the Graph Drawing conference (as well as numerous InfoVis papers) and even multiple books.

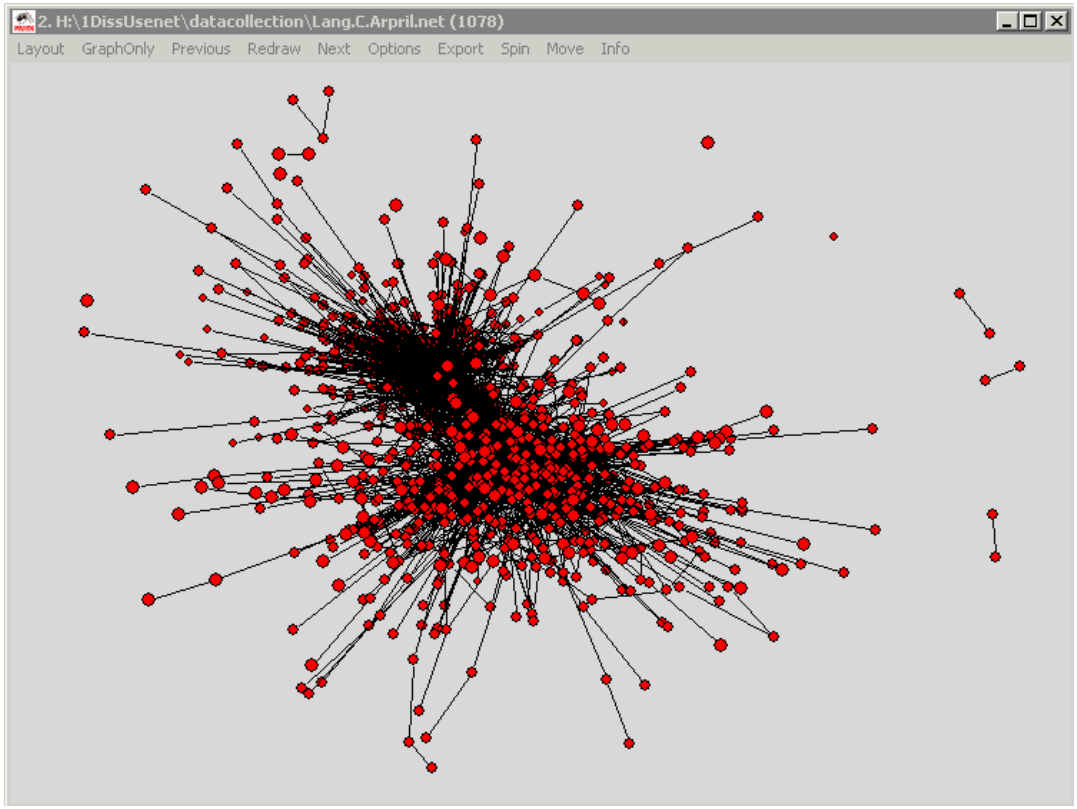
Slide adapted from Jeff Heer

Graphs: the problem

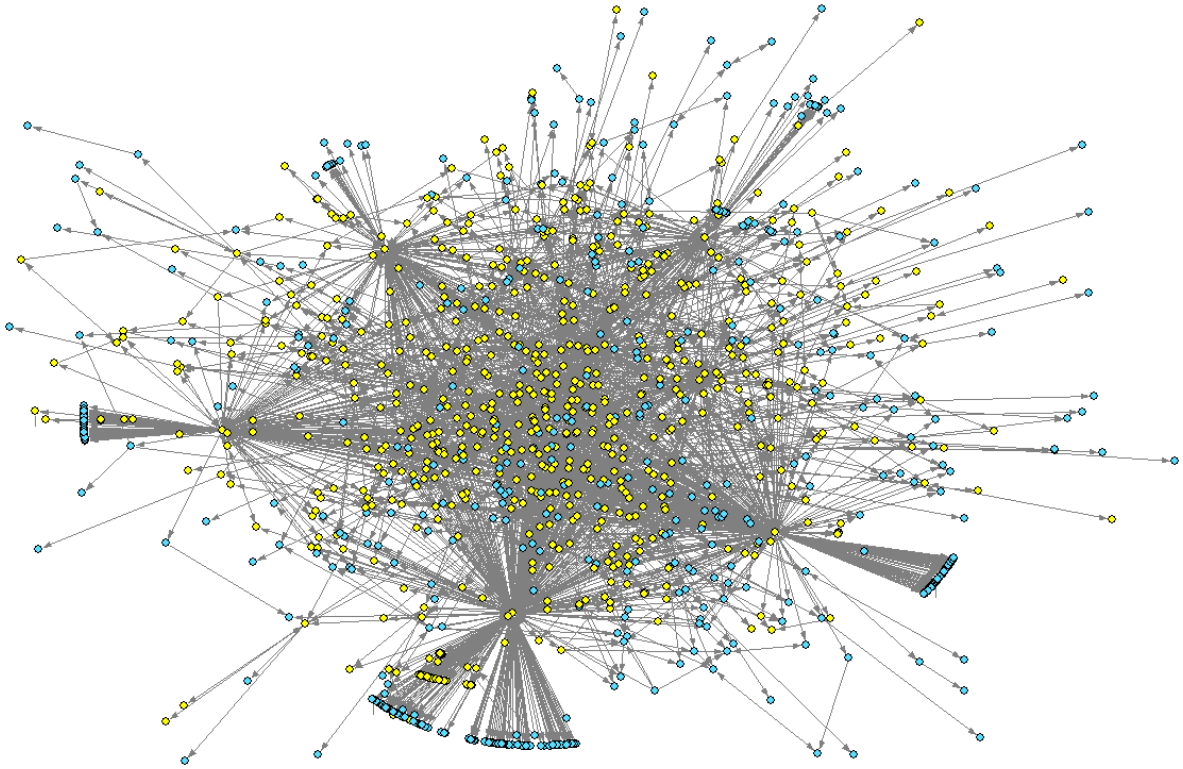


- YUK ……

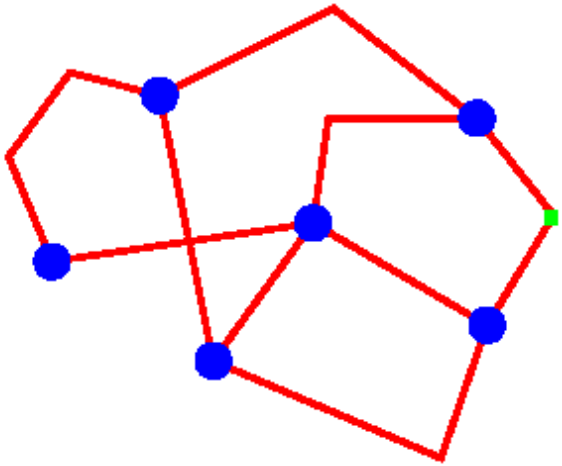
Graph Layout: The Problem



Graph Layout: The Problem

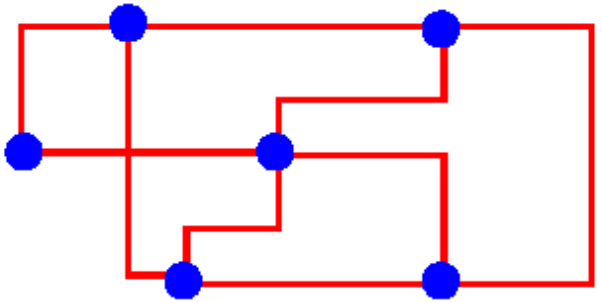


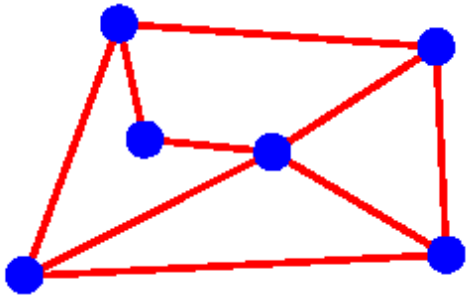
Traditional Graph Drawing



poly-line graphs
(includes bends)

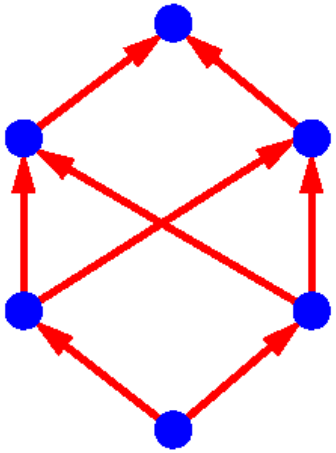
orthogonal
drawing





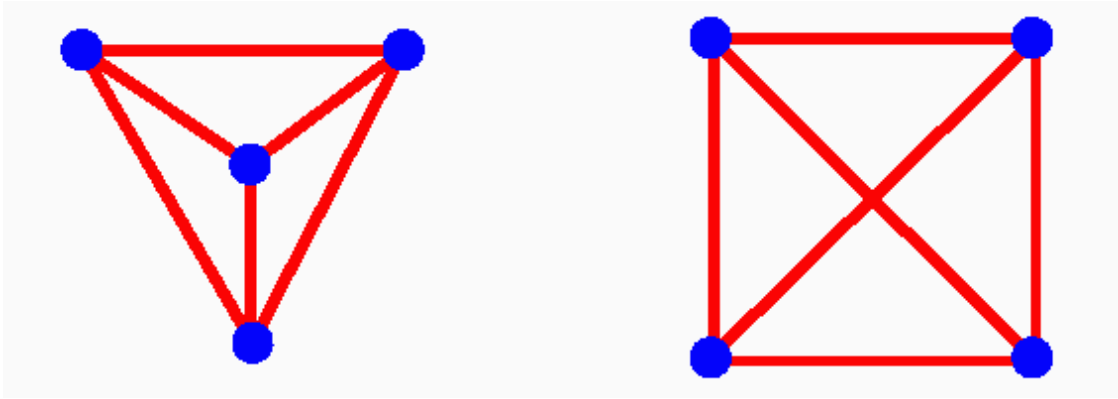
planar, straight-
line drawing

upward drawing of DAGs



Aesthetic constraints

- Minimize link crossings
- Minimize link lengths
- Minimize link bends
- Maximize symmetries
- Mathematically difficult to do everything



- Unsuited for interactive visualization
- Prioritise and precompute

Layout Approaches

- Tree-ify the graph - then use tree layout
- Hierarchical graph layout
- Radial graph layout
- Optimization-based techniques
  - Includes spring-embedding / force-directed layout
- Adjacency matrices
- Structurally-independent layout
- On-demand revealing of subgraphs
- Distortion-based views
  - Hyperbolic browser

- (this list is not meant to be exhaustive)

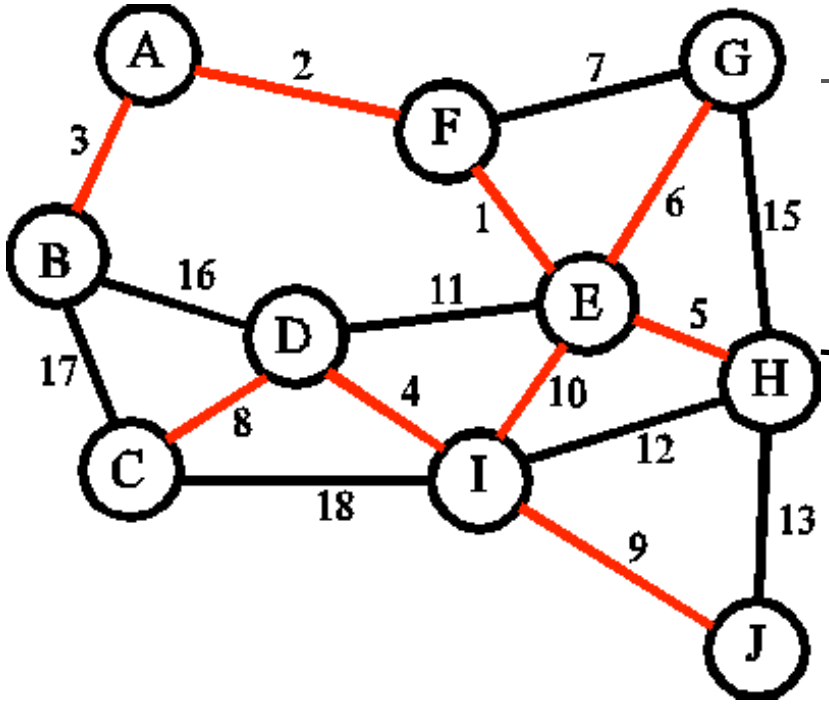
Slide adapted from Jeff Heer

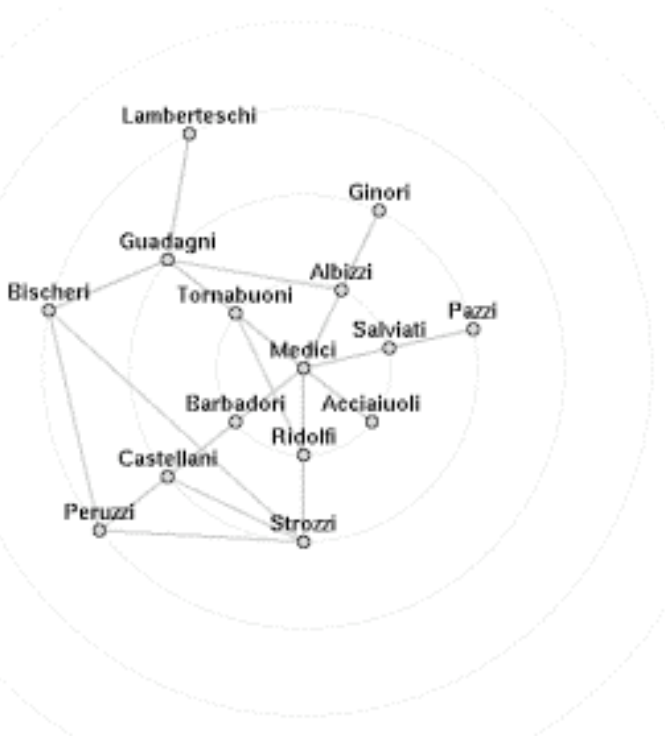
Tree-based graph layout

- Select a tree-structure out of the graph
  - Breadth-first-search tree
  - Minimum spanning tree
  - Other domain-specific structures
- Use a tree layout algorithm
- Benefits
  - Fast, supports interaction and refinement
- Drawbacks
  - Limited range of layouts

