## VISUALIZING TREES AND GRAPHS

#### Petra Isenberg



### RECAP

you have learned about

- simple plots
- multi-attribute data visualization

## DATA AND ITS STRUCTURE

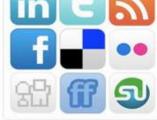
#### STRUCTURED DATA



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

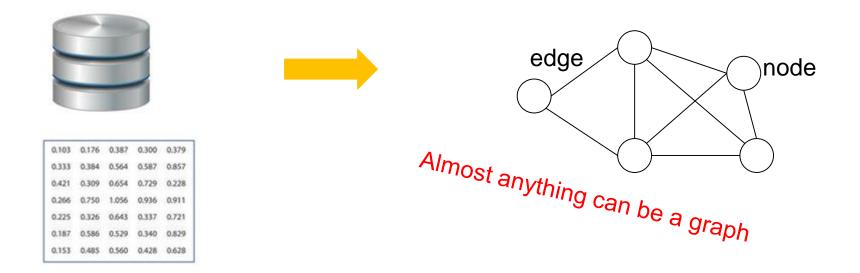
#### UNSTRUCTURED DATA



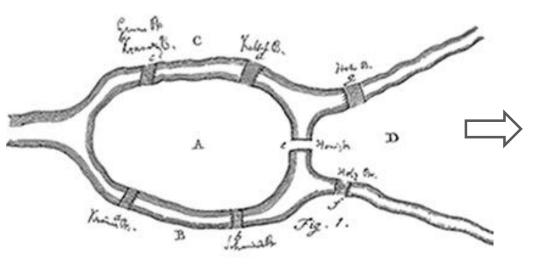


### STRUCTURED DATA

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



#### The 7 bridges of Königsberg



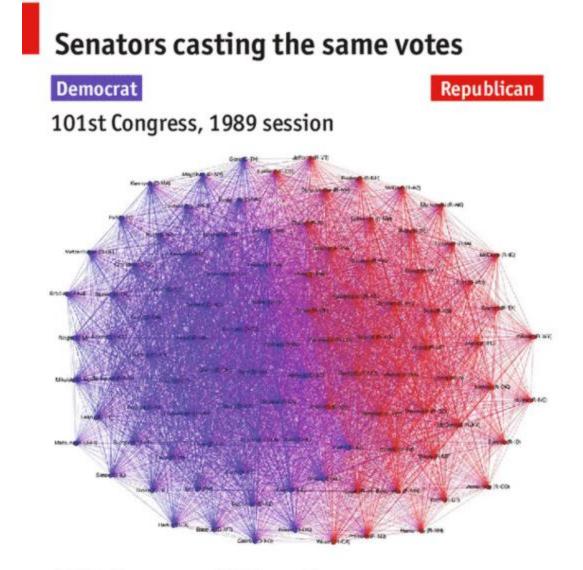
How can you cross all 7 bridges without crossing the same one twice?

A e D a b f B Euler's concentualization of the

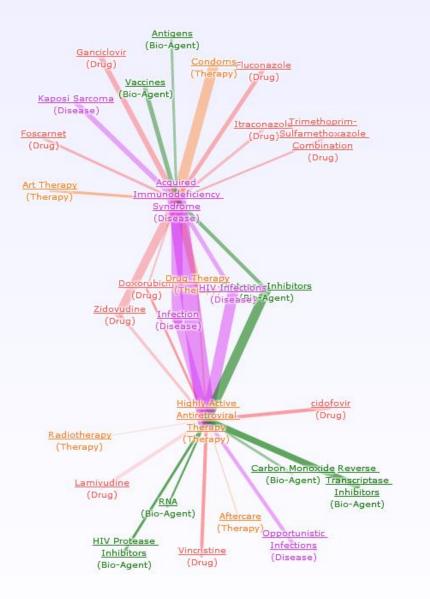
d

g

Euler's conceptualization of the same problem



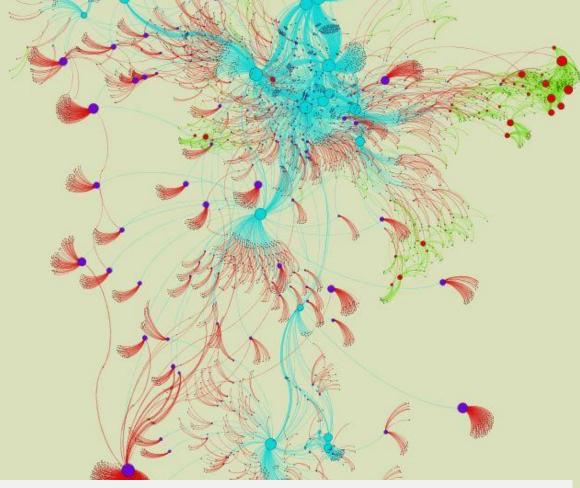
#### 107th Congress, 2002 session



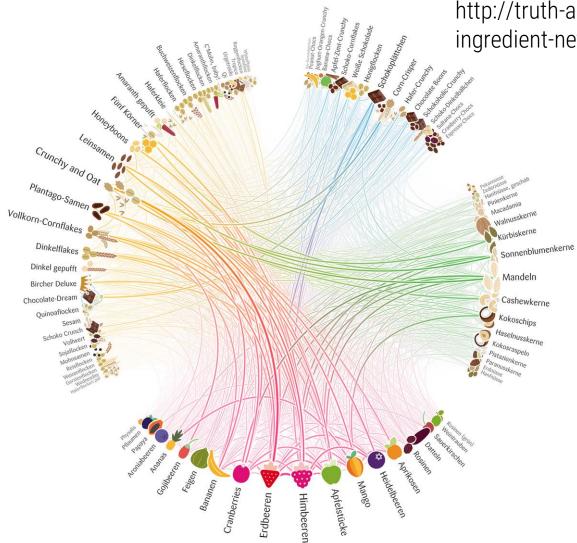
#### http://www.curehunter.com

### visual dictionary of drugs, diseases and therapies

https://dhs.stanford.edu/spatialhumanities/visualizing-databases/



Top Contributors to the Catalogue of Life and their associated species, references and databases

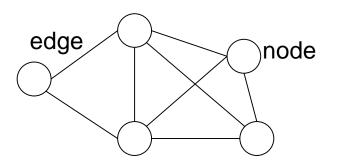


http://truth-and-beauty.net/projects/muesliingredient-network

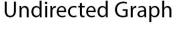
## GRAPHS

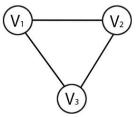
Graphs

- Describe relations among data items
- Using nodes and edges



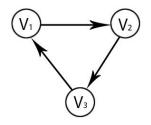
#### undirected graph: edges have no orientation



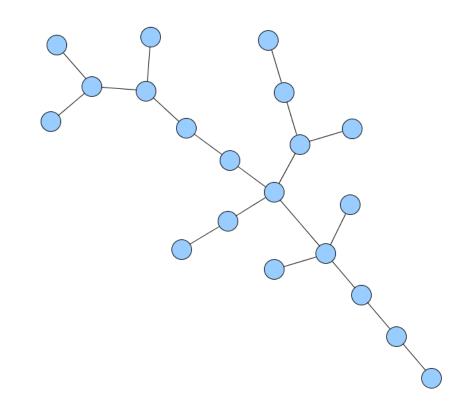


**Directed Graph** 

directed graph (digraph): edges have orientation

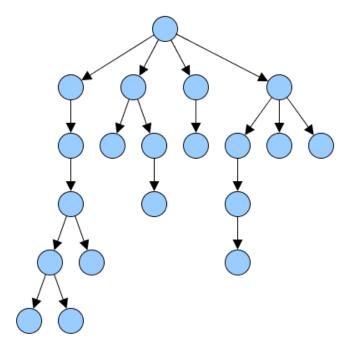


### a tree is a connected graph with no cycles

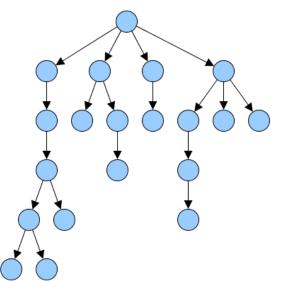


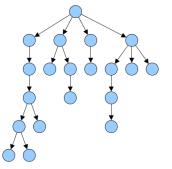
a directed tree is a digraph (directed graph) whose underlying graph is a tree

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



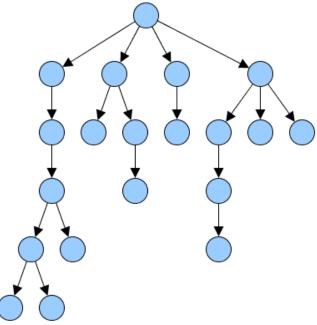
- the number of children of a node is called its degree
- leaf nodes are nodes without children





