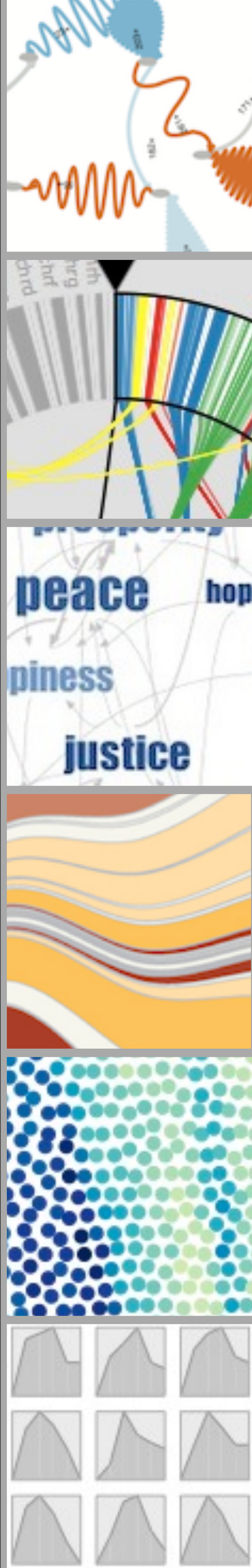
cs6964 | February 28 2012

TREES & GRAPHS

Miriah Meyer
University of Utah

slide acknowledgements:
Hanspeter Pfister, Harvard University
Jeff Heer, Stanford University