Slide adapted from Jeff Heer

Tree-ify the graph





Slide adapted from Jeff Heer

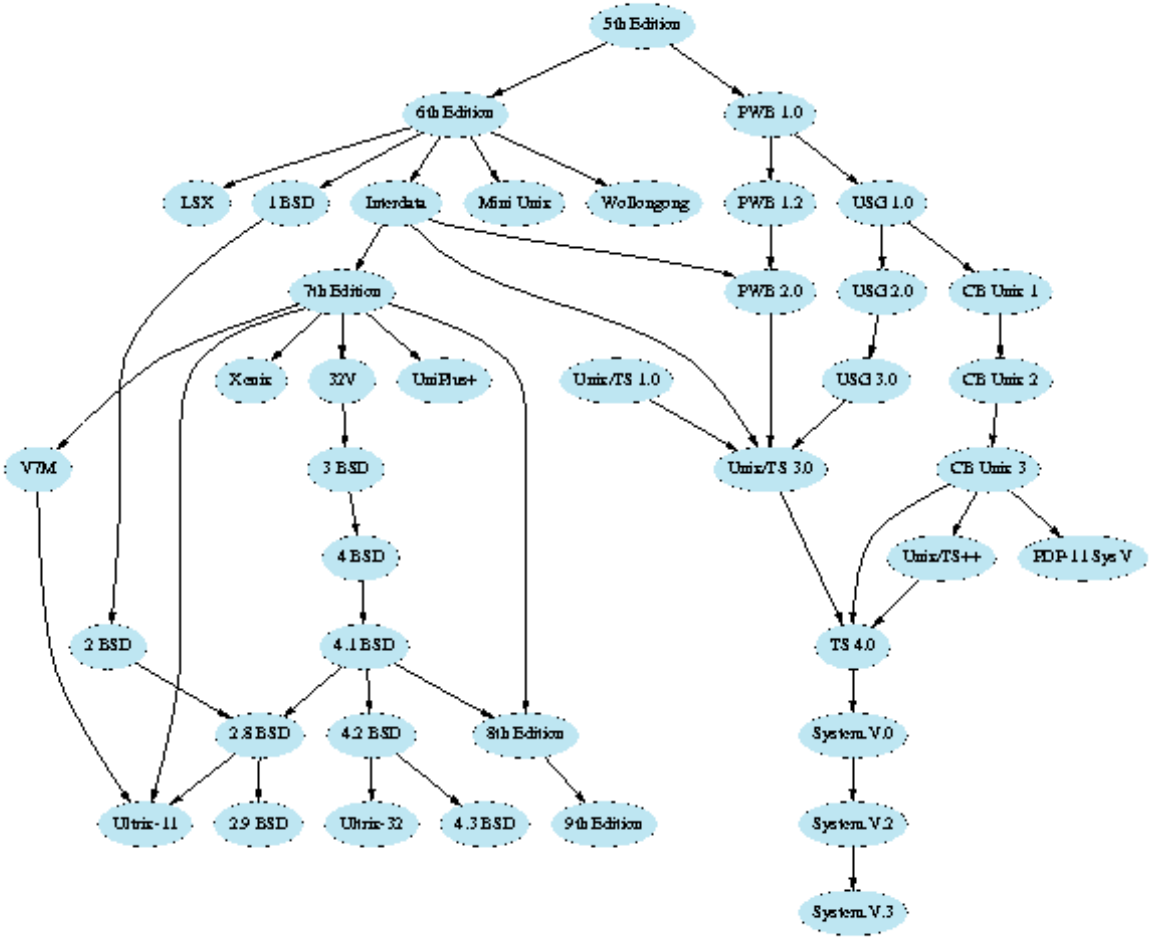
Hierarchical graph layout

- Use directed structure of graph to inform layout
- Order the graph into distinct levels
  - this determines one dimension
- Now optimize within levels
  - determines the second dimension
  - minimize edge crossings, etc
- The method used in graphviz's "dot" algorithm
- Great for directed acyclic graphs, but often misleading in the case of cycles

Slide adapted from Jeff Heer

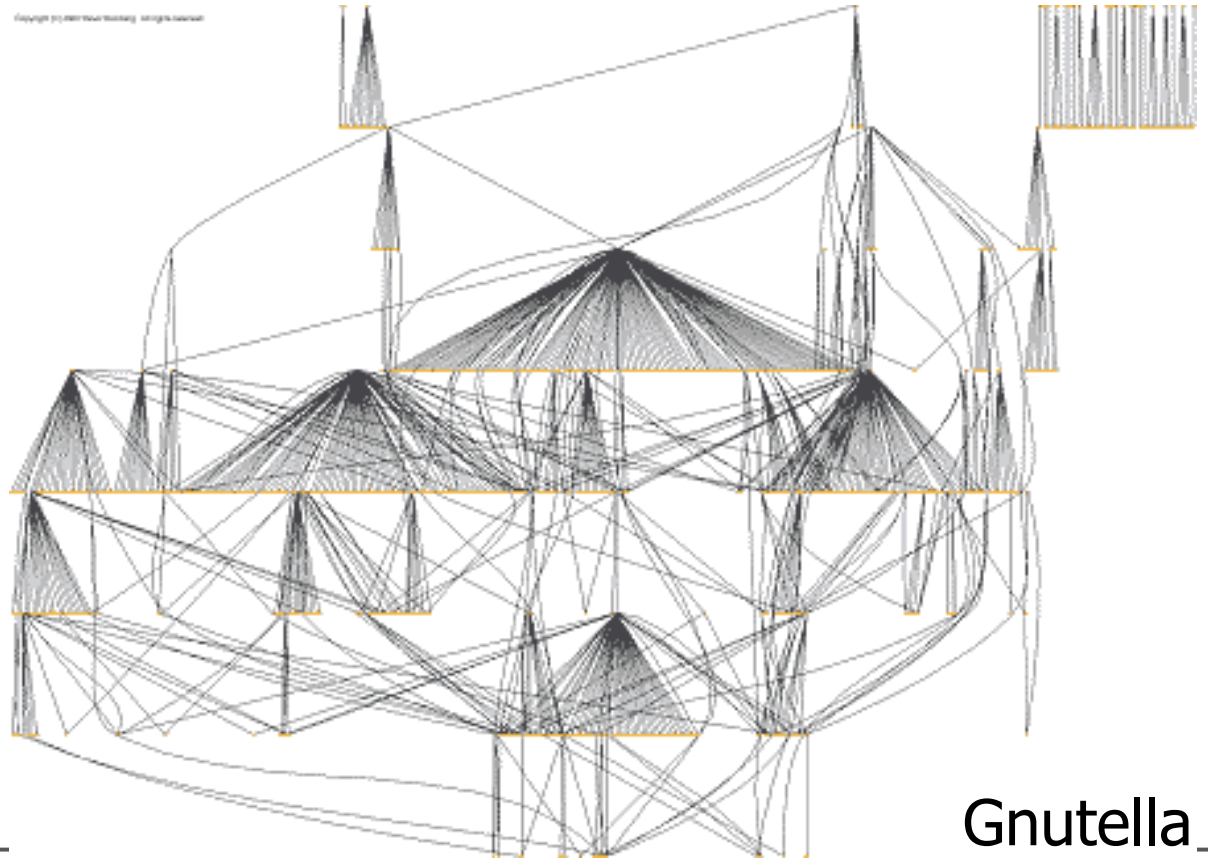
Hierarchical Graph Layout

- Evolution of the UNIX operating system
- Hierarchical layering based on descent



Slide adapted from Jeff Heer

Hierarchical graph layout



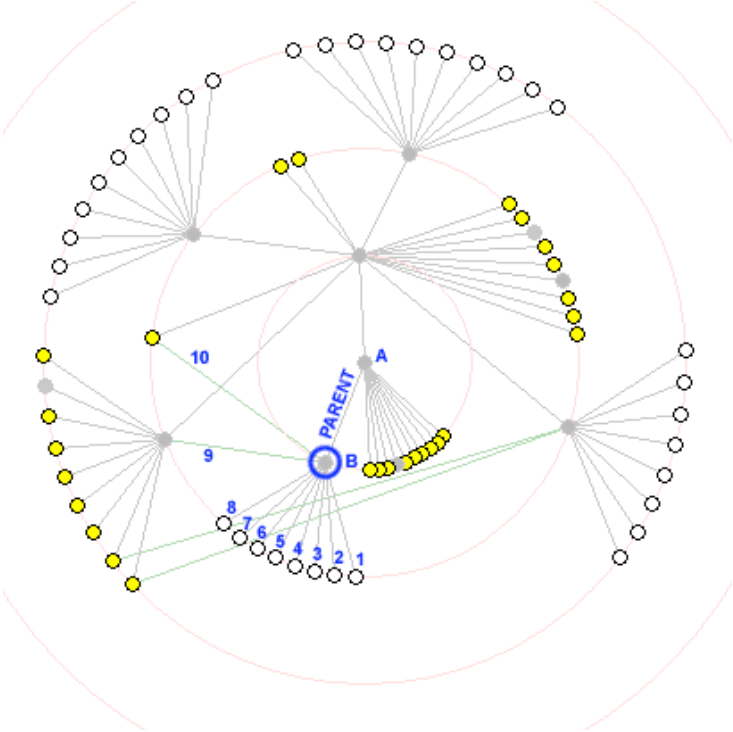
Gnutella network

Slide adapted from Jeff Heer

SCHOOL OF INTERACTIVE
ARTS + TECHNOLOGY

Radial Layout

- **Animated Exploration of Graphs with Radial Layout**, Yee et al., 2001
- Gnutella network

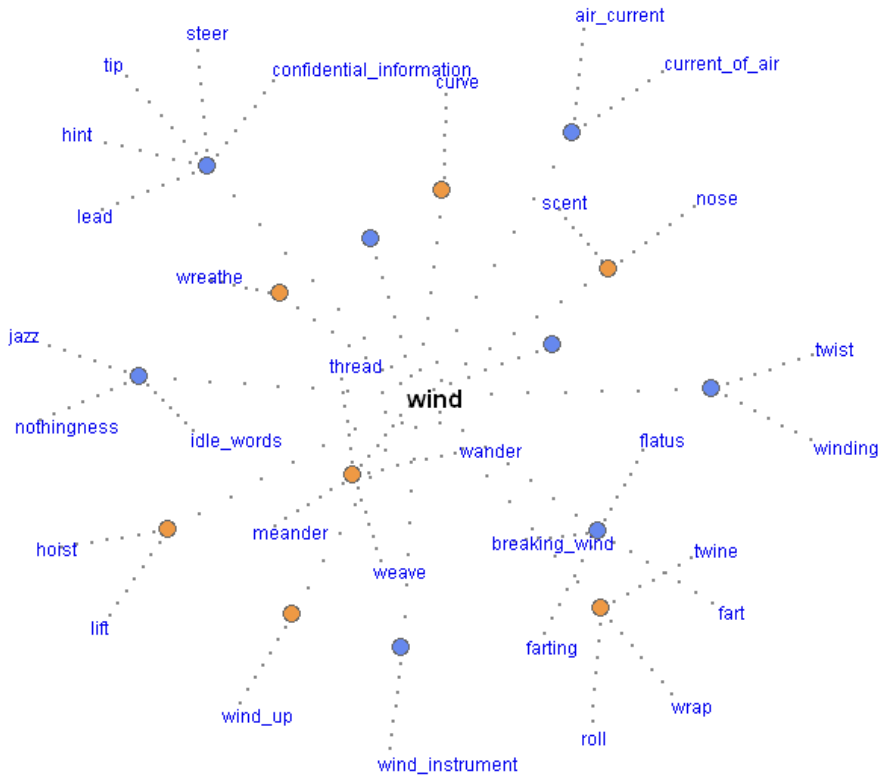


Optimization-based layout

- Specify constraints for layout
  - Series of mathematical equations
  - Hand to “solver” which tries to optimize the constraints
- Examples
  - Minimize edge crossings, line bends, etc
  - Multi-dimensional scaling (preserve multi-dim distance)
  - Force-directed placement (use physics metaphor)
- Benefits
  - General applicability
  - Often customizable by adding new constraints
- Drawbacks
  - Approximate constraint satisfaction
  - Running time; “organic” look not always desired

Slide adapted from Jeff Heer

Example: Force-Directed Layout



Uses physics model to layout graph,

Nodes repel each other, edges act as springs, and some amount of friction or drag force is used.