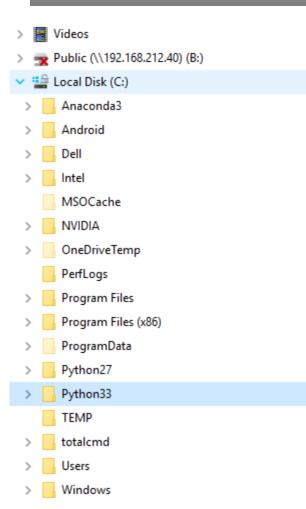
- a rooted tree is a directed tree with a distinguished vertex r, called the root, such that for every other vertex v there is directed path from r to v
- the root node is the only node with no parent
- any node may act as a root in undirected trees

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



### **EXAMPLES OF TREES**

### HIERARCHIES



Name
DLLs
Doc
include
Lib
libs
Scripts
tcl
Tools
🛃 ez_setup.py
LICENSE.txt
📔 NEWS.txt
🕹 python.exe
🕹 pythonw.exe
📔 README.txt
📕 setuptools-20.3.1.zip

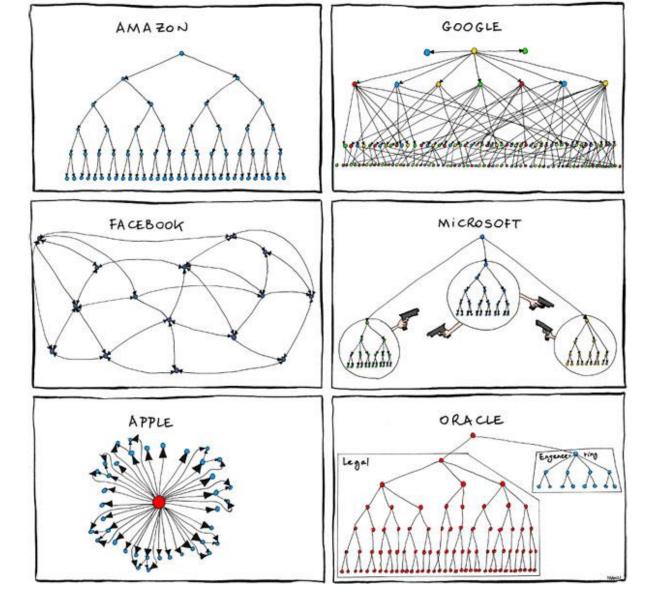
Date modified	Туре	Size
23-Mar-16 13:39	File folder	
23-Mar-16 15:20	File folder	
23-Mar-16 13:39	File folder	
23-Mar-16 13:39	File folder	
23-Mar-16 13:42	Python File	12 KB
09-Mar-14 10:37	TXT File	31 KB
09-Mar-14 10:27	TXT File	258 KB
09-Mar-14 10:35	Application	40 KB
09-Mar-14 10:35	Application	40 KB
09-Mar-14 10:27	TXT File	7 KB
23-Mar-16 13:43	Compressed (zipp	706 KB

### HIERARCHIES

### OrgOrgChart

Autodesk Research

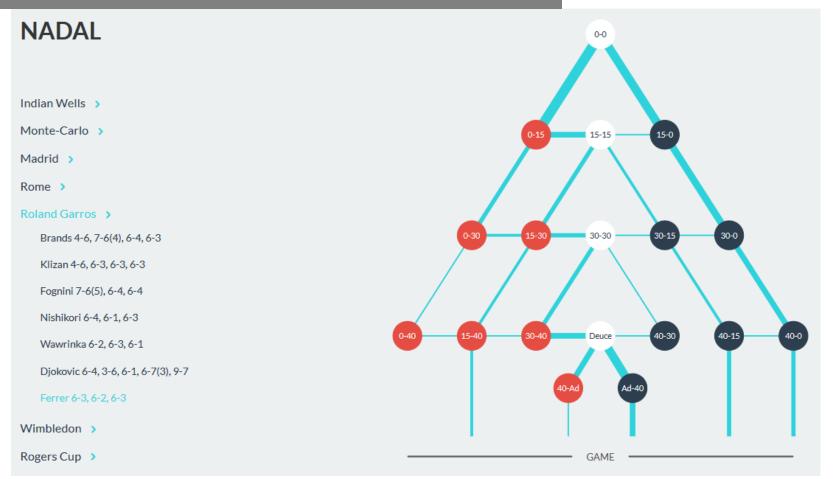
https://www.youtube.com/watch?v=mkJ-Uy5dt5g



org charts aren't always trees, though

http://www.bonkersworld.n et/organizational-charts/

### **DECISION PROCESS**



http://gamesetmap.com/nadalgametree/

### BRANCHING PROCESSES

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

### GeneaQuilts

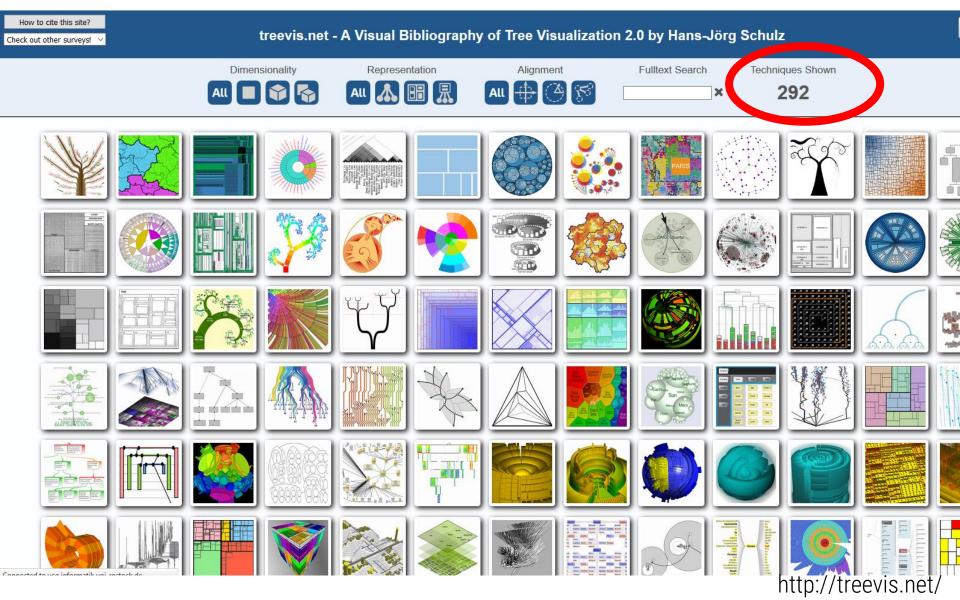
A System for Exploring Large Genealogies

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

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

### **TREE REPRESENTATION**

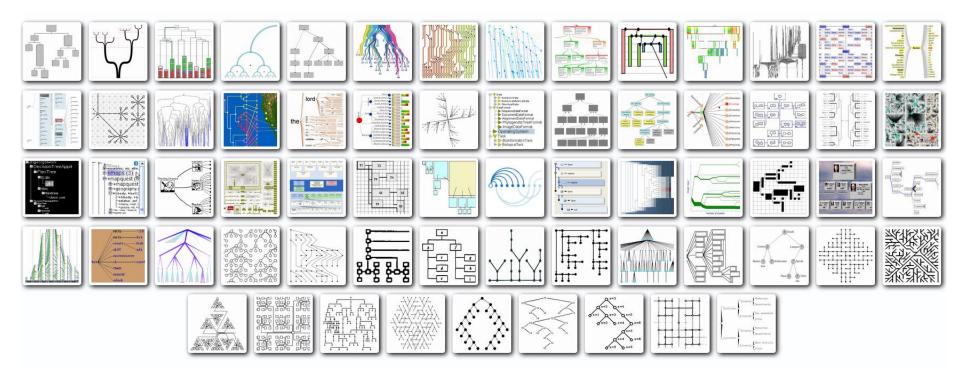
**TECHNIQUES** 



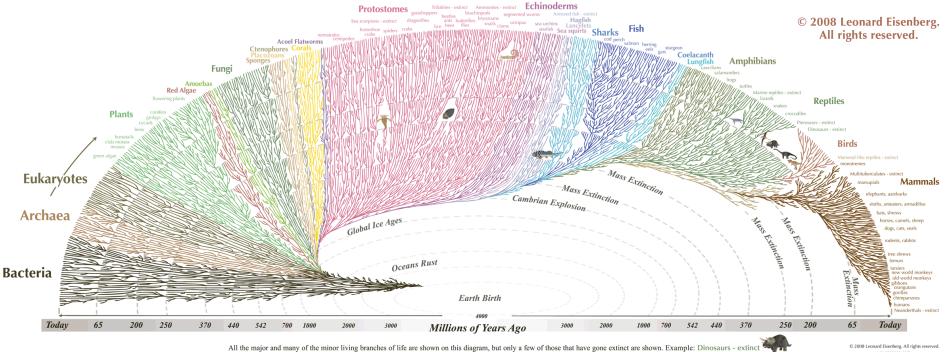
### CATEGORIZATIONS OF LAYOUTS

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

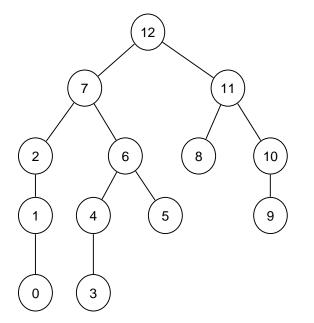
### 2D, AXIS-PARALLEL, EXPLICIT EDGES



### NODE-LINK



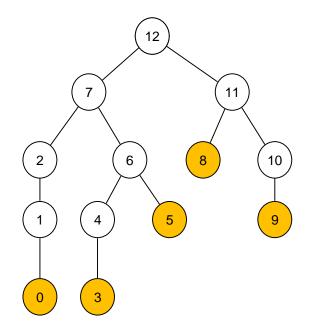
© 2008 Leonard Eisenberg. All rights reserved. evogeneao.com



