VISUALIZING TREES AND GRAPHS

Petra Isenberg

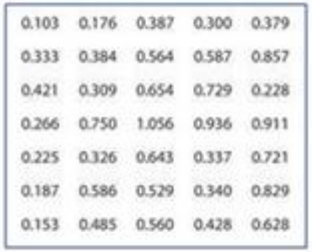
RECAP

you have learned about

- simple plots
- multi-attribute data visualization

DATA AND ITS STRUCTURE

STRUCTURED DATA



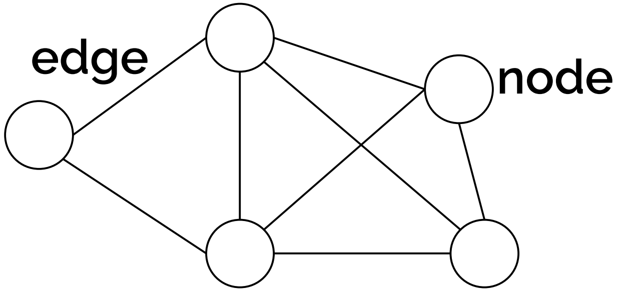
0.103	0.176	0.387	0.300	0.379
0.333	0.384	0.564	0.587	0.857
0.421	0.309	0.654	0.729	0.228
0.266	0.750	1.056	0.936	0.911
0.225	0.326	0.643	0.337	0.721
0.187	0.586	0.529	0.340	0.829
0.153	0.485	0.560	0.428	0.628

UNSTRUCTURED DATA

STRUCTURED DATA

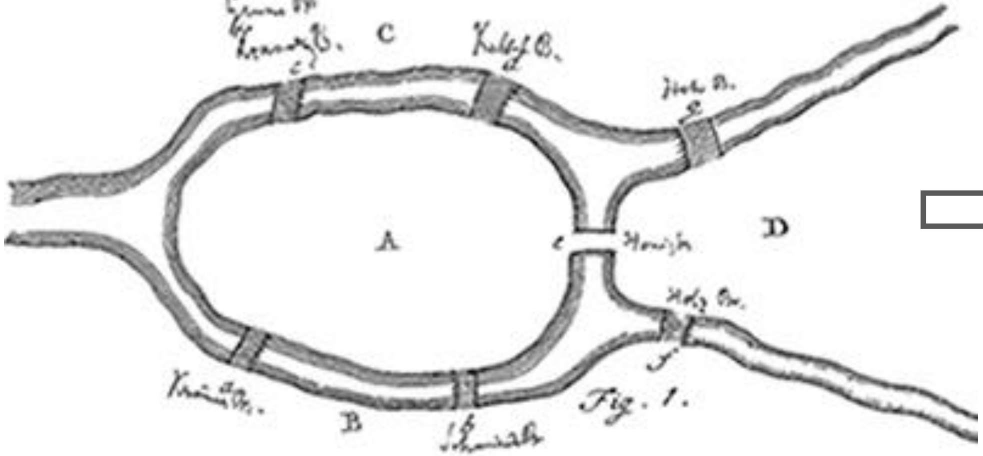
- there are relationships between the data items
- you can use a graph representation

0.103	0.176	0.387	0.300	0.379
0.333	0.384	0.564	0.587	0.857
0.421	0.309	0.654	0.729	0.228
0.266	0.750	1.056	0.936	0.911
0.225	0.326	0.643	0.337	0.721
0.187	0.586	0.529	0.340	0.829
0.153	0.485	0.560	0.428	0.628

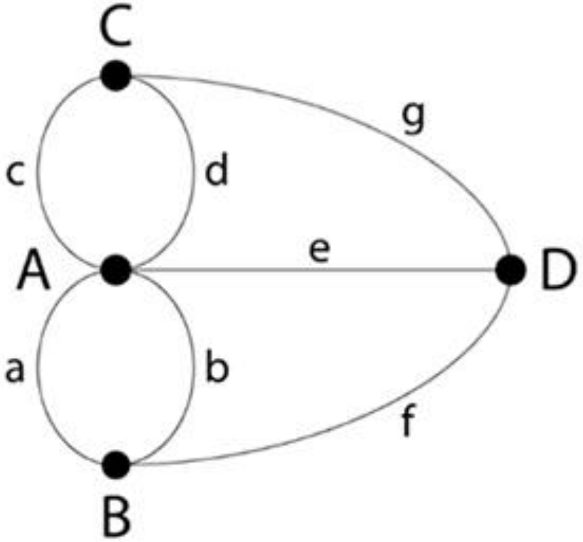


Almost anything can be a graph

The 7 bridges of Königsberg



How can you cross all 7 bridges without crossing the same one twice?



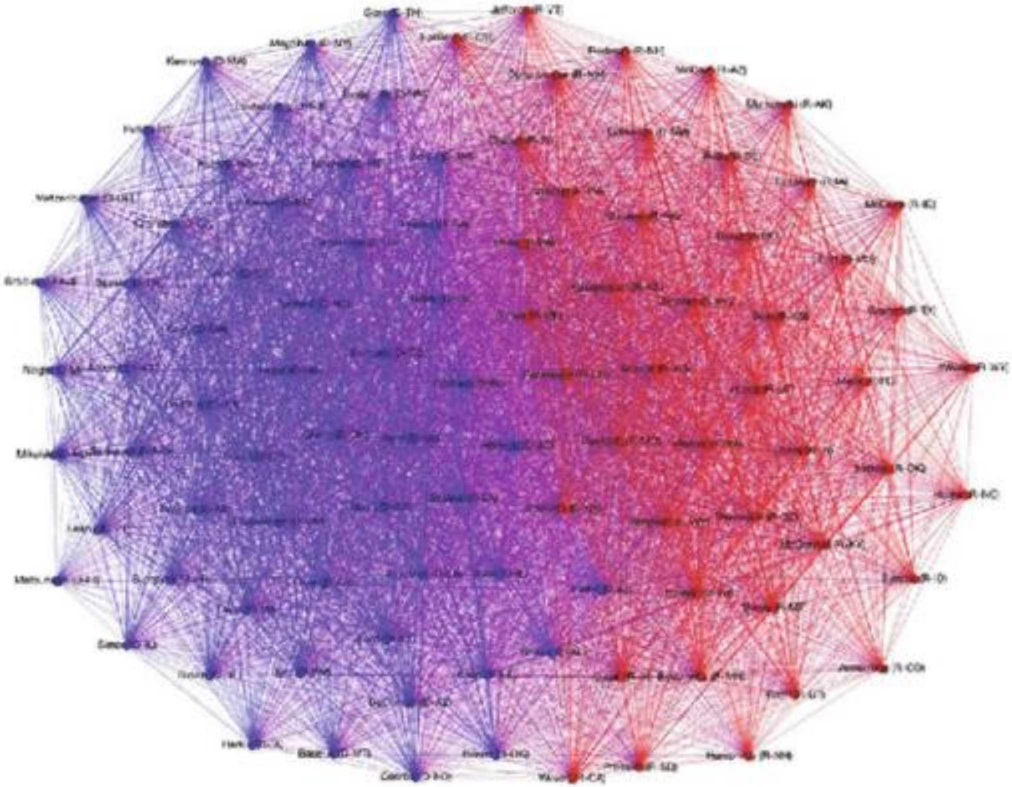
Euler’s conceptualization of the same problem

Senators casting the same votes

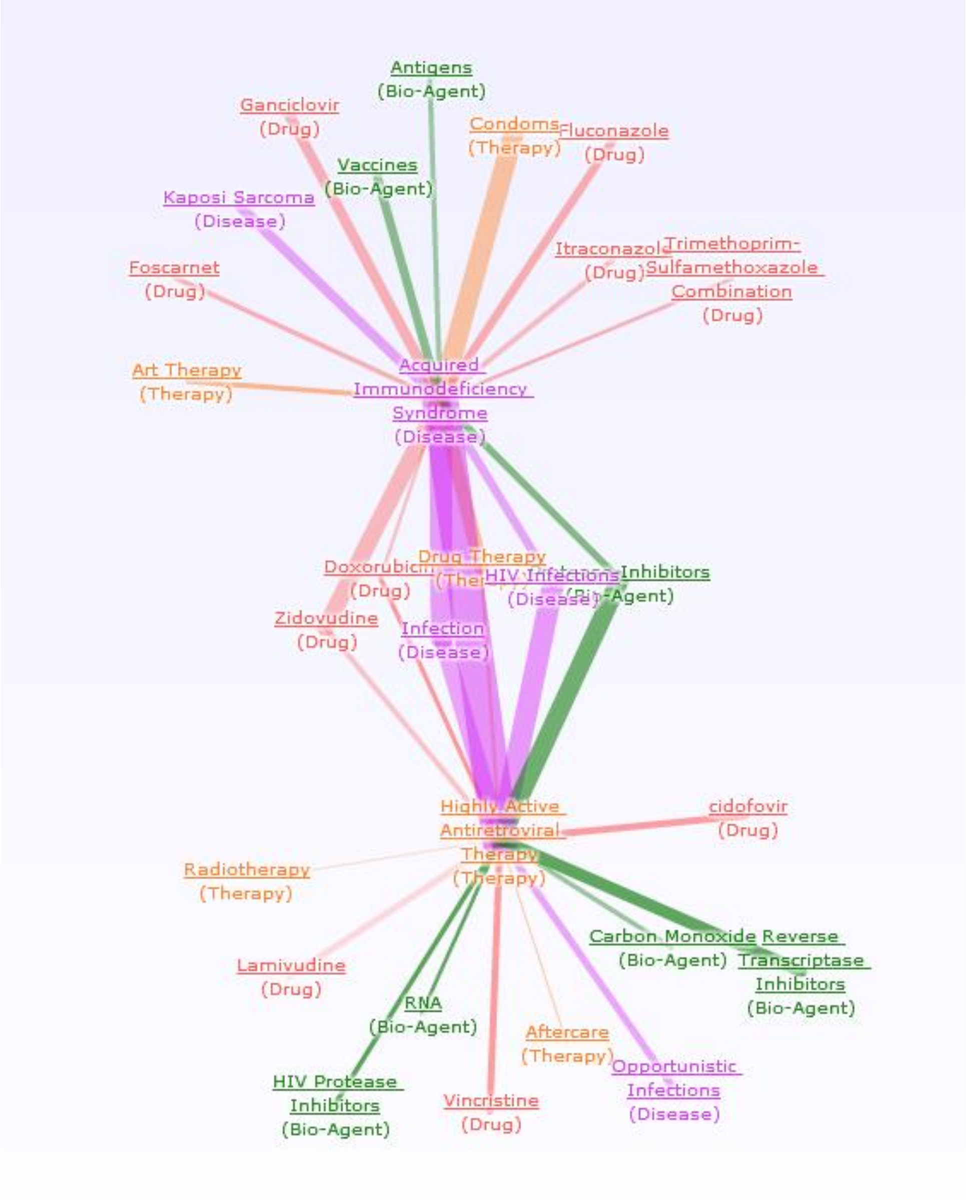
Democrat

Republican

101st Congress, 1989 session

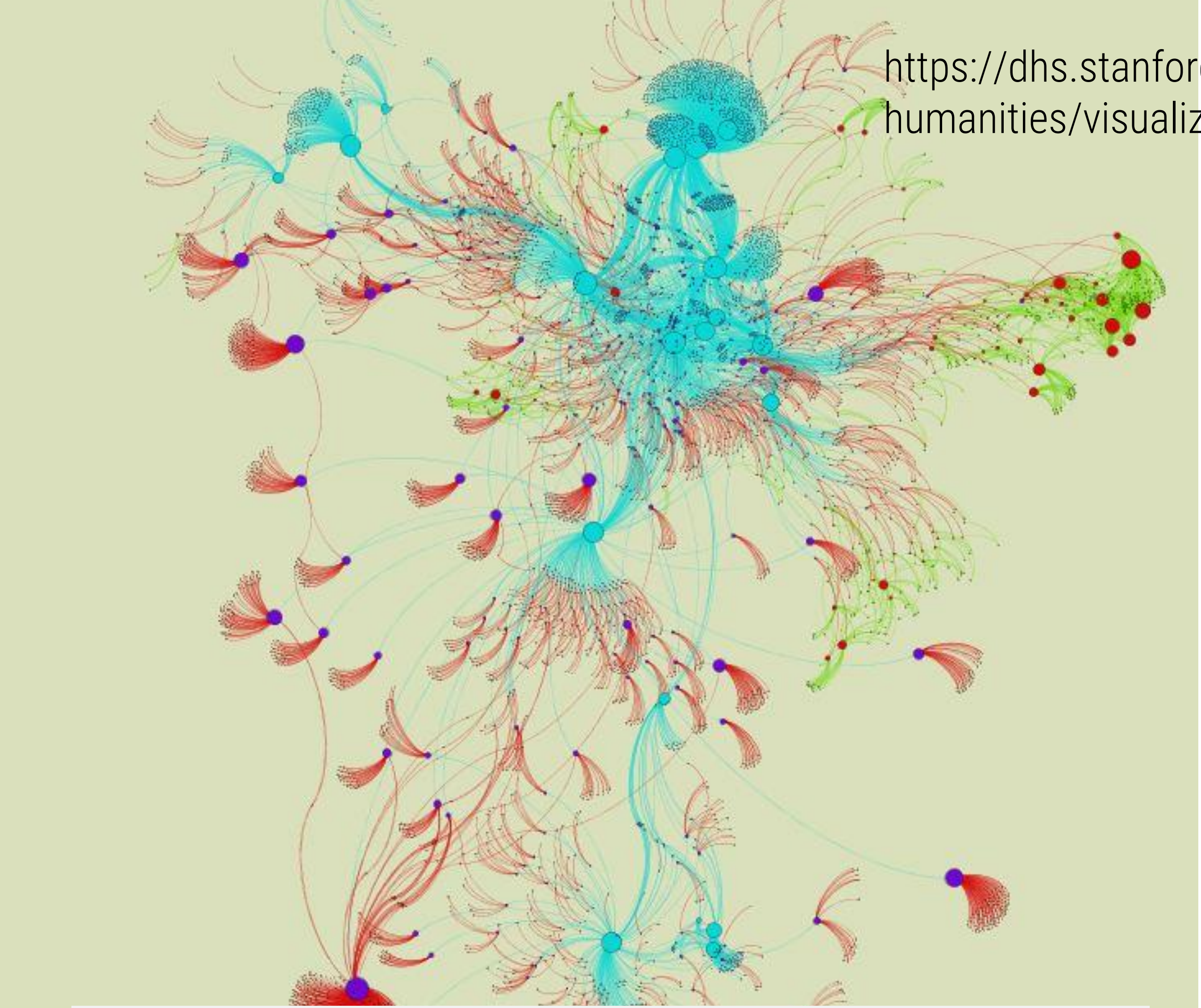


107th Congress, 2002 session

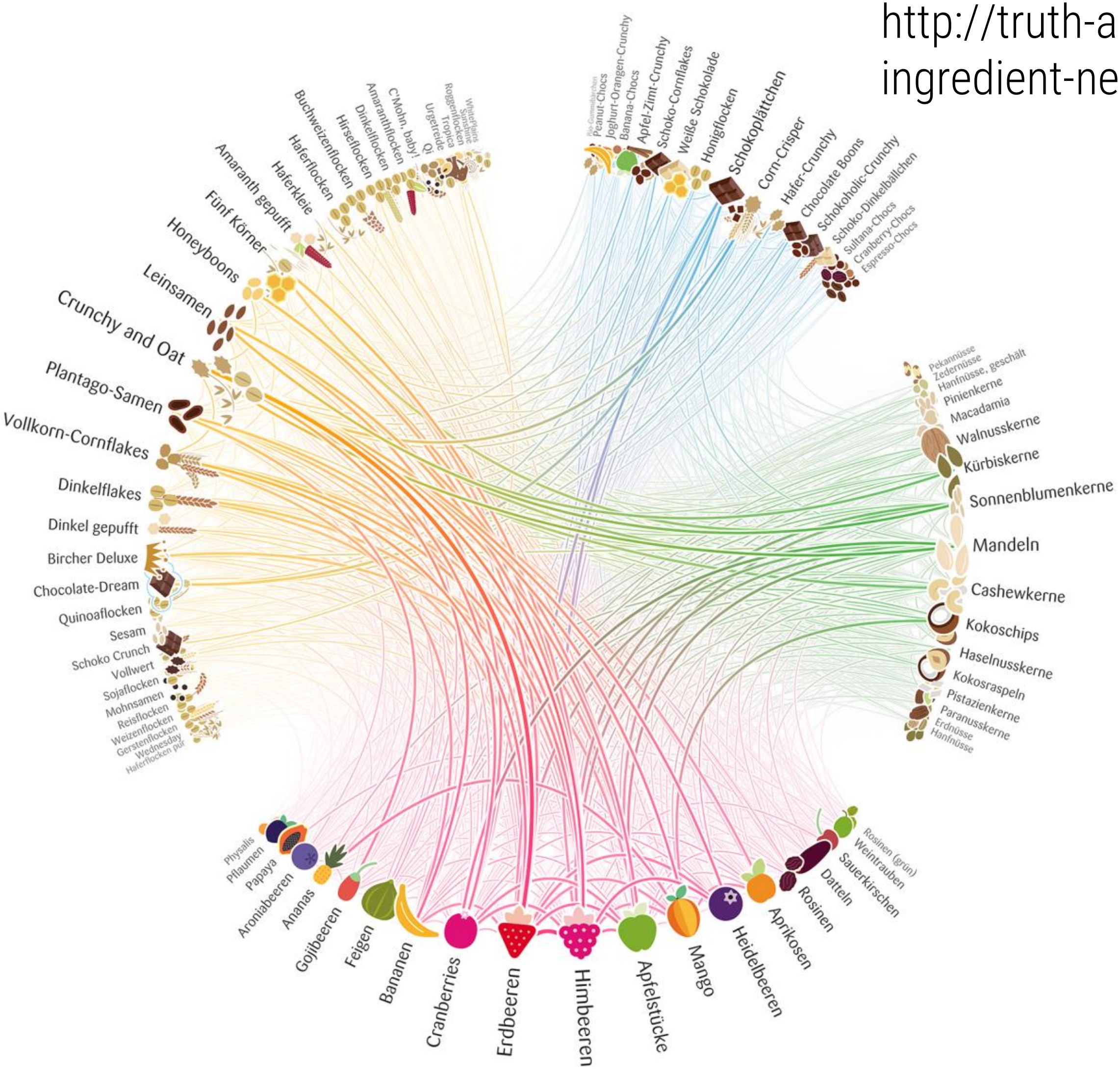


http://www.curehunter.com

visual dictionary of drugs, diseases and therapies https://dhs.stanford.edu/spatial-humanities/visualizing-databases/



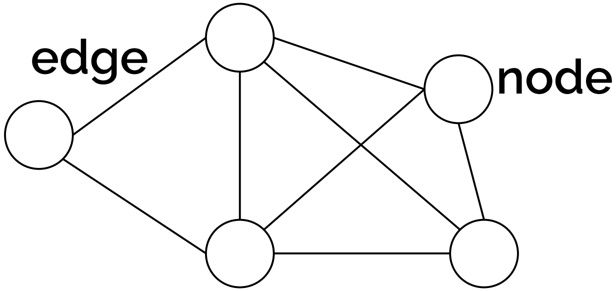
Top Contributors to the Catalogue of Life and their associated species, references and databases http://truth-and-beauty.net/projects/muesli-ingredient-network



GRAPHS

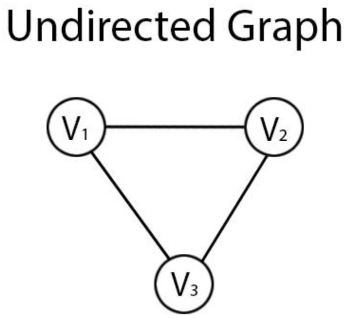
Graphs

- Describe relations among data items
- Using **nodes** and **edges**

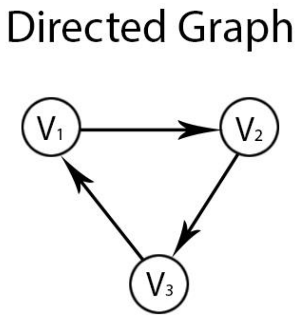


DEFINITIONS

undirected graph: edges have no orientation

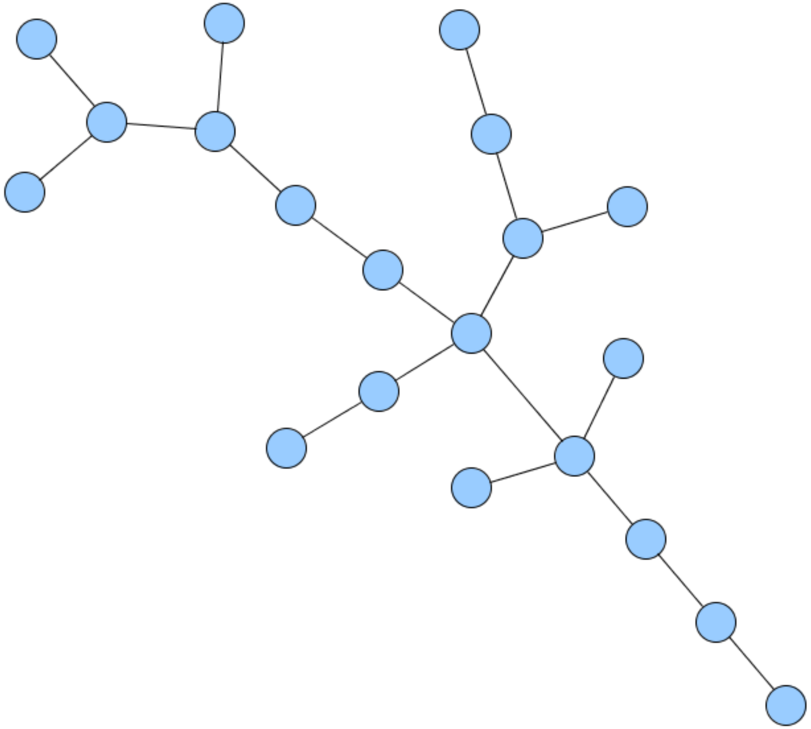


directed graph (digraph): edges have orientation



DEFINITIONS

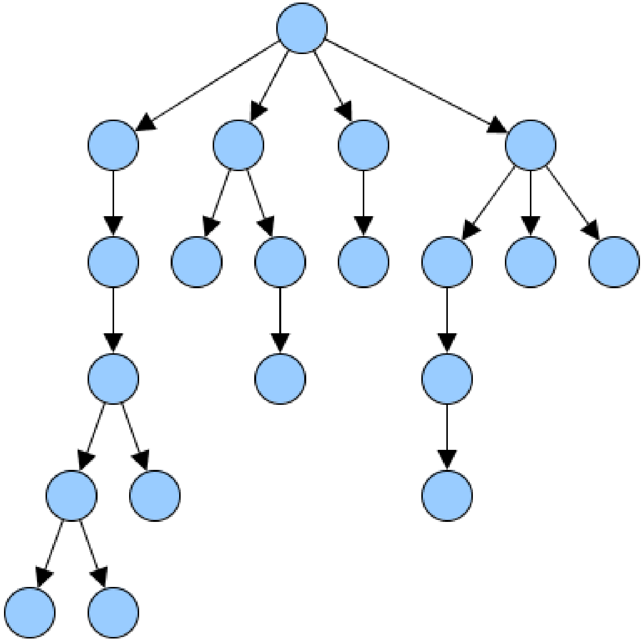
a tree is a **connected graph with no cycles**



DEFINITIONS

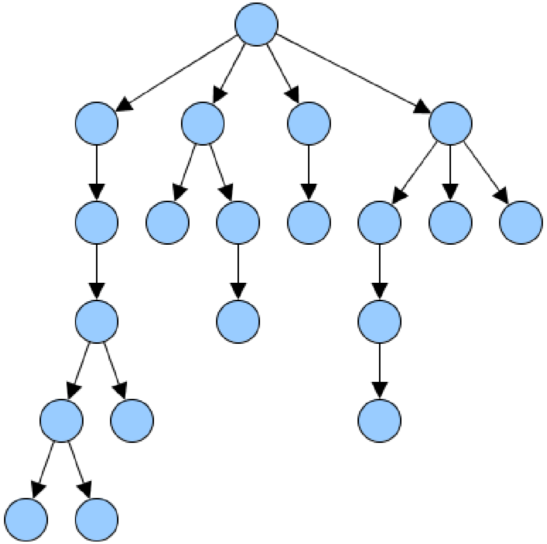
a directed tree is a digraph (directed graph) whose underlying graph is a tree

- a directed tree consists of a number of **nodes** and **parent-child relationships**
- every node has just one **parent** and any number of **children**
- *directed trees are the most common form in computer science*

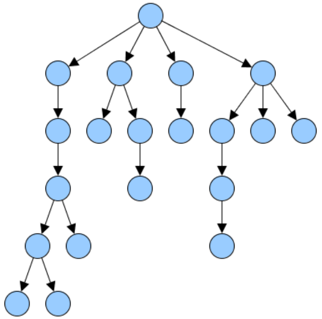


DEFINITIONS

- the number of children of a node is called its **degree**
- **leaf** nodes are nodes without children



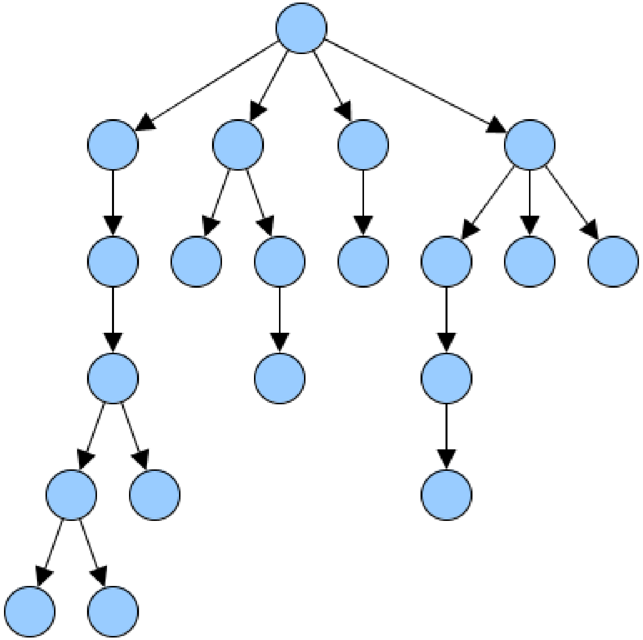
DEFINITIONS



- a **rooted tree** is a directed tree with a distinguished vertex r, called the **root**, such that for every other vertex v there is directed path from r to v
- the root node is the only node with no parent
- any node may act as a root in undirected trees