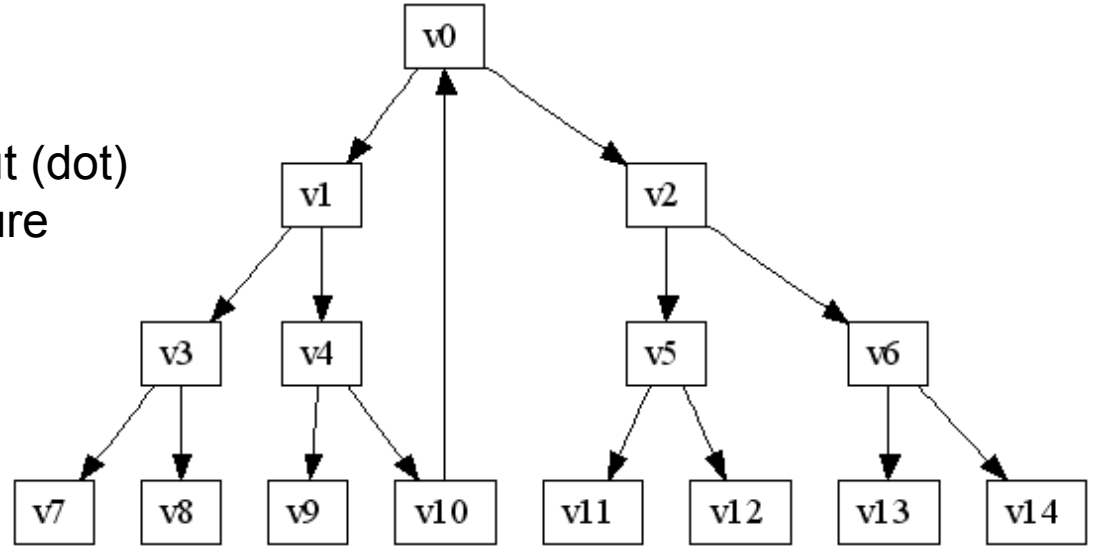
Special techniques to dampen “jitter”.

visual wordnet http://www.kylescholz.com/projects/wordnet

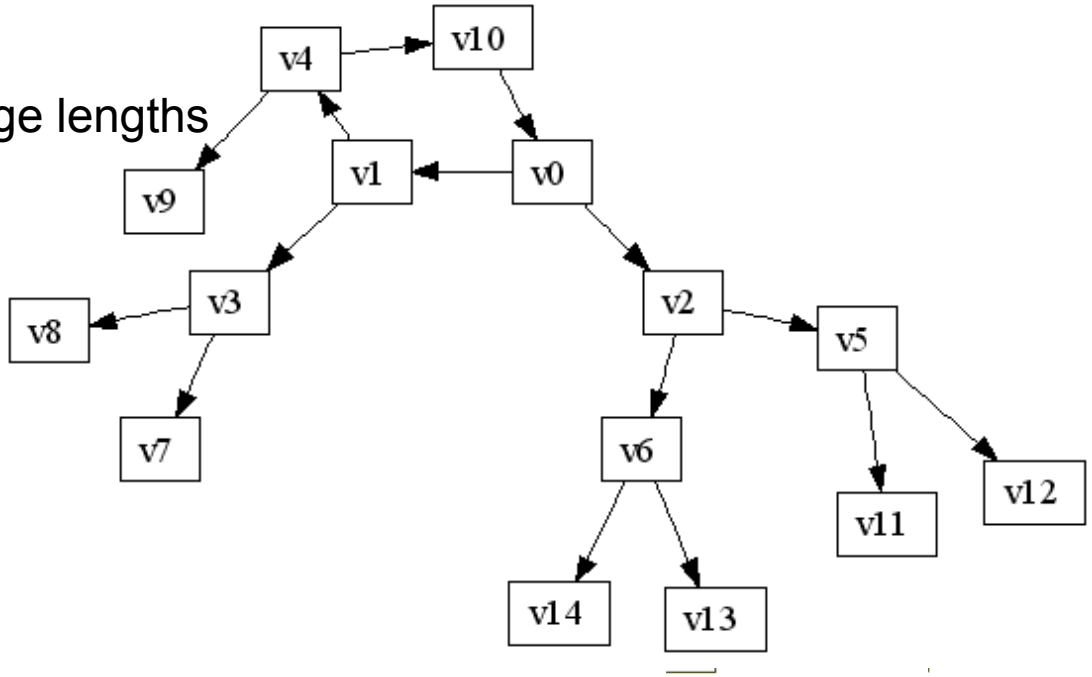
visuwords http://www.visuwords.com/

Slide adapted from Jeff Heer

Typical Sugiyama layout (dot)
- preserves tree structure

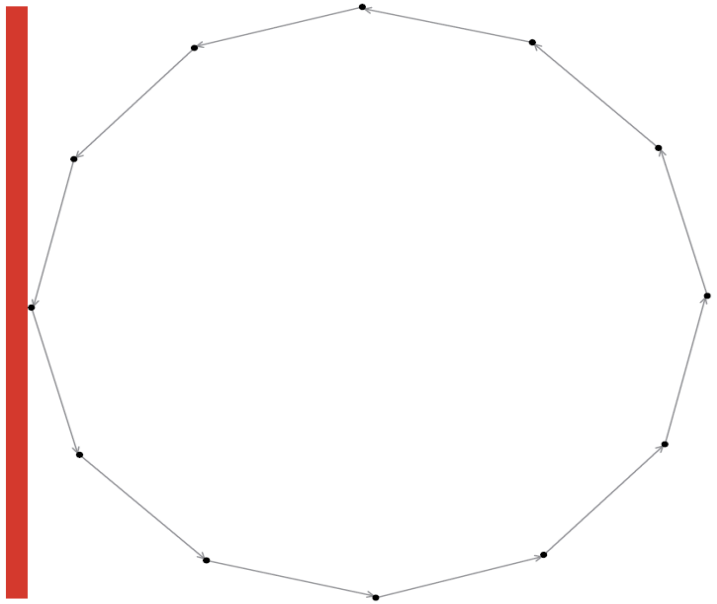


Alternative method
- preserves uniform edge lengths

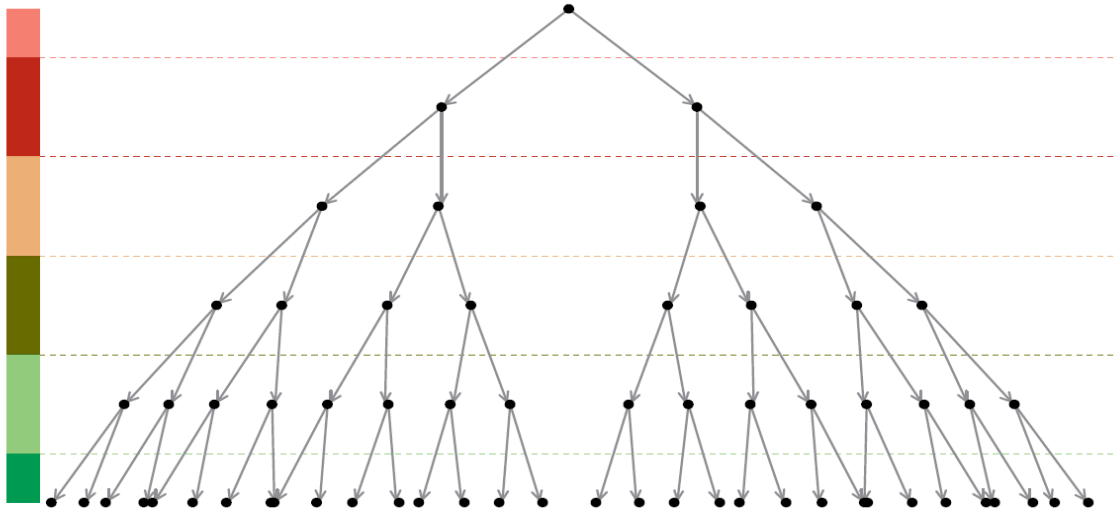


slide borrowed from Tim Dwyer

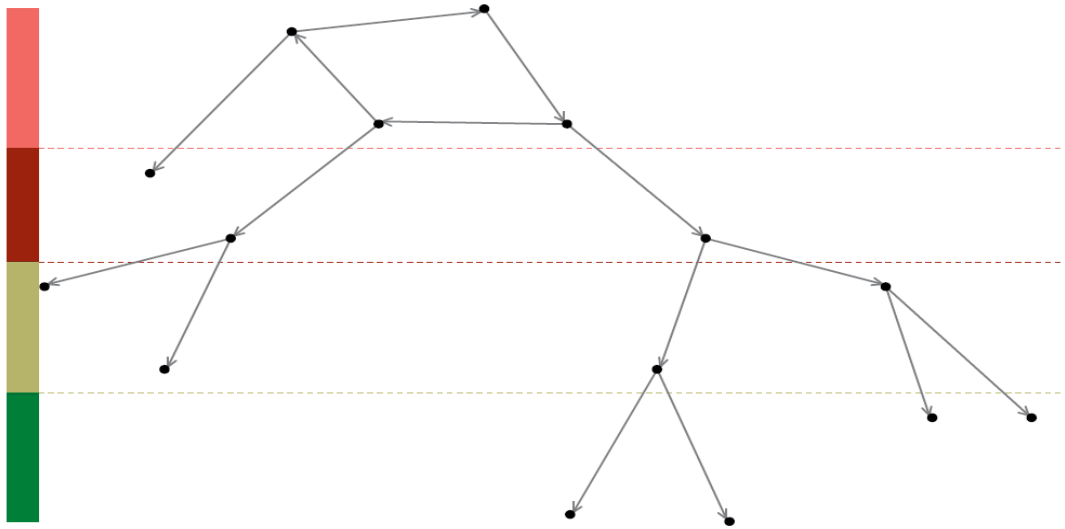
slide borrowed from Tim Dwyer



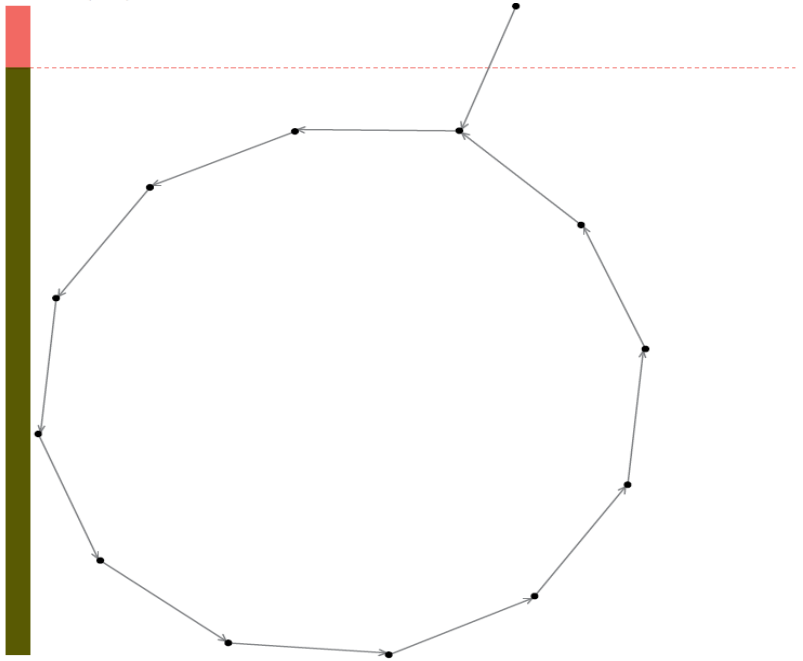
(a) 1 band



(b) 6 bands



(c) 4 bands

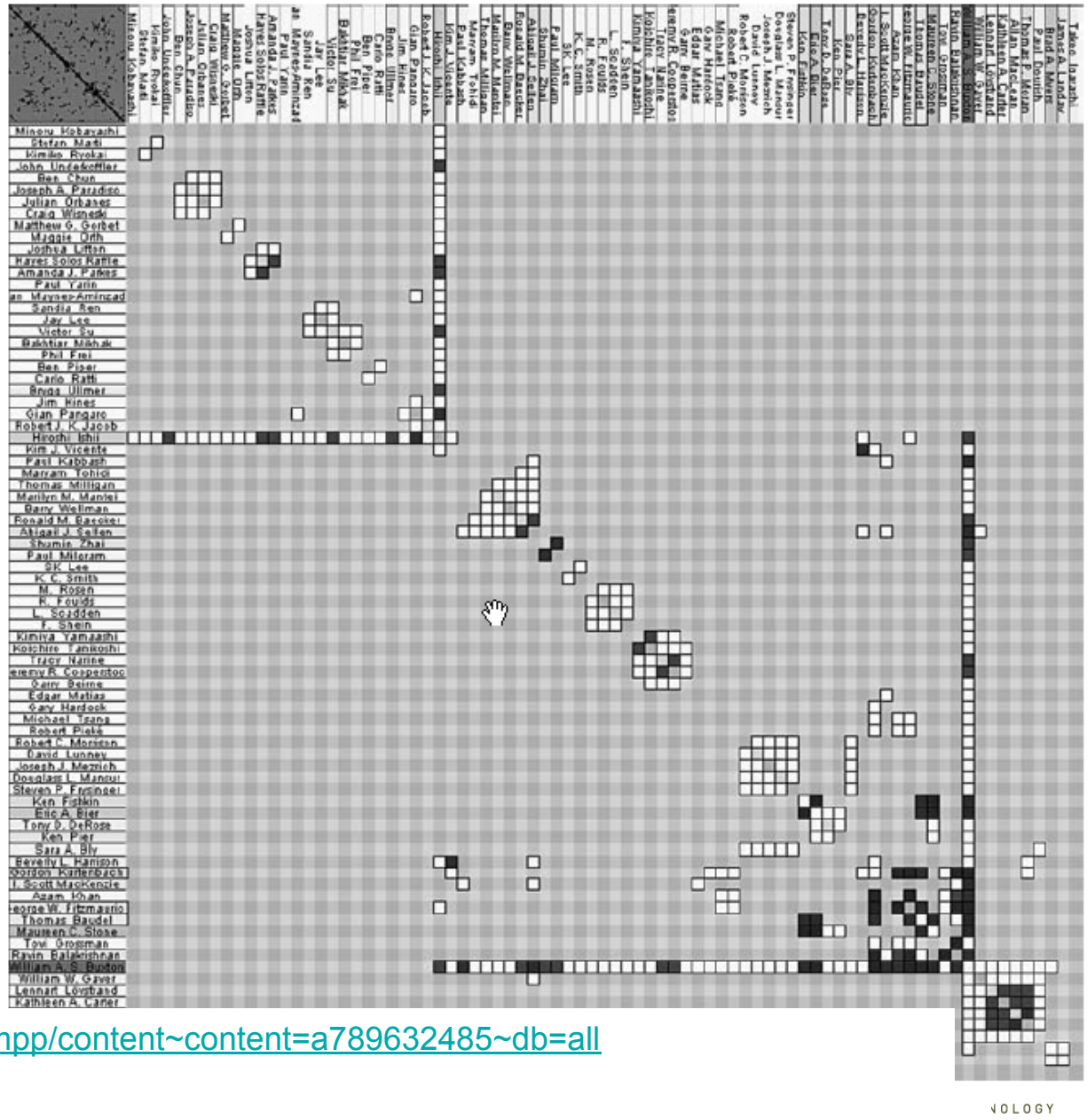


(d) 2 bands

Adjacency Matrices

- So far, only looked at node-link diagrams
- Often doesn't scale well due to edge-crossings, occlusion, etc. --> hard to read
- One solution: adjacency matrix
  - show graph as table
  - nodes as rows/columns
  - edges as table cells