#### THE LAYOUT WE WANT - HOW DO WE GET THERE?

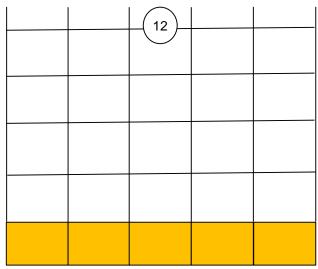
SIMPLE APPROACH

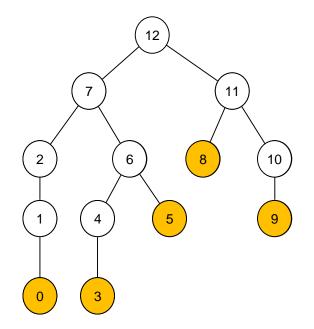
- 1) COUNT THE LEAVES
- 2) PLACE THE ROOT



SIMPLE APPROACH

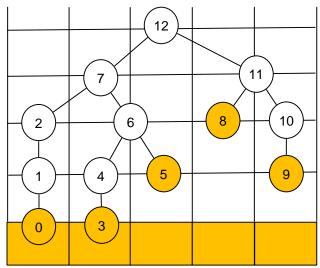
- 1) COUNT THE LEAVES
- 2) PLACE THE ROOT
- 3) RECURSIVELY DIVIDE

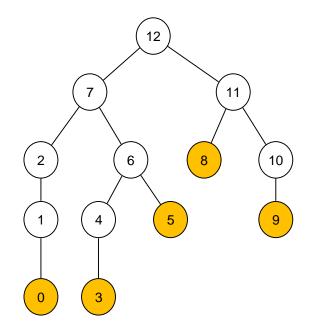




SIMPLE APPROACH

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



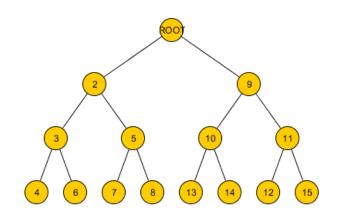


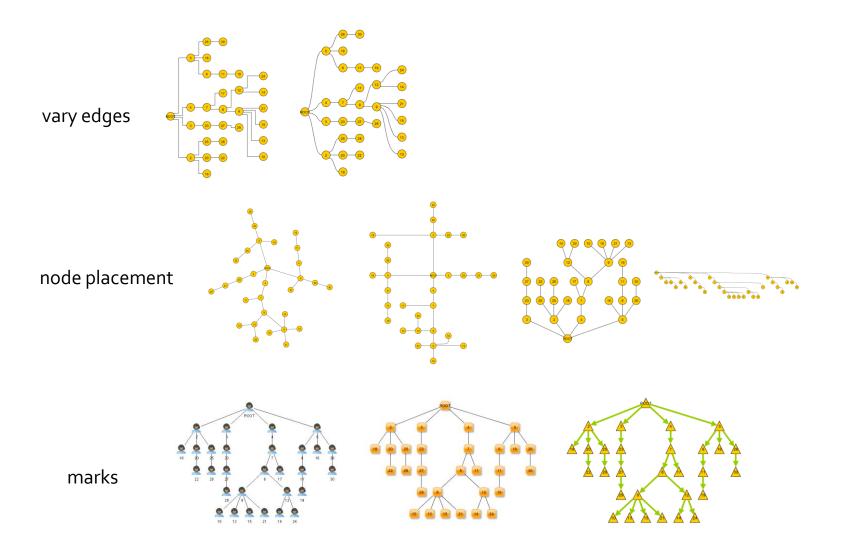
## PROS/CONS

- 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





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

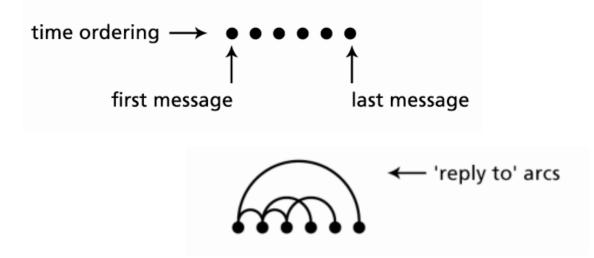
### AESTHETICS

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



#### http://threadvis.github.io/

**è o** 

**é è · ó** è

# THREADVIS

- time-scaling
- coloring people

From Tobias Isenberg

Subject Re: Slides, first draft

To Jian Chen🈭

. Cc. Torsten Mölleria, Michael Sedlmairia, Me < petra.isenberg@inria.fr>

と 🏩 🔺 Re: Slides, first draft	<ul> <li>Tobias Isenberg</li> </ul>	<ul> <li>14/10/2013 7:28 AN</li> </ul>
🟫 🕹 Re: Slides, first draft	Jian Chen	<ul> <li>14/10/2013 1:55 PN</li> </ul>
😭 🤌 🛛 Re: Slides, first draft	<ul> <li>Tobias Isenberg</li> </ul>	<ul> <li>14/10/2013 10:32 P</li> </ul>
🗅 - Re: Slides, first draft	<ul> <li>Torsten Möller</li> </ul>	<ul> <li>14/10/2013 11:27 P</li> </ul>
🖄 Next steps: QI	<ul> <li>Michael SedImair</li> </ul>	<ul> <li>22/10/2013 11:49 A</li> </ul>
🟫 - Re: Next steps: QI	<ul> <li>Torsten Möller</li> </ul>	<ul> <li>23/10/2013 1:58 AN</li> </ul>
🟫 Re: Next steps: QI	<ul> <li>Tobias Isenberg</li> </ul>	<ul> <li>24/10/2013 4:14 PN</li> </ul>
😭 🔶 🔶 Re: Next steps: QI	<ul> <li>Michael SedImair</li> </ul>	<ul> <li>24/10/2013 11:01 P</li> </ul>
From Tobias Isenbergin	r	🔦 Reply 🦚 Reply All 👻 🌩 Forward 🔯 Archive
Subject Re: Slides, first draft To Jian Chenie Cc Torsten Mölleric, Michael SedImairic, Me <petra.isenberg@inria.fr>int</petra.isenberg@inria.fr>	•	f. f

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

Cheers,

Tobias

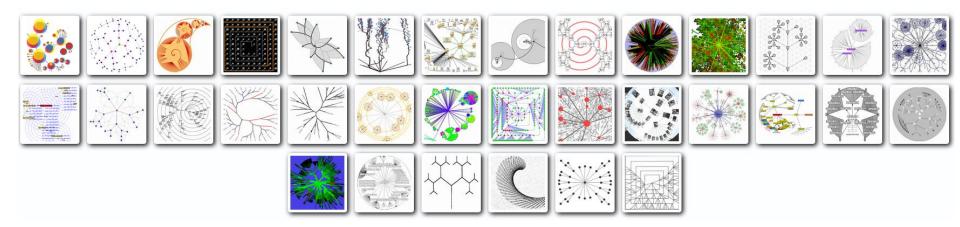
#### TREEJUXTAPOSER

Rectilinear layout and interaction for comparison of very large trees

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

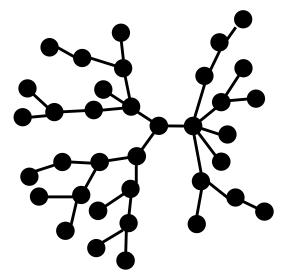
http://www.cs.ubc.ca/~tmm/papers/tj/#video

#### 2D, RADIAL, EXPLICIT EDGES



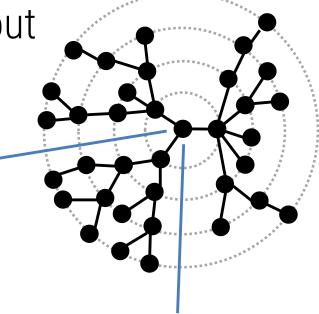
## RADIAL NODE-LINK DRAWING

# 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



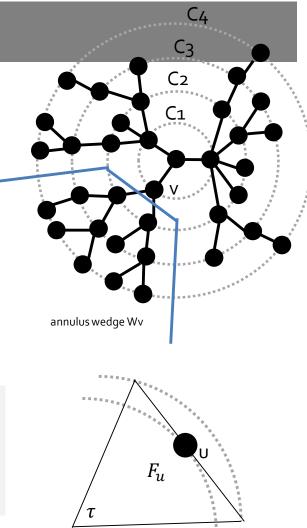
#### RADIAL NODE-LINK DRAWING

- radius of C<sub>i</sub> 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.