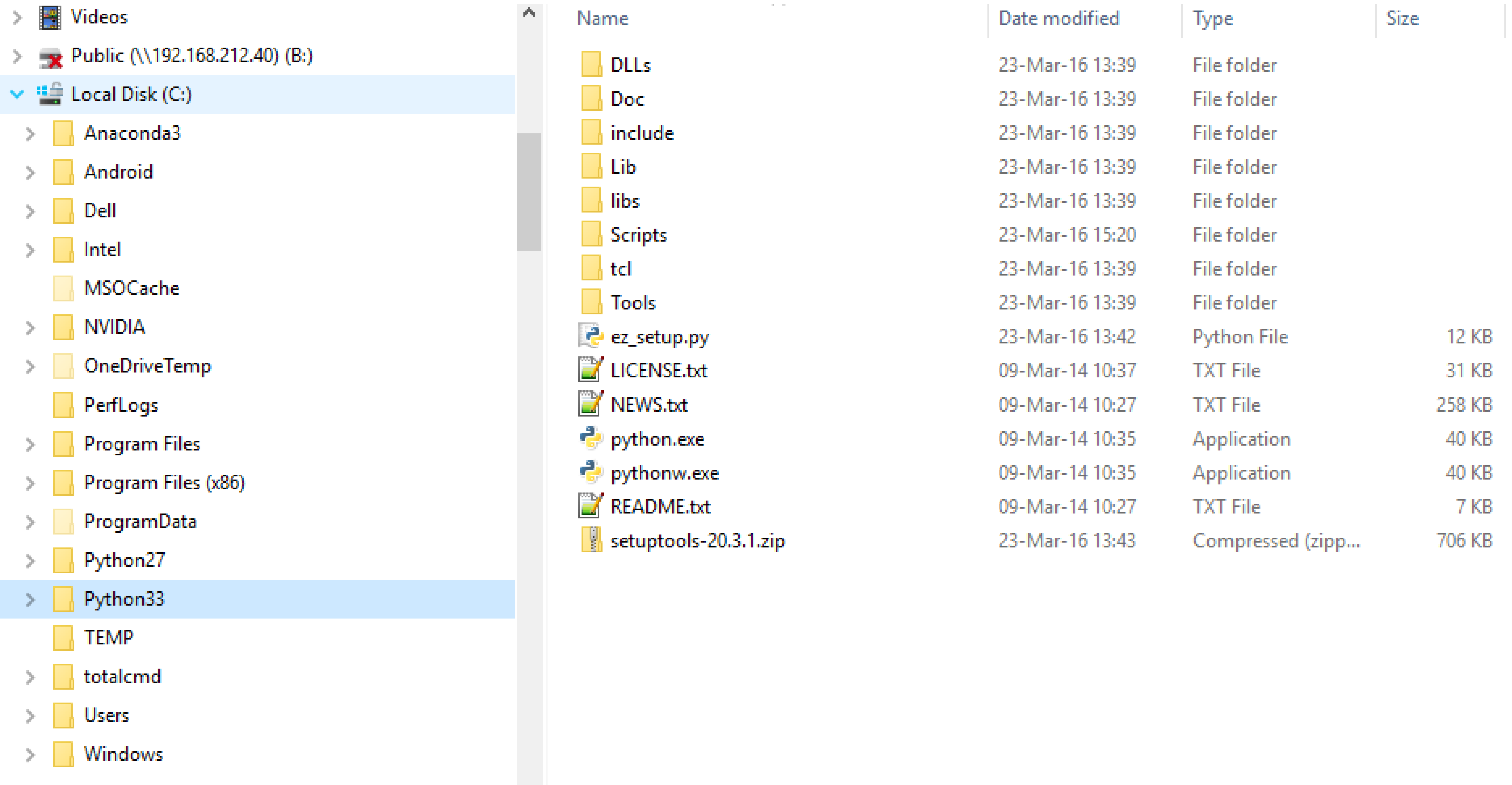
DEFINITIONS

the connection between parent and child nodes is called an **edge**



EXAMPLES OF TREES

HIERARCHIES

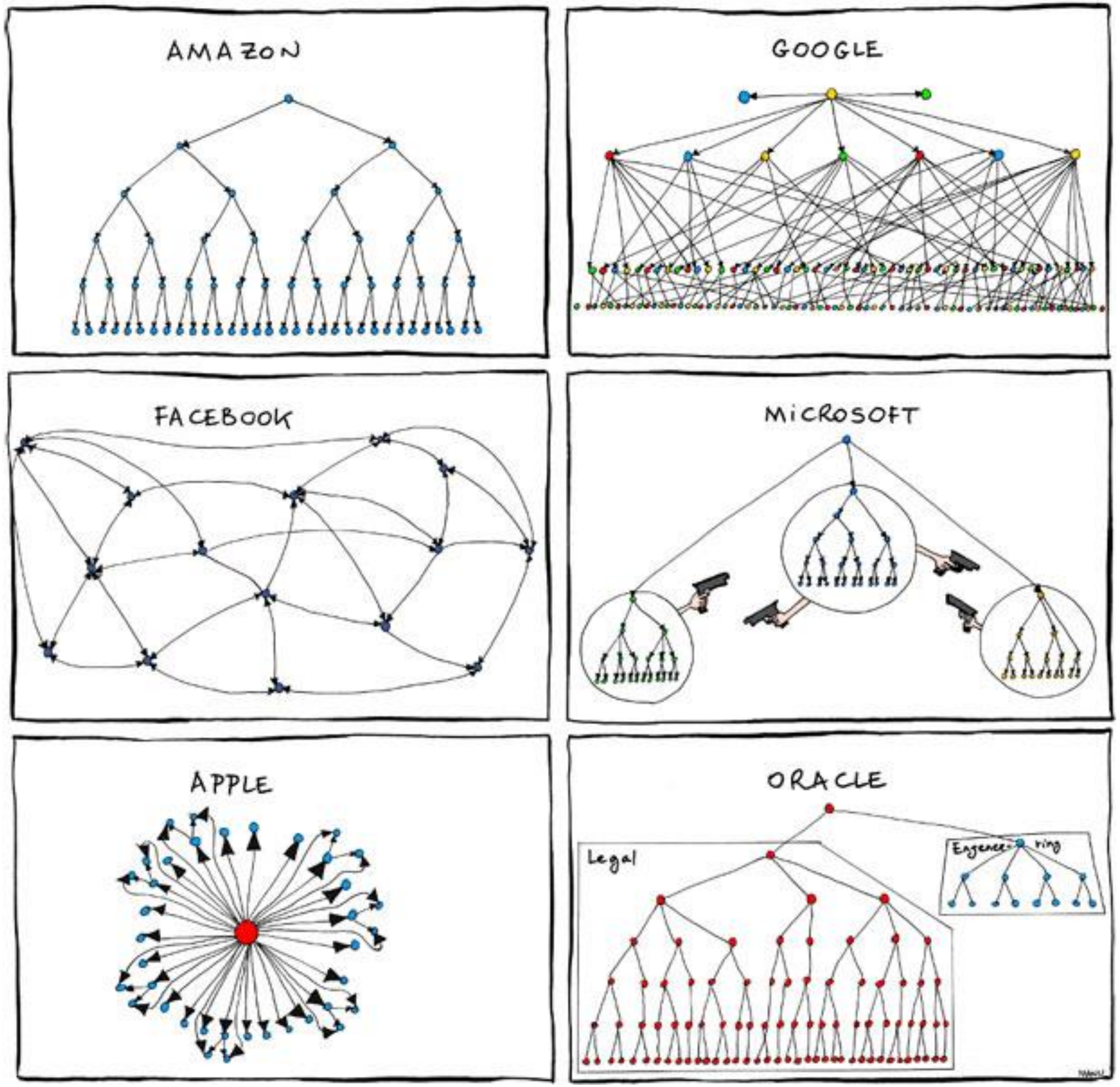


- Videos
- Public (\\192.168.212.40) (B:)
- Local Disk (C:)
  - Anaconda3
  - Android
  - Dell
  - Intel
  - MSOCache
  - NVIDIA
  - OneDriveTemp
  - PerfLogs
  - Program Files
  - Program Files (x86)
  - ProgramData
  - Python27
  - Python33
  - TEMP
  - totalcmd
  - Users
  - Windows

Name	Date modified	Type	Size
DLLs	23-Mar-16 13:39	File folder	
Doc	23-Mar-16 13:39	File folder	
include	23-Mar-16 13:39	File folder	
Lib	23-Mar-16 13:39	File folder	
libs	23-Mar-16 13:39	File folder	
Scripts	23-Mar-16 15:20	File folder	
tcl	23-Mar-16 13:39	File folder	
Tools	23-Mar-16 13:39	File folder	
ez_setup.py	23-Mar-16 13:42	Python File	12 KB
LICENSE.txt	09-Mar-14 10:37	TXT File	31 KB
NEWS.txt	09-Mar-14 10:27	TXT File	258 KB
python.exe	09-Mar-14 10:35	Application	40 KB
pythonw.exe	09-Mar-14 10:35	Application	40 KB
README.txt	09-Mar-14 10:27	TXT File	7 KB
setuptools-20.3.1.zip	23-Mar-16 13:43	Compressed (zipp...	706 KB

HIERARCHIES

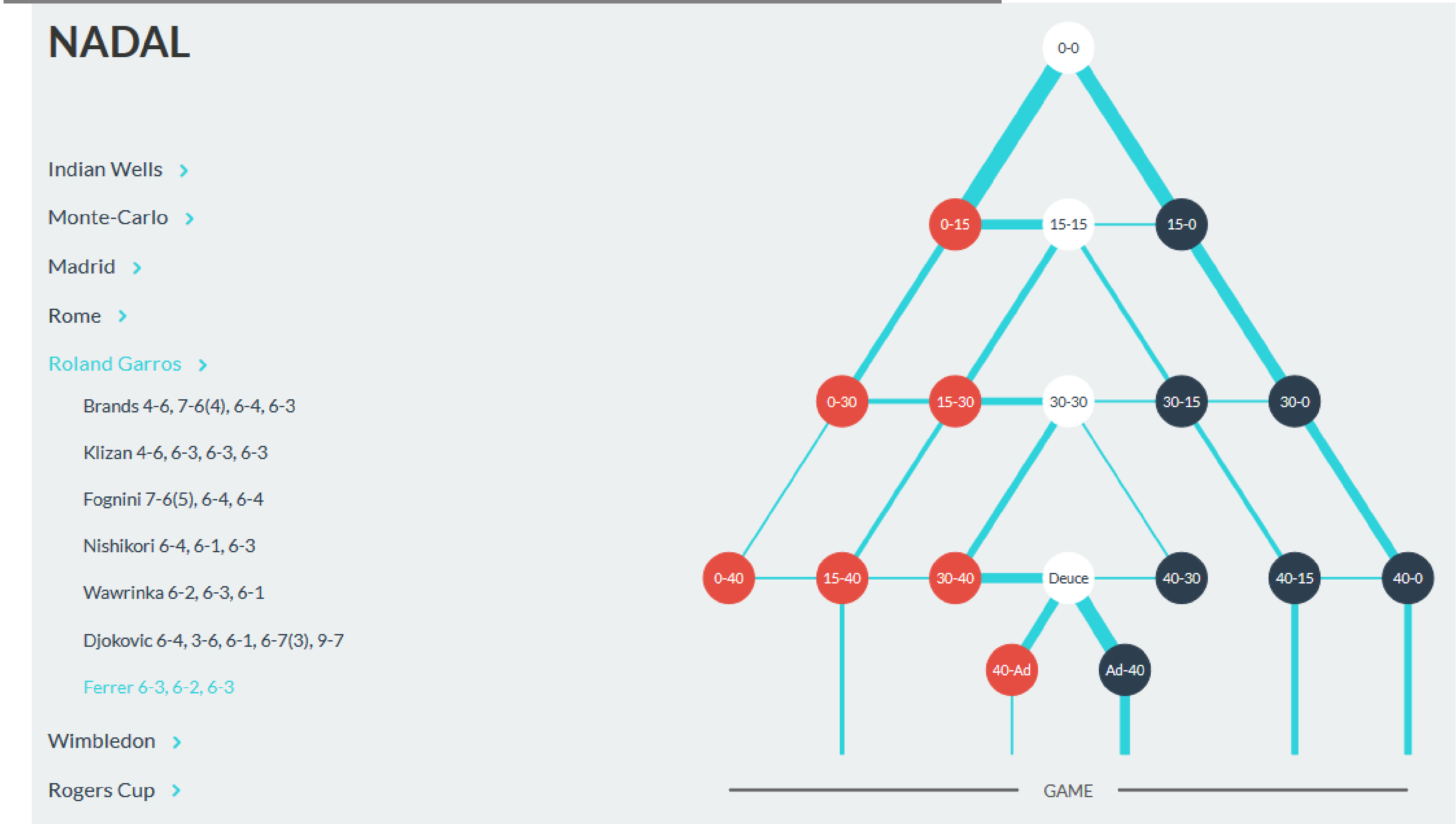
https://www.youtube.com/watch?v=mkJ-Uy5dt5g



org charts aren't always trees, though

http://www.bonkersworld.net/organizational-charts/

DECISION PROCESS



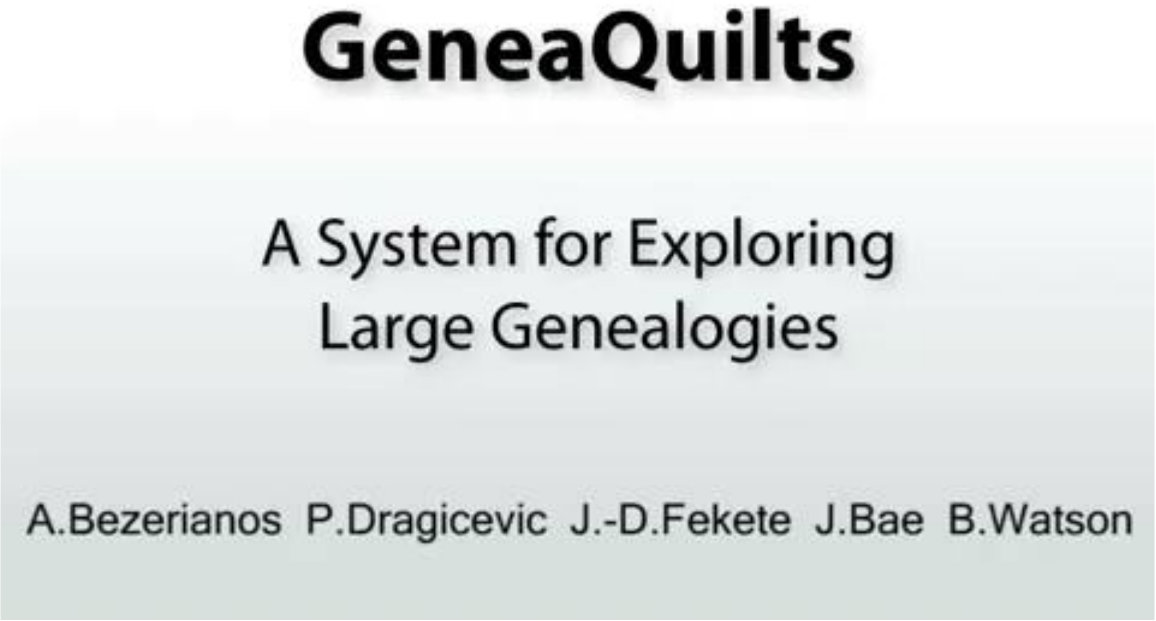
NADAL

- Indian Wells ›
- Monte-Carlo ›
- Madrid ›
- Rome ›
- Roland Garros ›
  - Brands 4-6, 7-6(4), 6-4, 6-3
  - Klizan 4-6, 6-3, 6-3, 6-3
  - Fognini 7-6(5), 6-4, 6-4
  - Nishikori 6-4, 6-1, 6-3
  - Wawrinka 6-2, 6-3, 6-1
  - Djokovic 6-4, 3-6, 6-1, 6-7(3), 9-7
  - Ferrer 6-3, 6-2, 6-3
- Wimbledon ›
- Rogers Cup ›

http://gamesetmap.com/nadalgametree/

BRANCHING PROCESSES

Think about it: is a family tree really a tree?



GeneaQuilts

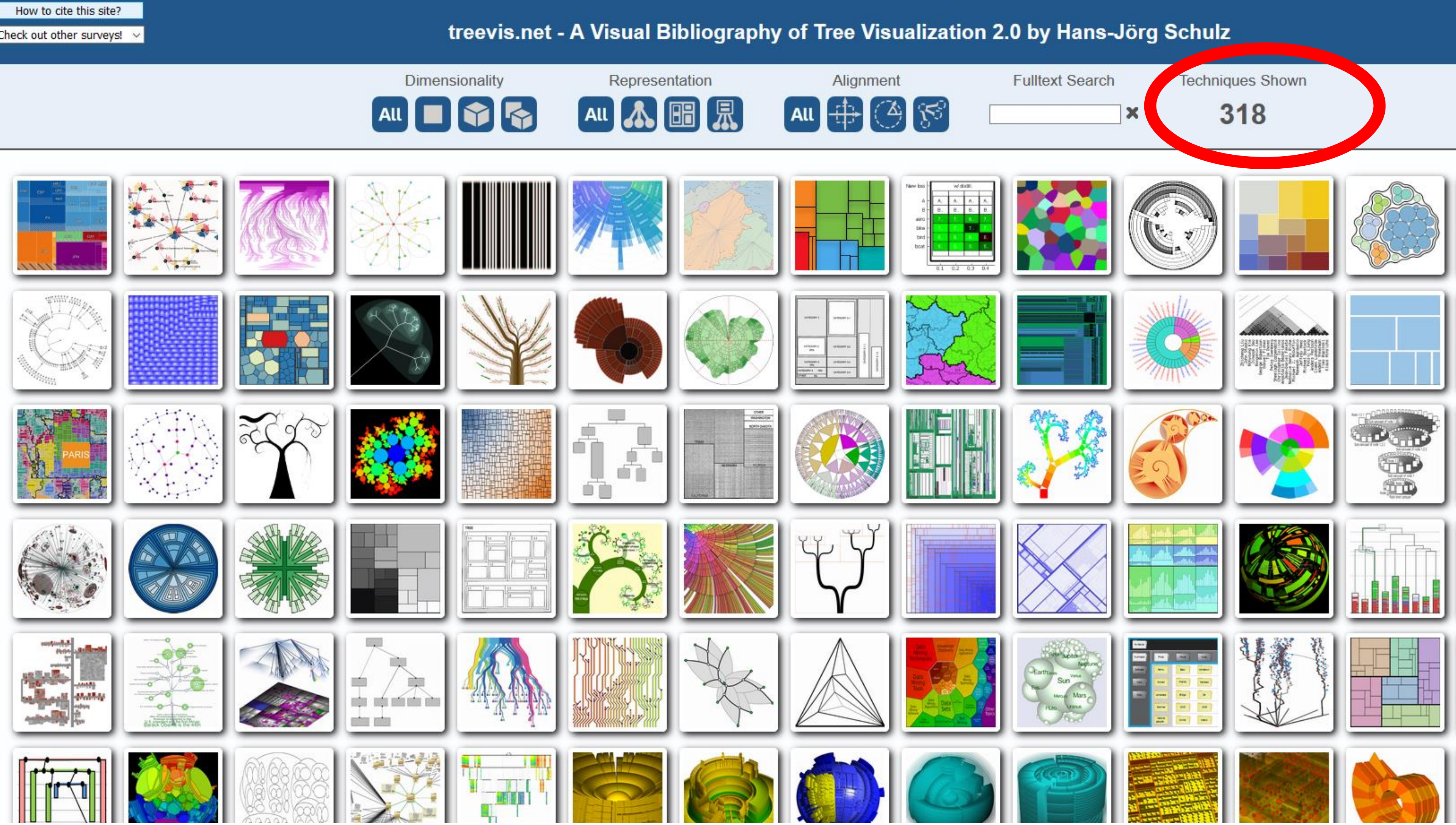
A System for Exploring Large Genealogies

A.Bezerianos P.Dragicevic J.-D.Fekete J.Bae B.Watson

https://www.youtube.com/watch?v=gncBzqI7R-Q

TREE REPRESENTATION

TECHNIQUES

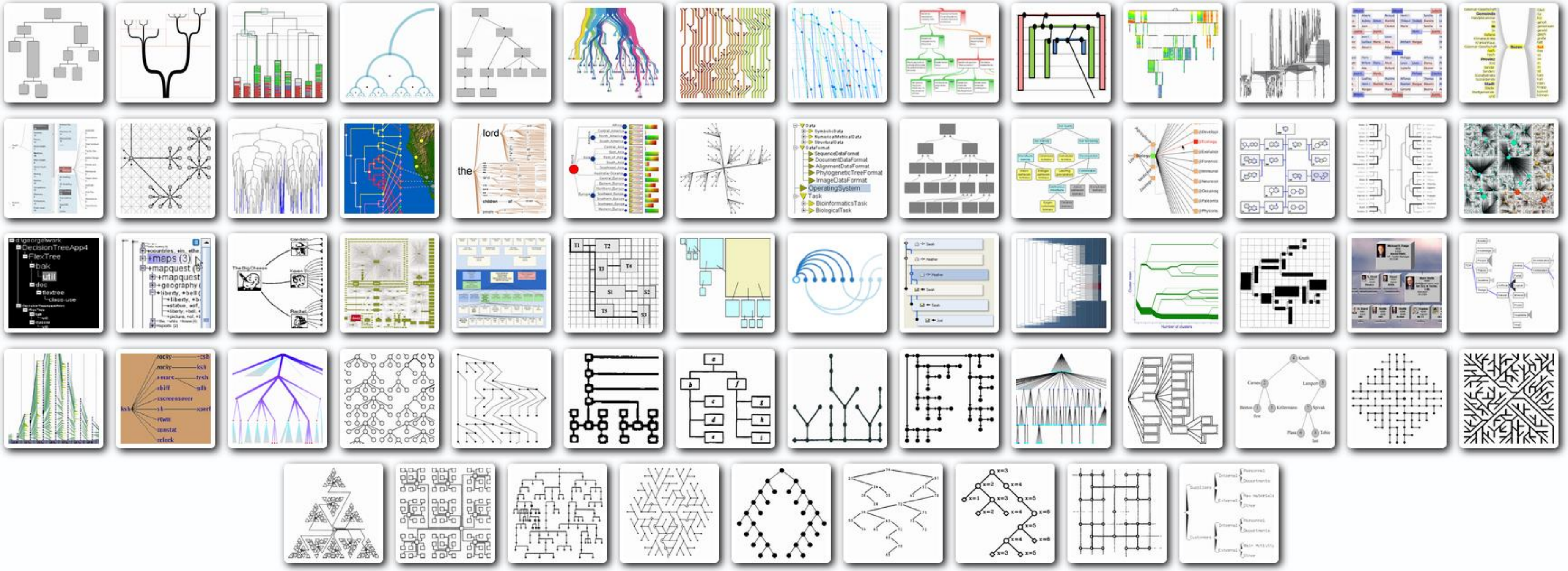


http://treevis.net/

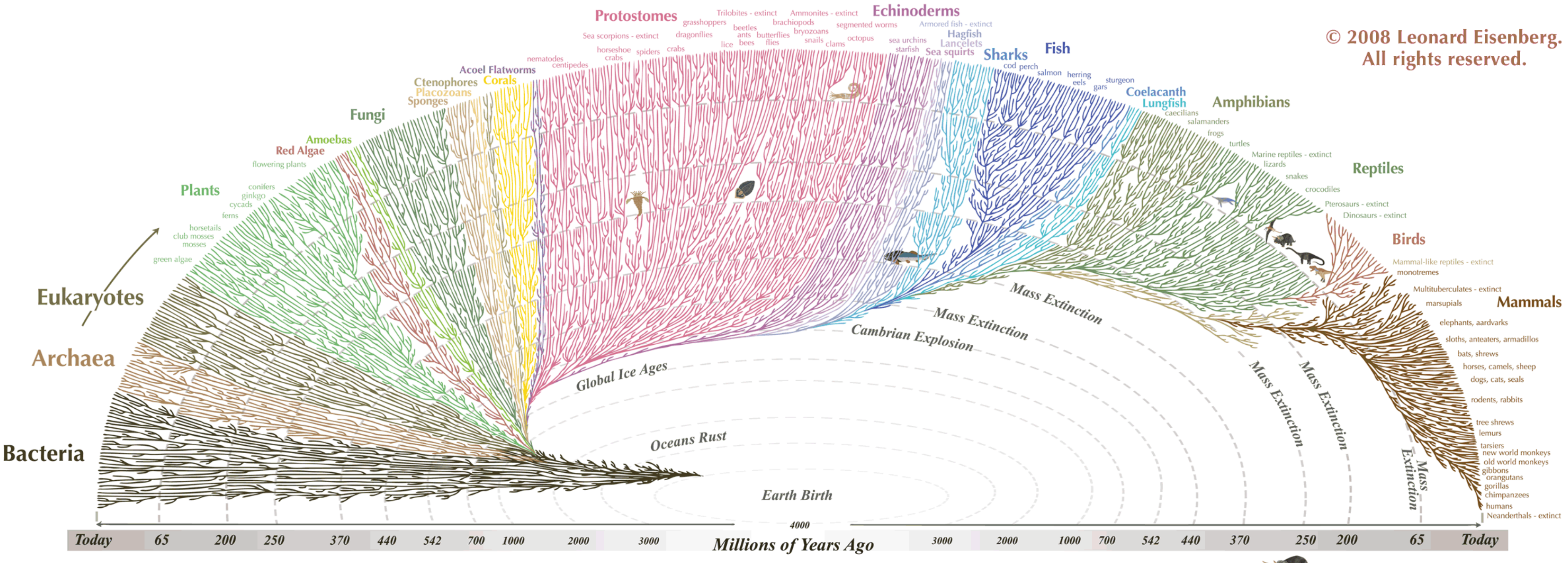
CATEGORIZATIONS OF LAYOUTS

- many possible
- here we follow the categorization on treevis.net:
  - Dimensionality of the layout
  - Representation type
  - Alignment of nodes in space

2D, AXIS-PARALLEL, EXPLICIT EDGES

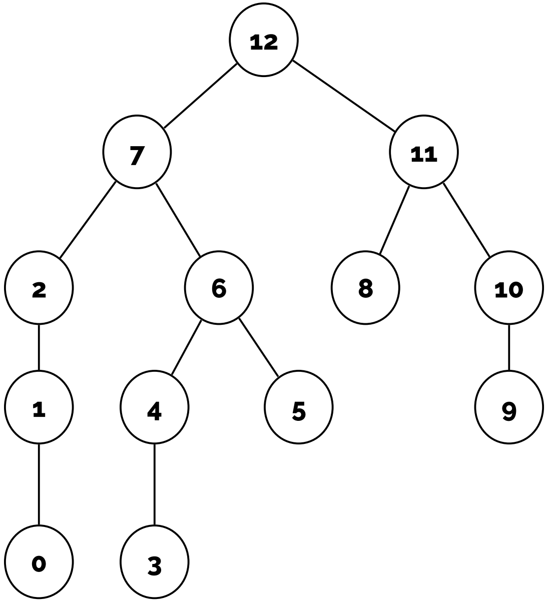


NODE-LINK



All the major and many of the minor living branches of life are shown on this diagram, but only a few of those that have gone extinct are shown. Example: Dinosaurs - extinct

NODE-LINK ALGORITHM

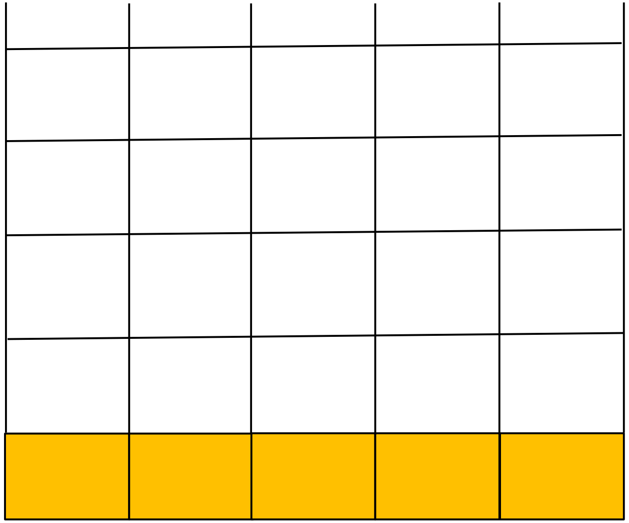


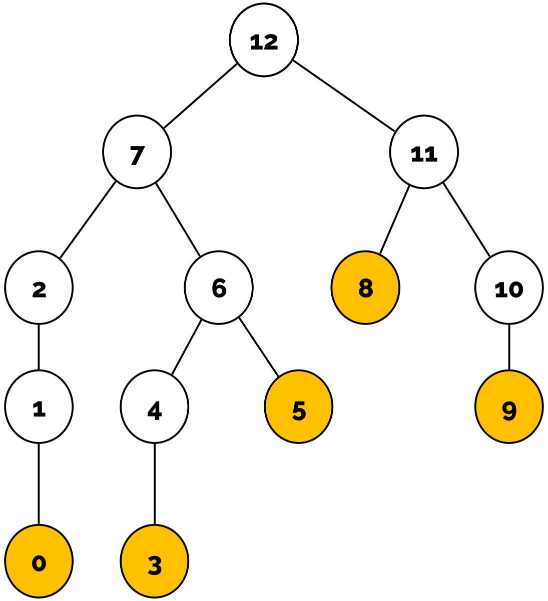
THE LAYOUT WE WANT – HOW DO WE GET THERE?