Matrices

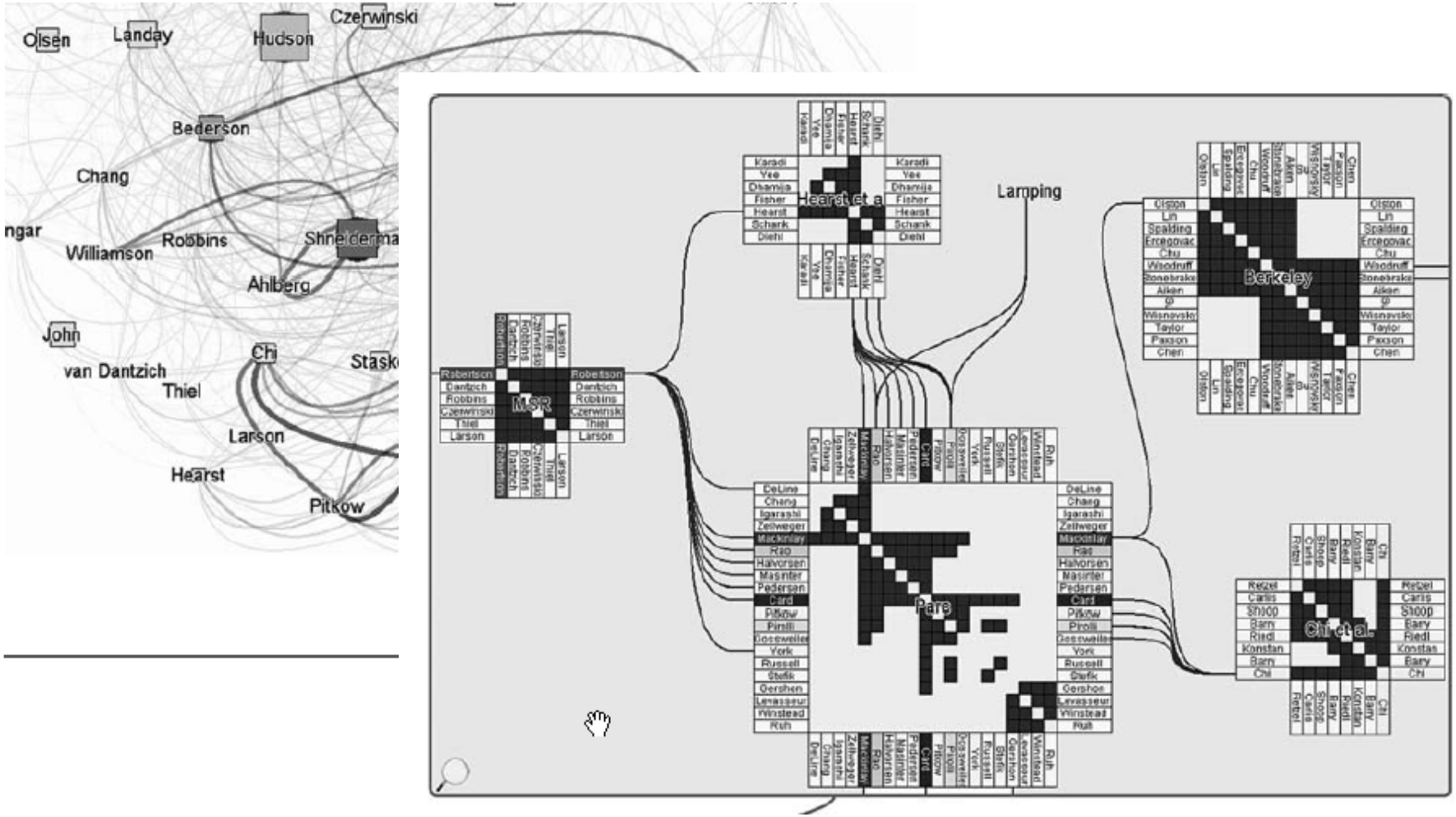


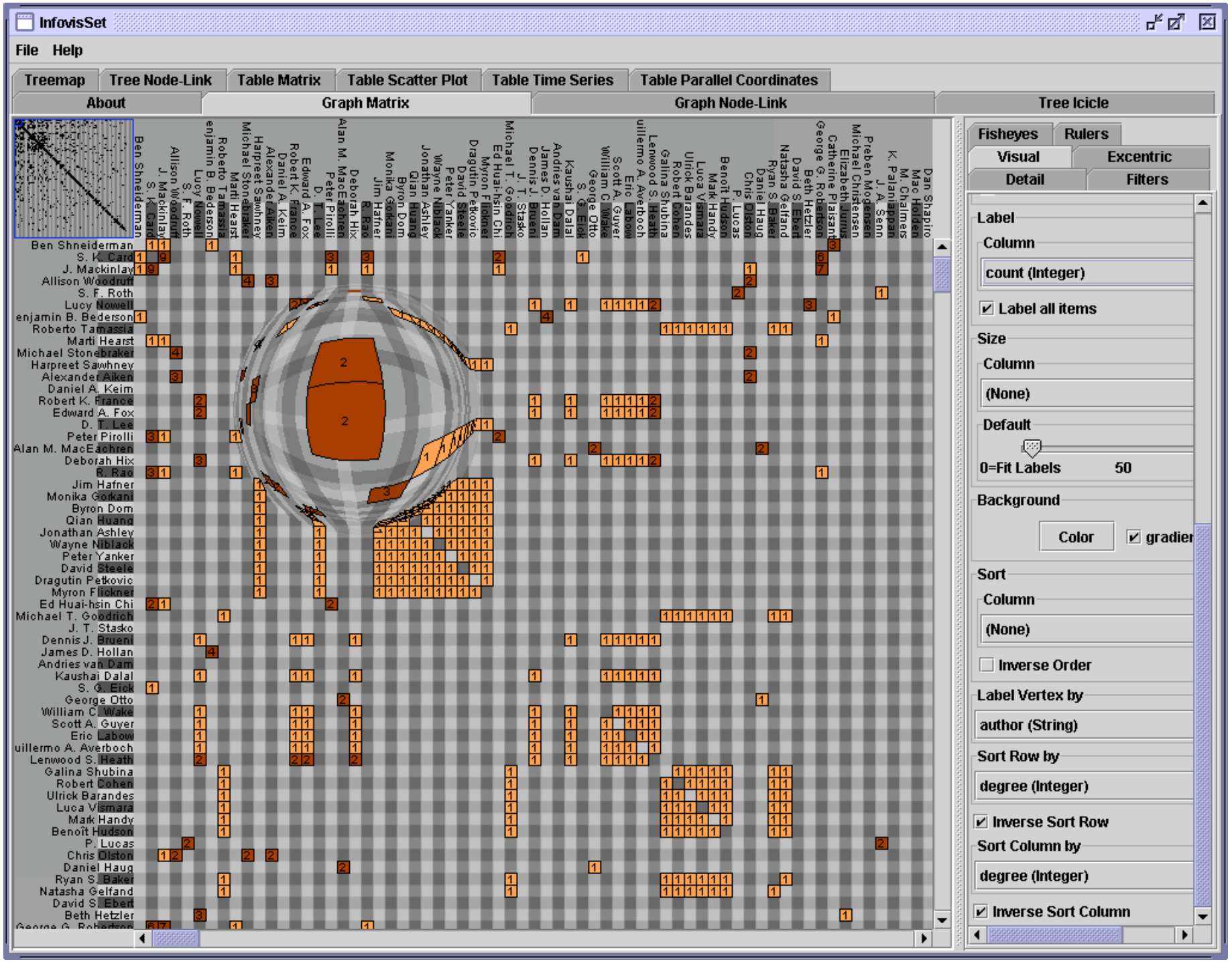
http://www.informaworld.com/smpp/content~content=a789632485~db=all

20 Years of Four HCI Conferences: A Visual Exploration
Henry et al. IJHCI 2007

Matrices with Submatrices

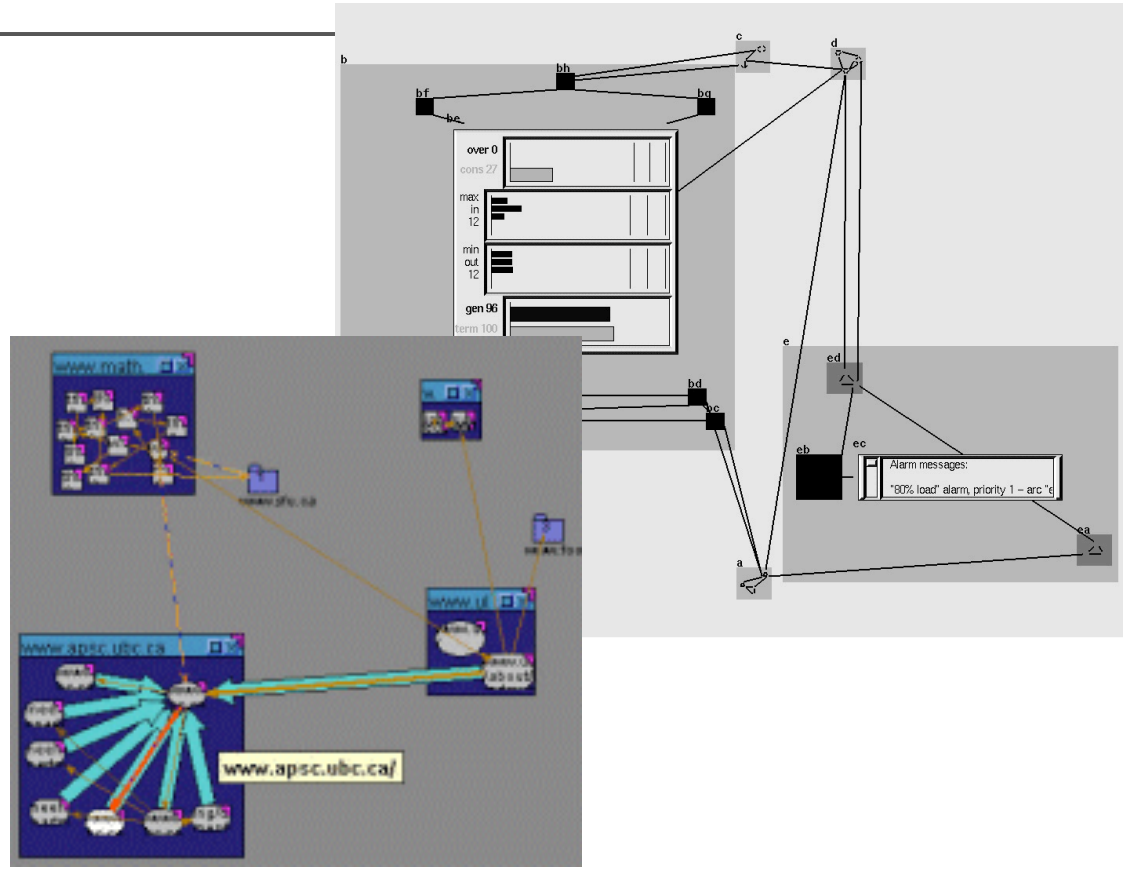
- http://www.informaworld.com/smpp/content~content=a789632485~db=all




InfovisSet
File
Help
Treemap
Tree Node-Link
Table Matrix
Table Scatter Plot
Table Time Series
Table Parallel Coordinates
About
Graph Matrix
Graph Node-Link
Tree Icicle
Fisheyes
Rulers
Visual
Excentric
Detail
Filters
Label
Column
count (Integer)
Label all items
Size
Column
(None)
Default
0=Fit Labels
50
Background
Color
gradier
Sort
Column
(None)
Inverse Order
Label Vertex by
author (String)
Sort Row by
degree (Integer)
Inverse Sort Row
Sort Column by
degree (Integer)
Inverse Sort Column

Focus+Context

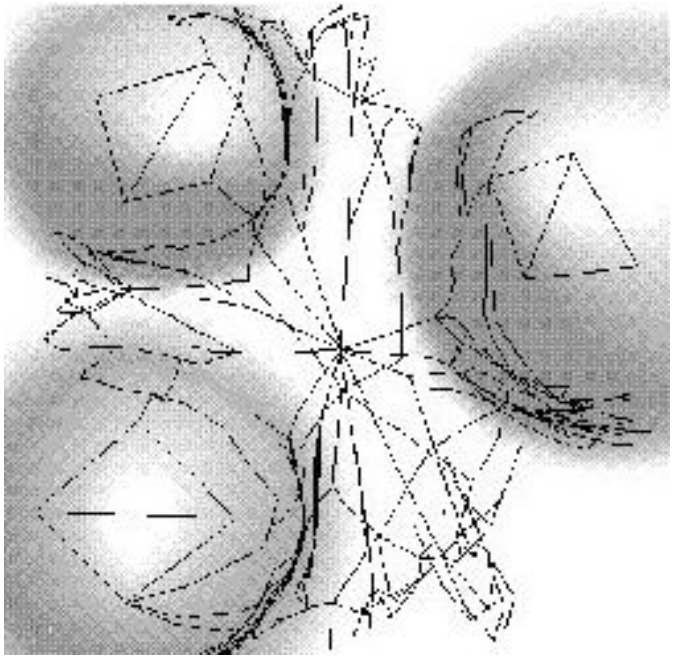
- Pioneered by Furnas
- Continuous Zoom (Bartram, Ho, Dill)
  - CZWeb
  - CZSaw



Slide adapted from Jeff Heer

Fisheye and Nested Approaches

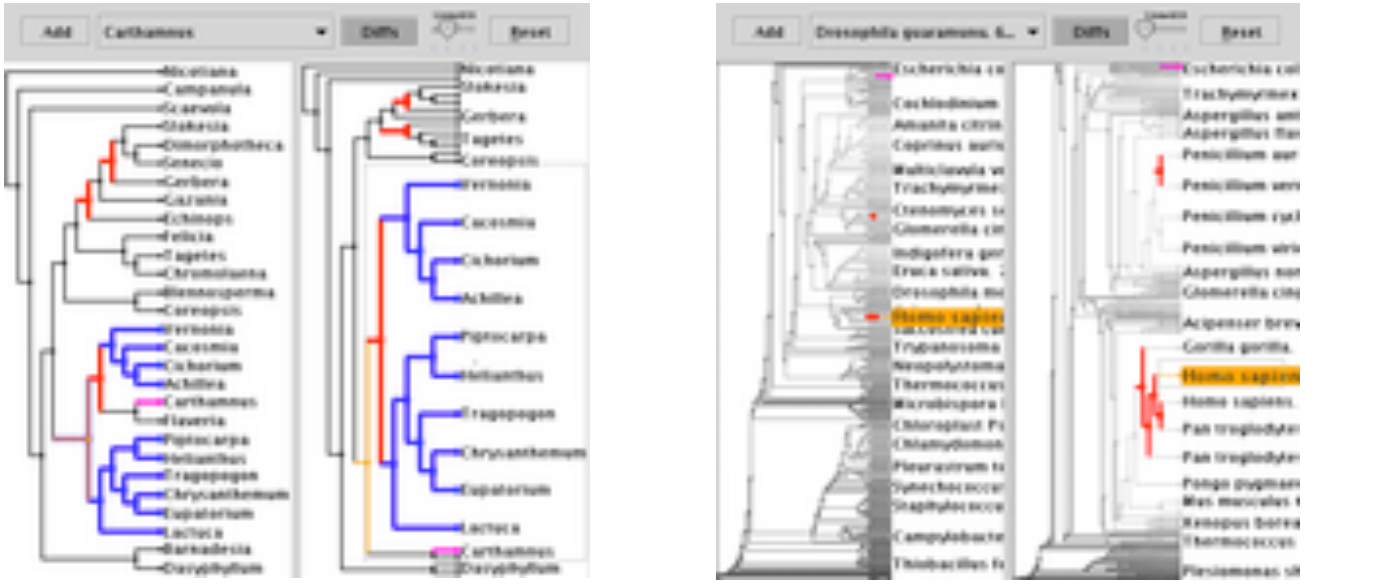
- Carpendale : Extended Pliable Surfaces