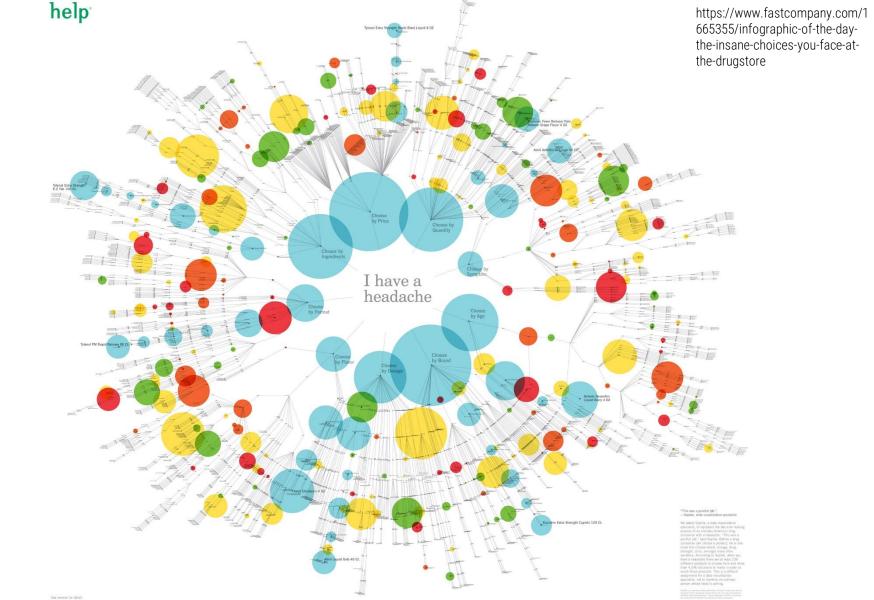
$$\beta_u = \left(\frac{l(u)\beta_v}{l(v)}, \tau\right)$$

For each child u of v:

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

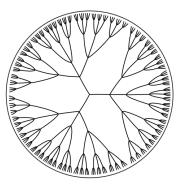


Resource: Battista et al. Graph Drawing



#### HYPERBOLIC BROWSER

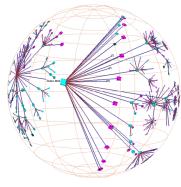
- uses hyperbolic geometry (not euclidean geometry)
- a hyperbolic plane can be displayed using the Poincaré disk model
  - a tree structure of any size fits within a finite area (circle)
  - node is displayed in center
  - all oder nodes move away from center and become exponentially smaller



(a) Uniform hyperbolic tree.



(b) StarTree by Inxight Software.



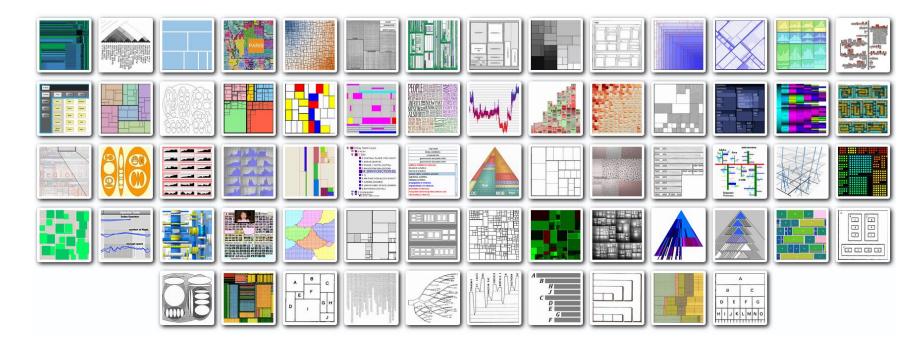
(c) H3 Browser.

#### 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.

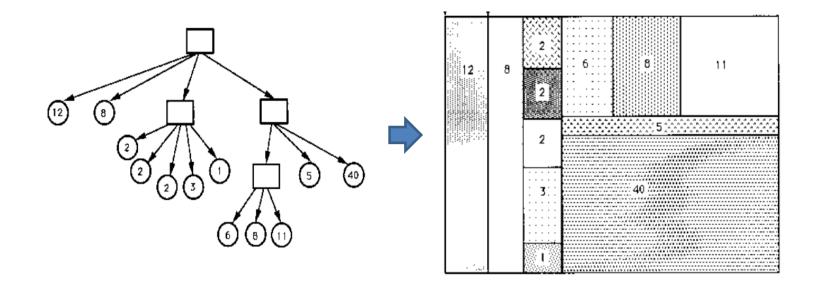
copyright 1995 MIT Media Laboratory

#### **2D, AXIS-PARALLEL, IMPLICIT EDGES**



#### A CLASSIC CONTAINMENT LAYOUT

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



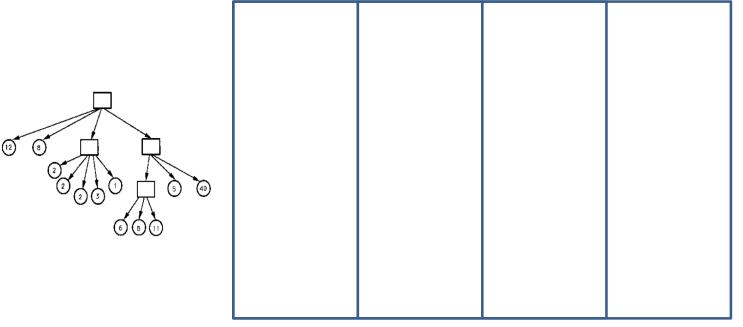
- Take a rectangular display area P1(x1,y1), Q1(x2,y1)
- This area represents the root of the tree



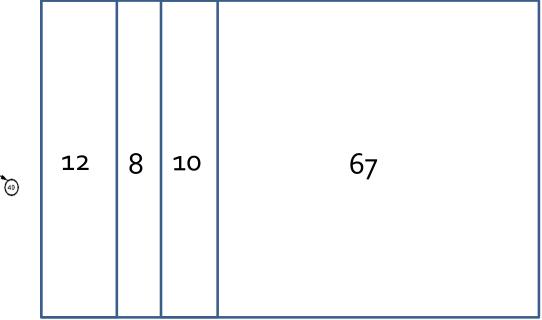
 $O_1$ 

#### recursive algorithm

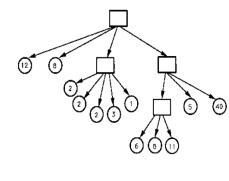
 the number of children of the current node define the number of partitions of the current node

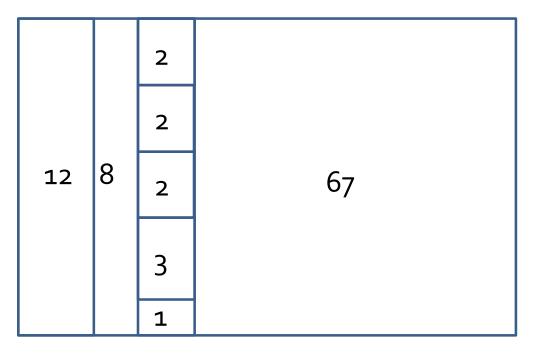


# the weight of each node determines the size of each partition

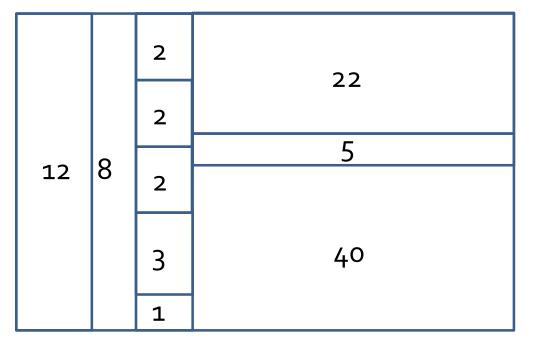


# at each change of level, rotate orientation of split by 90 degrees

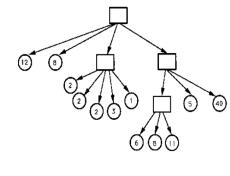


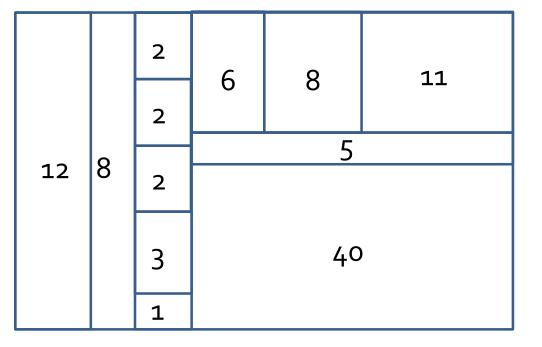


# at each change of level, rotate orientation of split by 90 degrees



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



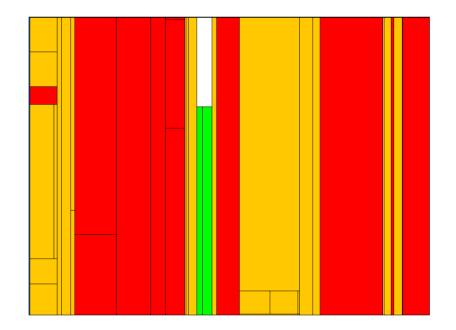


# TREEMAP

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

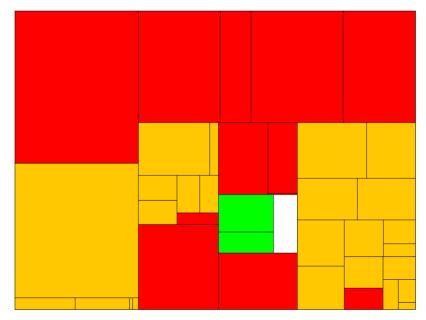
## TREEMAP VARIATIONS

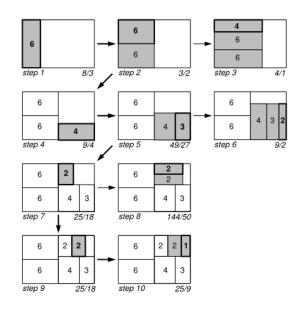
- problem with original treemap: lots of long stripes
- for long stripes the areas are difficult to compare