NODE-LINK ALGORITHM

SIMPLE APPROACH

1) COUNT THE LEAVES
2) PLACE THE ROOT

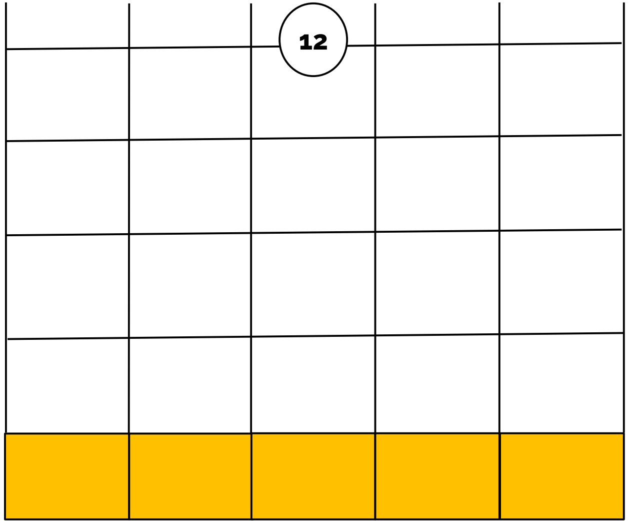


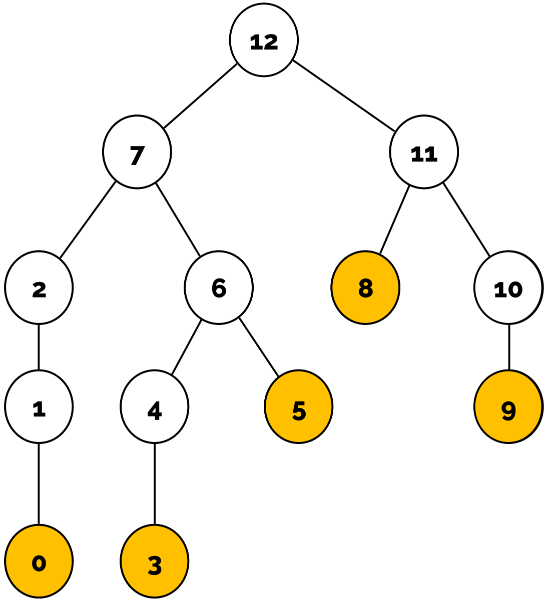


NODE-LINK ALGORITHM

SIMPLE APPROACH

1) COUNT THE LEAVES
2) PLACE THE ROOT
3) RECURSIVELY DIVIDE

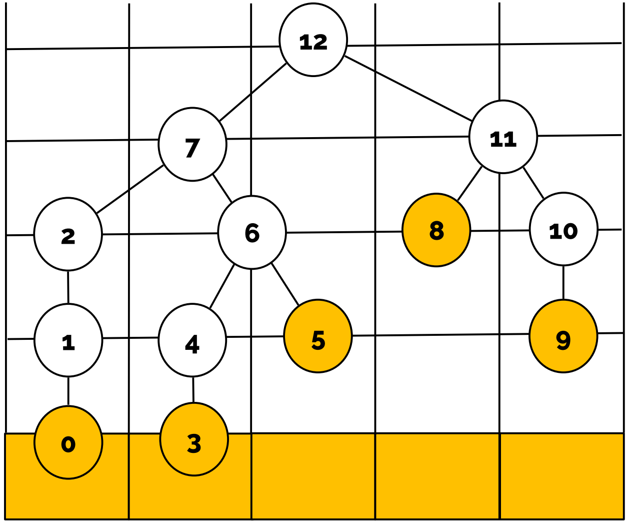


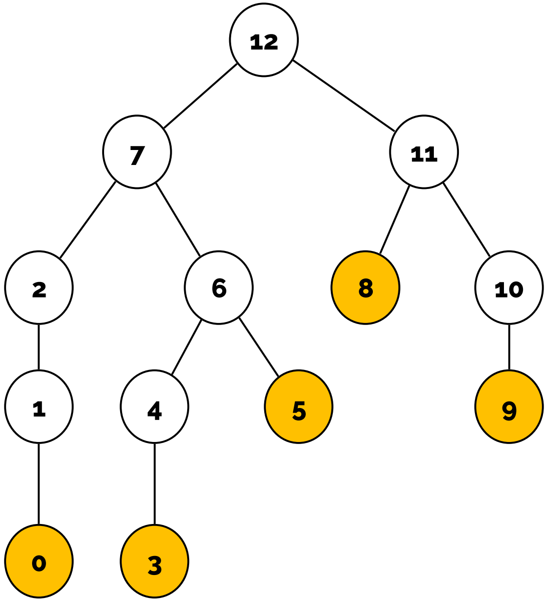


NODE-LINK ALGORITHM

SIMPLE APPROACH

1) COUNT THE LEAVES
2) PLACE THE ROOT
3) RECURSIVELY DIVIDE



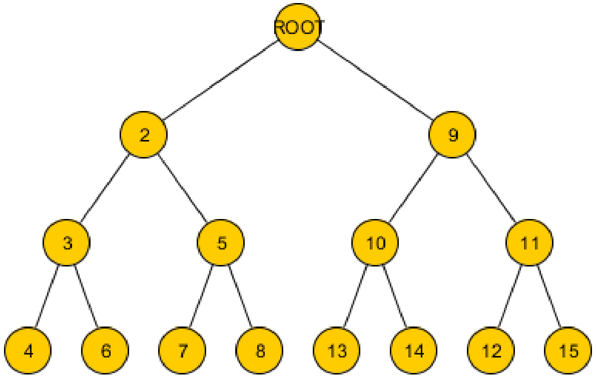


PROS/CONS

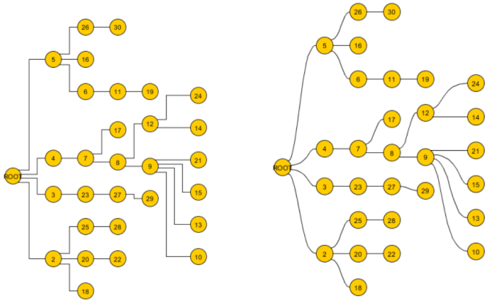
- nodes at the same distance from the root are horizontally aligned
- positive: simple to understand, clear symmetries
- negative: needs large area, often bad aspect ratio (much wider than tall)

what can we vary in this representation?

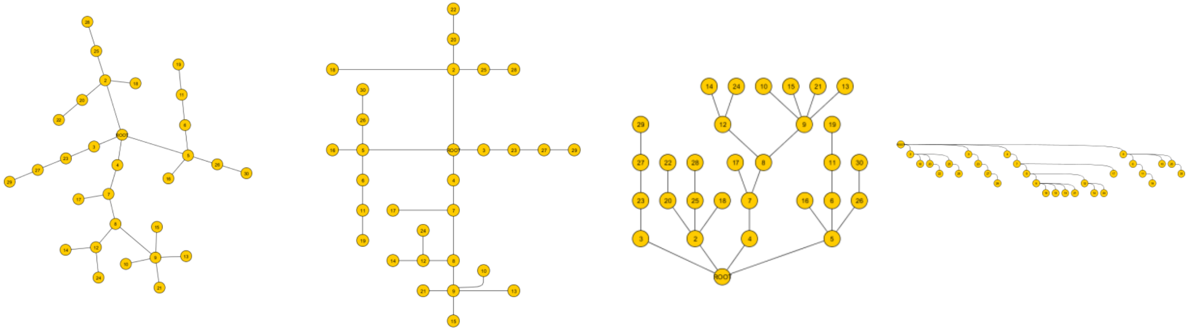
- marks that depict nodes
- visual variables used on marks to depict metadata
- type of links
- visual variables used on marks that depict the links
- placement of nodes



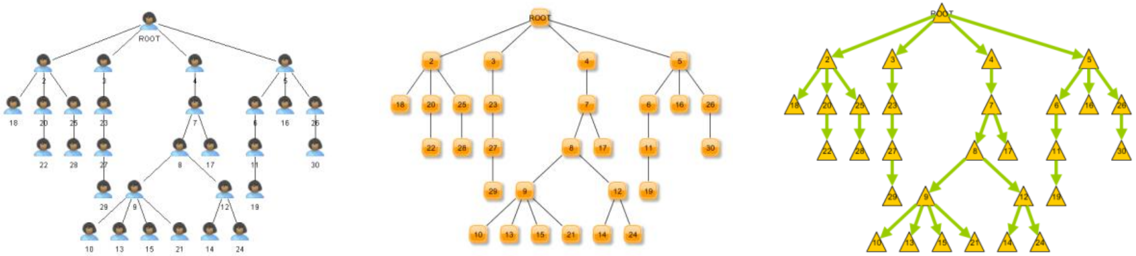
vary edges



node placement



marks



Images created with yEd: http://www.yworks.com

SPECIFIC ALGORITHMS

- usually described recursively
- well known: Reingold-Tilford algorithm
- lots of research in this direction:
  - Wetherell and Shannon 1978, Tidy Drawings of Trees
  - **Reingold and Tilford 1981**, Tidier Drawing of Trees
  - Walker 1990, A Node-positioning Algorithm for General Trees
  - Buchheim et al 2002, Improving Walker’s Algorithm to Run in Linear Time

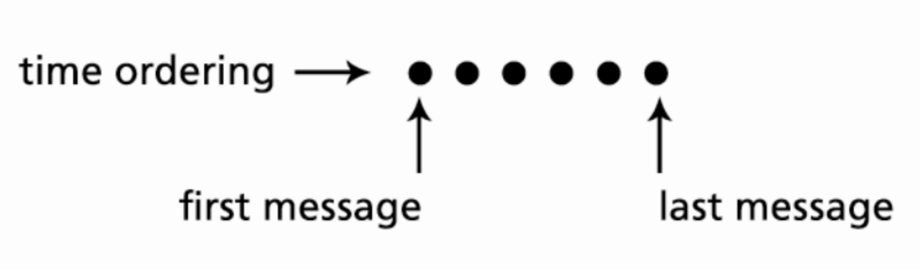
AESTHETICS

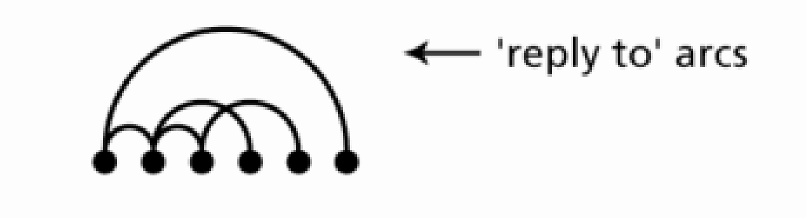
aesthetics of node-link tree algorithms describe properties that improve the perception of the data that is being layed out

- **area**: match area of your layout to the size of the display and data
- **aspect ratio**: usually optimal if close to 1
- **subtree separation**: try not to overlap subtrees
- **root-leaf distance**: minimize distance from root to leaves
- **edge lengths**: minimize total, average, maximum, edge lengths & try to make edge lengths uniform
- **angular resolution**: increase angles formed by edges
- **symmetry**: symmetric layouts usually considered pleasing

LAYOUT DIMENSIONALITY: 2D – THREAD ARCS

email visualization



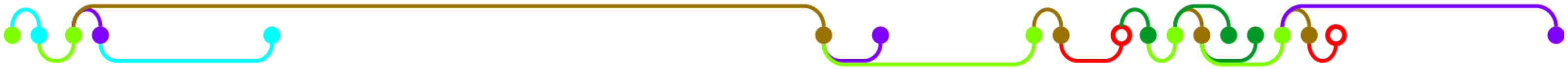


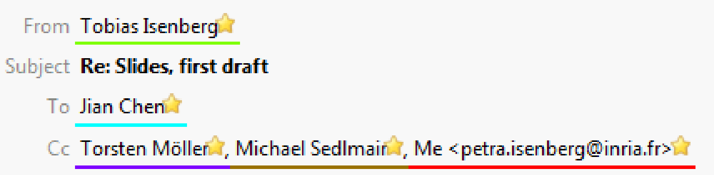
Thread Arcs: An email thread visualization
InfoVis'03: Proceedings of the IEEE Symposium on Information Visualization, pages 211-218.

THREADVIS

http://threadvis.github.io/

- time-scaling
- coloring people



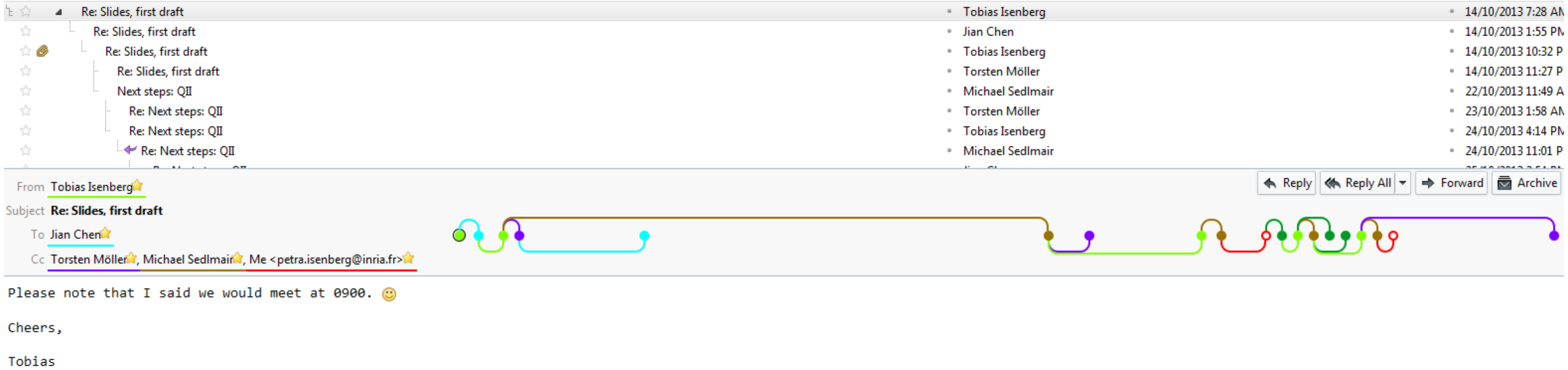


From Tobias Isenberg

Subject **Re: Slides, first draft**

To Jian Chen

Cc Torsten Möller, Michael Sedlmair, Me <petra.isenberg@inria.fr>



Re: Slides, first draft	Tobias Isenberg	14/10/2013 7:28 AM
Re: Slides, first draft	Jian Chen	14/10/2013 1:55 PM
Re: Slides, first draft	Tobias Isenberg	14/10/2013 10:32 P
Re: Slides, first draft	Torsten Möller	14/10/2013 11:27 P
Next steps: QII	Michael Sedlmair	22/10/2013 11:49 A
Re: Next steps: QII	Torsten Möller	23/10/2013 1:58 AM
Re: Next steps: QII	Tobias Isenberg	24/10/2013 4:14 PM
Re: Next steps: QII	Michael Sedlmair	24/10/2013 11:01 P

From Tobias Isenberg

Subject **Re: Slides, first draft**

To Jian Chen

Cc Torsten Möller, Michael Sedlmair, Me <petra.isenberg@inria.fr>

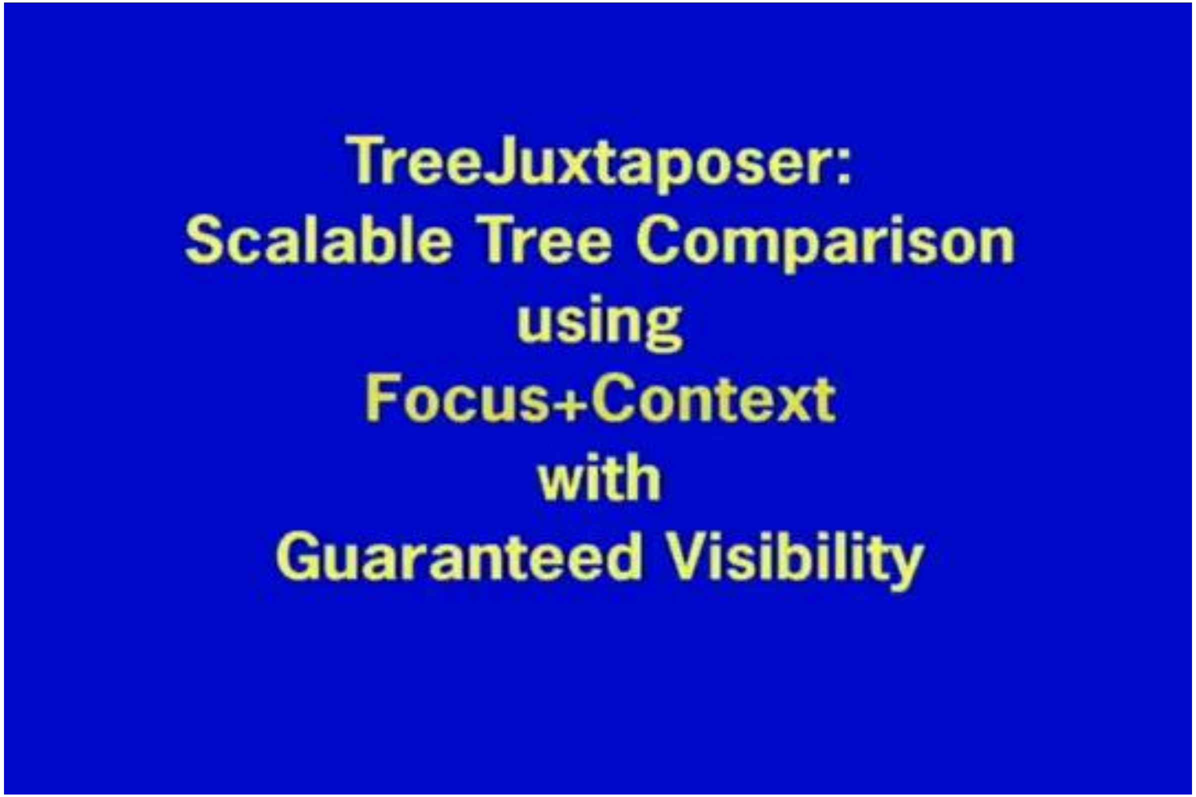
Please note that I said we would meet at 0900. ☺

Cheers,

Tobias

TREEJUXTAPOSER

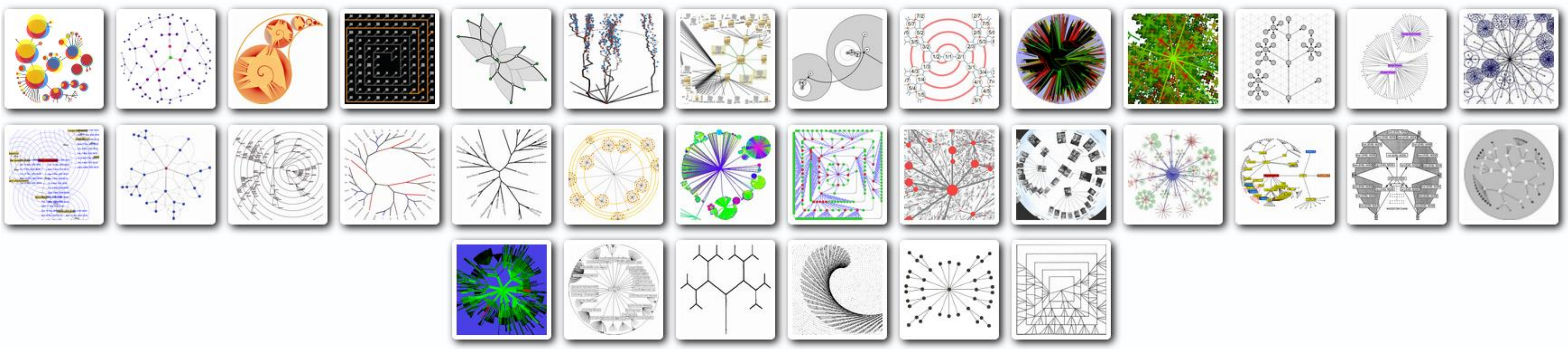
Rectilinear layout and interaction for comparison of very large trees



TreeJuxtaposer:
Scalable Tree Comparison
using
Focus+Context
with
Guaranteed Visibility

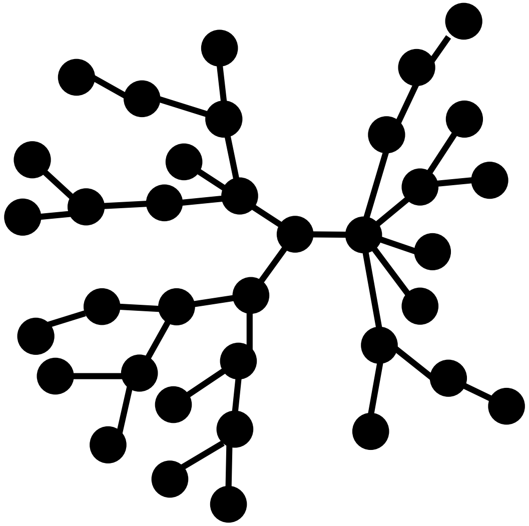
http://www.cs.ubc.ca/~tmm/papers/tj/#video

2D, RADIAL, EXPLICIT EDGES



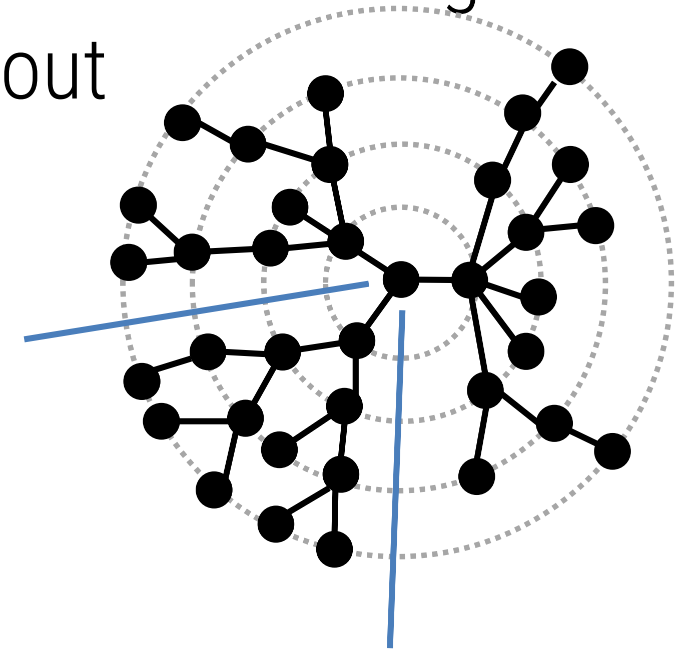
RADIAL NODE-LINK DRAWING

variation of layered drawing from beginning of lecture



RADIAL NODE-LINK DRAWING

- nodes drawn on concentric circles
- nodes drawn within wedges of the circular layout



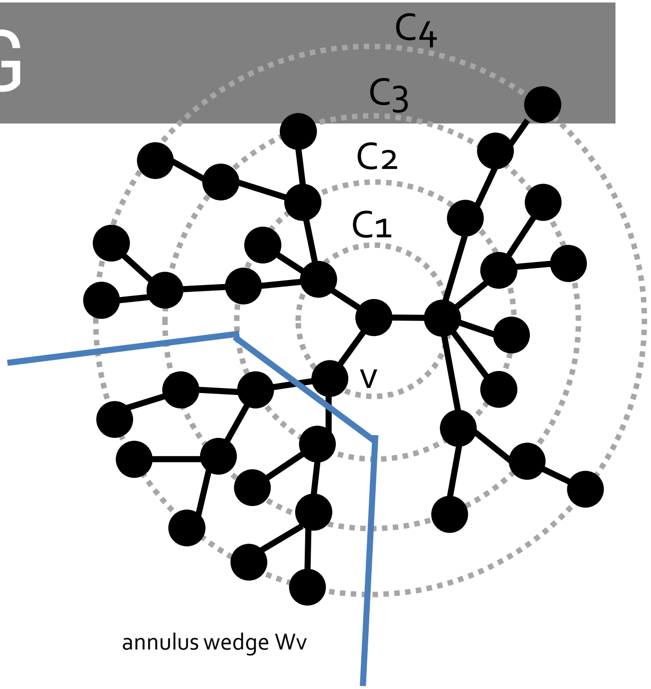
RADIAL NODE-LINK DRAWING

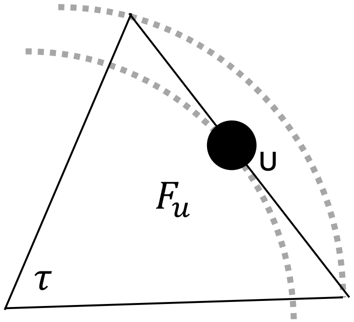
- radius of C_i given by function p(i)
- subtree of v drawn within W_v
- to guarantee planarity (no edge crossings), wedge has to be convex
- several algorithms exist for figuring out the correct angles, e.g.

$$\beta_u = \left(\frac{l(u)\beta_v}{l(v)}, \tau\right)$$

For each child u of v:

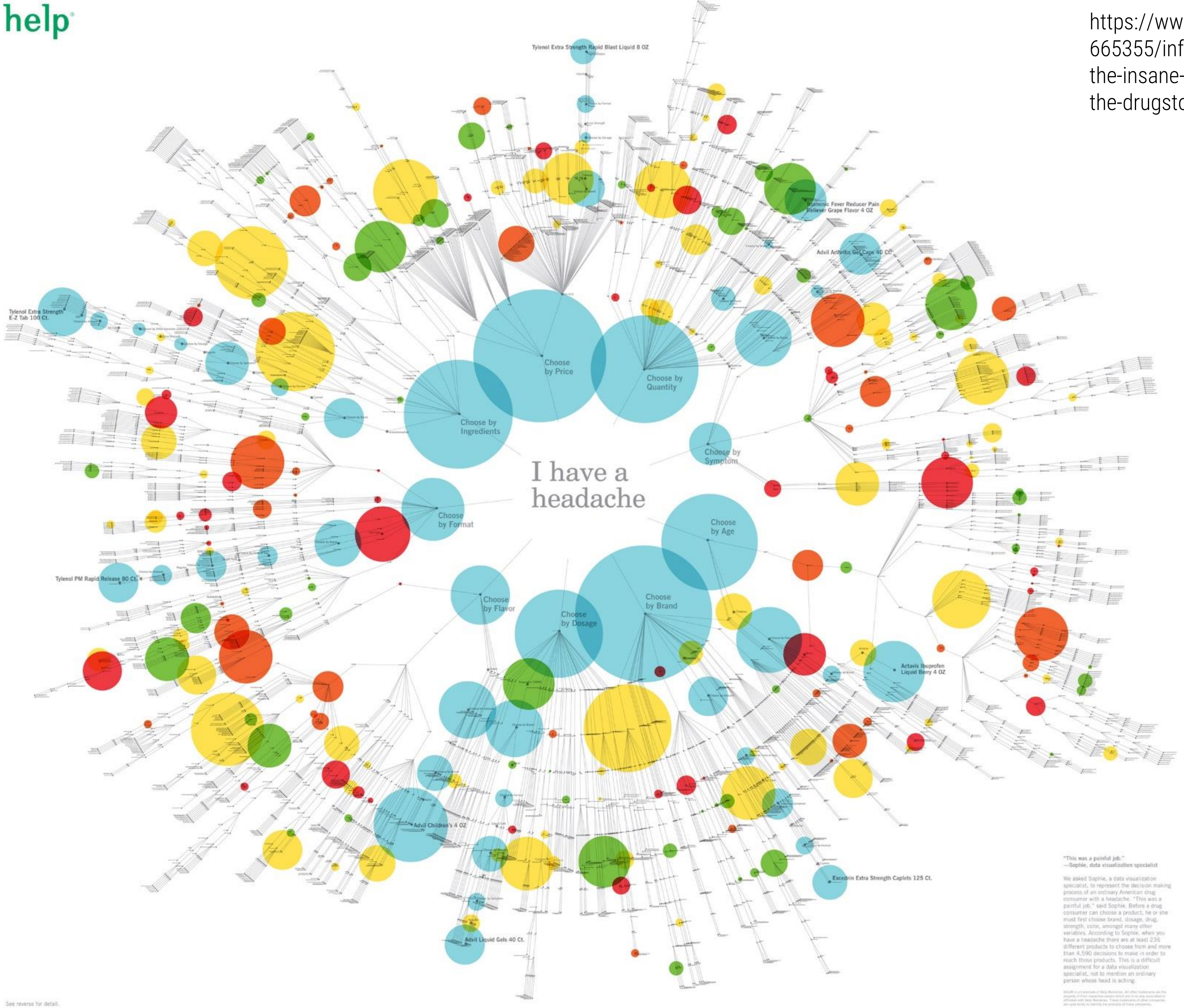
- is the angle of
- is the angle formed by region
- l(v): number of leaves in subtree rooted at v
- place u at center of





Resource: Battista et al. Graph Drawing

https://www.fastcompany.com/1665355/infographic-of-the-day-the-insane-choices-you-face-at-the-drugstore



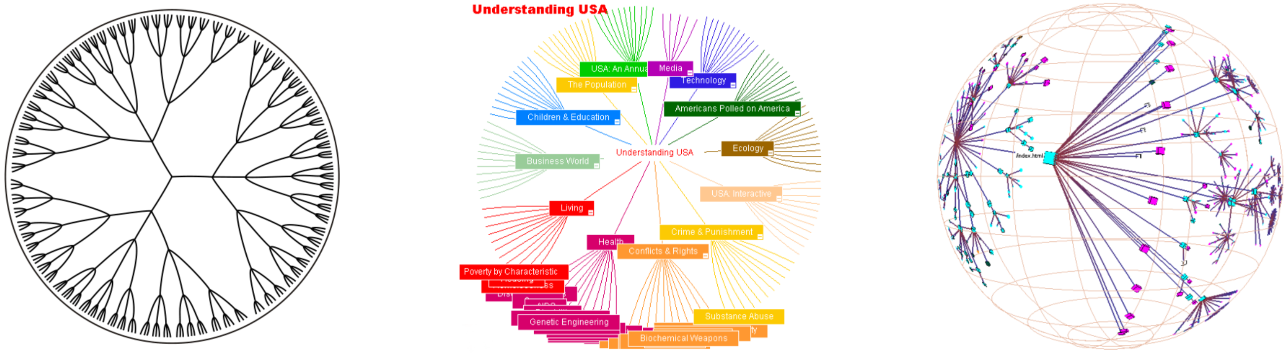
"This was a painful job."
—Sophie, data visualization specialist

We asked Sophie, a data visualization specialist, to represent the decision making process of an ordinary American drug consumer with a headache. "This was a painful job," said Sophie. Before a drug consumer can choose a product, he or she must first choose brand, dosage, drug, strength, color, amongst many other variables. According to Sophie, when you have a headache there are at least 236 different products to choose from and more than 4,590 decisions to make in order to reach those products. This is a difficult assignment for a data visualization specialist, not to mention an ordinary person whose head is aching.