Slide adapted from Jeff Heer

Fisheye and Nested Approaches

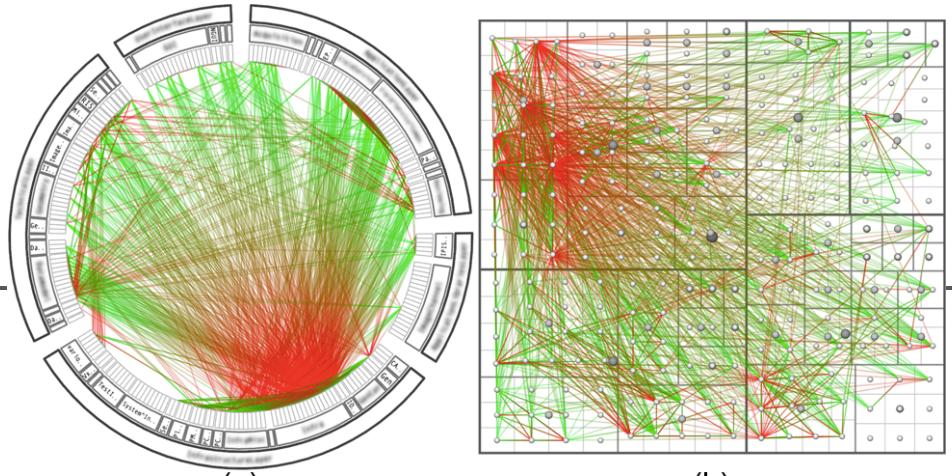
- TreeJuxtaposer (Munzner, 2006)

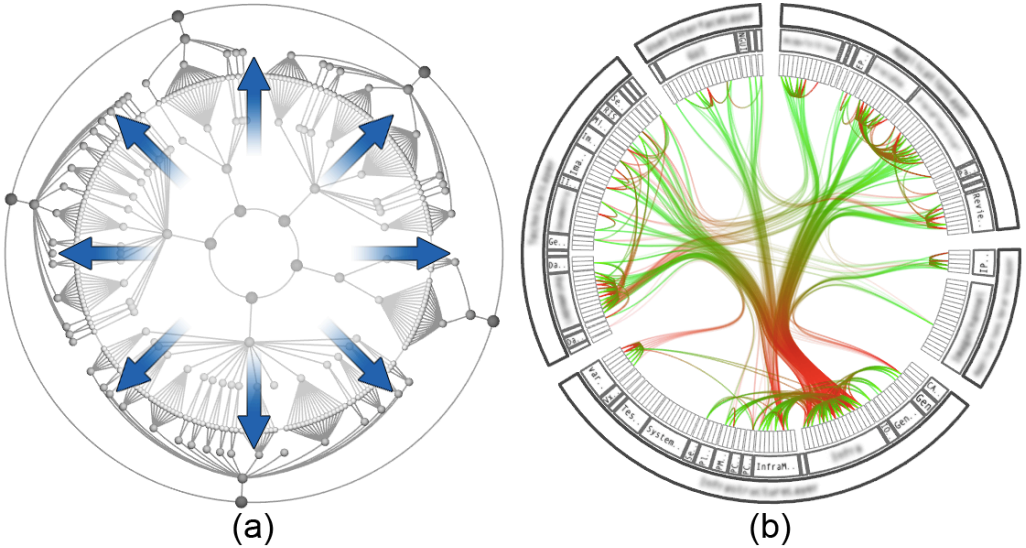


Slide adapted from Jeff Heer

Detail and Context with Links

- Edge Bundles (Holten)





Slide adapted from Jeff Heer

Hyperbolic Browser: Inspiration

Using Distortion and Focus + Context

- The Hyperbolic Tree Browser

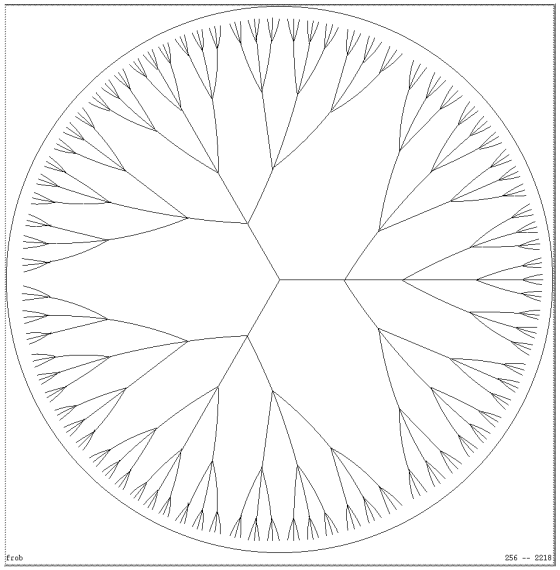
  The Hyperbolic Browser: A Focus + Context Technique for Visualizing Large Hierarchies, Lamping & Rao, CHI 1996.

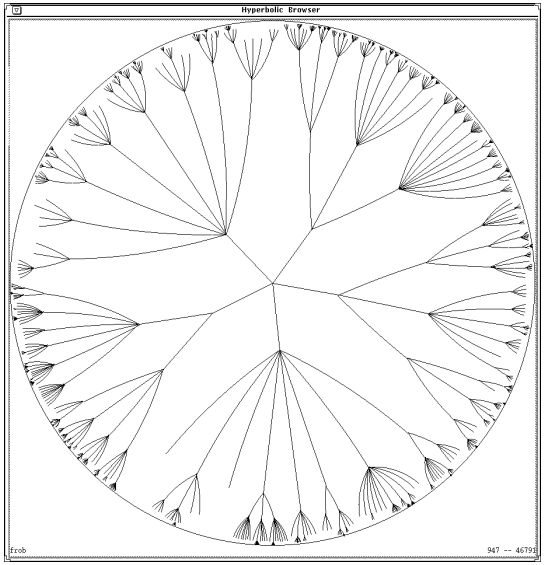
  - http://www.inxight.com/products/sdks/st/
  - Uses non-Euclidean geometry as basis of focus + context technique
    - The hyperbolic browser is a projection into a Euclidean space – a circle
  - The circumference of a circle increases at a linearly with radius (2 PI)
  - The circumference of a circle in hyperbolic space increases exponentially

- Exponential growth in space available with linear growth of radius
  - Makes tree layout easy
- Size of objects decreases with growth of radius
  - Reduces expense of drawing trees when cut-off at one pixel

Appearance of Initial Layout

- Root mapped at center
- Multiple generations of children mapped out towards edge of circle
- Drawing of nodes cuts off when less than one pixel



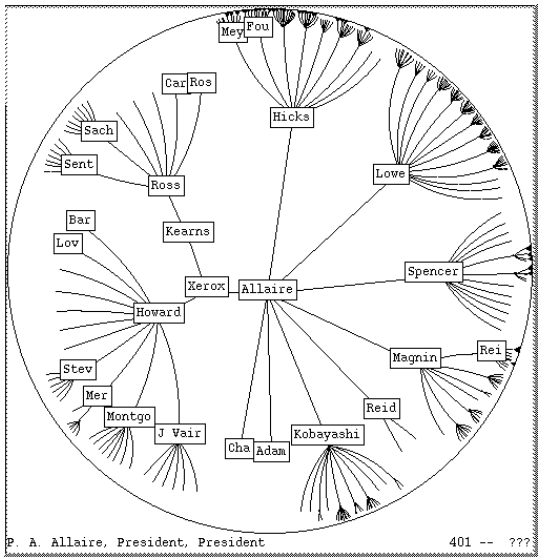


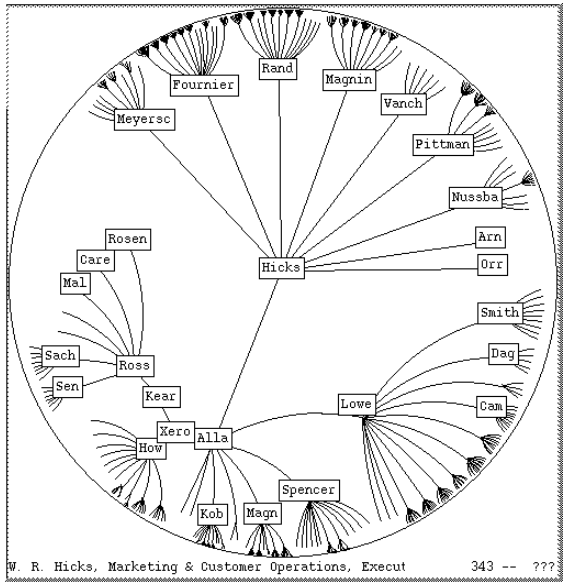
User orientation on refocus

- Problem
  - Hyperbolic Geometry can allow disorienting rotations of objects when refocusing
- Solution one:
  - Preserve initial angular orientation of parent to child nodes
- Solution two:
  - Preserve left to right orientation of parent to child nodes beginning with initial display

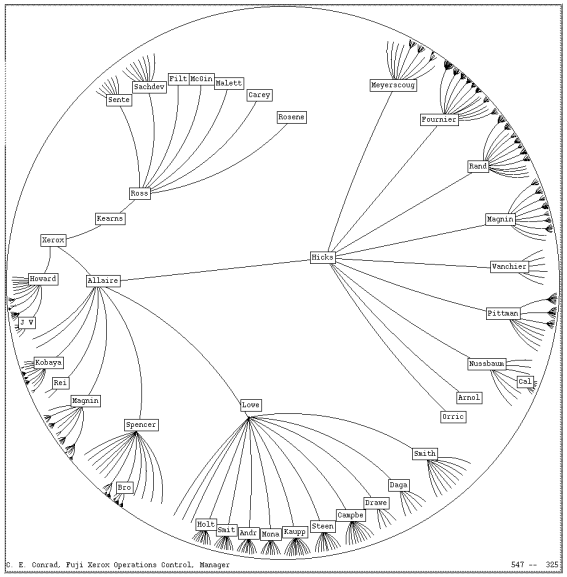
User orientations - Solutions

Preserving Angular Orientation





Left to Right Ordering



SCHOOL OF INTERACTIVE ARTS + TECHNOLOGY

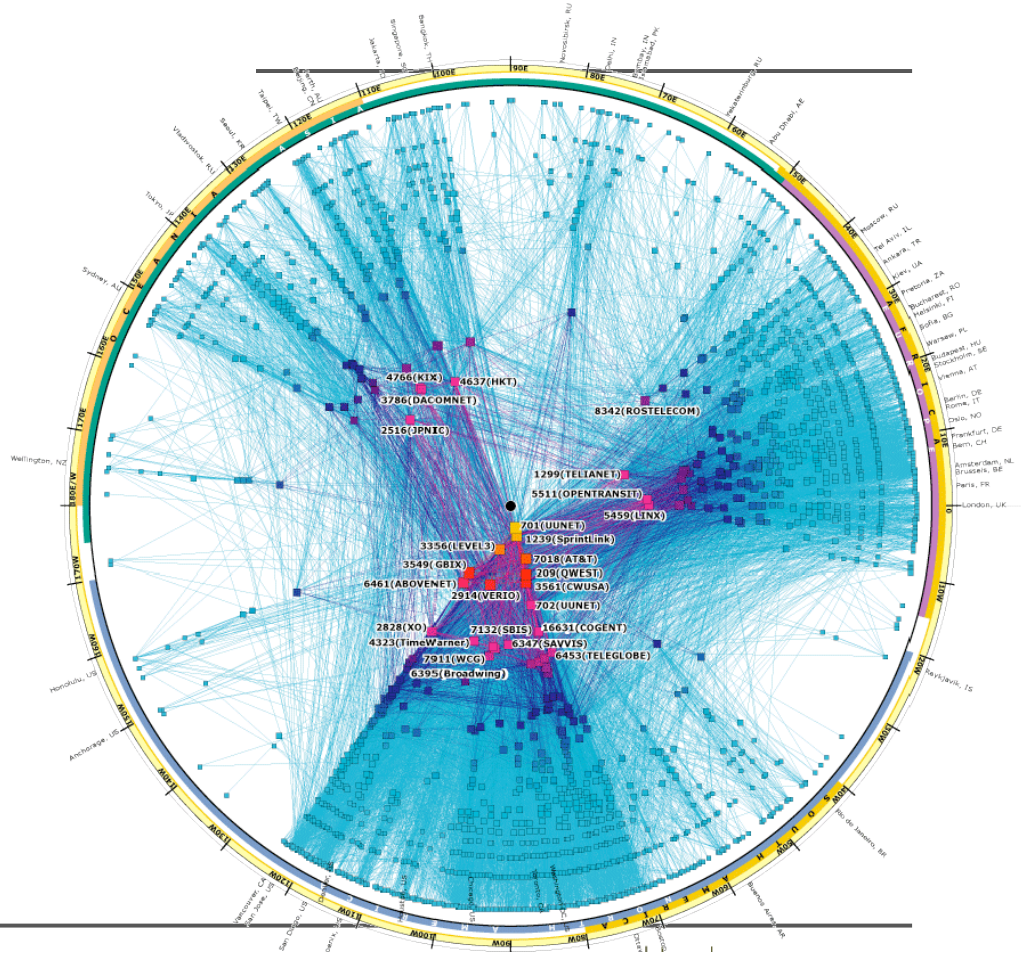
Structurally-Independent Layout

- Ignore the graph structure.
- Base the layout on other attributes of the data
- Examples:
  - Geography
  - Time
- Benefits
  - Often very quick layout
  - Optimizes communication of particular features
- Drawbacks
  - May or may not present structure well