## 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

1 2 3 4 5 6 7 8 9 10111213141516	1	2	3	4	1	2	3	4	1	2	5	10
	5	6	7	8	5	8	9	10	3	4	7	12
	9	10	11	12	6	11	13	14	6	8	9	14
	13	14	15	16	7	12	15	16	11	13	15	16

#### 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.

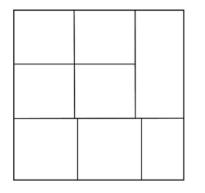
#### squarified

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.

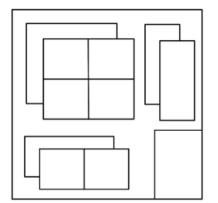
#### OTHER VARIATIONS OF TREEMAPS



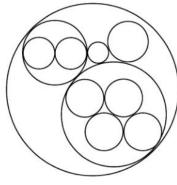
original squarified: emphasizes leafs and their attributes



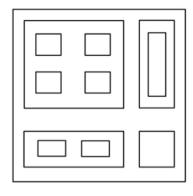
steptree: emphasizes structure with extrusion



cascaded layout: emphasizes structure with overlap

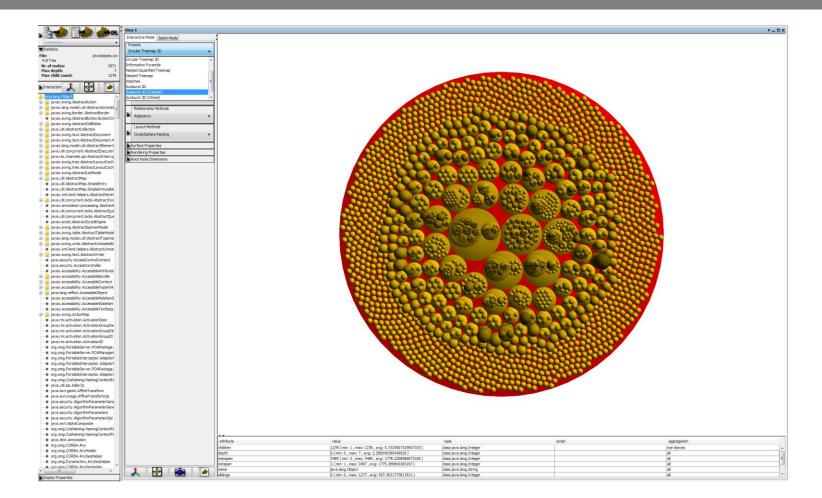


circular treemap: emphasizes structure with non-space-filling primitive



nested layout: emphasizes structure with whitespace

#### TRY THE IMPLICIT TREE VISUALIZATION TOOLKIT: HTTP://VCG.INFORMATIK.UNI-ROSTOCK.DE/~HS162/ITVTK/START.HTML



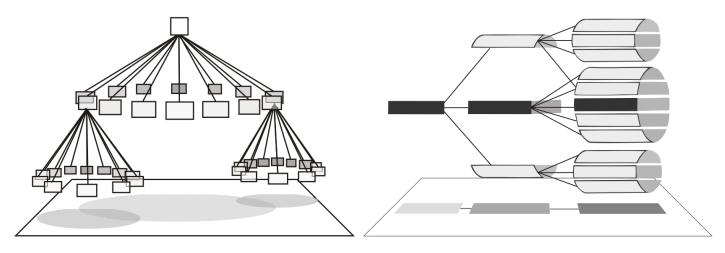
### HTTP://NEWSMAP.JP/

REGISTER LOGIN & CUSTOMIZE + SELECT ALL ARGENTINA AUSTRALIA	AUSTRIA	BRASIL CANADA FRA	ANCE	germany India	ITALIA MEXICO	NETHERLANDS	ZEALAND	SPAIN	<b>U.K.</b>	✓ U.S.	search	all	२ <b>-</b>
Bosnia says one	Amid India's Euphoria, Can Modi Live Up to Economic Expectations?			Modi mu on growt	Top Moments from the 2014				Jay Z and Beyonce hype tour with 'Run' movie trailer				
by floods, destruction				civil right					Kim Kardashian and Kanye West 'to many in France' now Stars' title tonight				
	Militi Libya	as attack		s Muslim Mi chensive Abo	Billboard Music Awards				Maya Rudolph looks to spice up NBC with a variety show software Kann to sue makers of				
	_	ament	into car	Toll in Svrian	Civil War Tops	IVIUSIC	<i>,                                    </i>	vai	นอ	Kahn to <sup>°</sup> sue mak Gerard Depardieı addiction film	Cers of U Sex Wash Aven Wash Aven	kins messing cut- trans revealing cut- trans revealing cut- casual Cl Casual Cl Cannes F Cronente	Lawrence Steps Out Rickling Thich Cannes Wilh Josh
	Ukraine crisis: No sign		16000	0: Human Rig	hts Group	Top Moments from th		at the					
		ia withdrawal, ito	Prosecutors Announce ( Against Mal Users of Had	Charges Maliki's S	Michael Jackson was brought back to life – sort of – at 2014 Billboard Music Awards at the MGM Grand Garder Arena in Las Vegas. A hologram of the King of Pop, surrounded by real-life backup dancers, performed "Sla to the Rhythm" Sunday n May 19 23:24   1782 related articles   ABC News				Football leam				
		Van Gaal brir	Software		s against FBI agent	AstraZe	neca	a	p	Credit Suisse to Ilead guilty in tax case	UPDATE 2- Campbell cuts sales forecast : US soup sales	S&P 500 and Nasdaq gai as investors ca s cool Wany Target Corp.	d Southwest Airlines Plants in, but judious (1997) But South Warren Park (1973) But South
to Permit Chrome to Wear Nasal Strip in the Belmont		experience,	iys	showd					6.1M to travel Memorial ay weekend; gas prices eady	; jump in reports of susp activity	rompt kickus Greek Saloff Show Recalling Orisis	or torrigie and tamino bedi hand Asiman Rijs Alengan is Rush for But Weitrolo UL	
		winning ways to United		to yet nearing uose		Justice Department charges 5 Chinese army officers with cyber-spying warming				Google Cele Anniversary Rubik's Cube Doodle Gam	Of The e With ne	Gasoli Over 3 Gallon	
		Blackhawks create		te Tiger Woods still can't swing a club, has no idea when he'li play again diverse statistic to Miss adauting to Miss adautin		NH official resigns after racist rant about		Sunny with Su	urfer charged with tempted murder of derly woman	Verizon Promise Speeds With XL	TE	Esa's Cryosat missio sees Antarctic ice lo double SP and Microsoft Doubling formative data	anes Dorrt Buaro Me DSSES and Antarolic ke raikelby Burgean task' bullan for
						President Obama	Families ask US Crast Guard to resume yacht search		the Rhode bland man tells police to was leaving bor when he noticed hefoloeen Earle works for state primate	Apple, Google re truce Bl, InterMune	1.8 million pound	dis of beef House backso	kouins mobile conserve like and the conserve like and the conserve like and the conserve like and the conserve like the conserve like the conserve like and the conserve like an
		Quick		ays end road trip ith loss to Angels	Brandon Marshall signs extension 2011 PL cell Rwelet Part 1	US judge to rule Tuesday in corruption case of ex- Virginia governor, wife	Remark has questions, few answers in 9-year-olds death	Vicion for Fary Oxfory Found Guily In New York. John Kerry References Also Baldwin, Deponce at EC Gaalastion	nunoff starts Wonday Wind, traine among reministers for Oklahoma studente, Supreme Dautto Hear Christikhithan	tout lung drugs	Uganda HV russe Rosernary Kampala court	y Kamubiru jalied by W Olici as Opa	doesity risk for children id Says Ank Foed Austas Bad encore
Mon May 19, 2014 20:32:39	ting by					+ SELECT ALL V WORLD	MATIONAL	USINESS	TECHNOLOGY	SPORTS	ENTERTAINMEN	IT 🖌 HEALTH	LESS THAN 10 MIN. AGO More than 10 Min. Ago