administrivia

feb 14-23 : proposal meetings

march 7 : presentation topics due

march 9 : proposals due

march 27-april 3 : project updates

april 5-24 : paper presentations

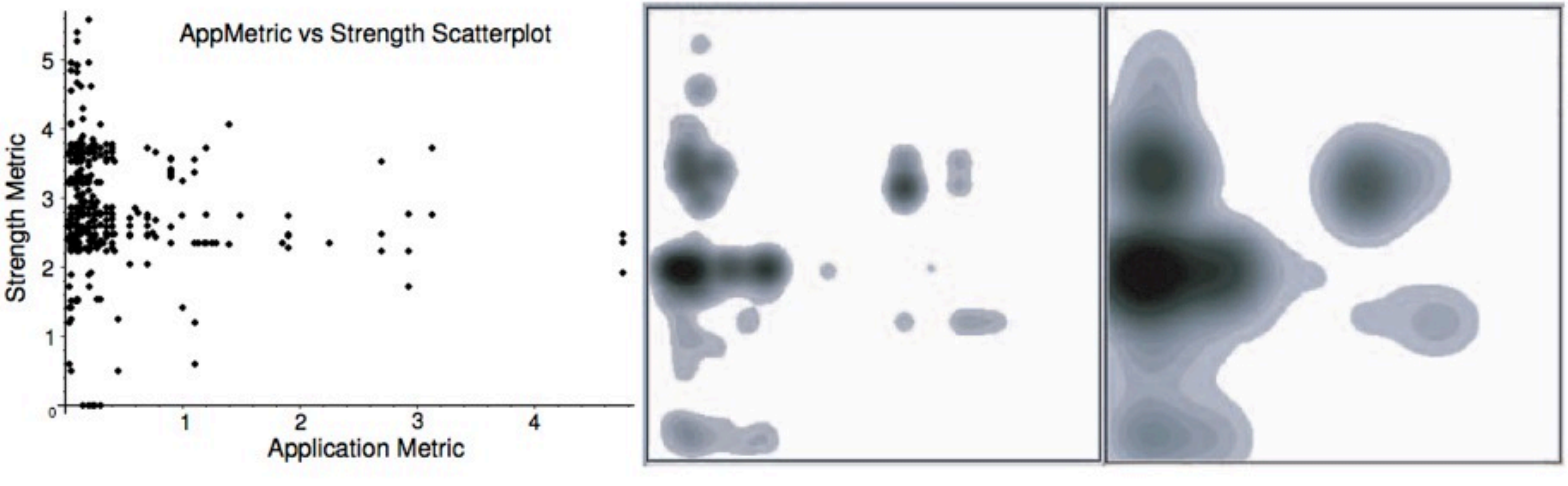
may 1 : final project presentations

may 3 : process books due

LAST TIME

MULTISCALE SCATTERPLOTS

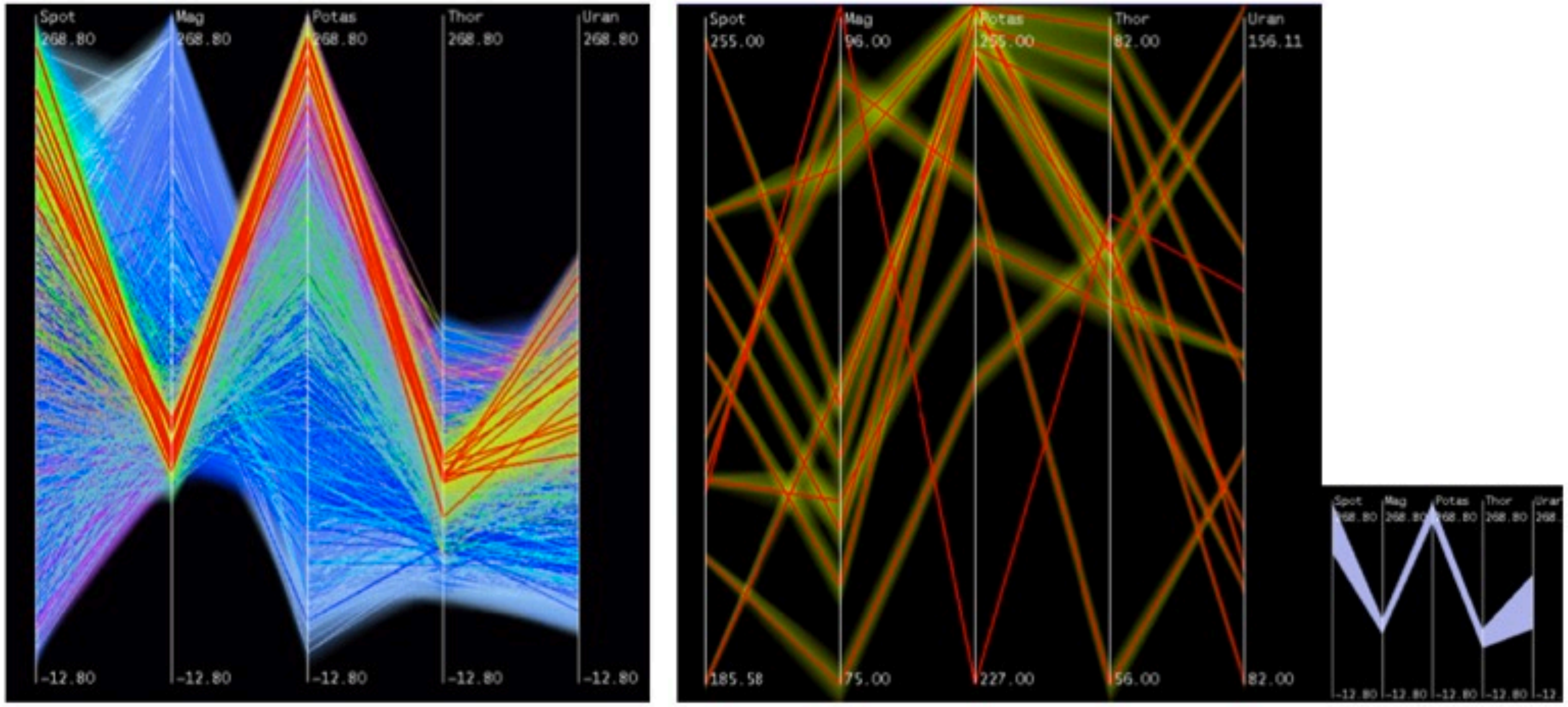
- **blur shows structure at multiple scales**
  - convolve with Gaussian
  - slider to control scale parameter interactively
- **easily selectable regions in quantized image**



Chiricota 2004

HIERARCHICAL PARALLEL COORDINATES

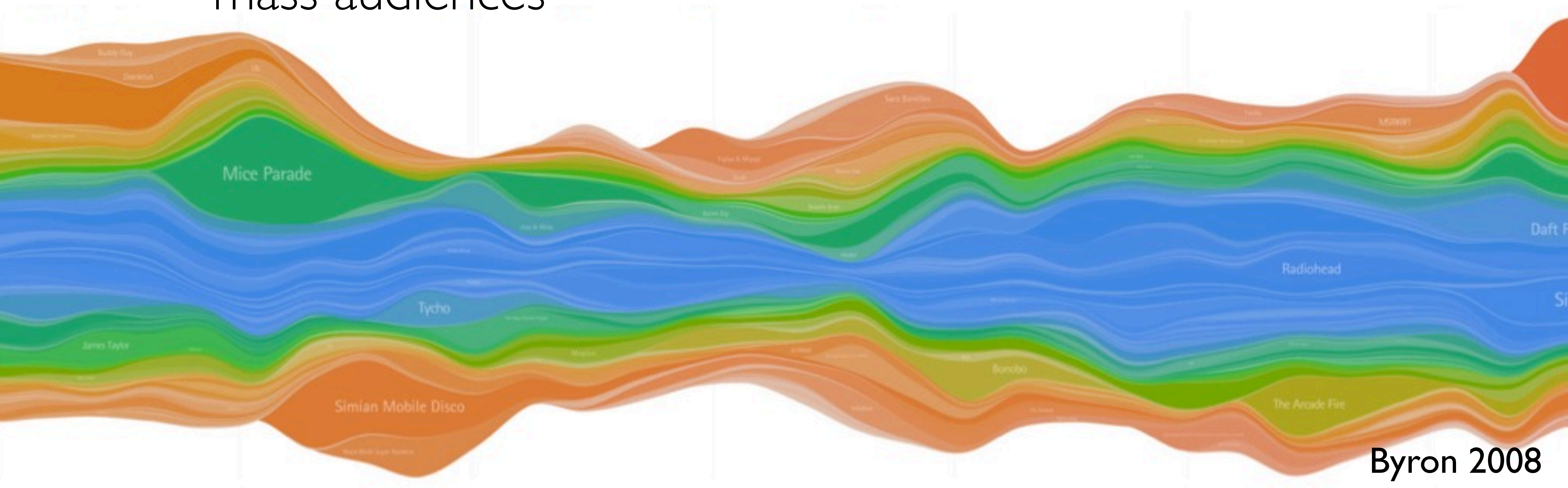
- **technique-driven paper**
  - no problem characterization
- **goal: scale up parallel coordinates to large datasets**
  - challenge: overplotting/occlusion