Slide adapted from Jeff Heer

Structurally Independent Layout

- The “Skitter” Layout
  - Internet Connectivity
- Angle = Longitude
  - geography
- Radius = Degree
  - # of connections



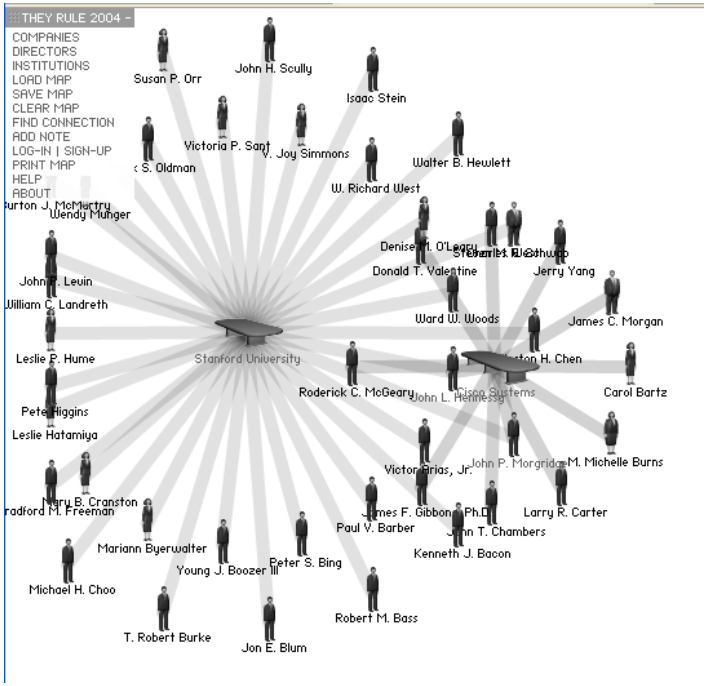
Skitter, www.caida.org

http://www.caida.org/research/topology/as_core_network/2007/images/ascore-simple.2007_big.png

Slide adapted from Jeff Heer

Progressive Disclosure

- Only show subsets that are currently selected
- http://www.theyrule.net/
- http://kylescholz.com/projects/wordnet/wordnet2.html



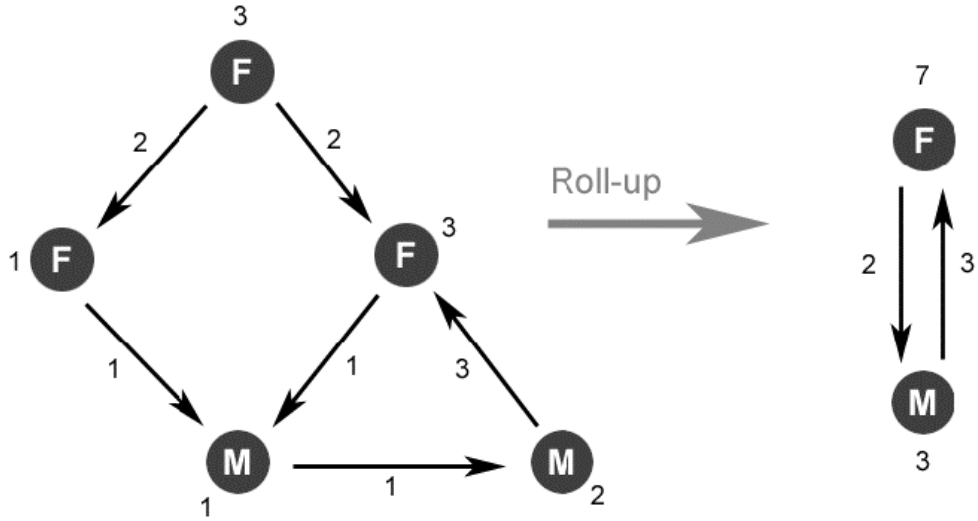
Problem: Multivariate Graphs

- What if you want to associate information with the nodes and edges?
- Typical approach: vary
  - Size of nodes
  - Color of nodes
  - Fatness of edges
  - Colors of edges
- However, it's hard to make quantitative comparisons when these retinal cues are spread throughout the graph.

Solution: Wattenberg's Pivot Graphs

- Use "roll-up" idea from OLAP to compress and re-express graph data.
  - Aggregate all nodes that have the same values on each of those dimensions, and aggregate edges accordingly.
- In graph below,

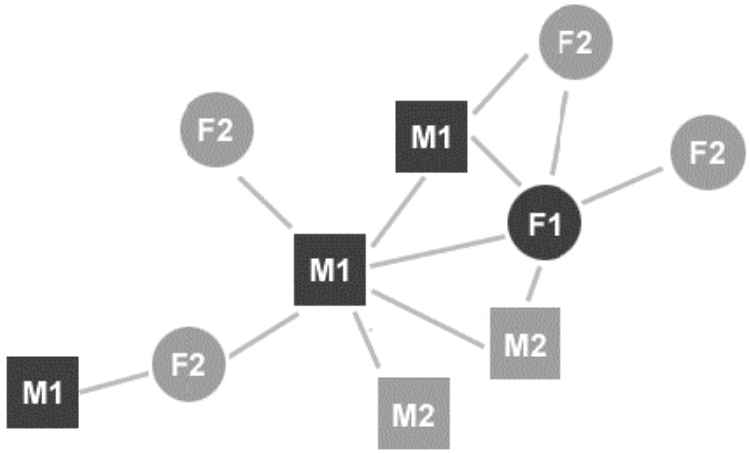
  F = Female, M = Male, Numbers mean counts



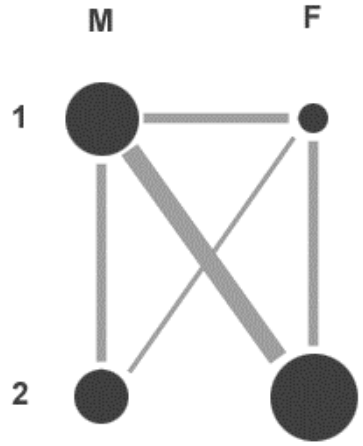
Visual Exploration of Multivariate Graphs, Wattenberg, IEEE Infoviz ???

Multidimensional Pivot Graphs

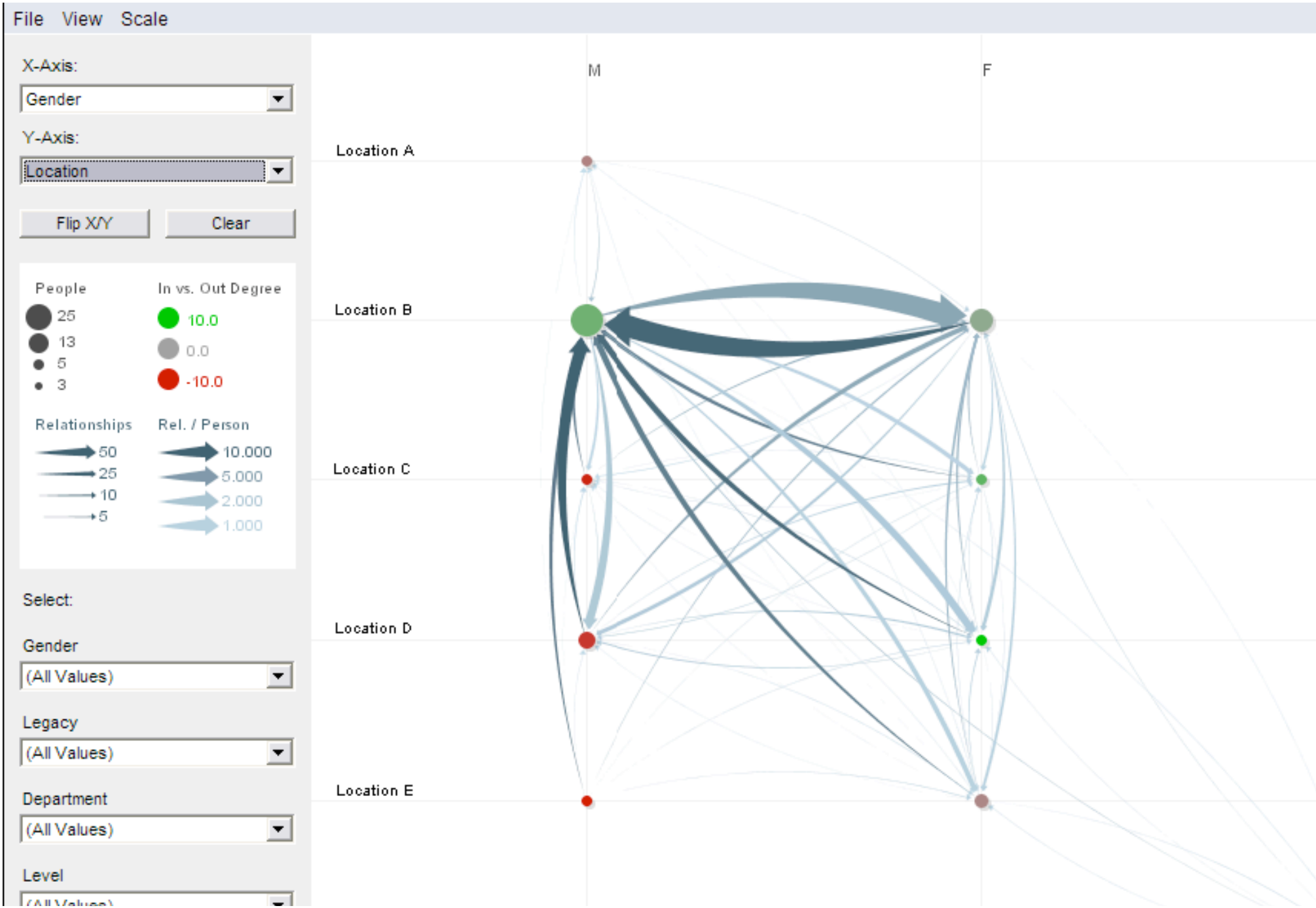
- What is added, and what is lost, from this transformation?



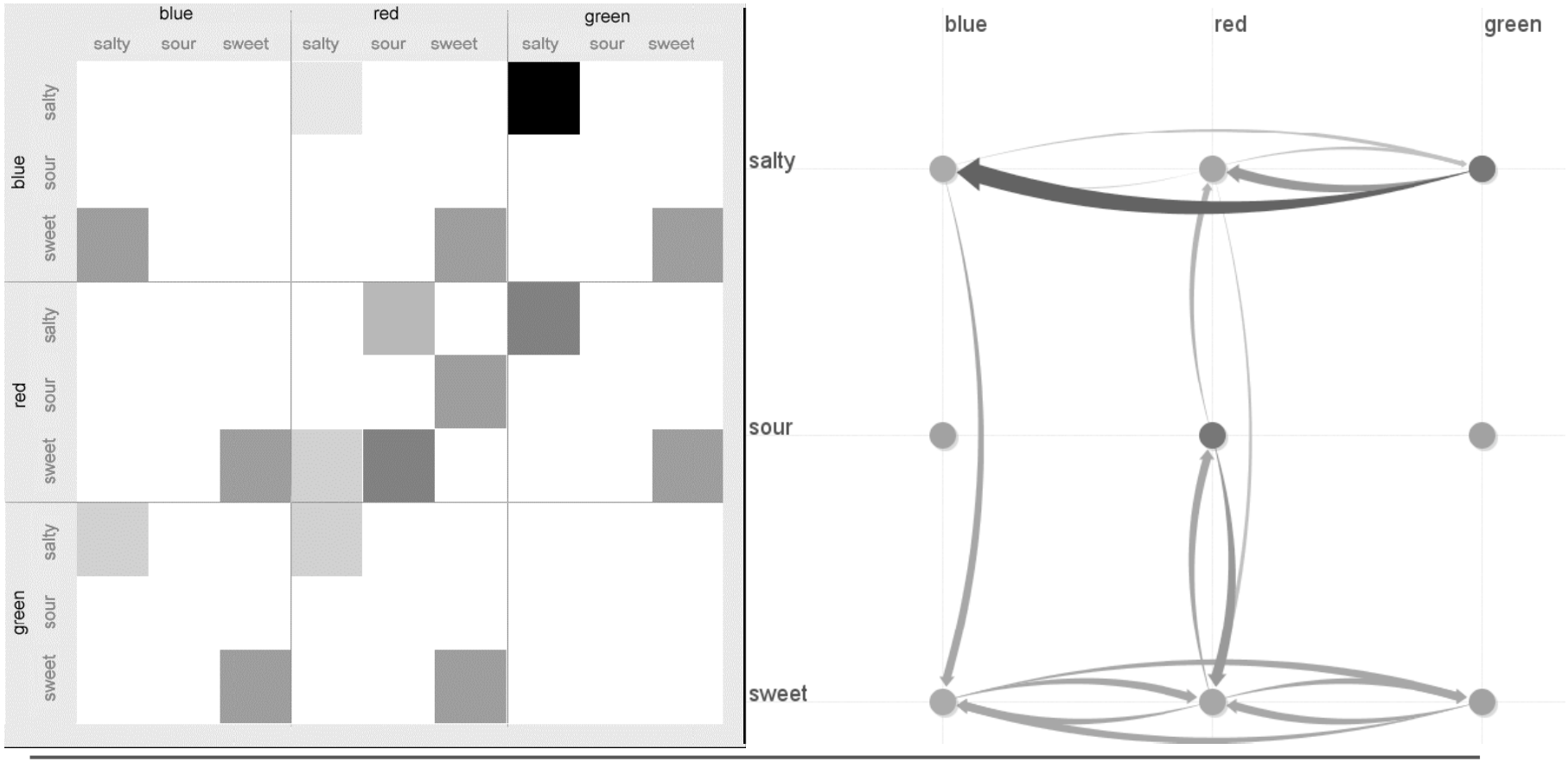
Node and Link Diagram



PivotGraph Roll-up


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
M
F
Location A
Location B
Location C
Location D
Location E

Compare 2D Pivot Graph with 2D Matrix



SCHOOL OF INTERACTIVE
ARTS + TECHNOLOGY

Issues with Pivot Graphs

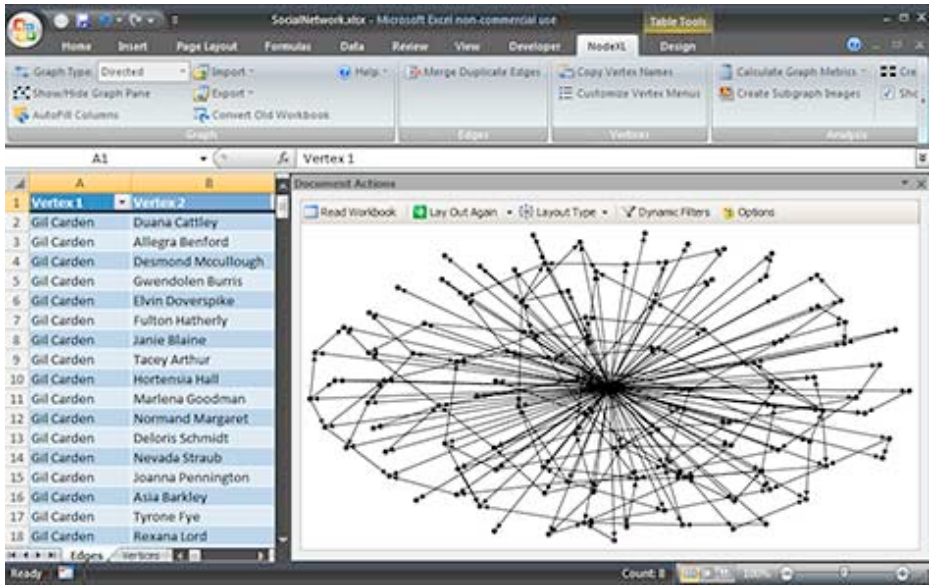
- Disconnected components may become connected
- Acyclic graphs may obtain cycles

New toolkits!