Help® is a trademark of Help Remedies. All other trademarks are the property of their respective owners which are in no way associated or affiliated with Help Remedies. These trademarks of other companies are used solely to identify the products of these companies.

HYPERBOLIC BROWSER

- uses hyperbolic geometry (not euclidean geometry)
- a hyperbolic plane can be displayed using the Poincaré disk model
  - a tree structure of any size fits within a finite area (circle)
  - node is displayed in center
  - all oder nodes move away from center and become exponentially smaller

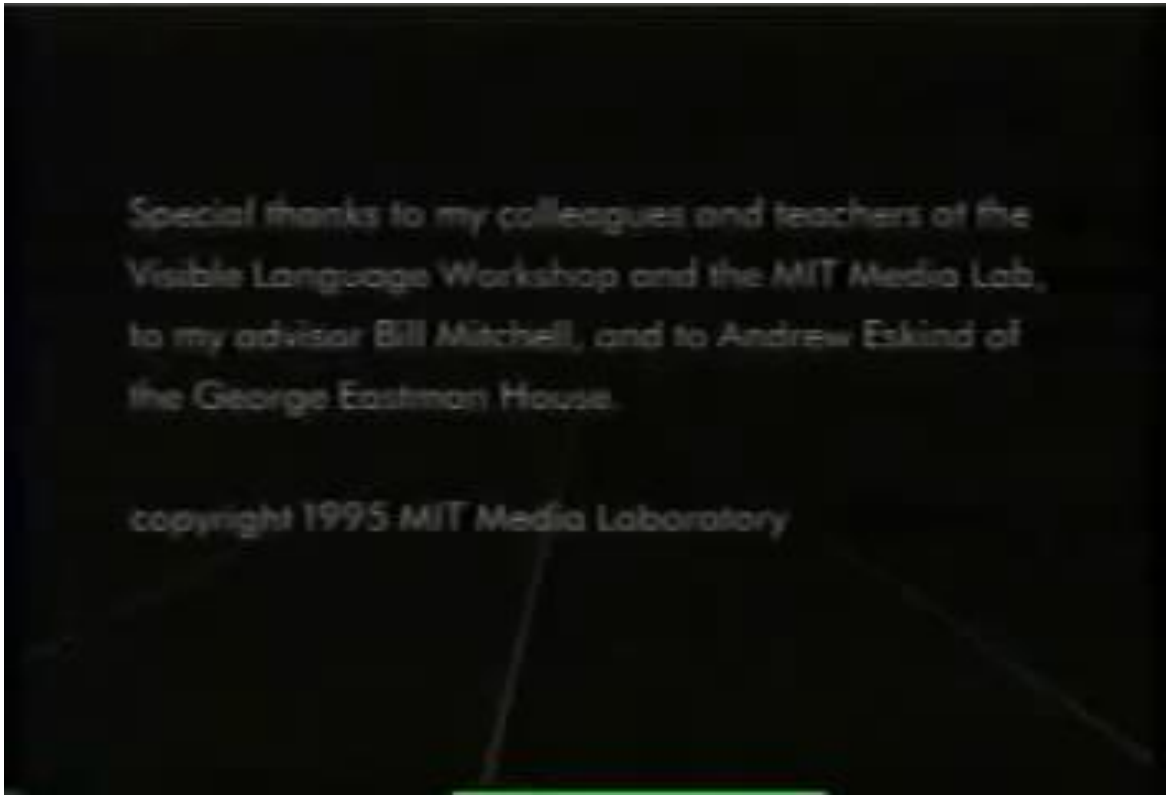


(a) Uniform hyperbolic tree.

(b) StarTree by Inxight Software.

(c) H3 Browser.

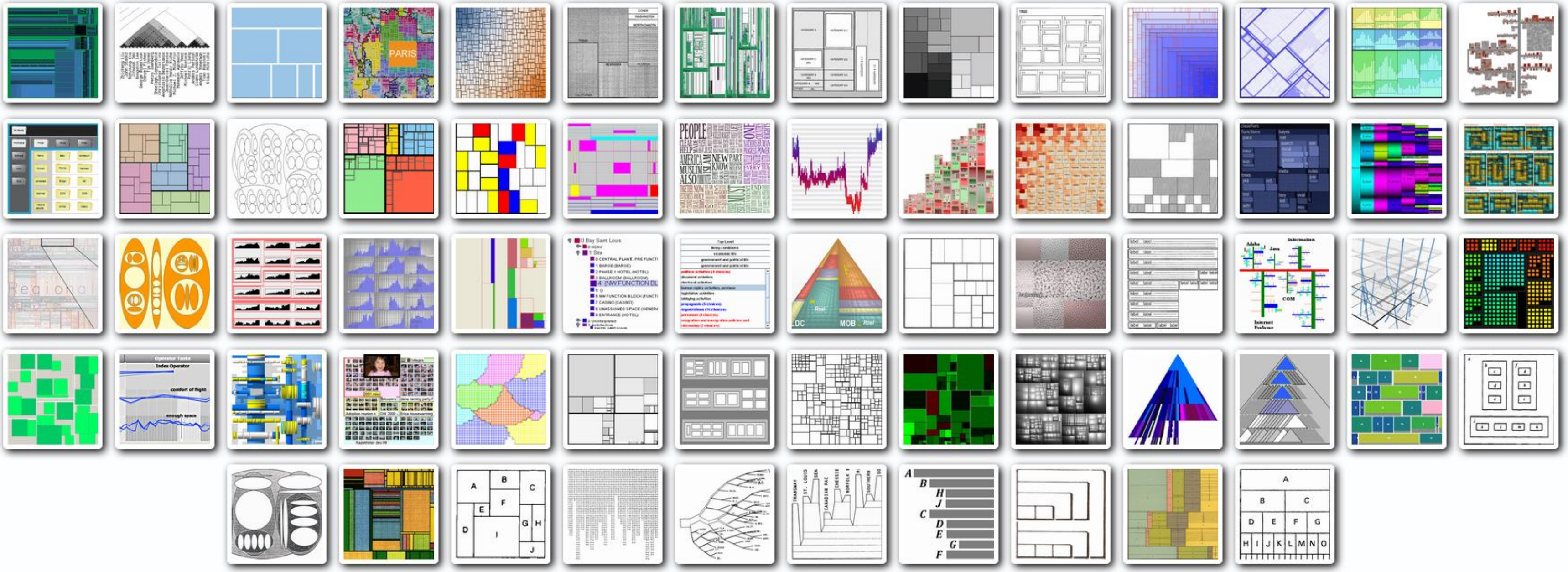
CHI 1996 VIDEO OF HYPERBOLIC BROWSER



Special thanks to my colleagues and teachers at the Visible Language Workshop and the MIT Media Lab, to my advisor Bill Mitchell, and to Andrew Eskind of the George Eastman House.

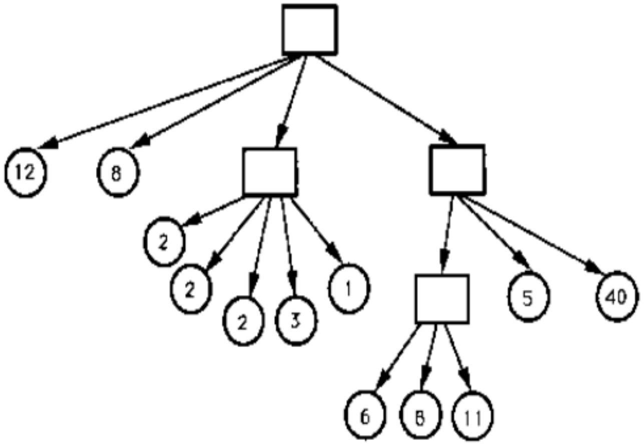
copyright 1995 MIT Media Laboratory

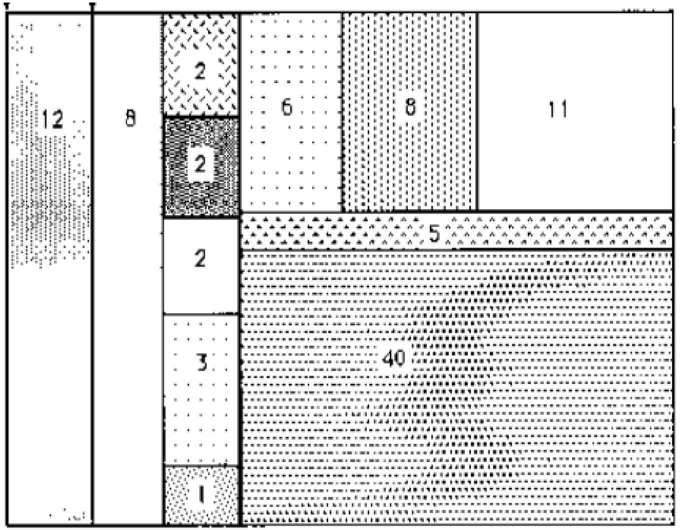
2D, AXIS-PARALLEL, IMPLICIT EDGES



A CLASSIC CONTAINMENT LAYOUT

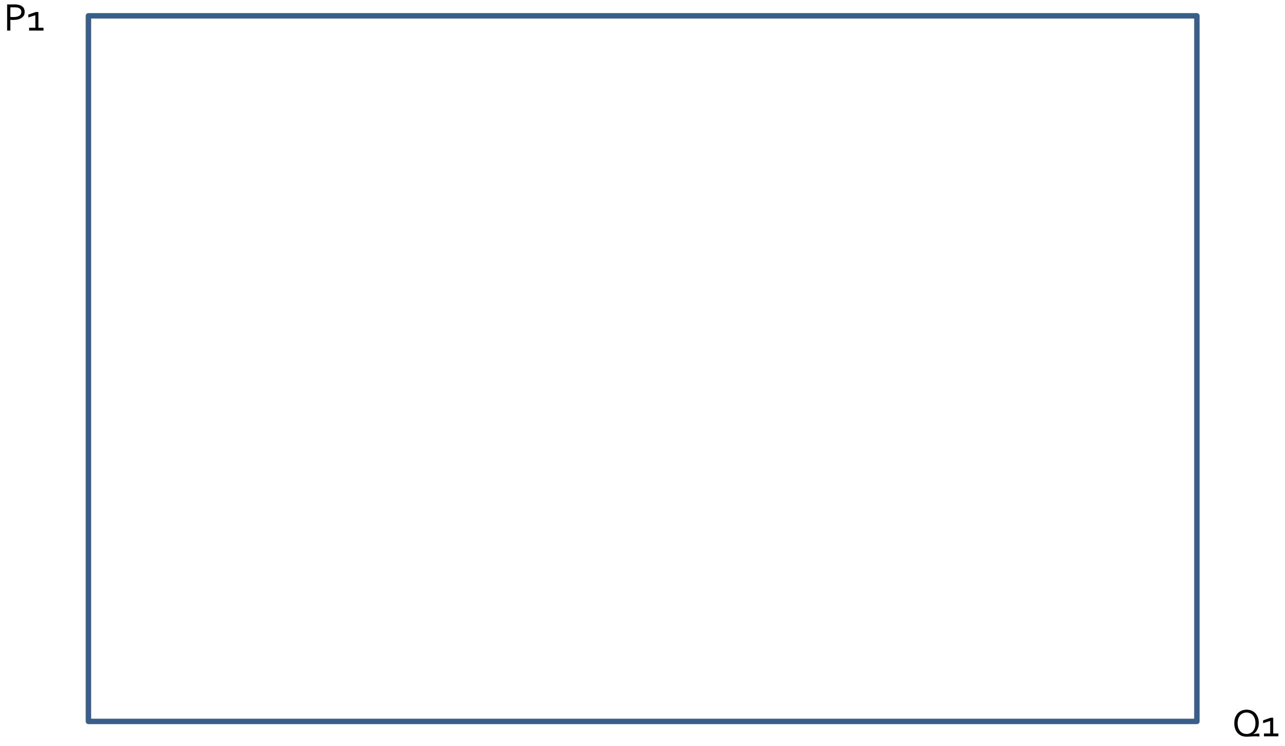
- example tree to rebuild with treemap algorithm
- size of each node as numbers in leaves





TREEMAP ALGORITHM

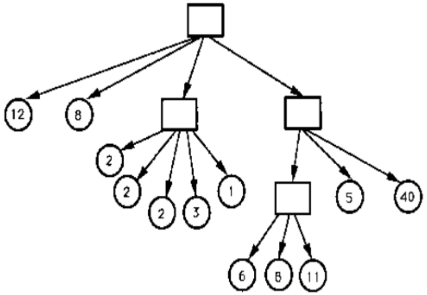
- Take a rectangular display area P1(x1,y1), Q1(x2,y1)
- This area represents the root of the tree

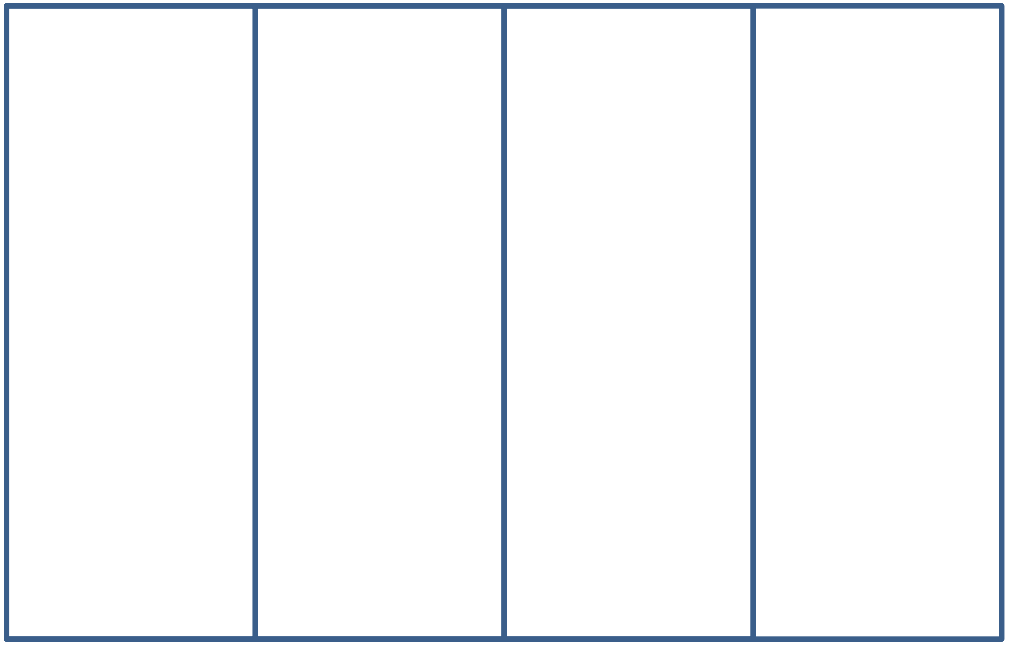


TREEMAP ALGORITHM

recursive algorithm

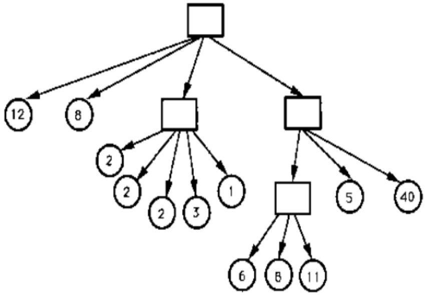
- the number of children of the current node define the number of partitions of the current node

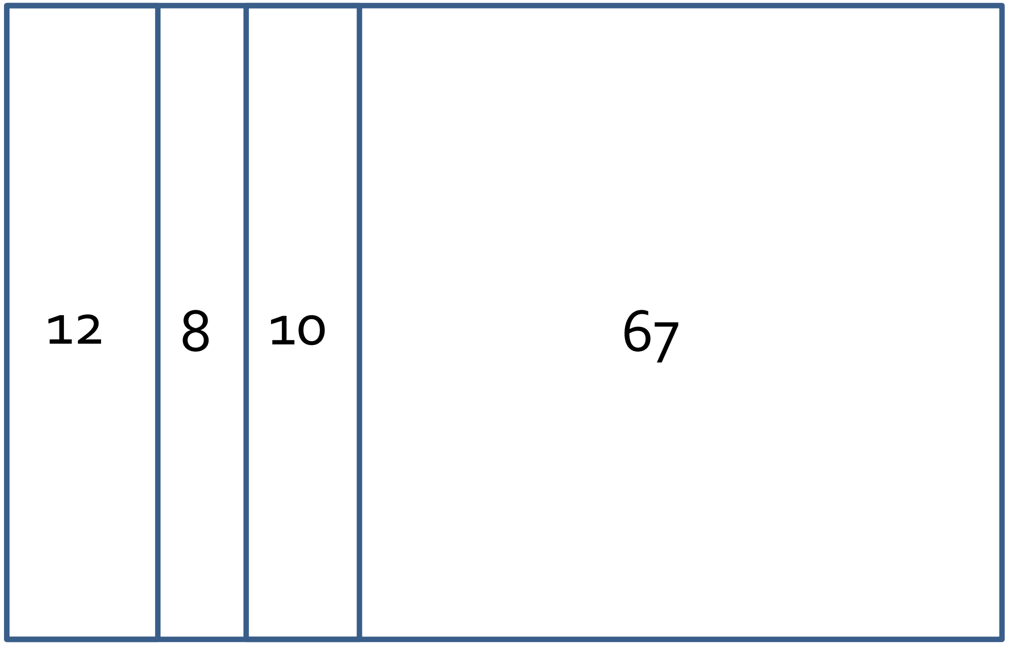




TREEMAP ALGORITHM

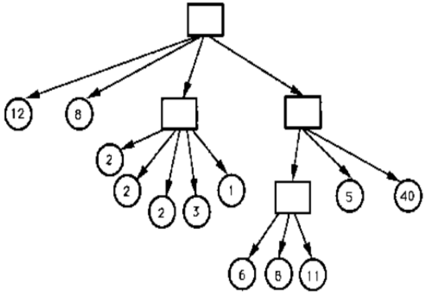
the weight of each node determines the size of each partition

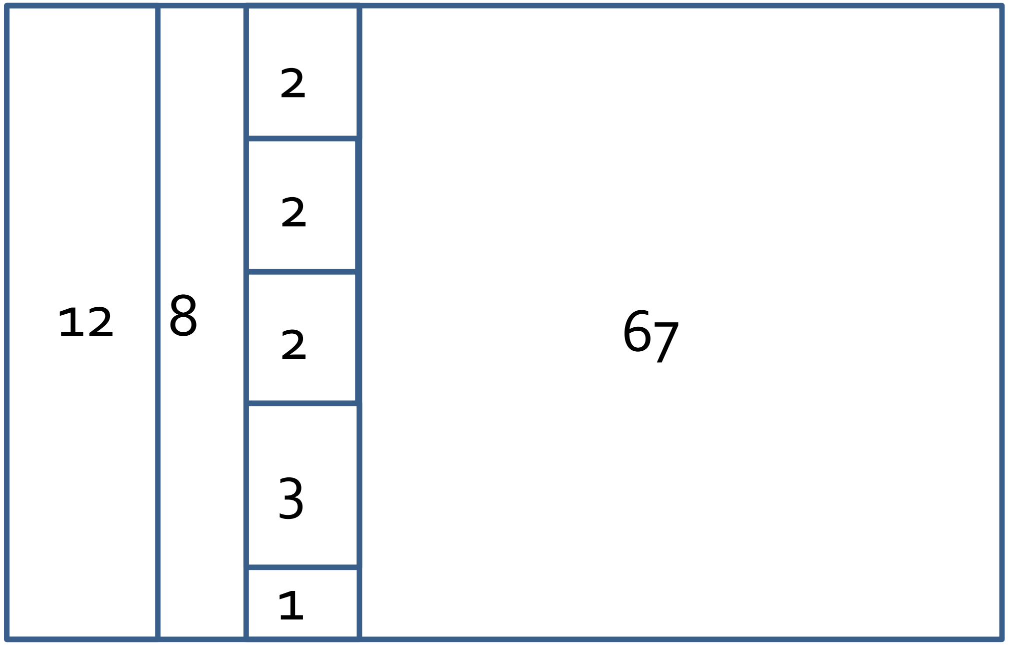




TREEMAP ALGORITHM

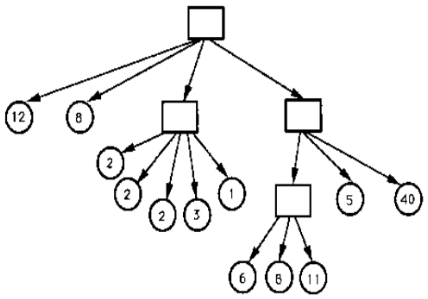
at each change of level, rotate orientation of split by 90 degrees

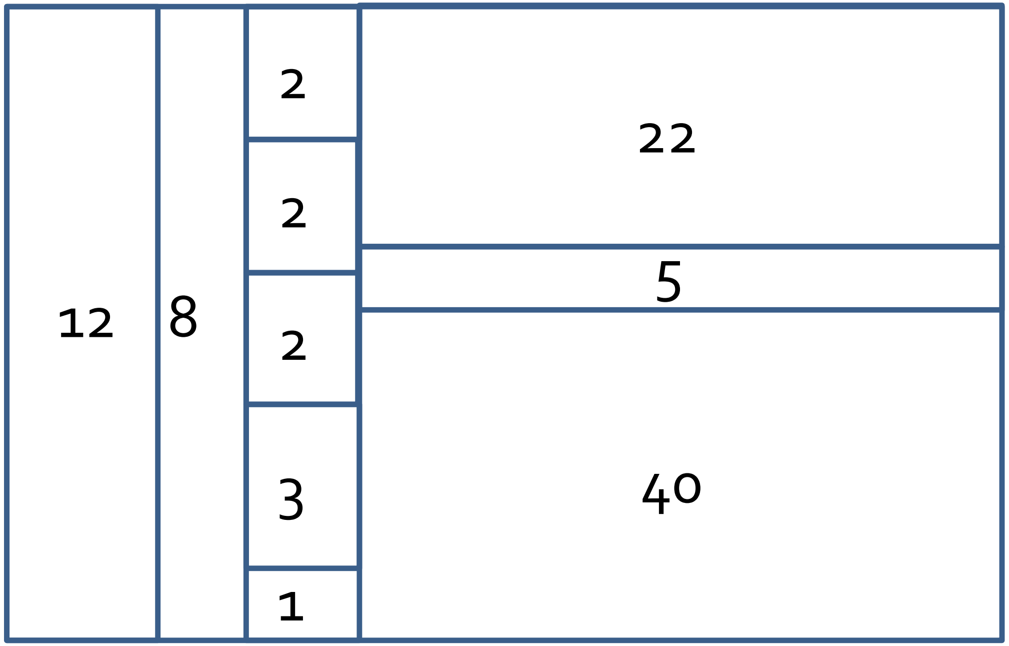




TREEMAP ALGORITHM

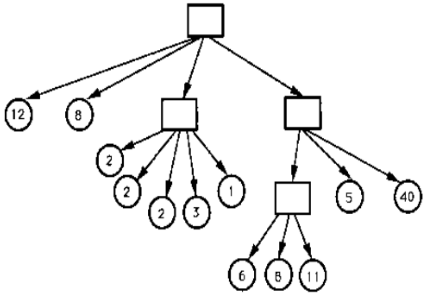
at each change of level, rotate orientation of split by 90 degrees

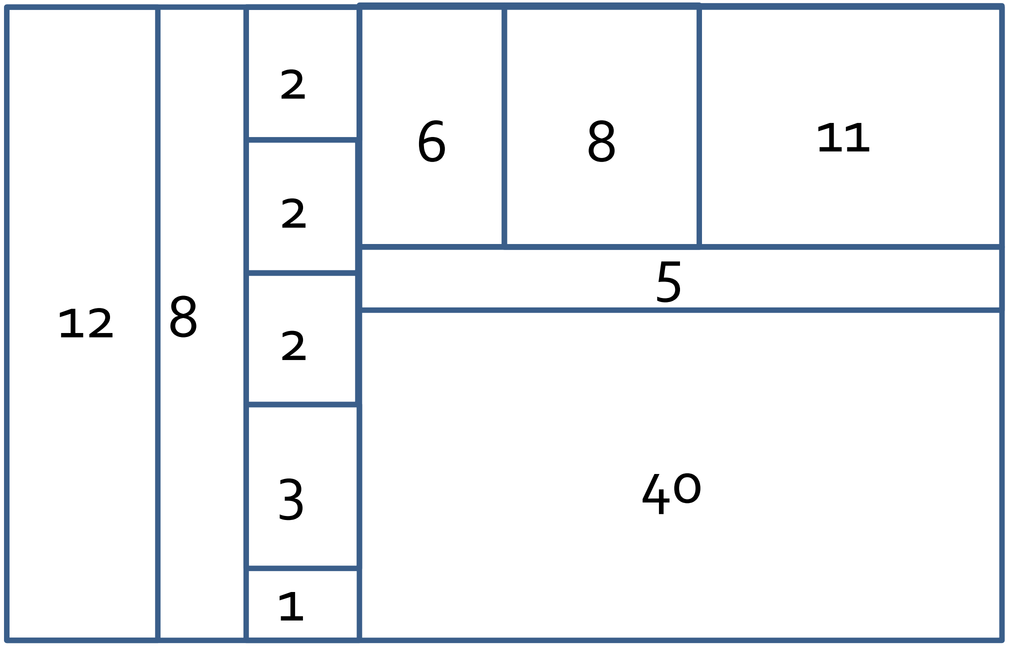




TREEMAP ALGORITHM

at each change of level, rotate orientation of split by 90 degrees



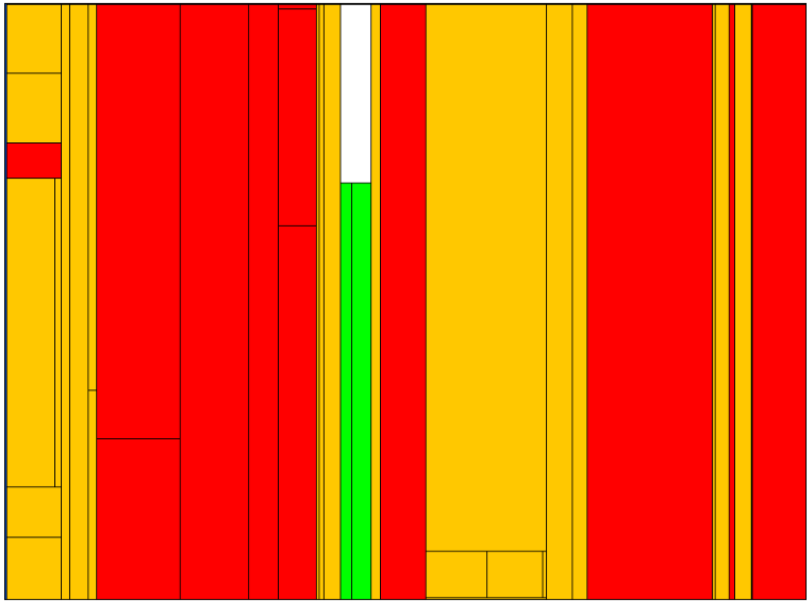


TREEMAP

- a 2-D space-filling layout
- for further references and to try out a treemap in various applications: http://www.cs.umd.edu/hcil/treemap-history/

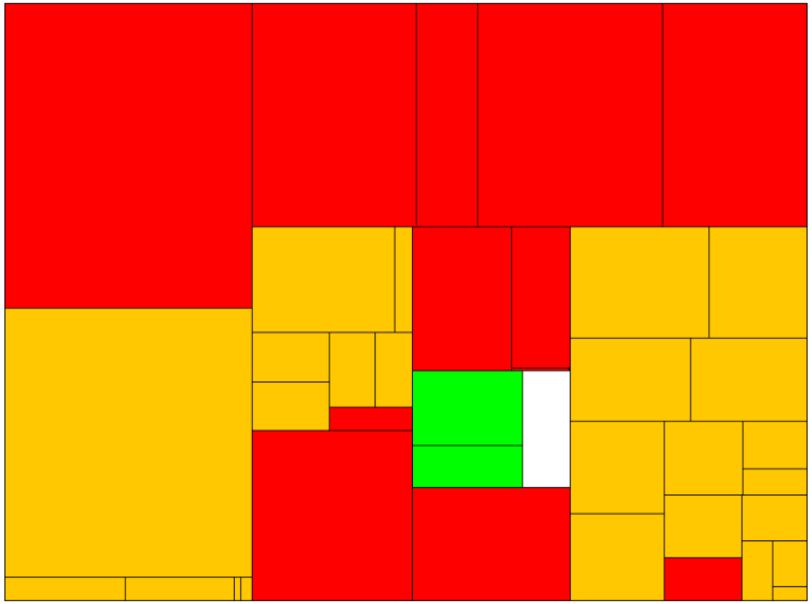
TREEMAP VARIATIONS

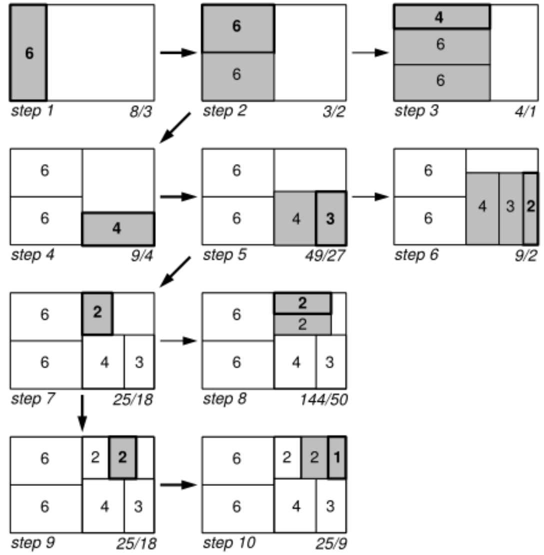
- problem with original treemap: lots of long stripes
- for long stripes the areas are difficult to compare