#### **3D LAYOUTS**



# HISTORIC EXAMPLE: CONETREE / CAMTREE



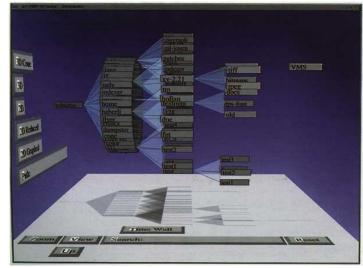
ConeTree

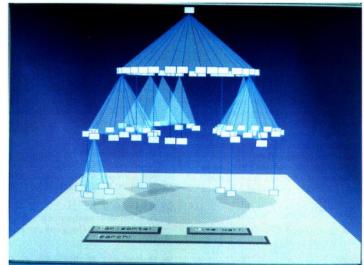
CamTree

[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* 

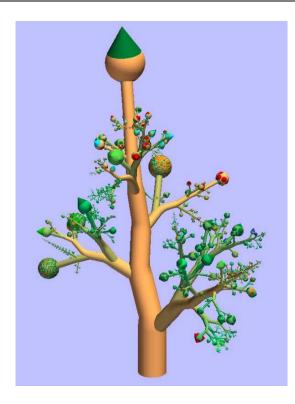
#### CONE/CAMTREE

- 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



# 3D LAYOUTS

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

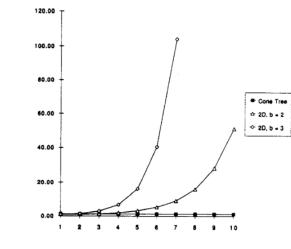


Figure 1: Aspect Ratio of 2D and 3D Trees.

### 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.):
    - <u>http://www.informatik.uni-rostock.de/~hs162/optreedemo/TestBed.htm#</u>
    - <u>http://w3.win.tue.nl/nl/onderzoek/onderzoek\_informatica/visualization/sequoia</u> <u>view/</u>
  - 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<sub>i</sub>}
- 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**

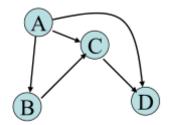
#### GRAPH MEASURES

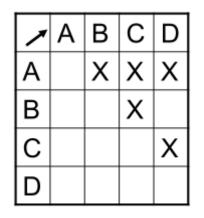
- SIZE = #nodes
- DENSITY = edges/vertices (roughly)
- PATH = sequence of edges connecting (different) vertices
- VERTEX DEGREE = #edge connections
- DISTANCE = #hops between vertices

# TWO CLASSICAL VISUAL REPRESENTATIONS

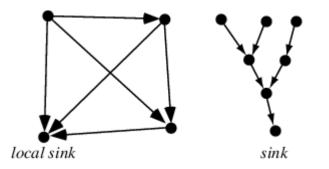
### Node-Link Diagram

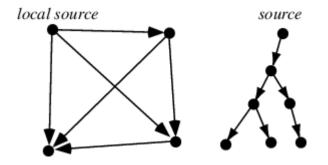
Adjacency Matrix





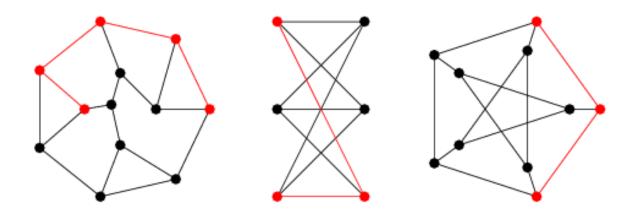
### Find # of neighbors of a vertex (e.g. source vs. sink)



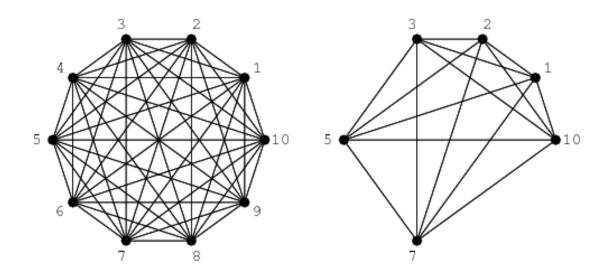


http://mathworld.wolfram.com/DigraphSink.html http://mathworld.wolfram.com/Source.html

### See paths (overviews, shortest, cycles)



### Identify Sub-graphs



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

### 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?

- 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

#### ABSTRACT

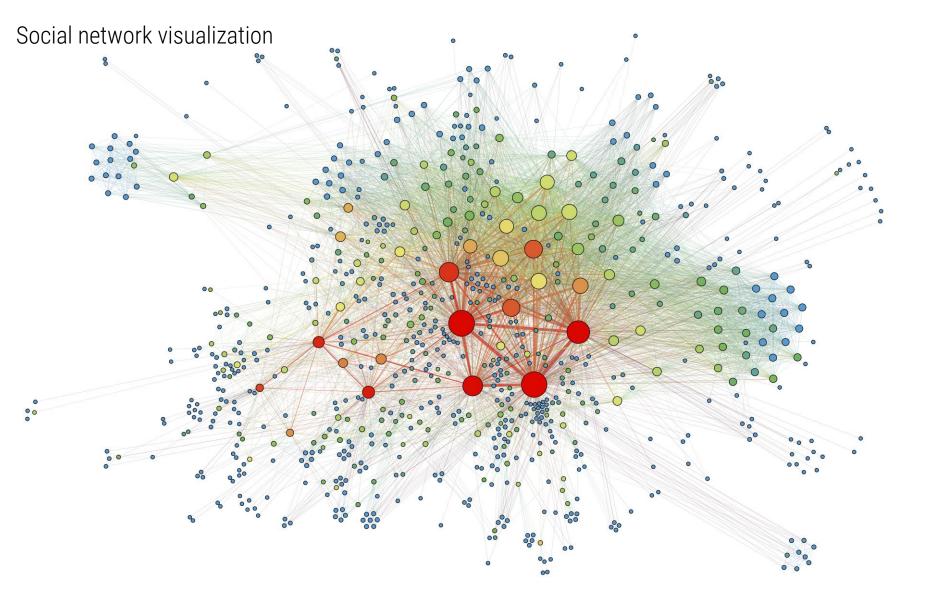
Our goal is to define a list of tasks for graph visualization that has enough detail and specificity to be useful to: 1) designers who want to improve their system and 2) to evaluators who want to compare graph visualization systems. In this paper, we suggest a list of tasks we believe are commonly encountered 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 Jean-Daniel Fekete, Nathalie Henry INRIA Futurs/LRI Bat. 490 Université Paris-Sud, 91405 ORSAY, France

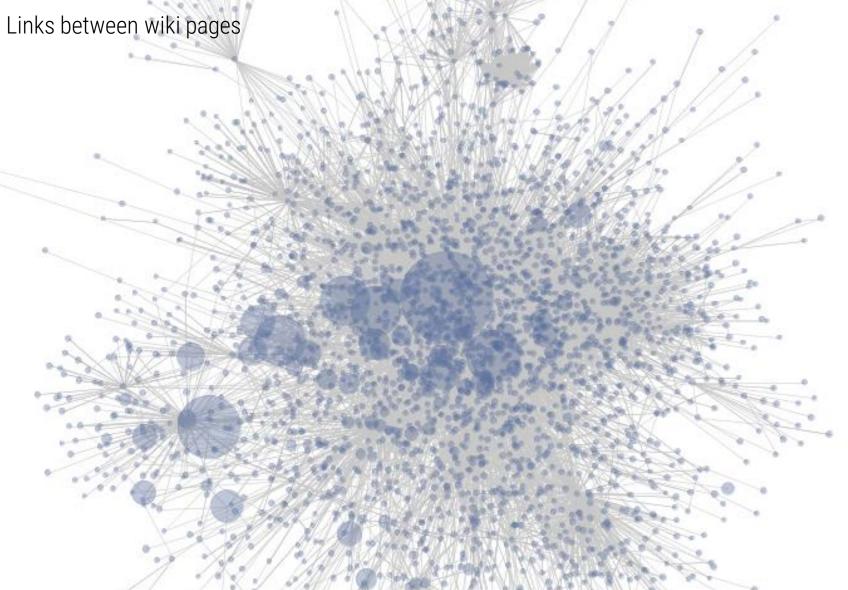
+33-1-69153460

Jean-Daniel.Fekete@inria.fr, nhenry@lri.fr

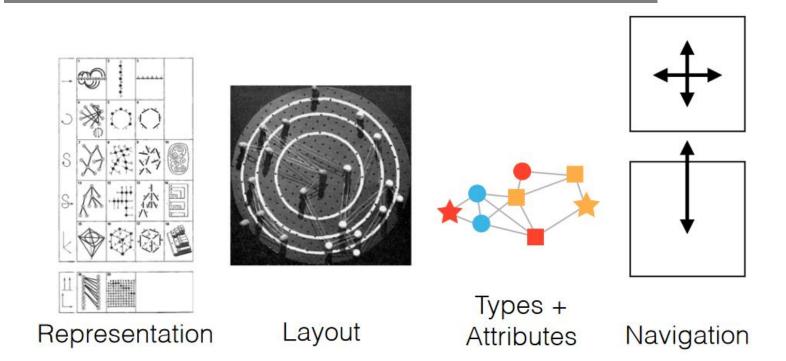
user studies of graph visualization techniques and extracted the tasks used in those studies.