Networks for excel by Marc Smith et al. at Microsoft research

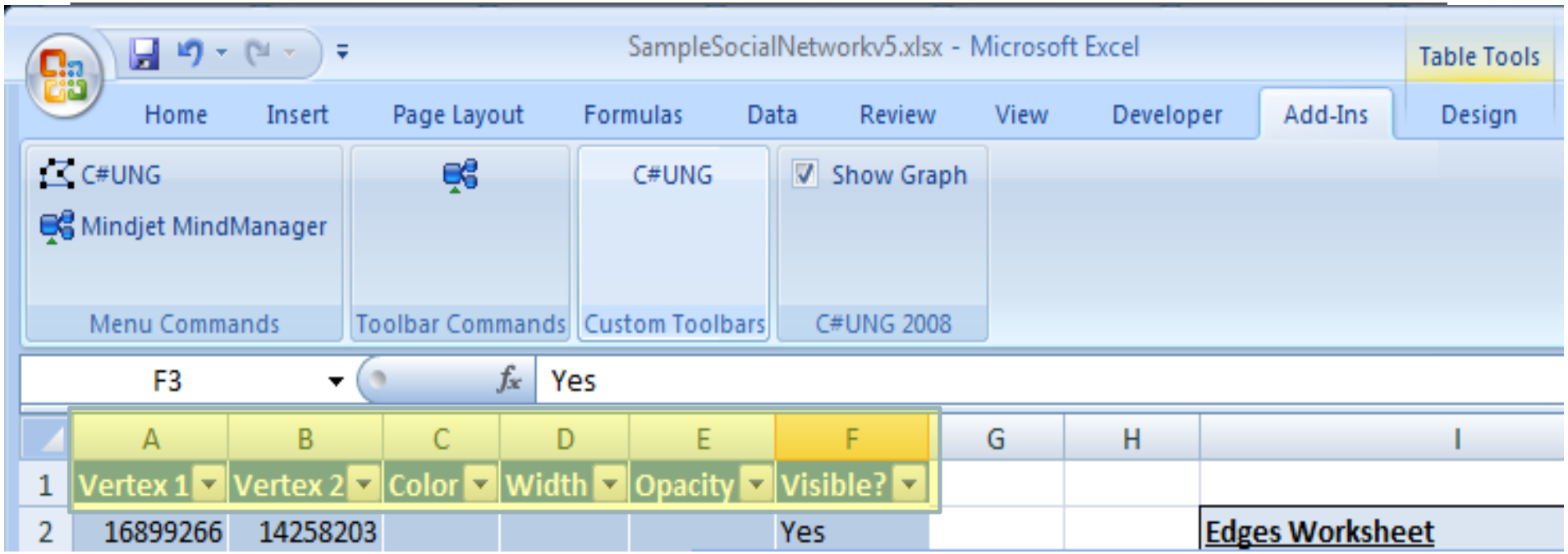
- Used to be called .Netmap
- Now called NodeXL
- Requires windows-specific software

- (Search on “excel NodeXL”)



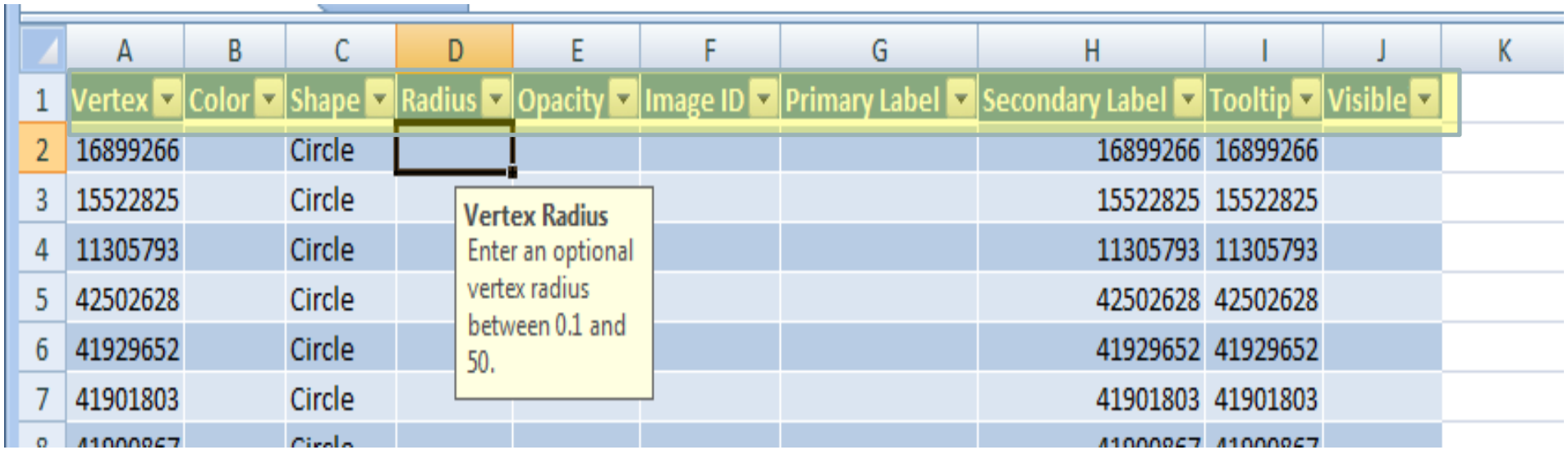
- Chart Tamer for Excel
- Stephen Few et al.

.NetMap: Edges Worksheet



Slide courtesy Marc Smith

.NetMap: Vertices Worksheet

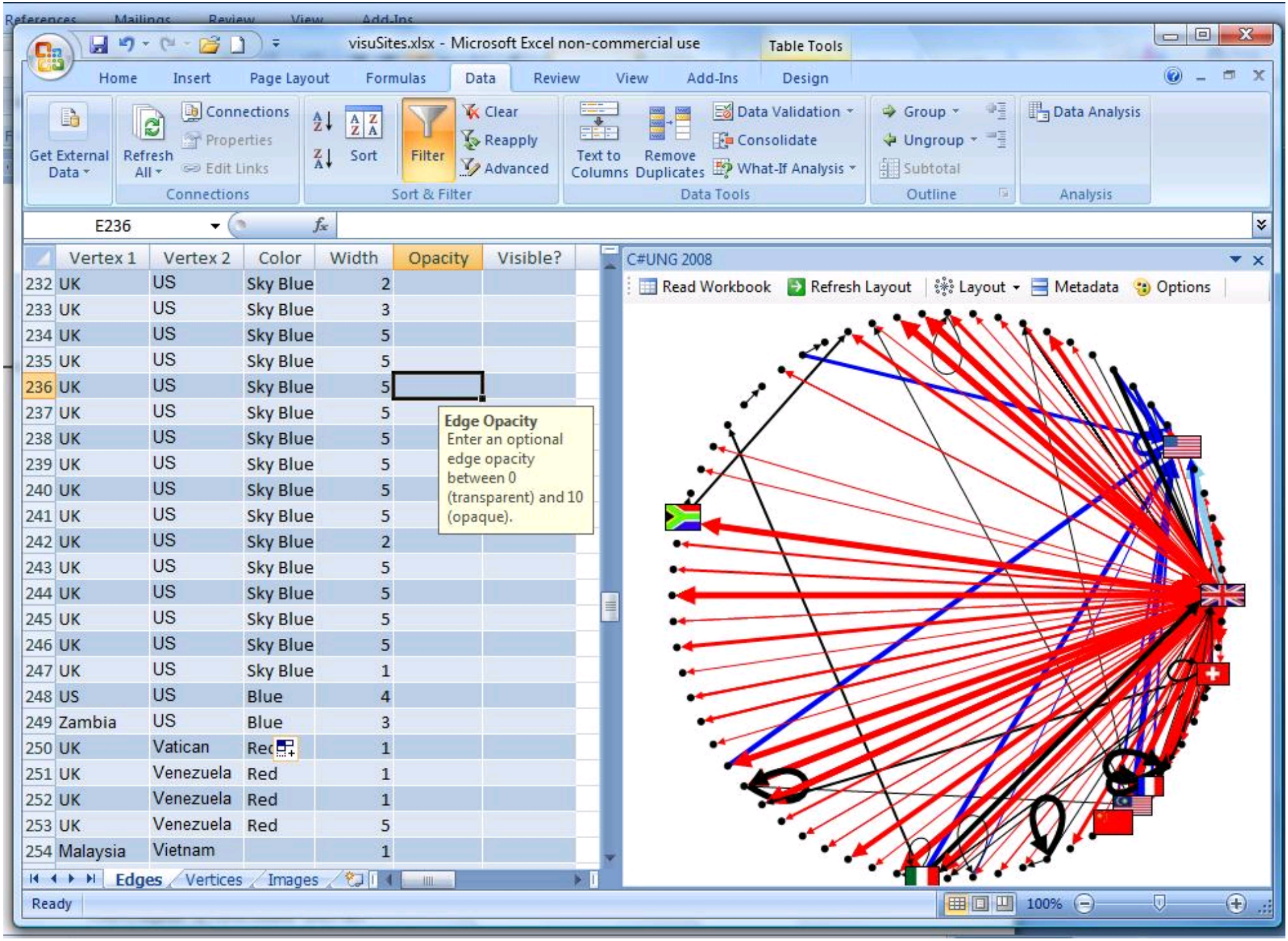


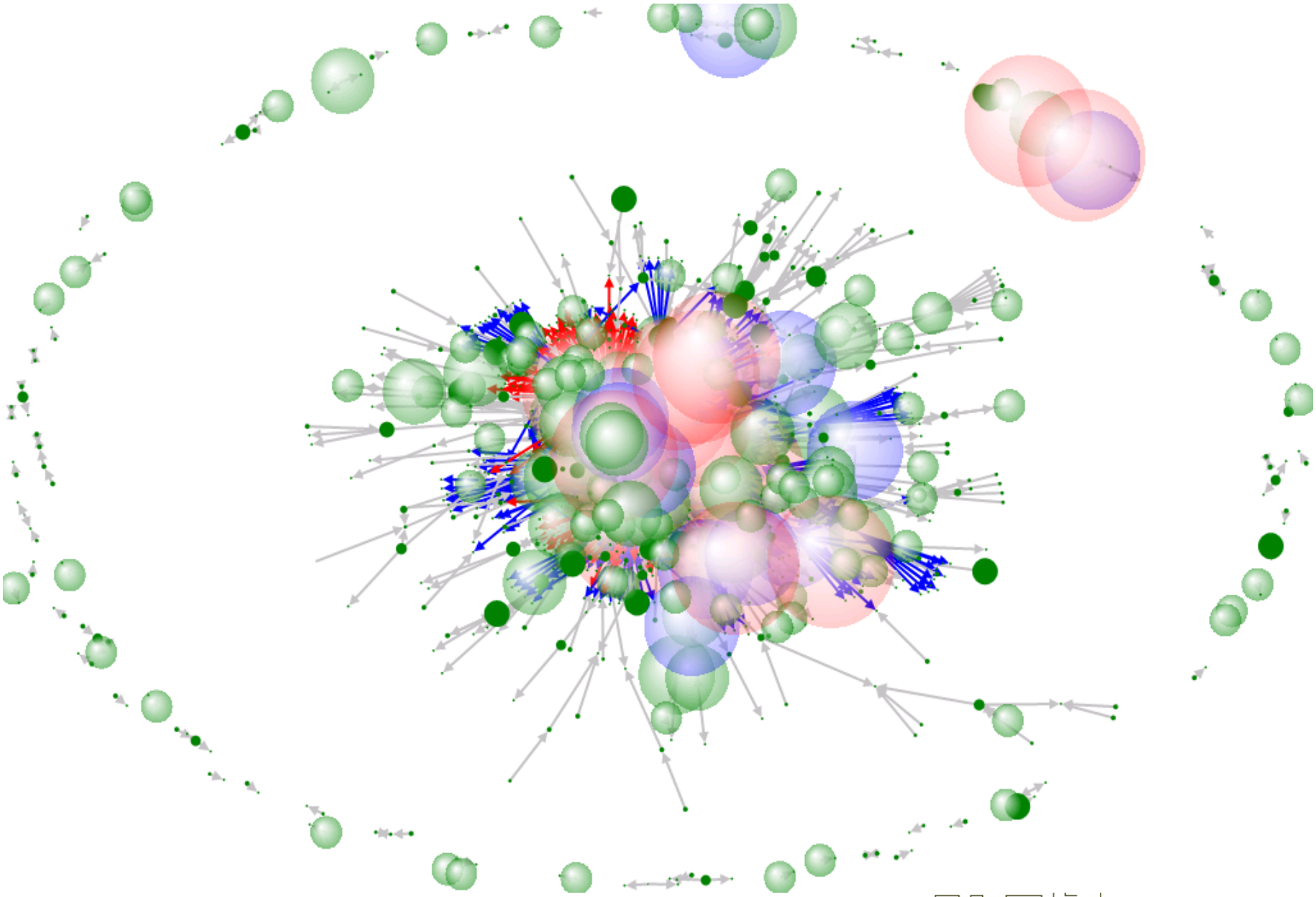
	A	B	C	D	E	F	G	H	I	J	K
1	Vertex	Color	Shape	Radius	Opacity	Image ID	Primary Label	Secondary Label	Tooltip	Visible	
2	16899266		Circle					16899266	16899266		
3	15522825		Circle					15522825	15522825		
4	11305793		Circle					11305793	11305793		
5	42502628		Circle					42502628	42502628		
6	41929652		Circle					41929652	41929652		
7	41901803		Circle					41901803	41901803		

Vertex Radius
Enter an optional vertex radius between 0.1 and 50.