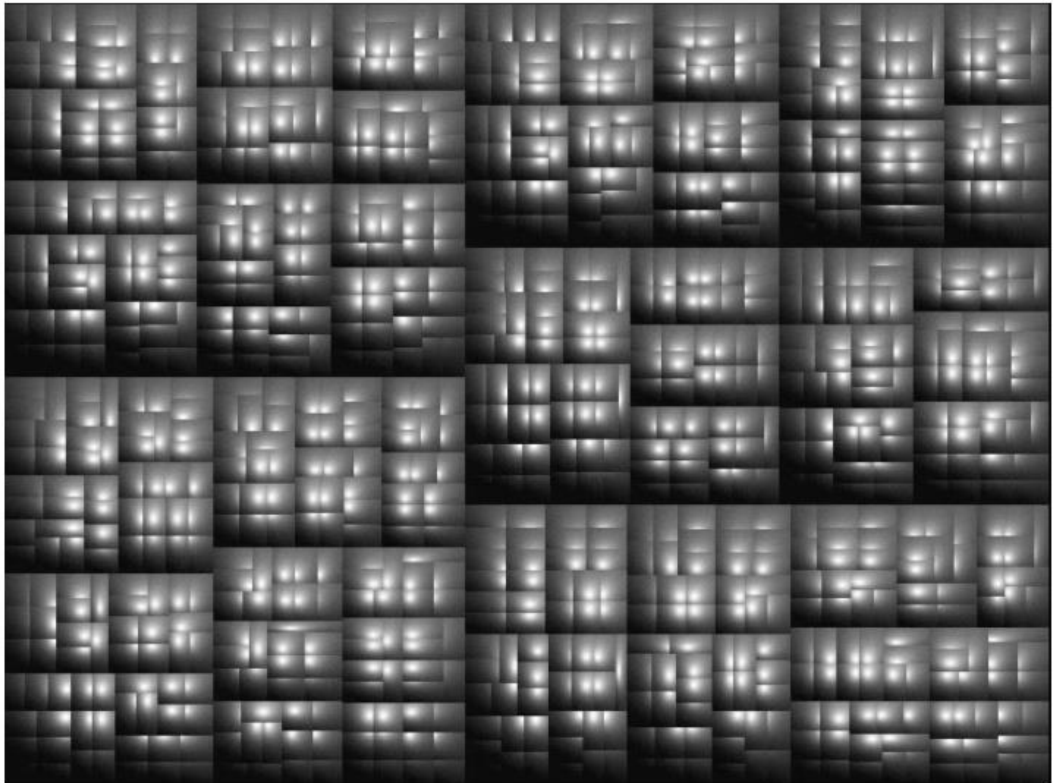
SQUARIFIED TREEMAP

- calculates more squared regions
- problem: order not as easily read, not very stable with dynamically changing data



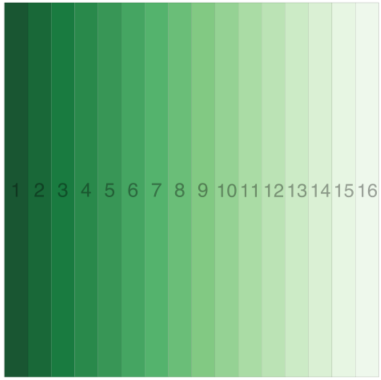


SQUARIFIED TREEMAP



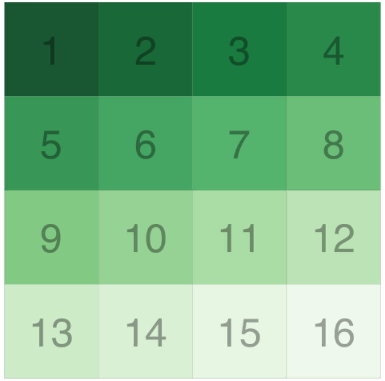
ORDERED TREEMAP

several algorithms in comparison



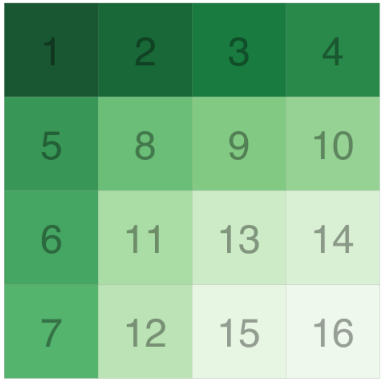
slice and dice

B. Shneiderman. Tree visualization with tree-maps: 2-d space-filling approach. ACM Transactions on Graphics, 11:92–99, 1992.



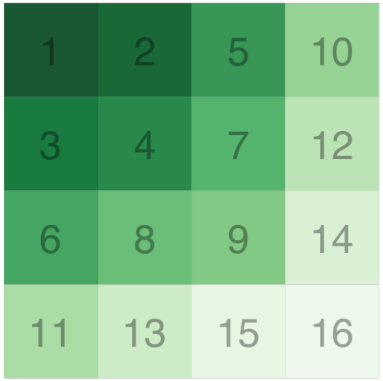
strip

B. B. Bederson, B. Shneiderman, and M. Wattenberg. Ordered and quantum treemaps: Making effective use of 2d space to display hierarchies. ACM Transactions on Graphics, 21:833–854, 2002.



squarified

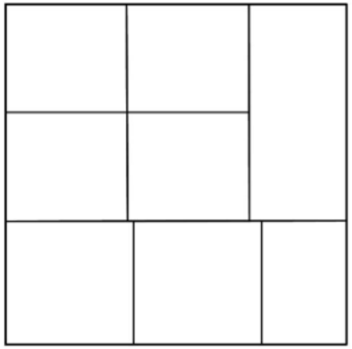
M. Bruls, K. Huizing, and J. van Wijk. Squarified treemaps. EuroVis, pages 33–42, 2000.



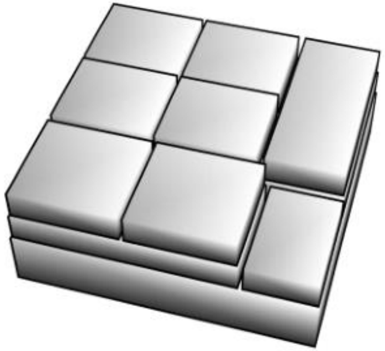
ordered squarified

B. Shneiderman and M. Wattenberg. Ordered treemap layouts. In Infovis01, pages 73–78, 2001.

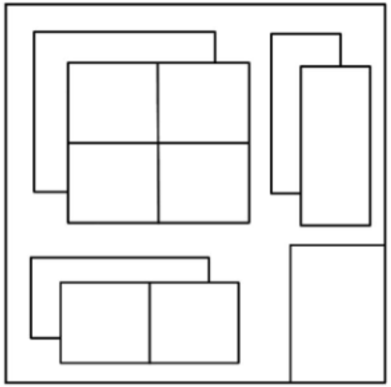
OTHER VARIATIONS OF TREEMAPS



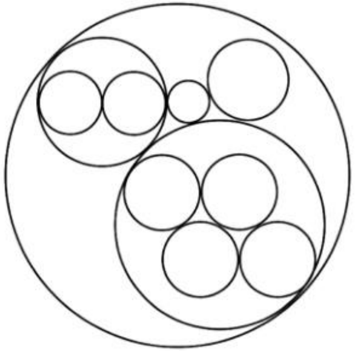
original squarified:
emphasizes leafs and their attributes



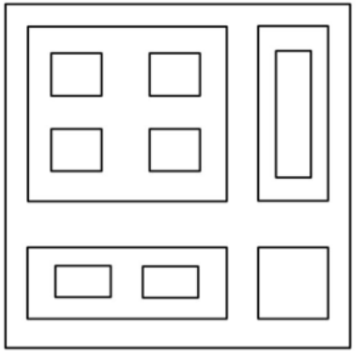
steptree:
emphasizes structure with extrusion



cascaded layout:
emphasizes structure with overlap



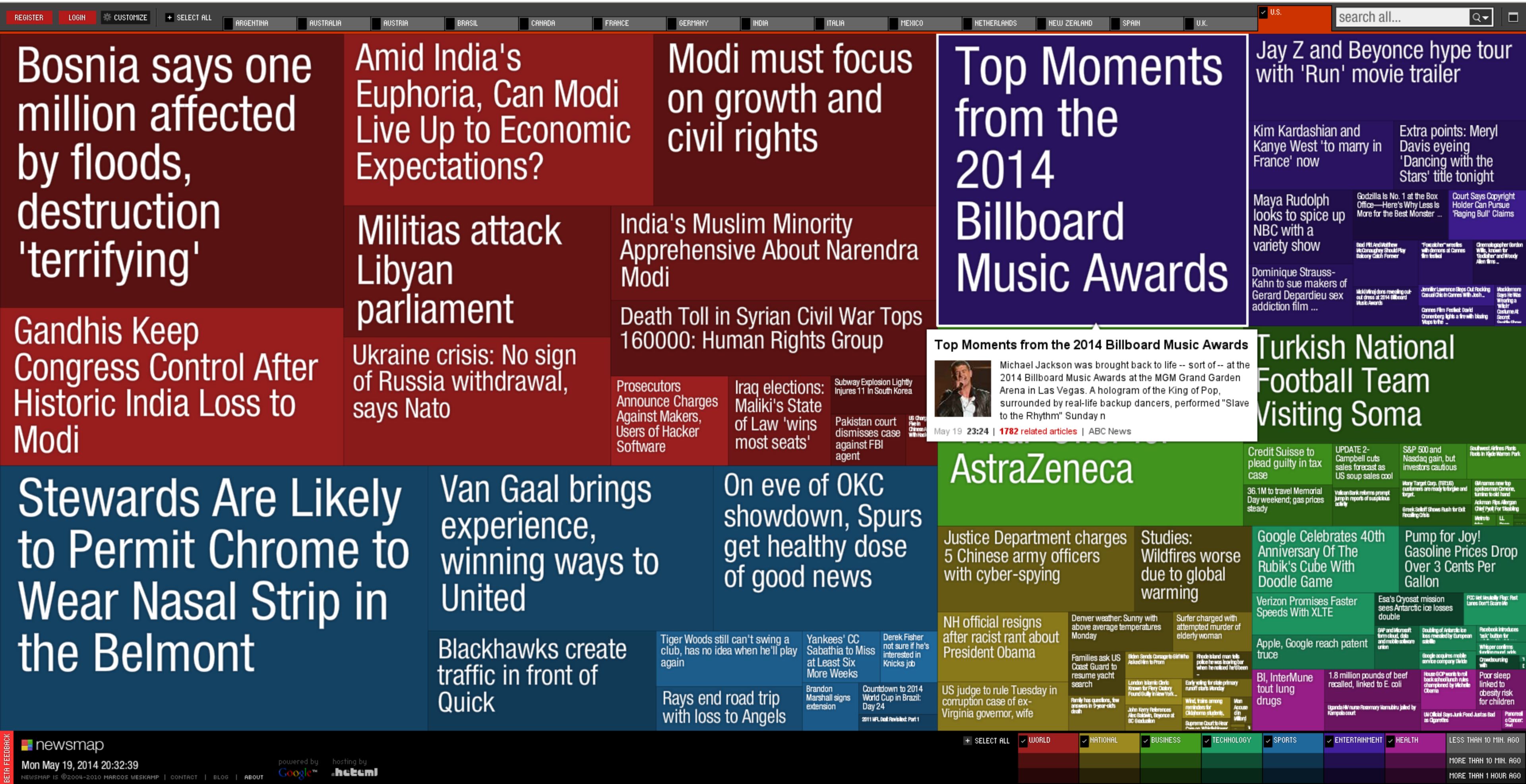
circular treemap:
emphasizes structure with
non-space-filling primitive



nested layout:
emphasizes structure with whitespace

TVCG, 17(4), The Design Space of Implicit Hierarchy Visualization: A Survey

HTTP://NEWSMAP.JP/



Bosnia says one million affected by floods, destruction 'terrifying'

Amid India's Euphoria, Can Modi Live Up to Economic Expectations?

Modi must focus on growth and civil rights

Gandhis Keep Congress Control After Historic India Loss to Modi

Militias attack Libyan parliament

India's Muslim Minority Apprehensive About Narendra Modi

Death Toll in Syrian Civil War Tops 160000: Human Rights Group

Ukraine crisis: No sign of Russia withdrawal, says Nato

Prosecutors Announce Charges Against Makers, Users of Hacker Software

Iraq elections: Maliki's State of Law 'wins most seats'

Subway Explosion Lightly Injures 11 In South Korea

Pakistan court dismisses case against FBI agent

Top Moments from the 2014 Billboard Music Awards

Jay Z and Beyonce hype tour with 'Run' movie trailer

Kim Kardashian and Kanye West 'to marry in France' now

Extra points: Meryl Davis eyeing 'Dancing with the Stars' title tonight

Maya Rudolph looks to spice up NBC with a variety show

Godzilla Is No. 1 at the Box Office—Here's Why Less Is More for the Best Monster ...

Court Says Copyright Holder Can Pursue 'Raging Bull' Claims

Brad Pitt And Matthew McConaughey Should Play Balcony Catch Forever

"Foxcatcher" wrestles with demons at Cannes film festival

Cinematographer Gordon Willis, known for 'Godfather' and Woody Allen films ...

Dominique Strauss-Kahn to sue makers of Gerard Depardieu sex addiction film ...

Nicki Minaj dons revealing cut-out dress at 2014 Billboard Music Awards

Jennifer Lawrence Steps Out Rocking Casual Chic In Cannes With Josh ...

Cannes Film Festival: David Cronenberg lights a fire with blazing 'Maps to the ...

Macklemore Says He Was Wearing a 'Witch' Costume At Secret

Top Moments from the 2014 Billboard Music Awards

Michael Jackson was brought back to life -- sort of -- at the 2014 Billboard Music Awards at the MGM Grand Garden Arena in Las Vegas. A hologram of the King of Pop, surrounded by real-life backup dancers, performed "Slave to the Rhythm" Sunday n

May 19 **23:24** | 1782 related articles | ABC News

Turkish National Football Team Visiting Soma

AstraZeneca

Credit Suisse to plead guilty in tax case

UPDATE 2-Campbell cuts sales forecast as US soup sales cool

S&P 500 and Nasdaq gain, but investors cautious

Southwest Airlines Plants Roots In Klyde Warren Park

36.1M to travel Memorial Day weekend; gas prices steady

Vatican Bank reforms prompt jump in reports of suspicious activity

Many Target Corp. (TGT.US) customers are ready to forgive and forget.

GM names new top spokesman Cervone, turning to old hand

Greek Selloff Shows Rush for Exit Recalling Crisis

Ackman Rips Allergan Chief Pyott For 'Disabling

Stewards Are Likely to Permit Chrome to Wear Nasal Strip in the Belmont

Van Gaal brings experience, winning ways to United

On eve of OKC showdown, Spurs get healthy dose of good news

Blackhawks create traffic in front of Quick

Tiger Woods still can't swing a club, has no idea when he'll play again

Yankees' CC Sabathia to Miss at Least Six More Weeks

Derek Fisher not sure if he's interested in Knicks job

Rays end road trip with loss to Angels

Brandon Marshall signs extension

Countdown to 2014 World Cup in Brazil: Day 24

2011 NFL Draft Revisited: Part 1

Justice Department charges 5 Chinese army officers with cyber-spying

Studies: Wildfires worse due to global warming

NH official resigns after racist rant about President Obama

Denver weather: Sunny with above average temperatures Monday

Surfer charged with attempted murder of elderly woman

Families ask US Coast Guard to resume yacht search

Biden Sends Corsage to Girl Who Asked Him to Prom

Rhode Island man tells police he was leaving bar when he noticed he'd been

US judge to rule Tuesday in corruption case of ex-Virginia governor, wife

Family has questions, few answers in 9-year-old's death

London Islamic Cleric Known for Fiery Oratory Found Guilty in New York ...

Early voting for state primary runoff starts Monday

John Kerry References Alec Baldwin, Beyonce at BC Graduation

Wind, trains among reminders for Oklahoma students,

Supreme Court to Hear

Google Celebrates 40th Anniversary Of The Rubik's Cube With Doodle Game

Pump for Joy! Gasoline Prices Drop Over 3 Cents Per Gallon

Verizon Promises Faster Speeds With XLTE

Esa's Cryosat mission sees Antarctic ice losses double

FCC Net Neutrality Flap: Fast Lanes Don't Scare Me

Apple, Google reach patent truce

SAP and Microsoft form cloud, data and mobile software union

Doubling of Antarctic ice loss revealed by European satellite

Facebook introduces 'ask' button for

Whisper confirms

Google acquires mobile service company Divide

Crowdsourcing with

BI, InterMune tout lung drugs

1.8 million pounds of beef recalled, linked to E. coli

House GOP wants to roll back school lunch rules championed by Michelle Obama

Poor sleep linked to obesity risk for children

Uganda HIV nurse Rosemary Namubiru jailed by Kampala court

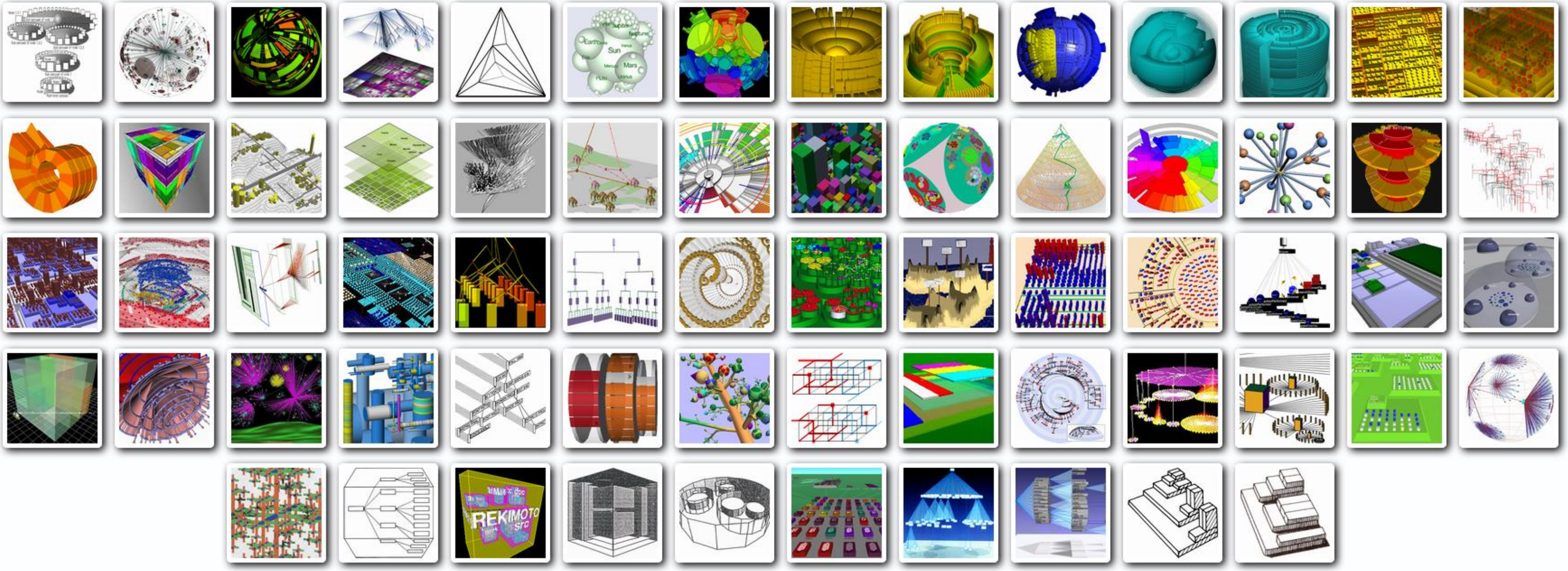
UN Official Says Junk Food Just as Bad as Cigarettes

newsmap

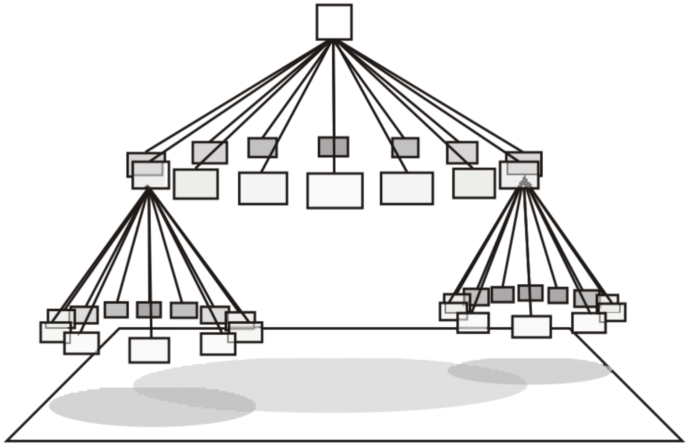
Mon May 19, 2014 20:32:39

NEWSMAP IS ©2004-2010 MARCOS WESKAMP | CONTACT | BLOG | ABOUT

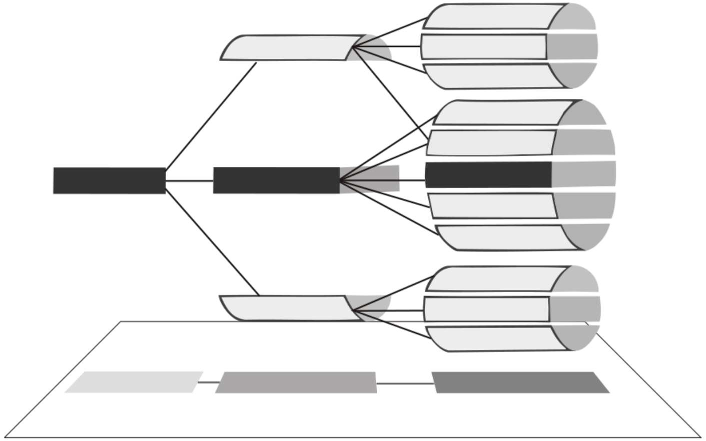
3D LAYOUTS



HISTORIC EXAMPLE: CONETREE / CAMTREE



ConeTree

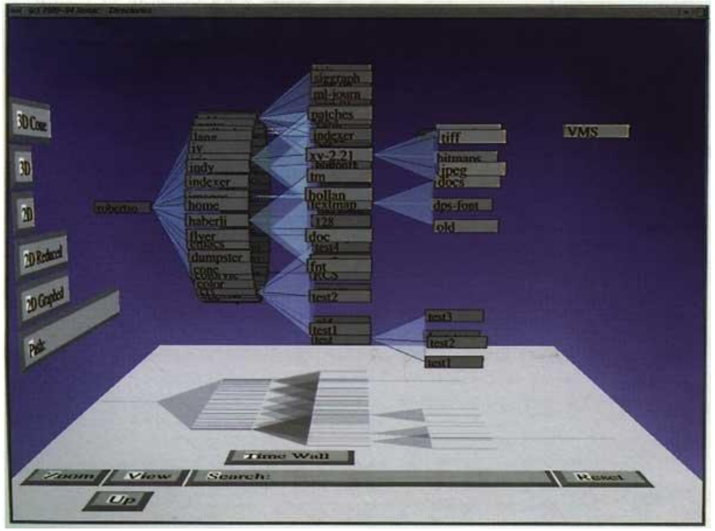


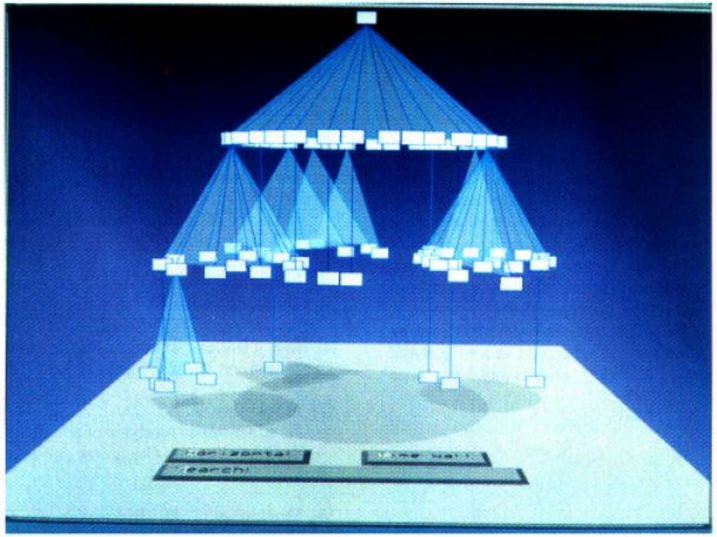
CamTree

[Robertson et al., 1991] George Robertson, Jock D. Mackinlay, Stuart Card. Cone Trees: Animated 3D Visualizations of Hierarchical Information. In *Proceedings of the ACM CHI 91*

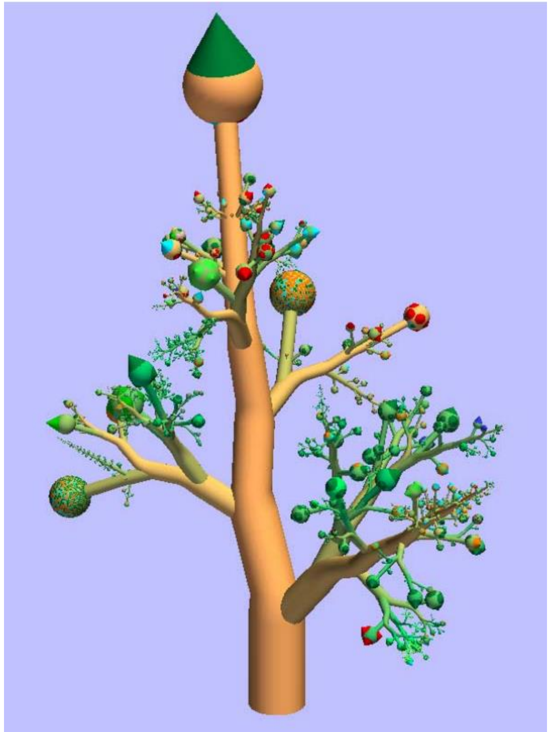
CONE/CAMTREE

- children of a node are laid out in a cylinder “below” the parent
- siblings located on the same 2D circle
- use of animation
- shadows to enhance structure





BOTANICAL VISUALIZATION OF HUGE HIERARCHIES



InfoVis 2001

3D LAYOUTS

- advantages
  - fit more data into same aspect ratio
  - aesthetically pleasing
- disadvantages
  - occlusion
  - requires interaction or animation
  - no overviews

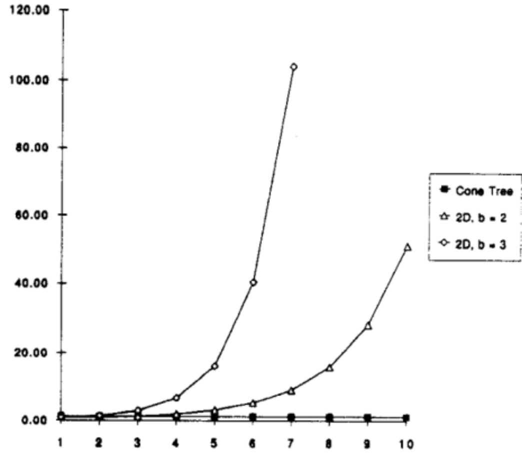


Figure 1: Aspect Ratio of 2D and 3D Trees.

TREE VISUALIZATION SUMMARY

- there are lots of tree visualizations
  - there is also lots of free software, try it out, (see links earlier in the lecture plus e.g.):
    - http://www.informatik.uni-rostock.de/~hs162/optreedemo/TestBed.htm#
    - http://w3.win.tue.nl/nl/onderzoek/onderzoek_informatica/visualization/sequoiaview/
  - there are a few overview articles, e.g.:
    - **A Generative Layout Approach for Rooted Tree Drawings** by Hans-Jörg Schulz, Zabed Akbar, and Frank Maurer - at IEEE PacificVis 2013
    - **The Design Space of Implicit Hierarchy Visualization: A Survey** by Hans-Jörg Schulz, Steffen Hadlak, and Heidrun Schumann - in IEEE TVCG 17(4)

TREE VISUALIZATIONS

- can be categorized by
  - edge representations (implicit, explicit)
  - dimensionality of layout
  - radial vs. axis-parallel

- can be modified by
  - layout parameters
  - which marks are used
  - visual variables on marks (which meta-data is represented?)