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





# GRAPH VISUALIZATION CHALLENGES



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

• determine if <u>Greenpeace</u> was or could become a significant disruptive geopolitical force

• first: identify who/what to concentrate resources on

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

	Data Scientist Insights EXPLORING THE DARKEST PLACES ON EARTH							
â	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  $\cdot \mathcal{O}(0)$ 



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



### 1) get Facebook data using Netvizz

#### Studying Facebook via Data Extraction: The Netvizz Application

#### **Bernhard Rieder**

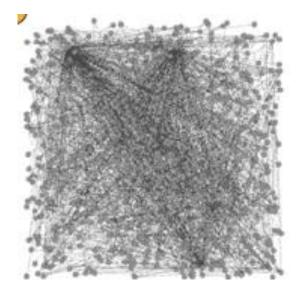
University of Amsterdam Turfdraagsterpad 9 1012TX Amsterdam rieder@uva.nl

#### 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, postdemographical 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.

### 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?"

# 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

### 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)

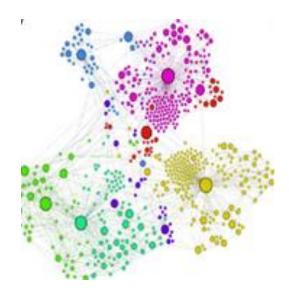


Bigger nodes are more central to behavioral dynamics.

Several nodes become central figures in the overall network.

### 5) highlight communities

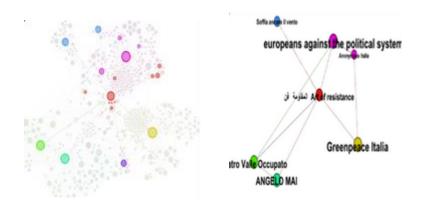
color nodes by modularity / clusters



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.

### 6) filter, explore, label



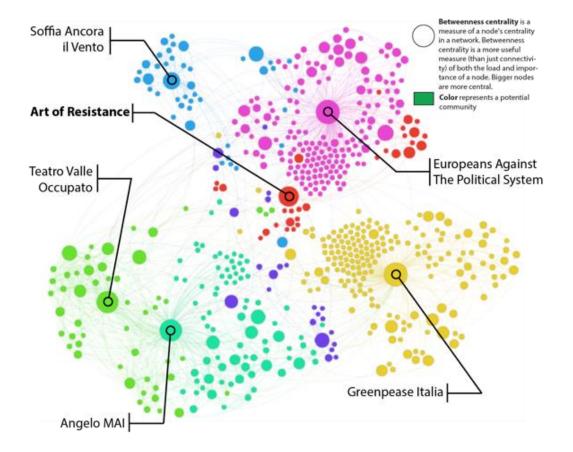
Five outlying nodes in the network (blue, maroon, yellow, dark green, and light green).

Center: an equally important red node

Emergence of a previously un-recognized activism player: <u>Art of Resistance</u>.



7) communicate & explain



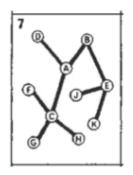
### LAYOUTS

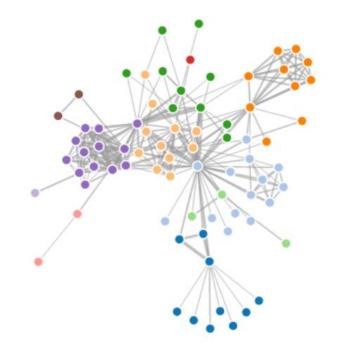
#### Important to the success of your analysis

# LAYOUT FREE

- Physical forces
- Proximity based
- Spring Model
- Kamanda&Kawai
- Frucherman&Reingold
- Davidson&Harel
- LinLog

#### **Force Directed**







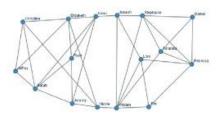
# LAYOUT FREE

### Aesthetic Criteria

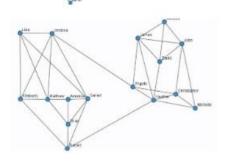
- Reduce number of edge crossing
- Foster Symmetry
- Uniform edge length
- Aspect Ratio
- Equal Angles

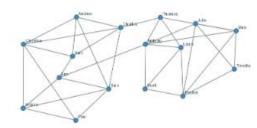
### **GRAPH DRAWING**

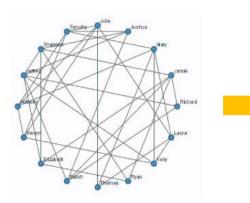
# LAYOUT FREE



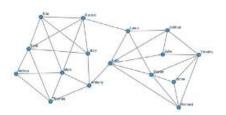
### HAND MADE

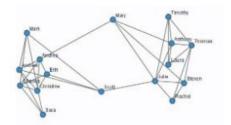


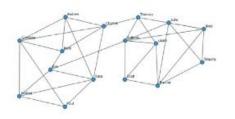


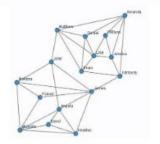


ptual organization in user-generated graph layouts am, F.J.J.; Rogowitz, B.