Slide courtesy Marc Smith

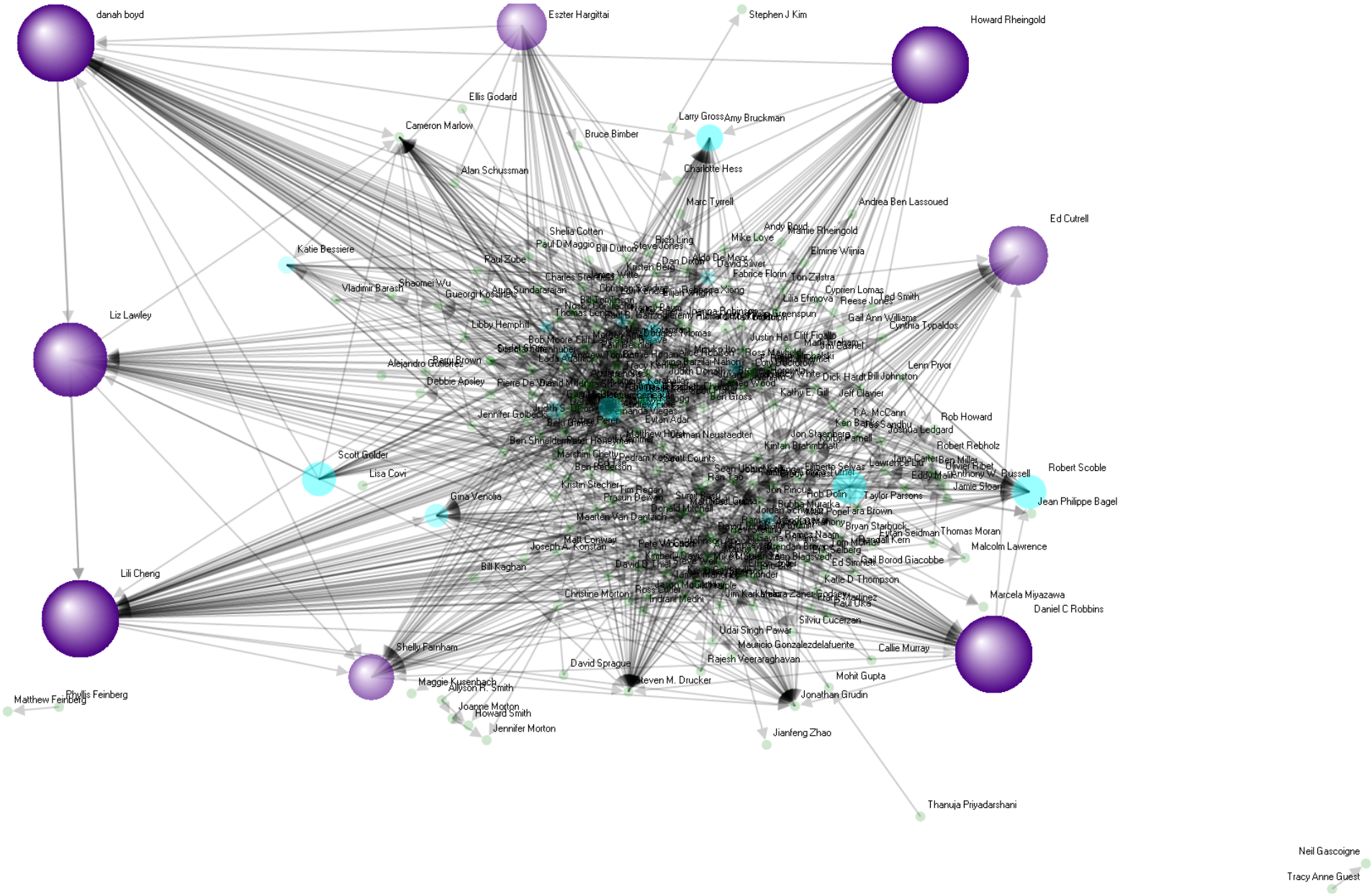
SCHOOL OF INTERACTIVE ARTS + TECHNOLOGY


visuSites.xlsx - Microsoft Excel non-commercial use
Table Tools
Home
Insert
Page Layout
Formulas
Data
Review
View
Add-Ins
Design
Get External Data
Refresh All
Connections
Properties
Edit Links
Connections
Sort
Filter
Clear
Reapply
Advanced
Sort & Filter
Text to Columns
Remove Duplicates
Data Validation
Consolidate
What-If Analysis
Data Tools
Group
Ungroup
Subtotal
Outline
Data Analysis
Analysis
E236
Vertex 1 Vertex 2 Color Width Opacity Visible?
232 UK US Sky Blue 2
233 UK US Sky Blue 3
234 UK US Sky Blue 5
235 UK US Sky Blue 5
236 UK US Sky Blue 5
237 UK US Sky Blue 5
238 UK US Sky Blue 5
239 UK US Sky Blue 5
240 UK US Sky Blue 5
241 UK US Sky Blue 5
242 UK US Sky Blue 2
243 UK US Sky Blue 5
244 UK US Sky Blue 5
245 UK US Sky Blue 5
246 UK US Sky Blue 5
247 UK US Sky Blue 1
248 US US Blue 4
249 Zambia US Blue 3
250 UK Vatican Red 1
251 UK Venezuela Red 1
252 UK Venezuela Red 1
253 UK Venezuela Red 5
254 Malaysia Vietnam 1
Edge Opacity
Enter an optional edge opacity between 0 (transparent) and 10 (opaque).
Edges
Vertices
Images
Ready
C#UNG 2008
Read Workbook
Refresh Layout
Layout
Metadata
Options
100%



Slide courtesy Marc Smith

Marc's Facebook Graph



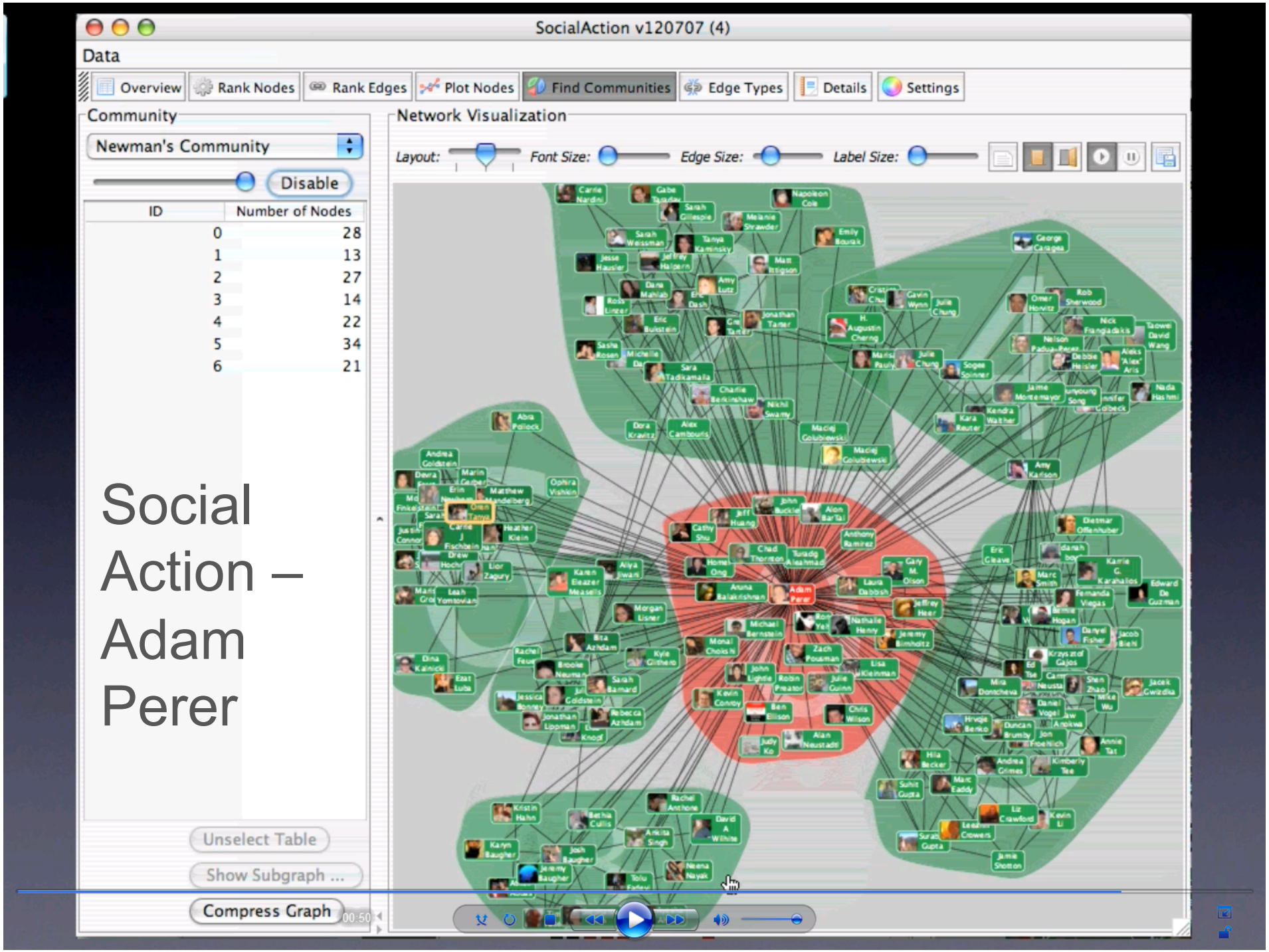
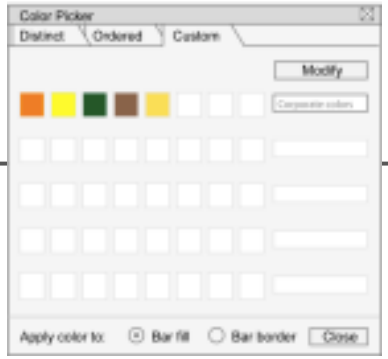
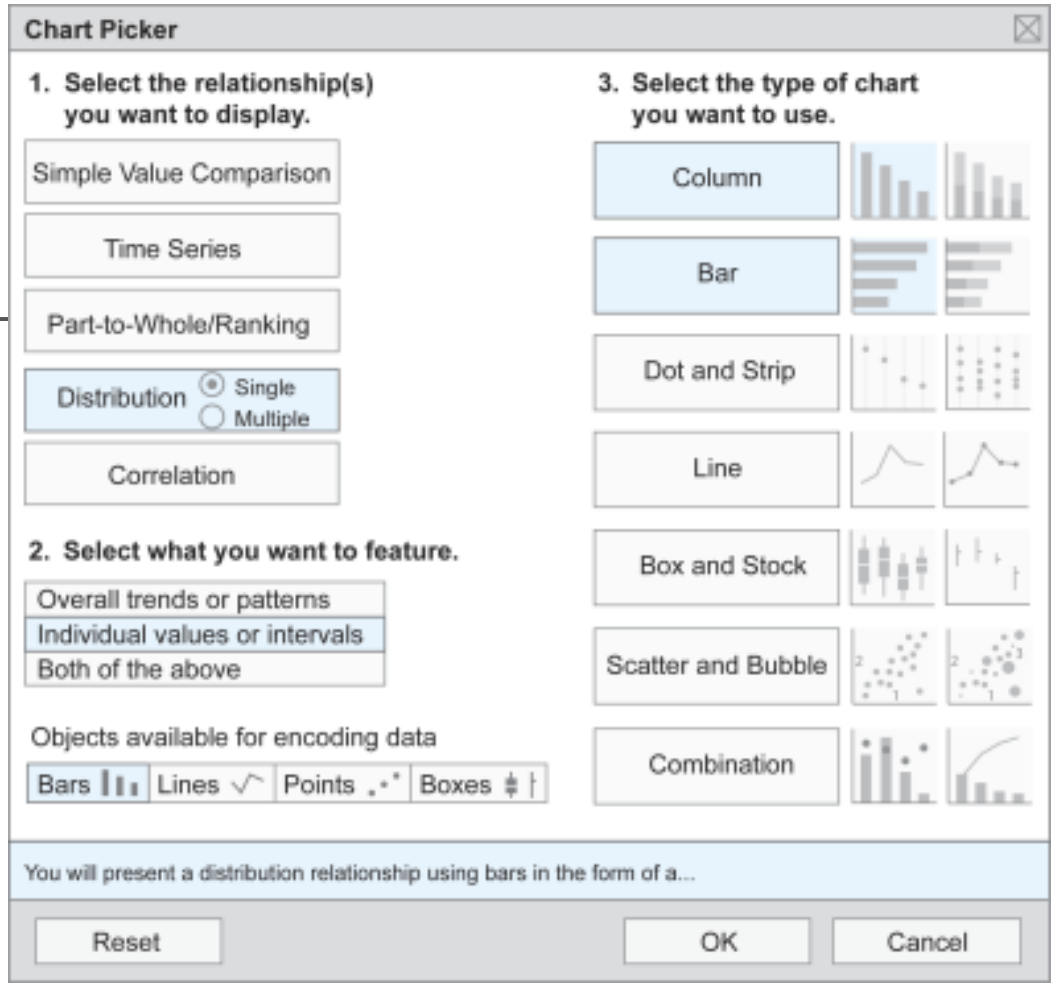

SocialAction v120707 (4)
Data
Overview
Rank Nodes
Rank Edges
Plot Nodes
Find Communities
Edge Types
Details
Settings
Community
Newman's Community
Disable
ID
Number of Nodes
0 28
1 13
2 27
3 14
4 22
5 34
6 21
Unselect Table
Show Subgraph ...
Compress Graph
Network Visualization
Layout:
Font Size:
Edge Size:
Label Size:
Social Action – Adam Perer

Chart Tamer





- Stephen Few + XL^3
- Function-based

SCHOOL OF INTERACTIVE ARTS + TECHNOLOGY