GRAPHS / NETWORKS

DEFINITION GRAPH

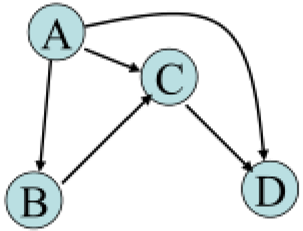
- A set of **vertices** V = $\{v_i\}$
- A set of **edges** E = $\{e_{ij}\}$ with $e_{ij} = \{v_i, v_j\}$
- When the order of vertices of an edge is meaningful, the graph is **directed**

GRAPH MEASURES

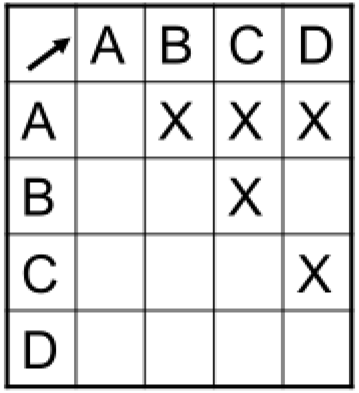
- SIZE = #nodes
- DENSITY = edges/vertices (roughly)
- PATH = sequence of edges connecting (different) vertices
- VERTEX DEGREE = #edge connections
- DISTANCE = #hops between vertices

TWO CLASSICAL VISUAL REPRESENTATIONS

Node-Link Diagram



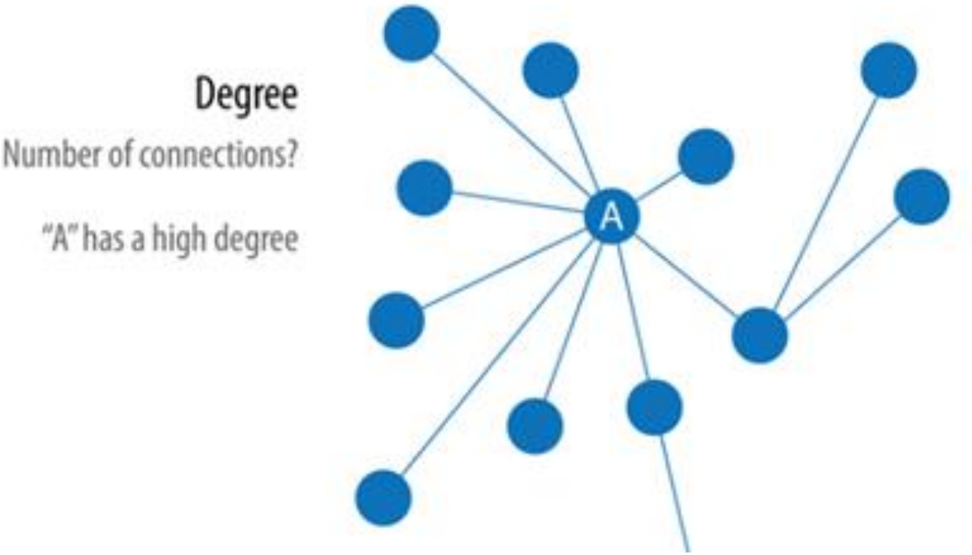
Adjacency Matrix



↗	A	B	C	D
A		X	X	X
B			X	
C				X
D				

TASKS

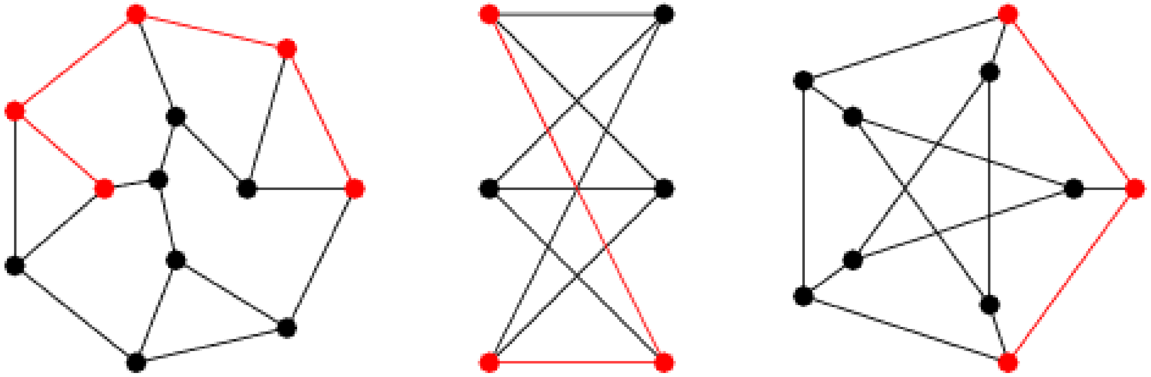
Find # of neighbors of a vertex



https://combattb.org/combat-tb-neodb/images/centalityalgos.png

TASKS

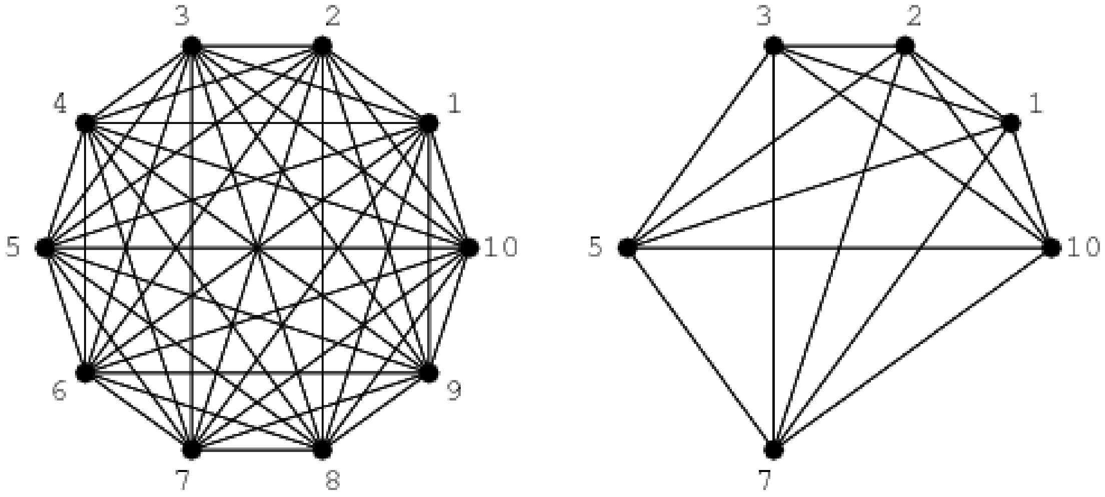
See paths (overviews, shortest, cycles)



http://mathworld.wolfram.com/GraphGeodesic.html

TASKS

Identify Sub-graphs



http://mathworld.wolfram.com/Vertex-InducedSubgraph.html

TASKS

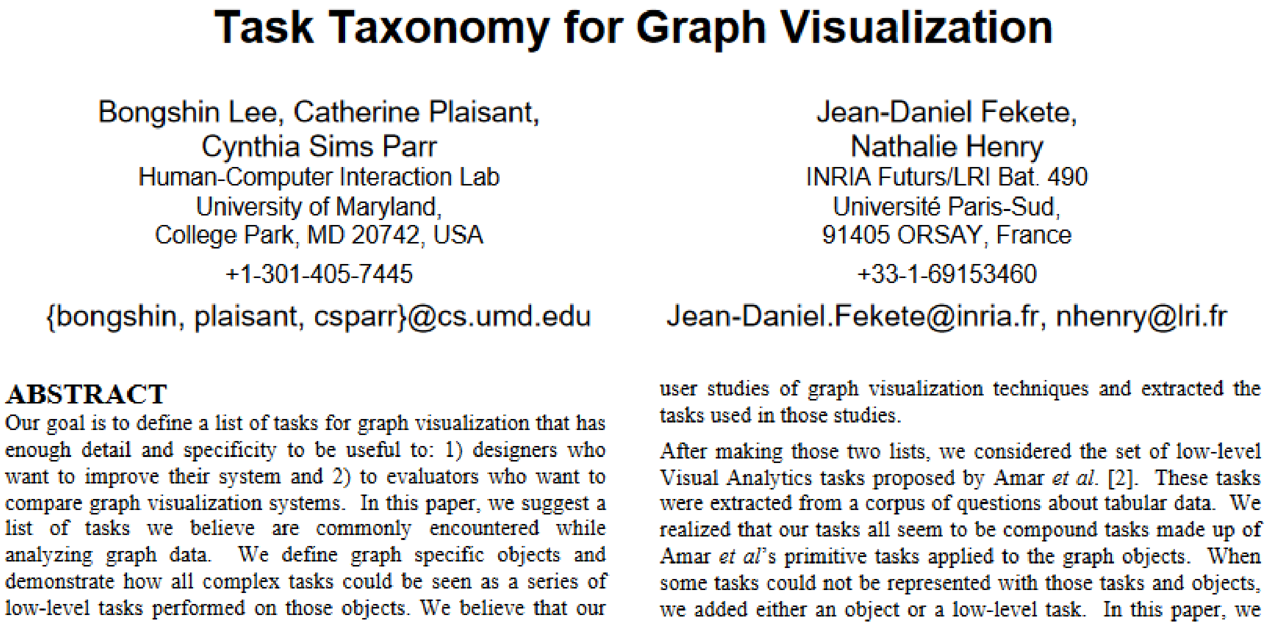
HIGHER-LEVEL

involves many elements
involves more human judgment

- which nodes are important?
- where are clusters?
- what are attribute and connection correlations?
- how does the network change over time?

TASKS

- Many many more specific tasks
- Each application domain will add more



Task Taxonomy for Graph Visualization

Bongshin Lee, Catherine Plaisant,
Cynthia Sims Parr
Human-Computer Interaction Lab
University of Maryland,
College Park, MD 20742, USA
+1-301-405-7445
{bongshin, plaisant, csparr}@cs.umd.edu

Jean-Daniel Fekete,
Nathalie Henry
INRIA Futurs/LRI Bat. 490
Université Paris-Sud,
91405 ORSAY, France
+33-1-69153460
Jean-Daniel.Fekete@inria.fr, nhenry@lri.fr

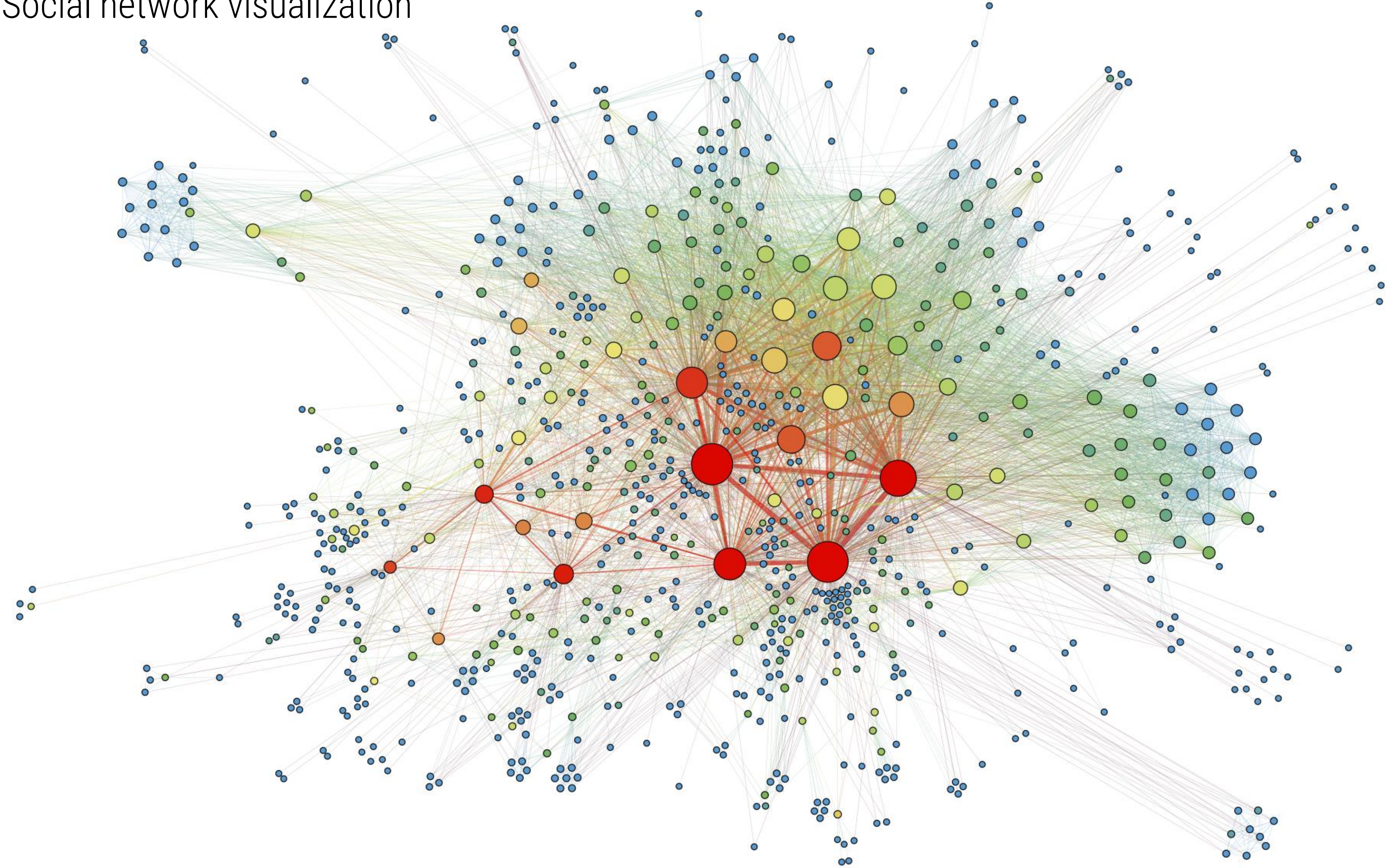
ABSTRACT

Our goal is to define a list of tasks for graph visualization that has enough detail and specificity to be useful to: 1) designers who want to improve their system and 2) to evaluators who want to compare graph visualization systems. In this paper, we suggest a list of tasks we believe are commonly encountered while analyzing graph data. We define graph specific objects and demonstrate how all complex tasks could be seen as a series of low-level tasks performed on those objects. We believe that our

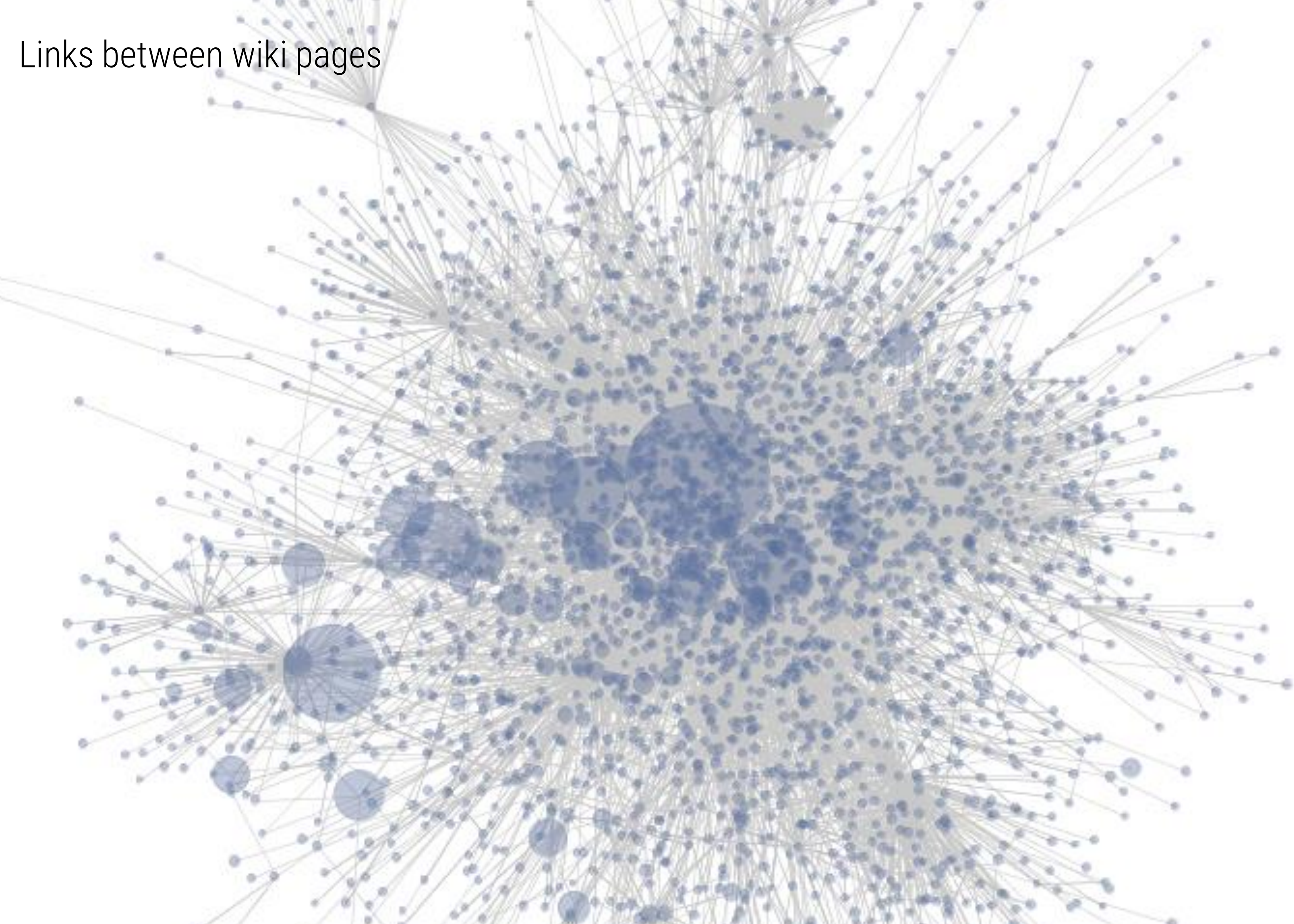
user studies of graph visualization techniques and extracted the tasks used in those studies.

After making those two lists, we considered the set of low-level Visual Analytics tasks proposed by Amar *et al*. [2]. These tasks were extracted from a corpus of questions about tabular data. We realized that our tasks all seem to be compound tasks made up of Amar *et al*'s primitive tasks applied to the graph objects. When some tasks could not be represented with those tasks and objects, we added either an object or a low-level task. In this paper, we

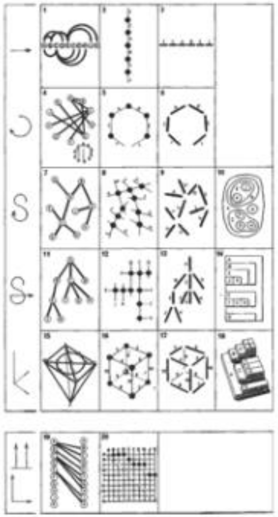
Social network visualization



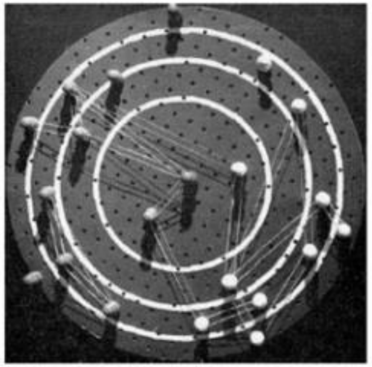
Links between wiki pages



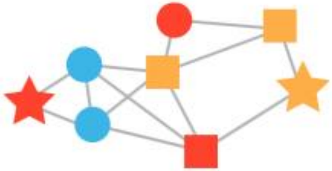
GRAPH VISUALIZATION CHALLENGES



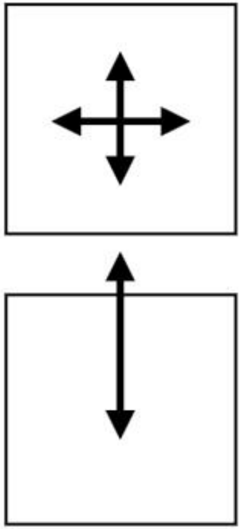
Representation



Layout



Types + Attributes



Navigation

SOCIAL NETWORK ANALYSIS

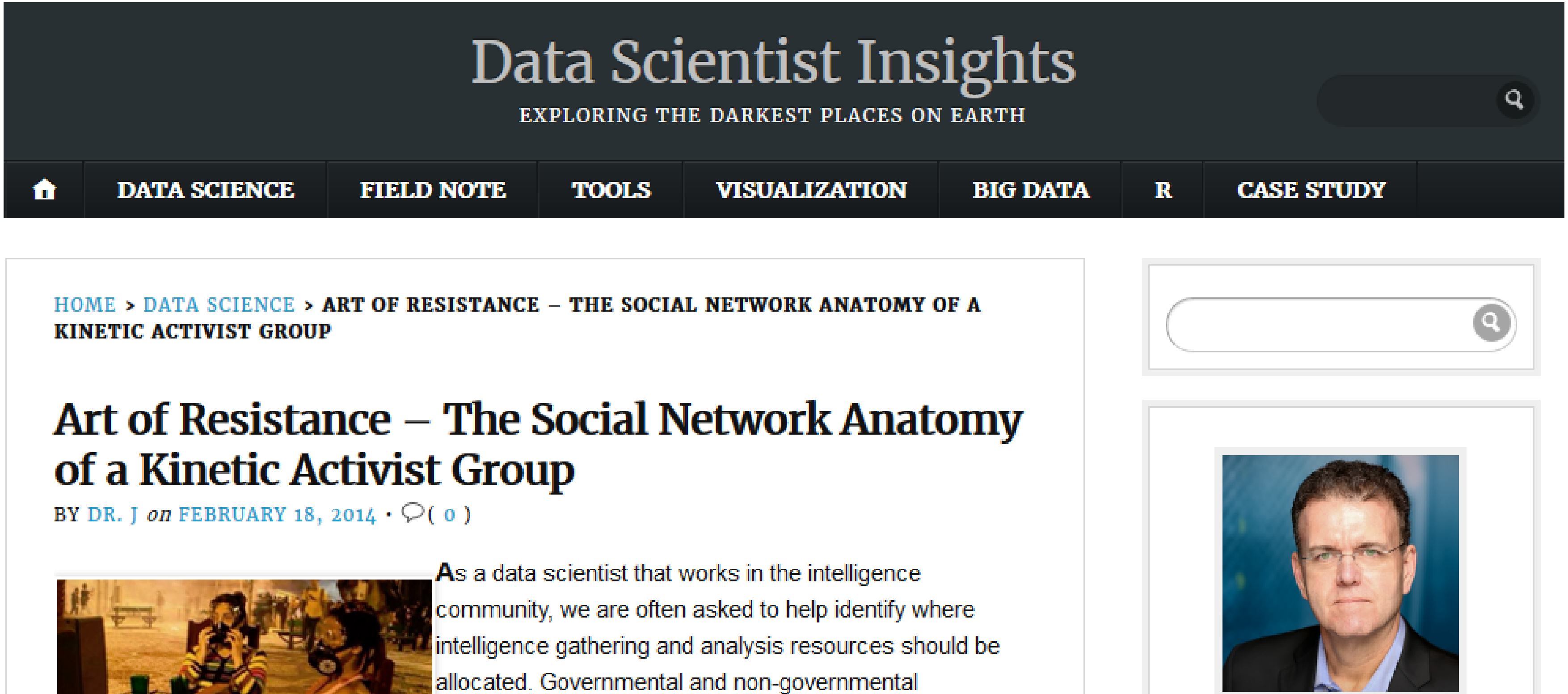
https://datascientistinsights.com/2014/02/18/art-of-resistance-the-social-network-anatomy-of-a-kinetic-activist-group/

- determine if Greenpeace was or could become a significant disruptive geopolitical force

- first: identify who/what to concentrate resources on

SOCIAL NETWORK ANALYSIS

https://datascientistinsights.com/2014/02/18/art-of-resistance-the-social-network-anatomy-of-a-kinetic-activist-group/



Data Scientist Insights
EXPLORING THE DARKEST PLACES ON EARTH
DATA SCIENCE
FIELD NOTE
TOOLS
VISUALIZATION
BIG DATA
R
CASE STUDY
HOME > DATA SCIENCE > ART OF RESISTANCE – THE SOCIAL NETWORK ANATOMY OF A KINETIC ACTIVIST GROUP

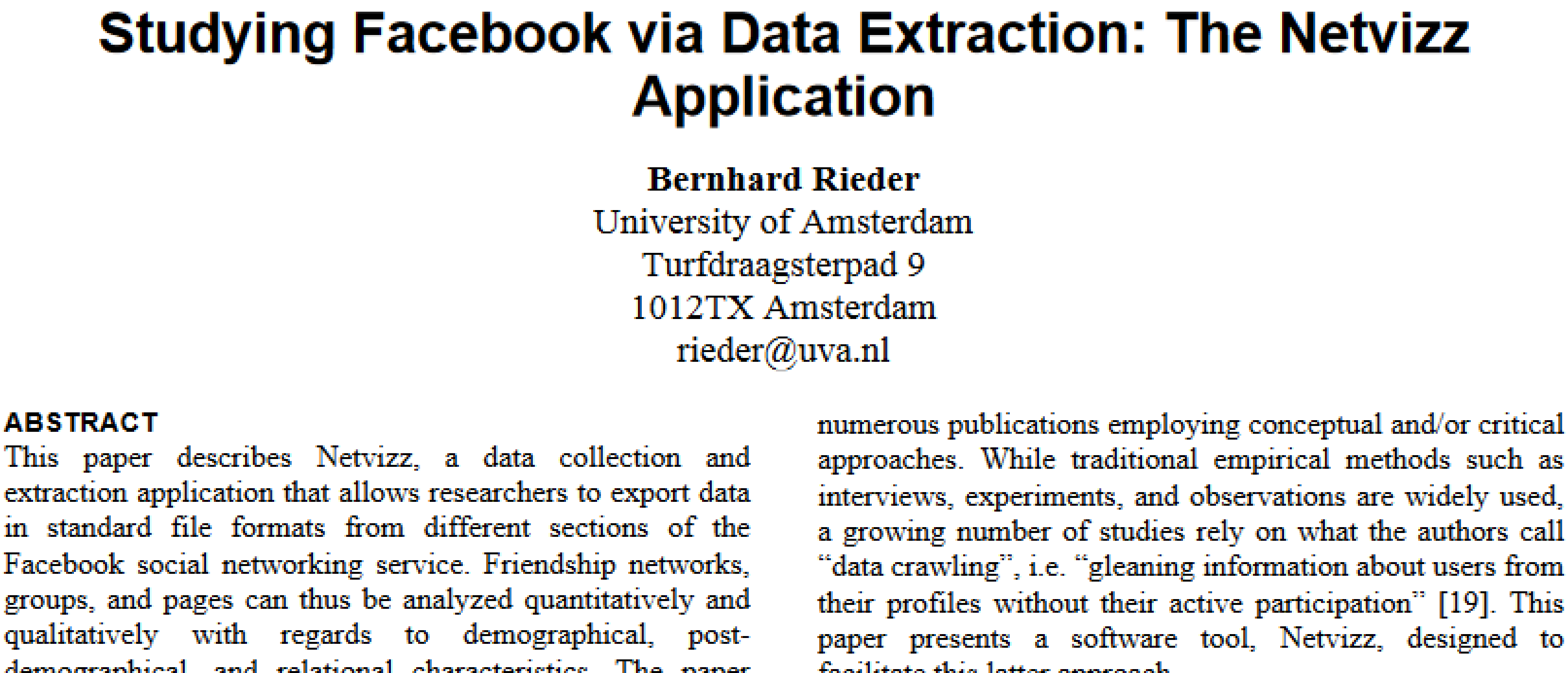
Art of Resistance – The Social Network Anatomy of a Kinetic Activist Group

BY DR. J *on* FEBRUARY 18, 2014 • (0)

As a data scientist that works in the intelligence community, we are often asked to help identify where intelligence gathering and analysis resources should be allocated. Governmental and non-governmental

SOCIAL NETWORK ANALYSIS

1) get Facebook data using Netvizz



Studying Facebook via Data Extraction: The Netvizz Application

Bernhard Rieder
University of Amsterdam
Turfdraagsterpad 9
1012TX Amsterdam
rieder@uva.nl

ABSTRACT

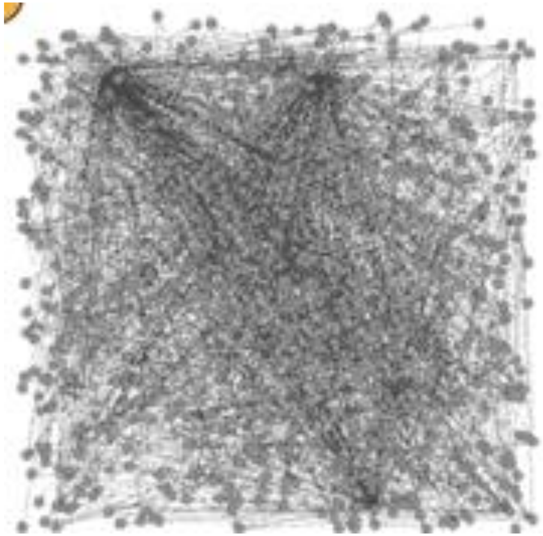
This paper describes Netvizz, a data collection and extraction application that allows researchers to export data in standard file formats from different sections of the Facebook social networking service. Friendship networks, groups, and pages can thus be analyzed quantitatively and qualitatively with regards to demographical, post-demographical, and relational characteristics. The paper

numerous publications employing conceptual and/or critical approaches. While traditional empirical methods such as interviews, experiments, and observations are widely used, a growing number of studies rely on what the authors call "data crawling", i.e. "gleaning information about users from their profiles without their active participation" [19]. This paper presents a software tool, Netvizz, designed to facilitate this latter approach.