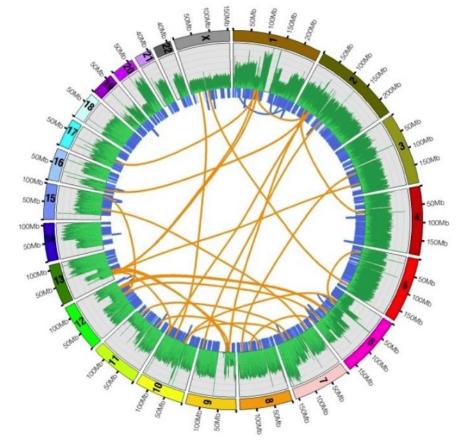


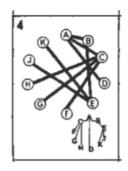




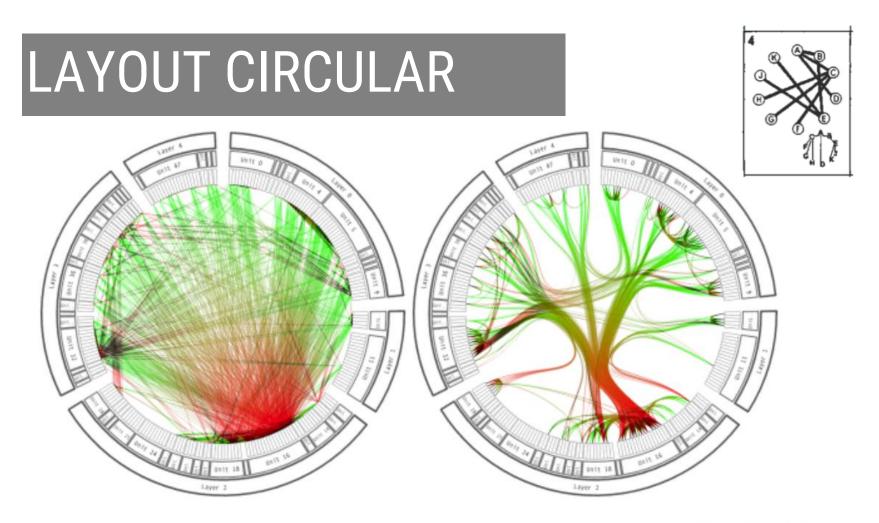


# LAYOUT CIRCULAR

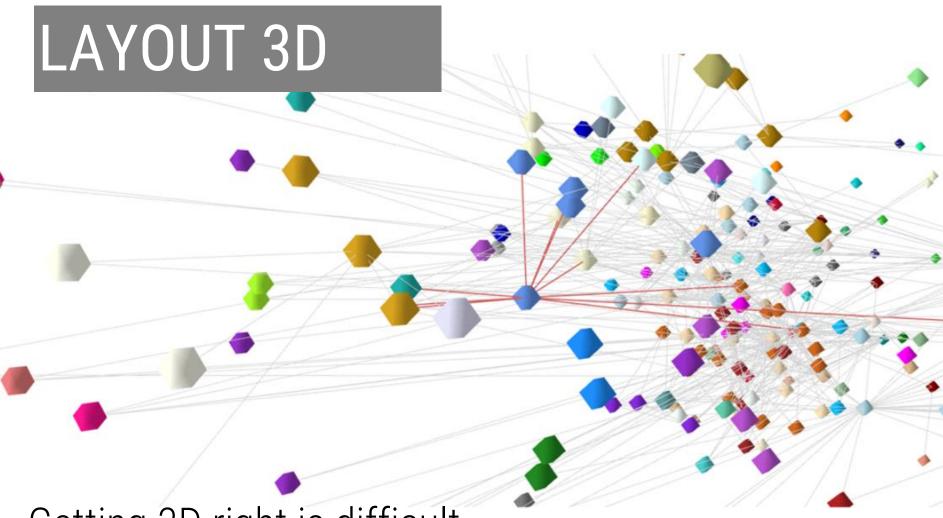




- Edges on the inside
- Vertices & attributes on the outside
- Ordering possible



Edge Bundling Holten 2006

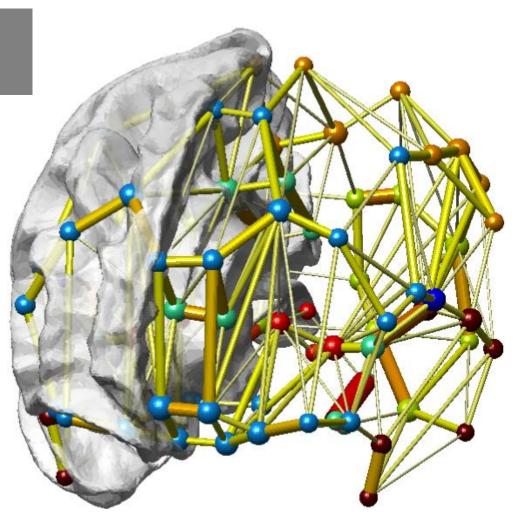


Getting 3D right is difficult

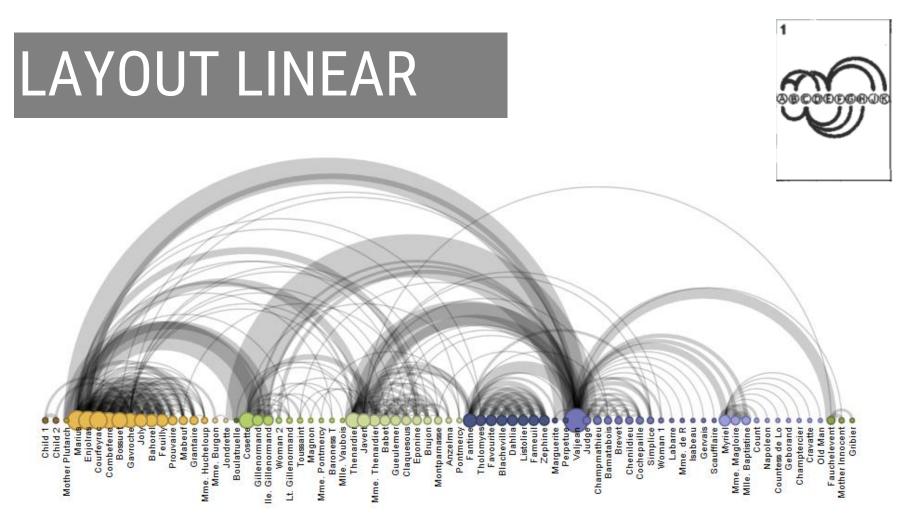
https://fwaris.wordpress.com/2012/07/08/a-simple-technique-for-creating-3dgraphs-from-2d-ones/

# LAYOUT 3D

# Sometimes necessary (!?)



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

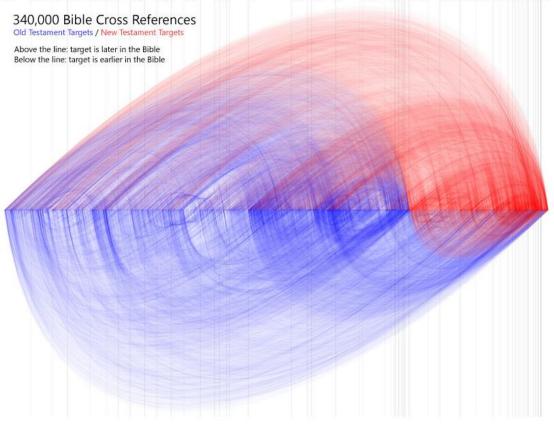


http://mbostock.github.io/protovis/ex/arc-full.html

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

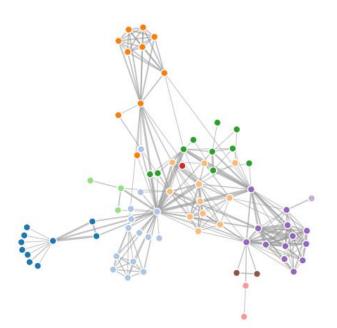
# LAYOUT LINEAR

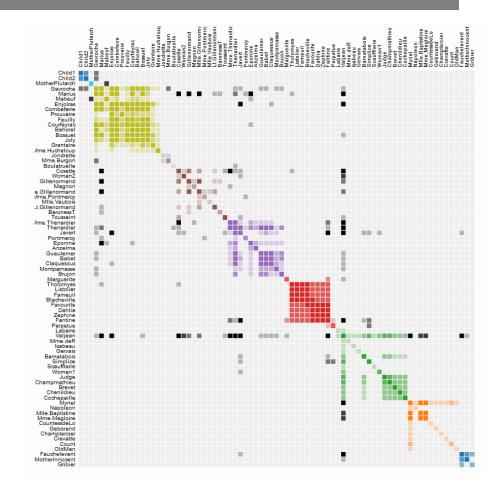
Gen Exod Lev Num Deut Josh 1Sam 1Kgs 1Chr Ezra Ps Prov Isa Jer Ezek Hos Matt Luke JohnActs Rom Gal Hb Re



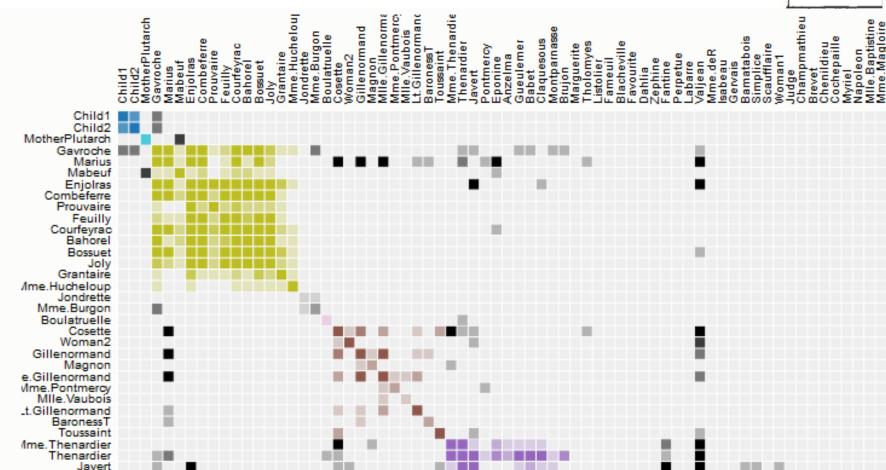
https://www.openbible.info/labs/ cross-references/

### LAYOUT ADJACENCY MATRIX





# ADJACENCY MATRIX



# PROS/CONS

### Matrix

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

### Node-Link

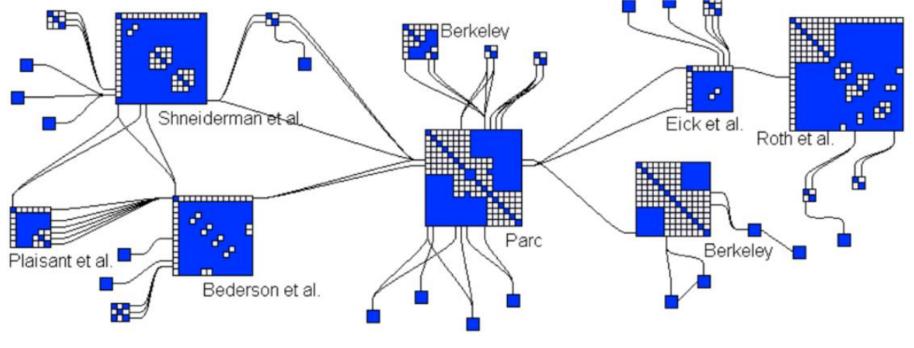
- Familiar
- Compact
- Path following easier
- Effective for small and sparse graphs

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

- 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	
I Adam Blackwell	
I Alameda	
I Aquatics	
I Arthur Swordane	
l Assan	1
I Brandow Tropical Fish	1
Catherine Carnes	
l Cesar Gil	
I Collie Carnes	
I Cr Wharton	)
Don Rabinowitz	
Donna Ghostley	
I Ed Parker	
Edward Abbey	
I Eva Berrima	
Faron Gardner	6
Gardner	
Green	_
I Griffin Vulture	
Jeri Ryan	
Jessica Alba	
John Burton Wade	
John Wharton	
Kim	
Kim Basinger	
I Lesie	
Lily	-
Luella Vedric	-
Madhi Kim	
I Marcus James	
Melissa Ethridge	-
I Michael Jackson	
Oogies	
Paul McCartney	
r'Bear	
Richmond	
Richmond Shire	
Singer	
songwriter Jimmy Buffett	
I Spix	
Terry Mulley	1
Tony Jones	

#### Туре В

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

#### Type C



**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)