SOCIAL NETWORK ANALYSIS

2) load the data into Gephi

https://gephi.org/

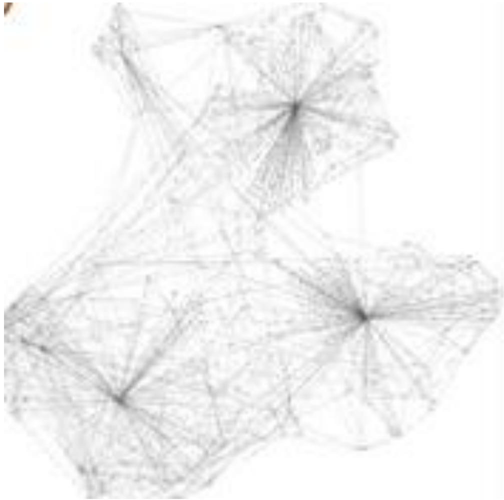


585 nodes, interconnected by 1788 edges.
“Somewhere in that spaghetti is a potential bad guy, but where?”

SOCIAL NETWORK ANALYSIS

3) choose a layout algorithm that makes sense for social networks

Force Atlas 2

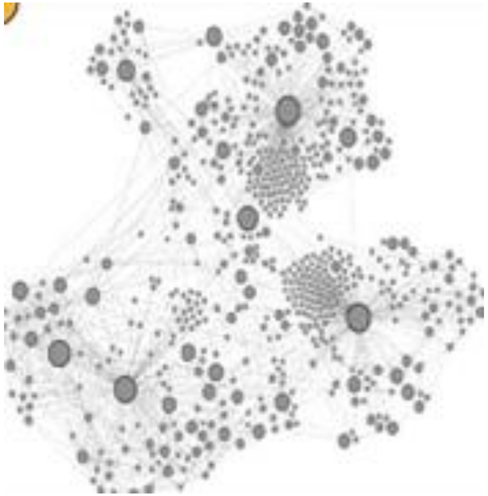


provides some transparency into the network but still lacks any real clarity around behavioral importance

SOCIAL NETWORK ANALYSIS

4) map an attribute to size of the nodes

betweenness centrality (number of shortest paths from all vertices to all others that pass through that node)



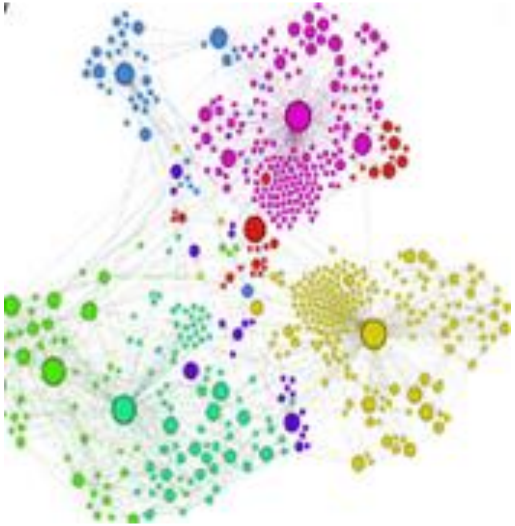
Bigger nodes are more central to behavioral dynamics.

Several nodes become central figures in the overall network.

SOCIAL NETWORK ANALYSIS

5) highlight communities

color nodes by modularity / clusters

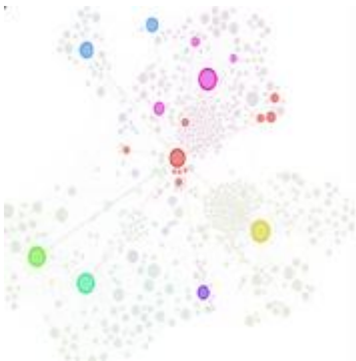


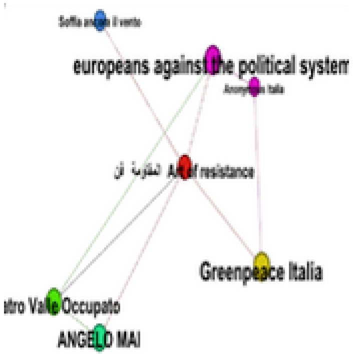
We now begin to see a clearer picture of who is doing what with whom.

What becomes really interesting at this stage is understanding some of the more nuanced relationships.

SOCIAL NETWORK ANALYSIS

6) filter, explore, label





Five outlying nodes in the network (blue, maroon, yellow, dark green, and light green).

Center: an equally important red node

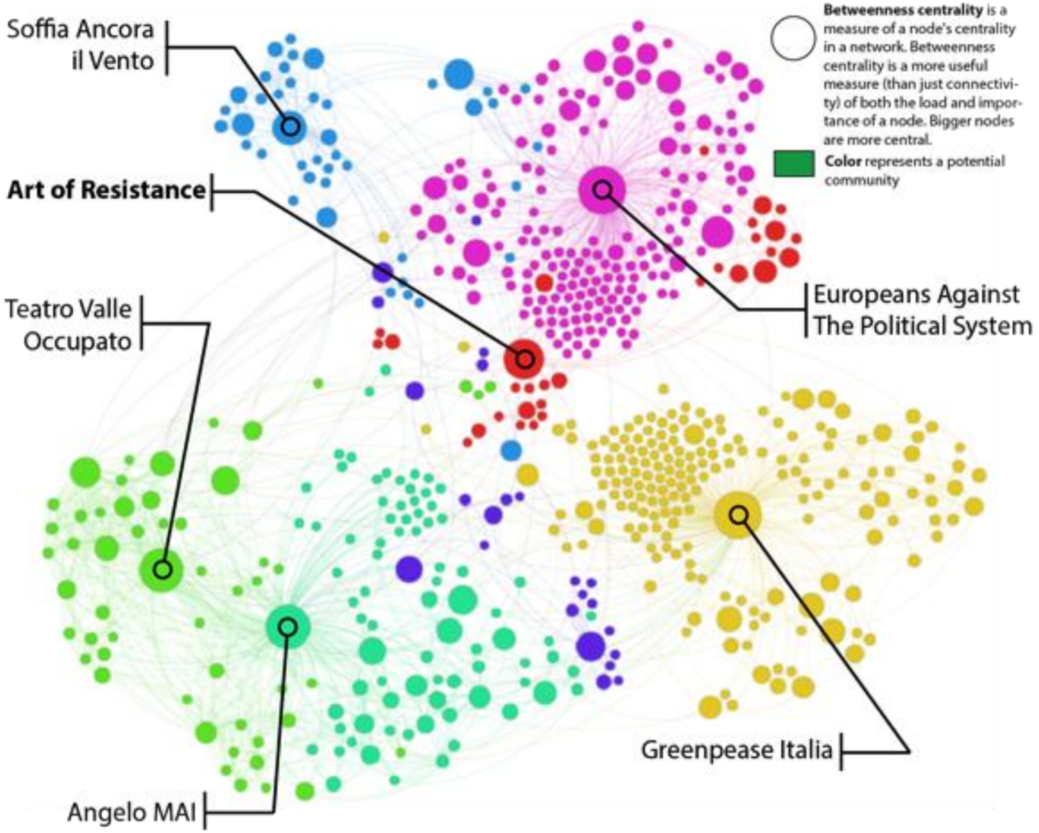
Emergence of a previously un-recognized activism player: [Art of Resistance](#).

فن المقاومة Art of resistance

63,663 likes · 13,905 talking about this

SOCIAL NETWORK ANALYSIS

7) communicate & explain

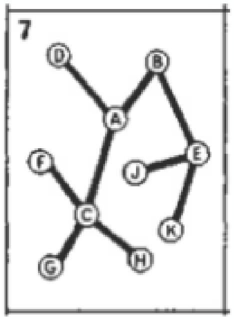

Soffia Ancora il Vento
Art of Resistance
Teatro Valle Occupato
Angelo MAI
Europeans Against The Political System
Greenpease Italia
Betweenness centrality is a measure of a node's centrality in a network. Betweenness centrality is a more useful measure (than just connectivity) of both the load and importance of a node. Bigger nodes are more central.
Color represents a potential community

LAYOUTS

Important to the success of your analysis

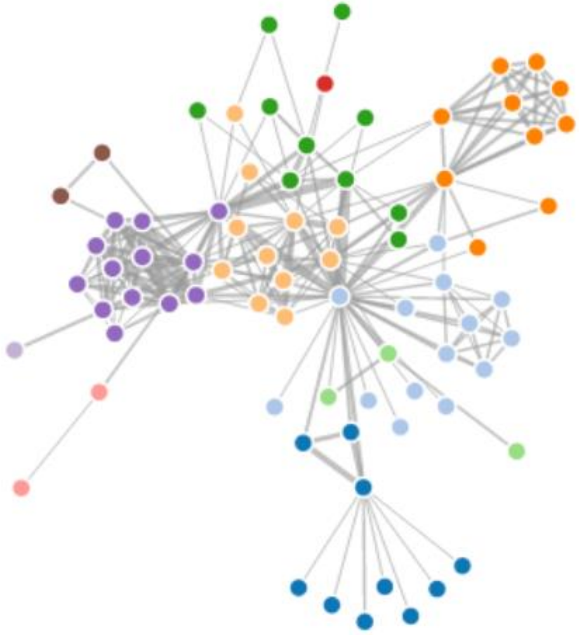
LAYOUT FREE

Force Directed



- Physical forces
- Proximity based

- Spring Model
- Kamanda&Kawai
- Frucherman&Reingold
- Davidson&Harel
- LinLog



[redacted]

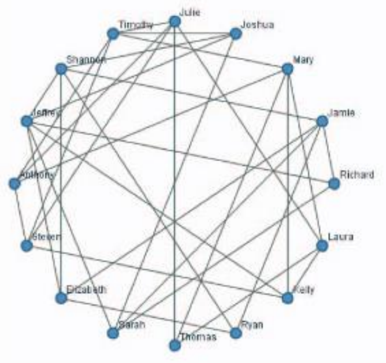
LAYOUT FREE

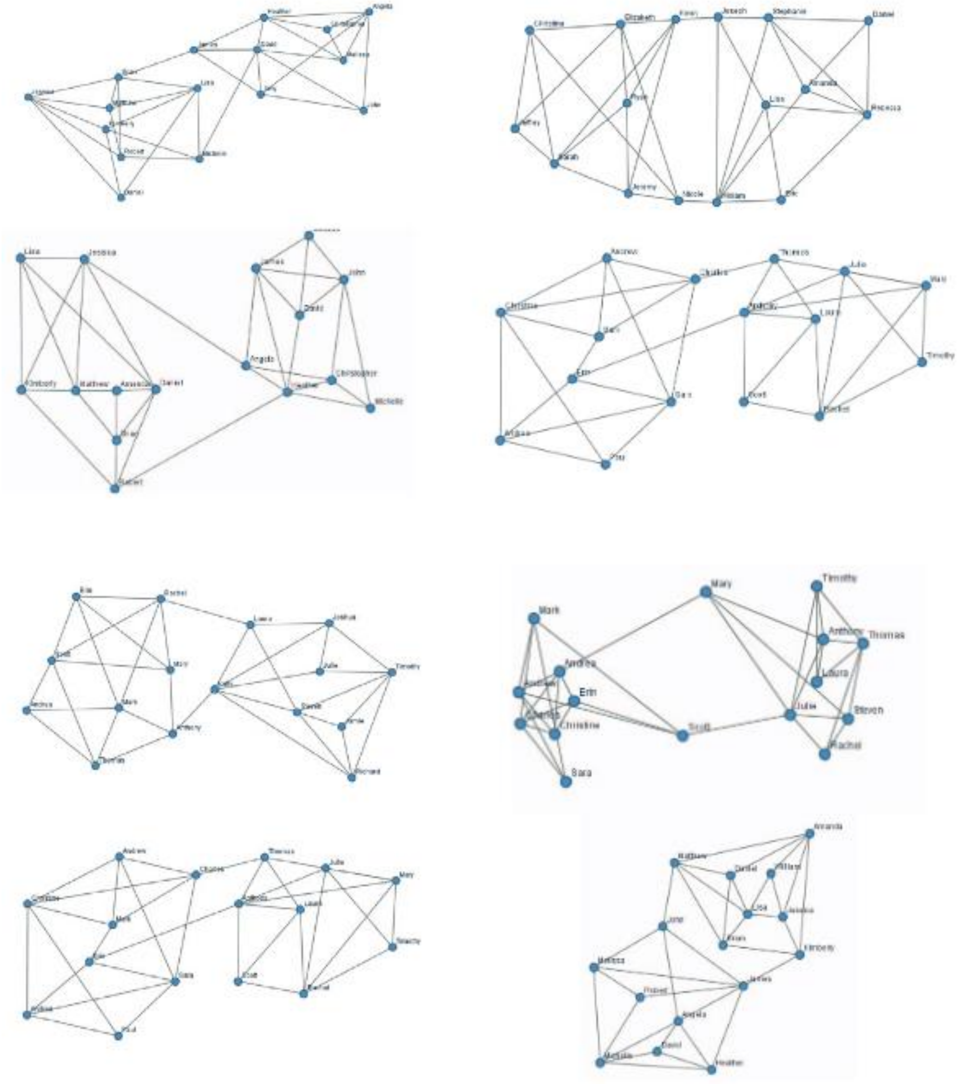
Aesthetic Criteria

- Reduce number of edge crossing
- Foster Symmetry
- Uniform edge length
- Aspect Ratio
- Equal Angles
- ...

LAYOUT FREE

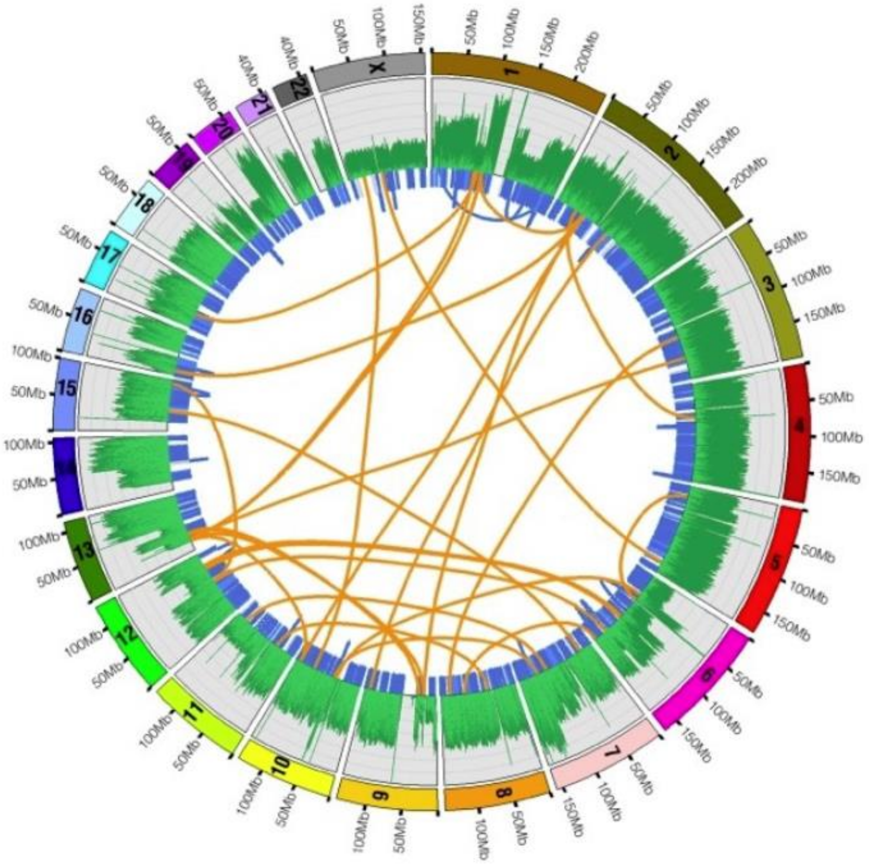
HAND MADE

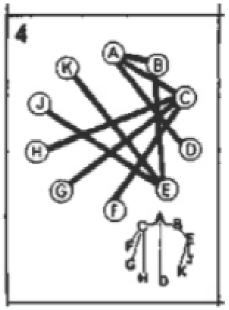




eptual organization in user-generated graph layouts
lam, F.J.J.; Rogowitz, B.

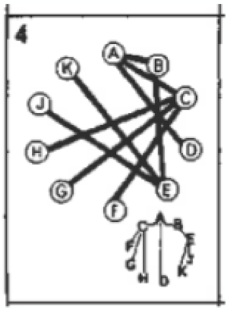
LAYOUT CIRCULAR

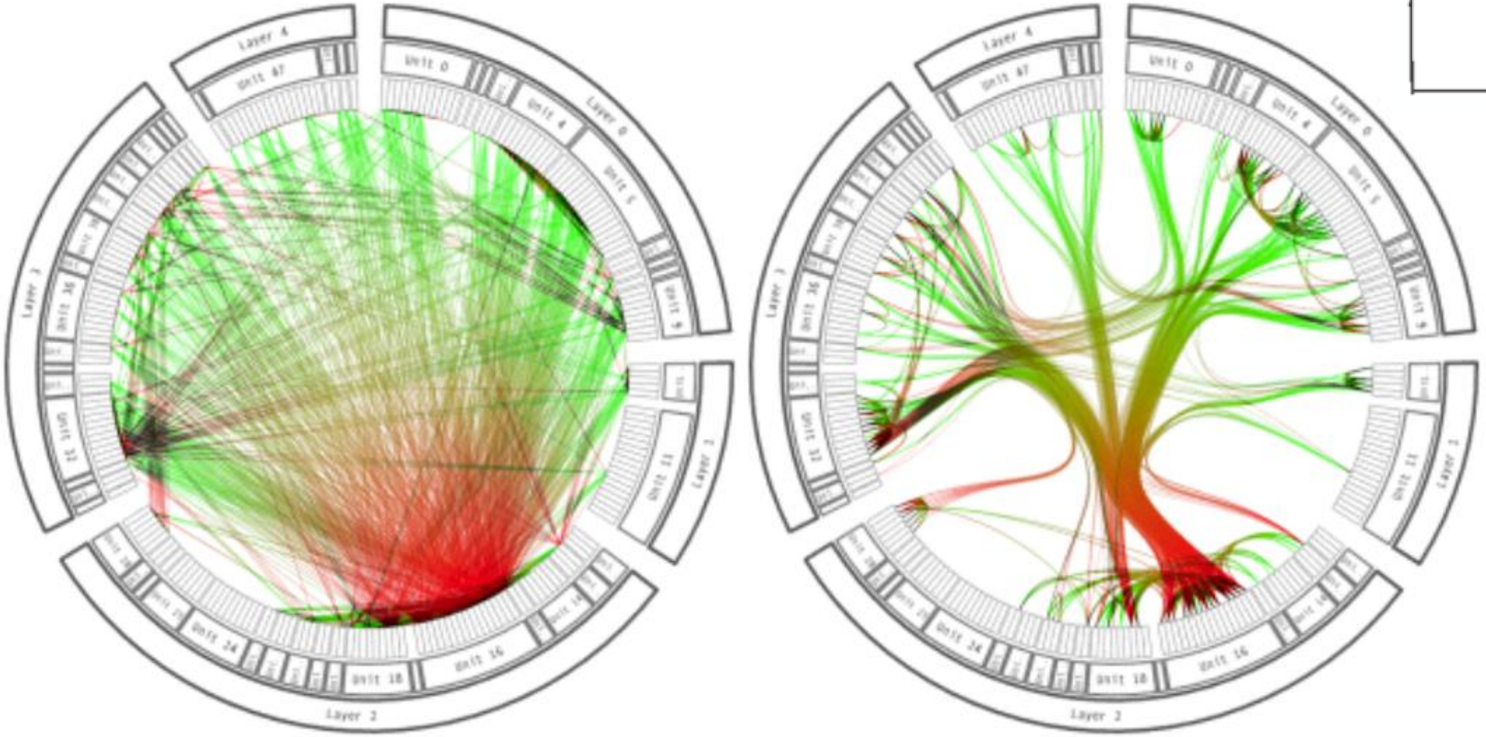




- Edges on the inside
- Vertices & attributes on the outside
- Ordering possible

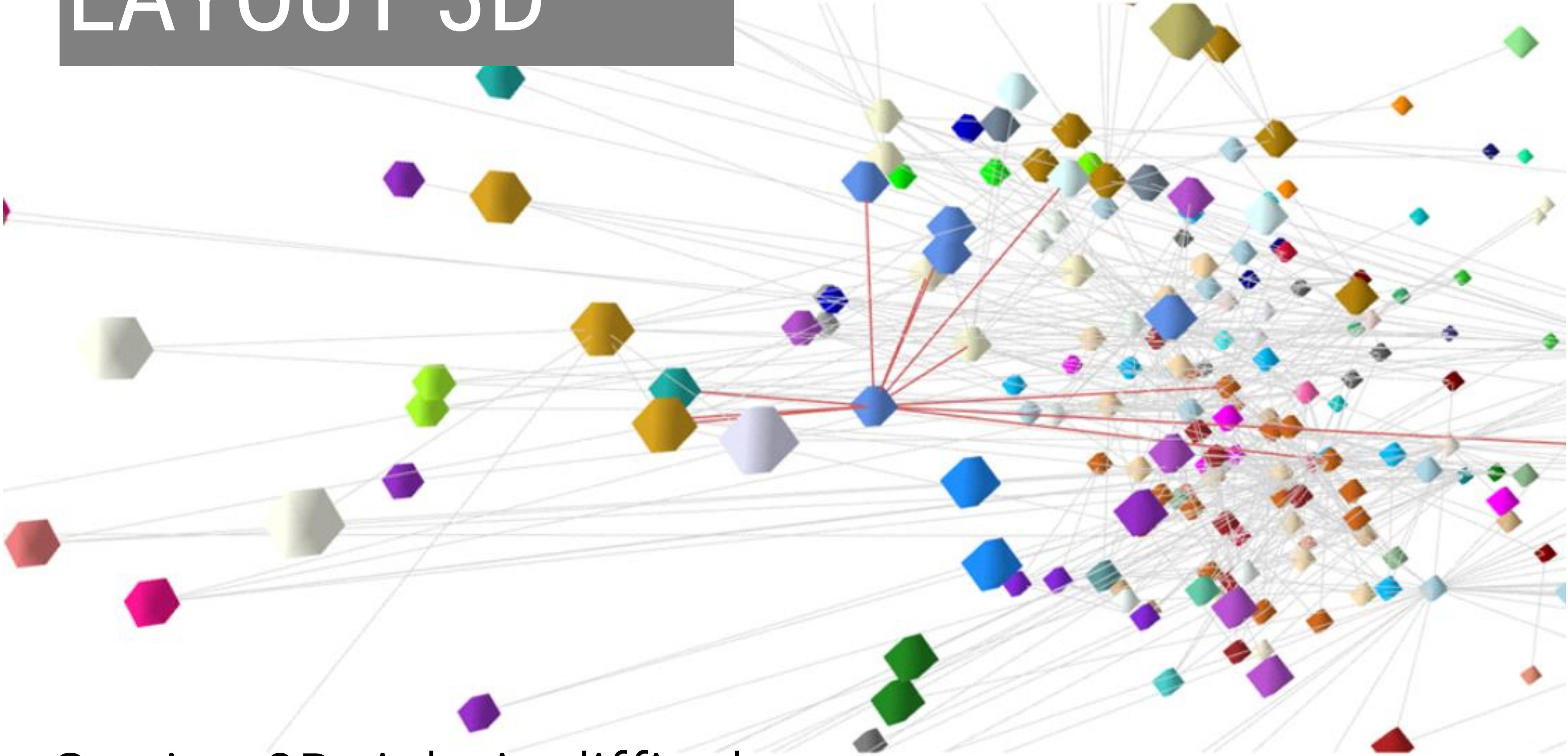
LAYOUT CIRCULAR





Edge Bundling
Holten 2006

LAYOUT 3D

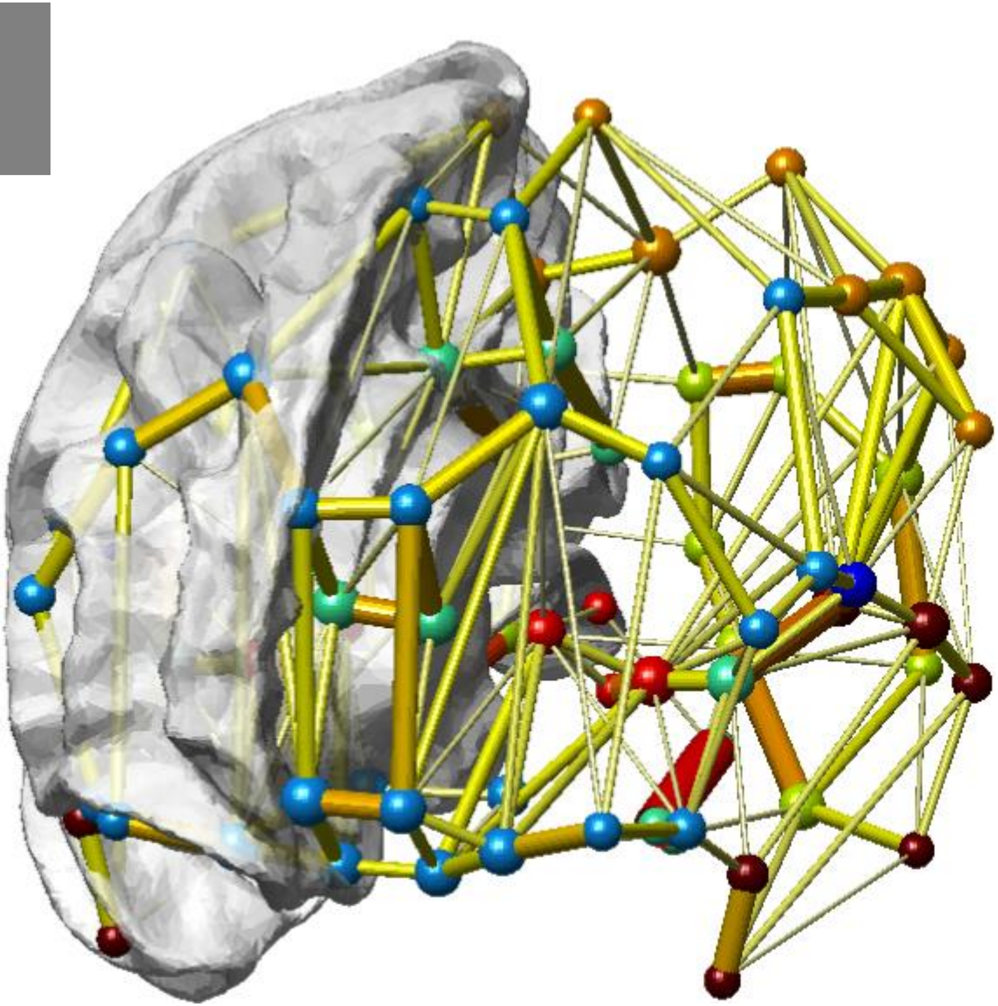


Getting 3D right is difficult

https://fwaris.wordpress.com/2012/07/08/a-simple-technique-for-creating-3d-graphs-from-2d-ones/

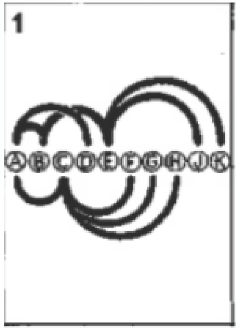
LAYOUT 3D

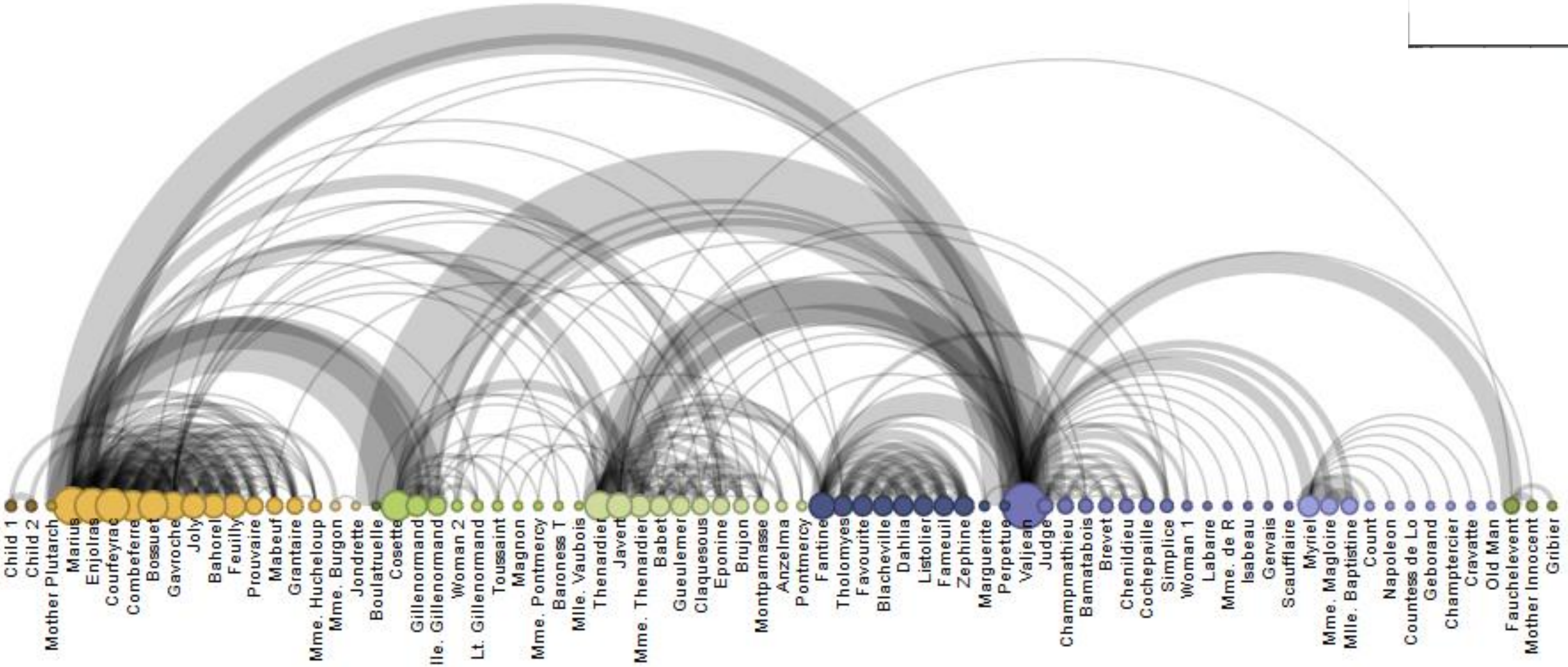
Sometimes necessary (!?)



http://gephi.forumatic.com/viewtopic.php?f=26&t=2616

LAYOUT LINEAR

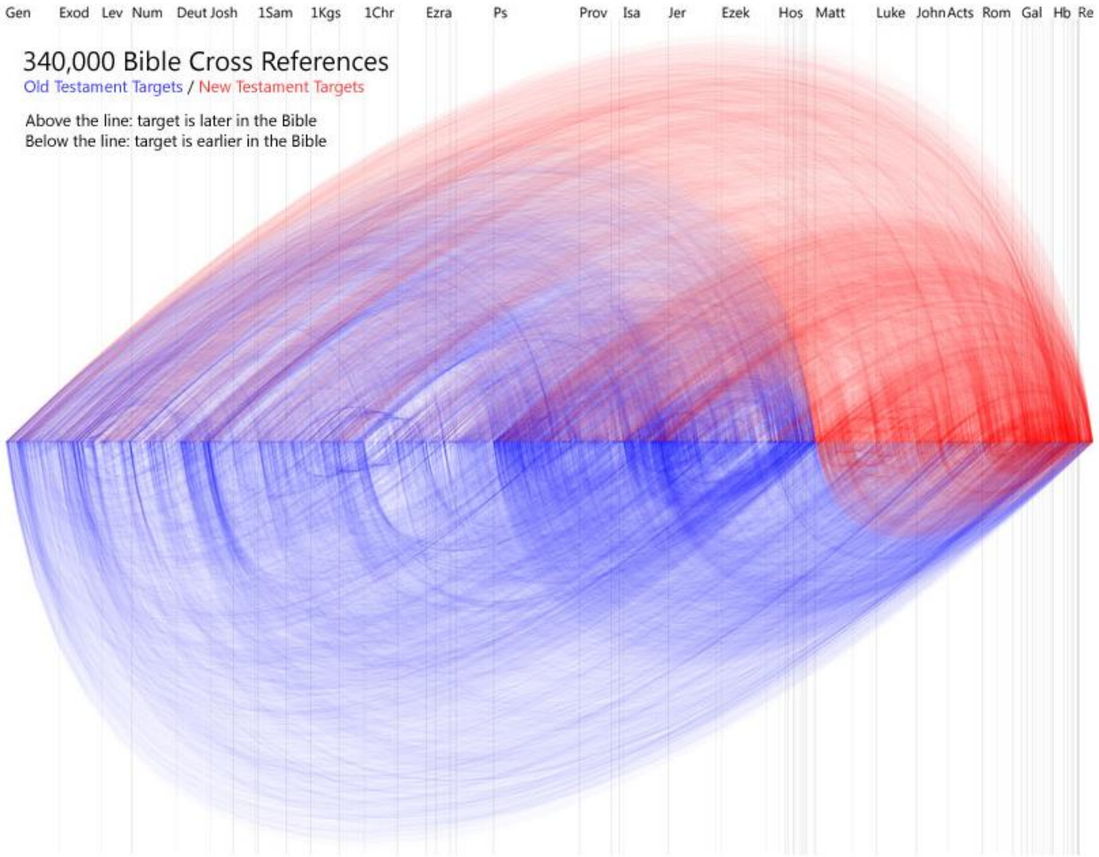




http://mbostock.github.io/protovis/ex/arc-full.html

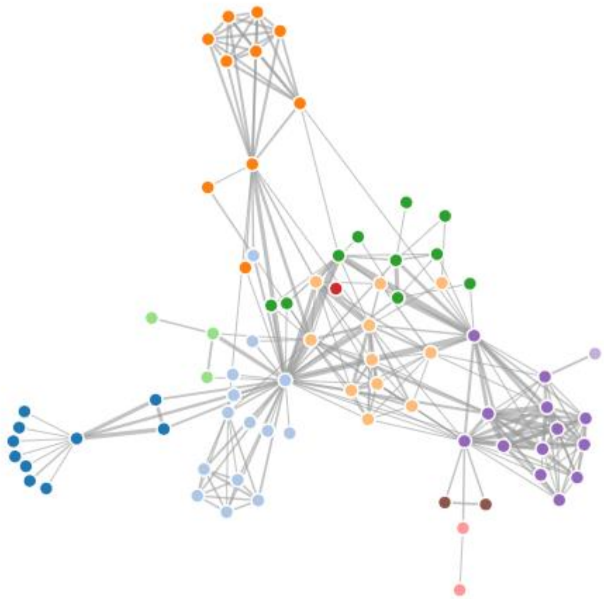
http://www.bewitched.com/song.html

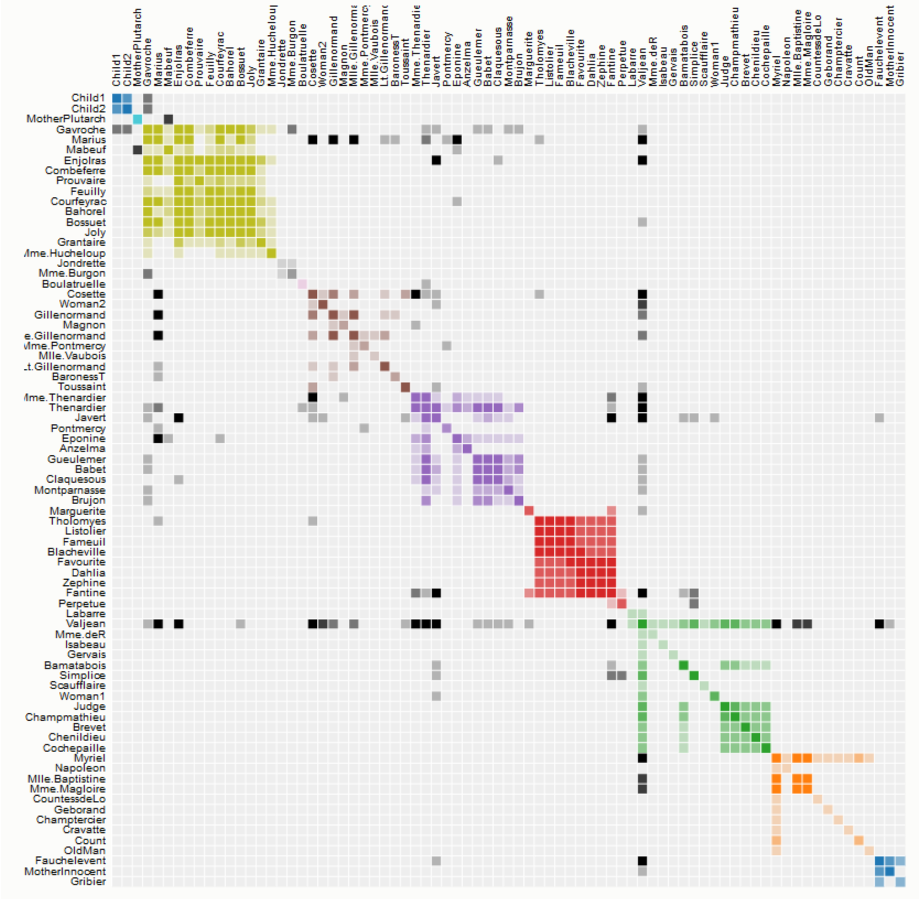
LAYOUT LINEAR



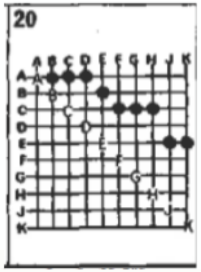
https://www.openbible.info/labs/cross-references/

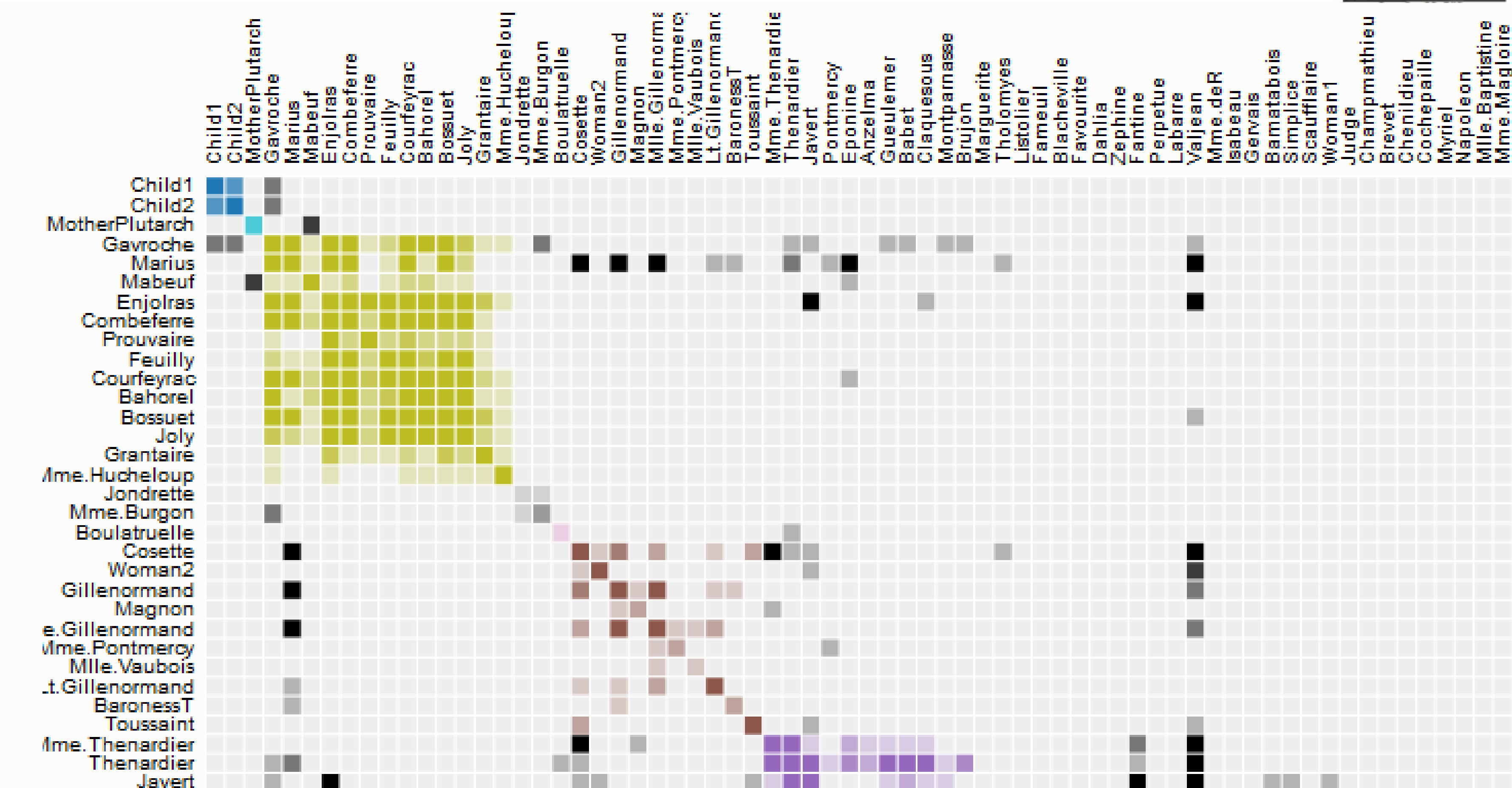
LAYOUT ADJACENCY MATRIX





ADJACENCY MATRIX





PROS/CONS

Matrix

- No vertex/edge overlap or crossings
- Readable for dense graph
- Fast navigation

- Less familiar
- Space intensive
- Weak for path following tasks

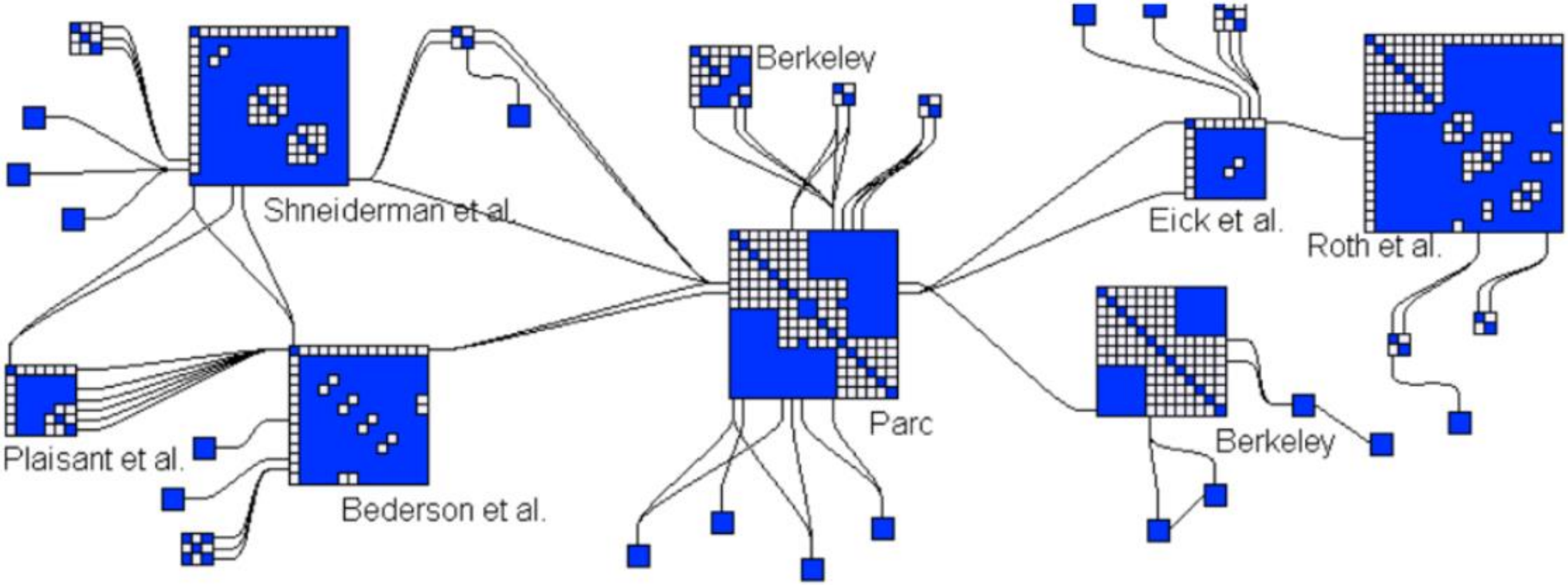
Node-Link

- Familiar
- Compact
- Path following easier
- Effective for small and sparse graphs

- Useless without layout
- Not readable for dense graphs
- Manipulation requires layout computation

HYBRID

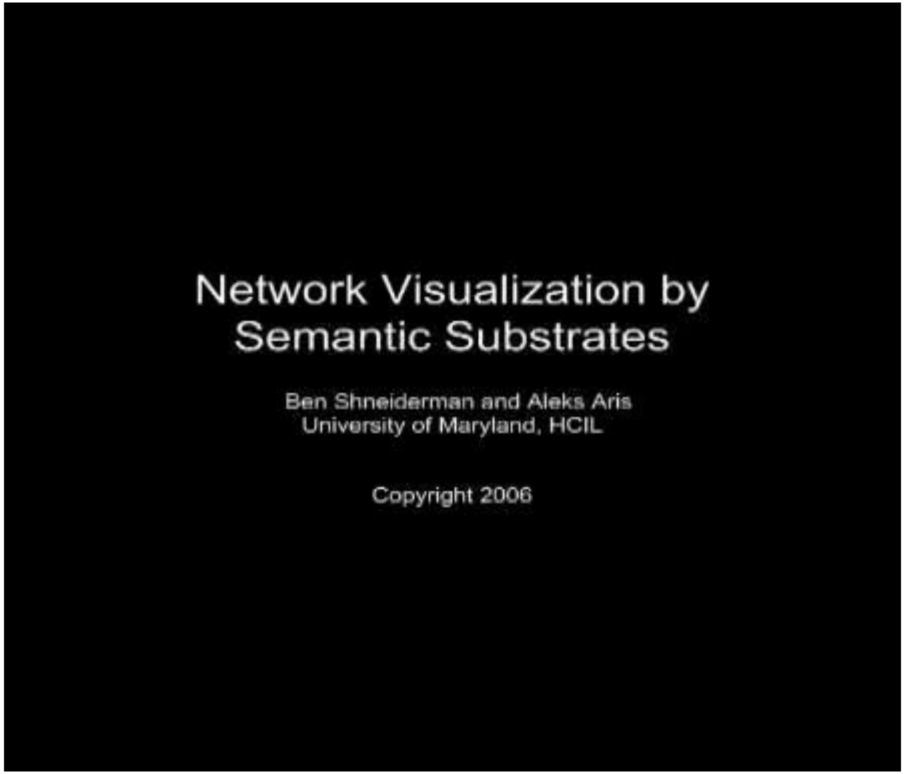
Henry et al., NodeTrix



Infovis Coauthorship (133 actors)

Dense = matrices, sparse = node-link

MULTIVARIATE NETWORKS



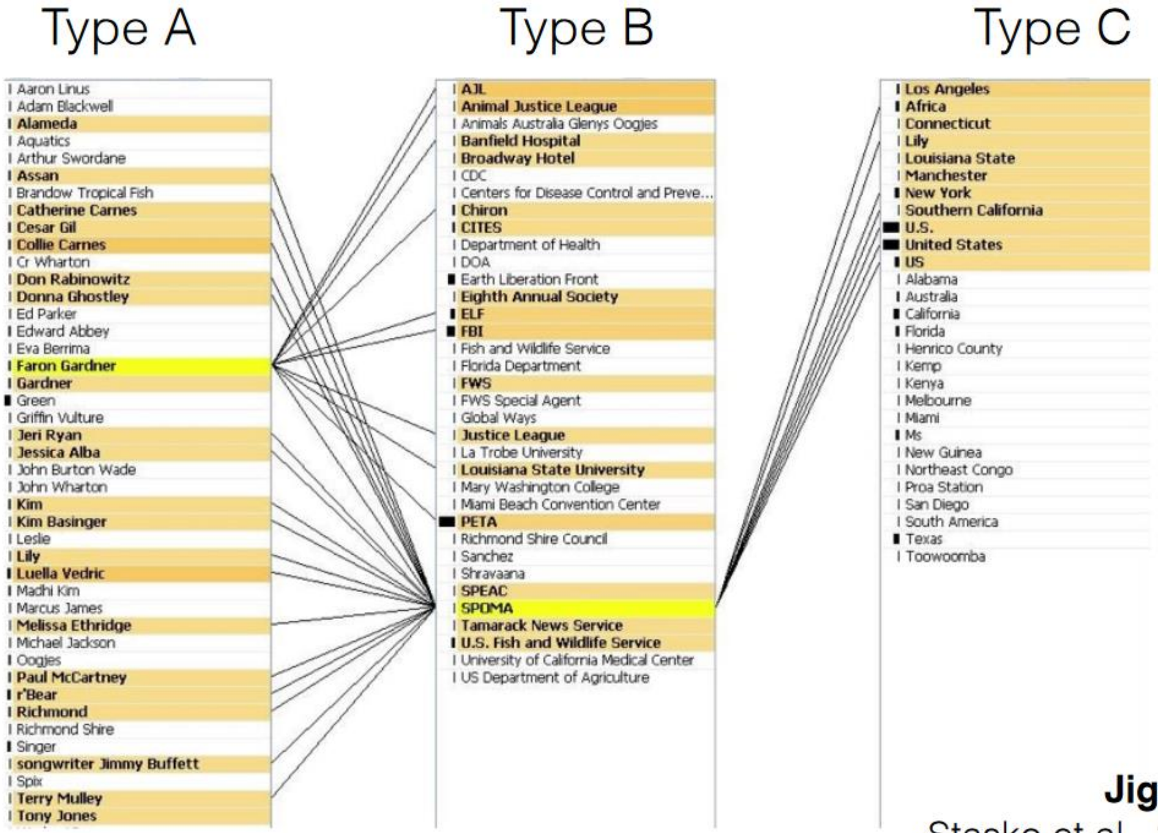
Network Visualization by
Semantic Substrates

Ben Shneiderman and Aleks Aris
University of Maryland, HCIL

Copyright 2006

https://www.youtube.com/watch?v=f3hmn7gvocQ

MULTIVARIATE NETWORKS



Type A

Aaron Linus
Adam Blackwell
Alameda
Aquatics
Arthur Swordane
Assan
Brandow Tropical Fish
Catherine Carnes
Cesar Gil
Collie Carnes
Cr Wharton
Don Rabinowitz
Donna Ghostley
Ed Parker
Edward Abbey
Eva Berrima
Faron Gardner
Gardner
Green
Griffin Vulture
Jeri Ryan
Jessica Alba
John Burton Wade
John Wharton
Kim
Kim Basinger
Leslie
Lily
Luella Vedric
Madhi Kim
Marcus James
Melissa Ethridge
Michael Jackson
Oogjes
Paul McCartney
r'Bear
Richmond
Richmond Shire
Singer
songwriter Jimmy Buffett
Spix
Terry Mulley
Tony Jones

Type B

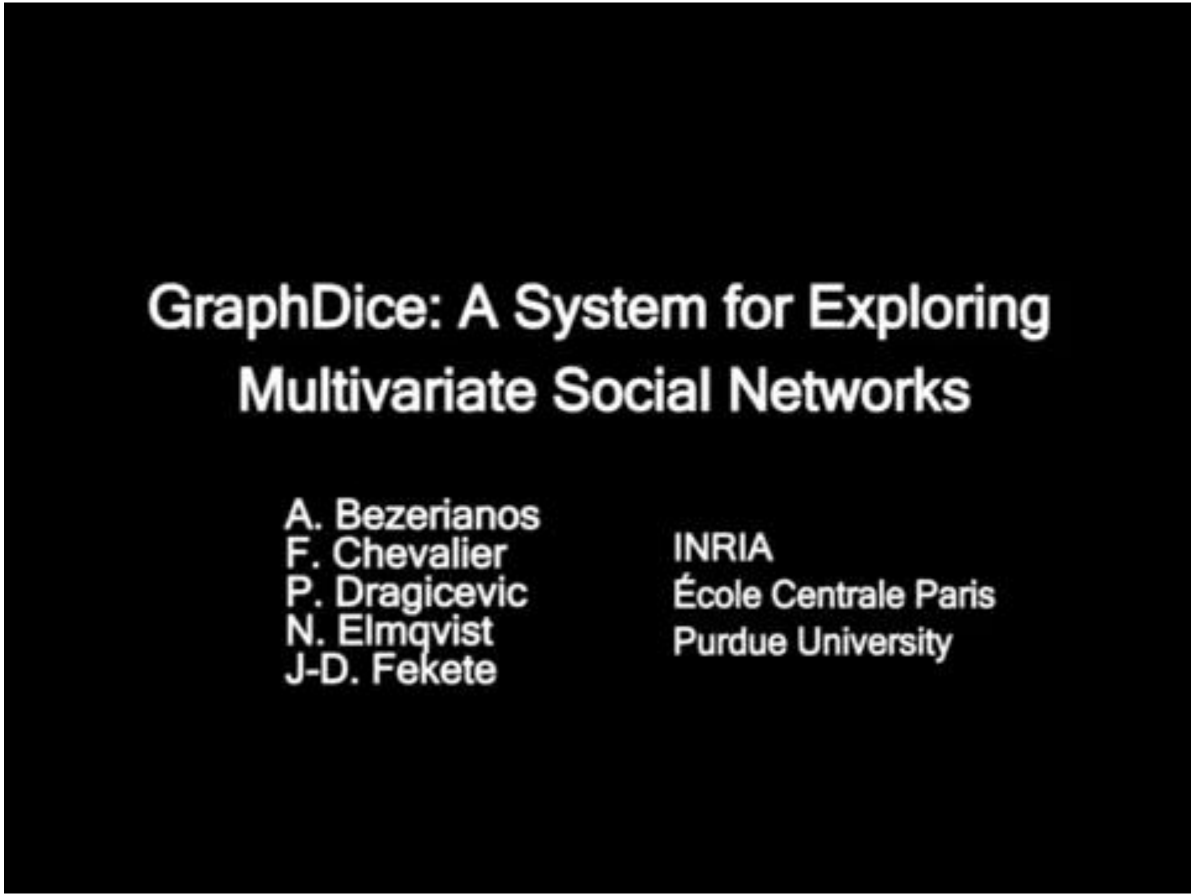
AJL
Animal Justice League
Animals Australia Glenys Oogjes
Banfield Hospital
Broadway Hotel
CDC
Centers for Disease Control and Preve...
Chiron
CITES
Department of Health
DOA
Earth Liberation Front
Eighth Annual Society
ELF
FBI
Fish and Wildlife Service
Florida Department
FWS
FWS Special Agent
Global Ways
Justice League
La Trobe University
Louisiana State University
Mary Washington College
Miami Beach Convention Center
PETA
Richmond Shire Council
Sanchez
Shravaana
SPEAC
SPOMA
Tamarack News Service
U.S. Fish and Wildlife Service
University of California Medical Center
US Department of Agriculture

Type C

Los Angeles
Africa
Connecticut
Lily
Louisiana State
Manchester
New York
Southern California
U.S.
United States
US
Alabama
Australia
California
Florida
Henrico County
Kemp
Kenya
Melbourne
Miami
Ms
New Guinea
Northeast Congo
Proa Station
San Diego
South America
Texas
Toowoomba

Jigsaw,
Stasko et al., 2008

MULTIVARIATE NETWORKS



GraphDice: A System for Exploring Multivariate Social Networks

A. Bezerianos
F. Chevalier
P. Dragicevic
N. Elmqvist
J-D. Fekete

INRIA
École Centrale Paris
Purdue University

https://youtu.be/NLiHw5B0Mco

ADDITIONAL CHALLENGES

- TIME
- INTERACTION
- EDGE DIRECTION

NETWORK TOOLS

- **Gephi**
- **Cytoscape**
- **Pajek**
- **Java** Jung toolkit
- **D3** + Cola.js

ACKNOWLEDGEMENTS

Slides in were inspired and adapted from slides by

- Sheelagh Carpendale (University of Calgary)
- Pat Hanrahan (Stanford University)
- Benjamin Bach (University of Edinburgh)
- Jean-Daniel Fekete (Inria)