Fua 1999

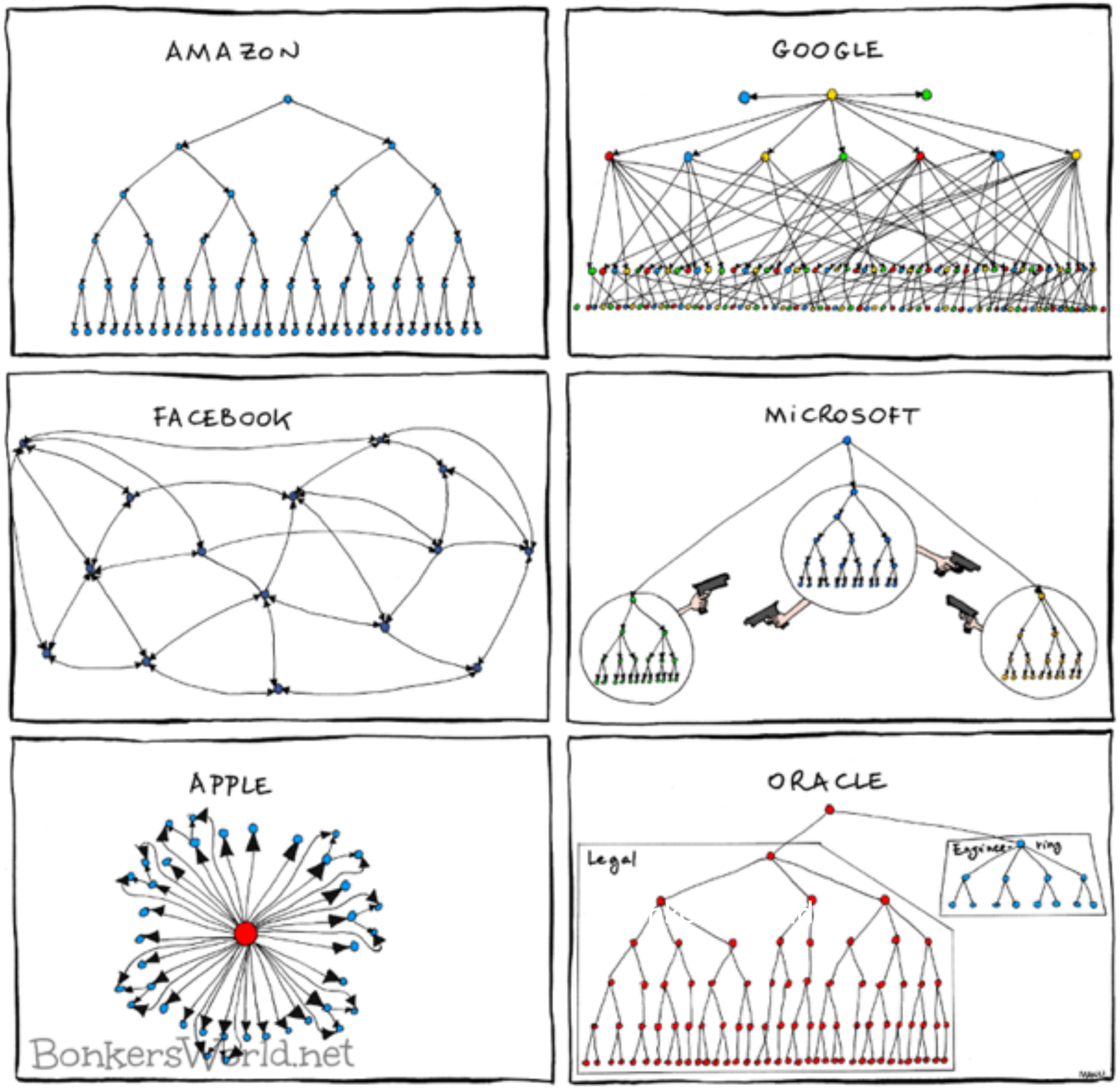
STREAMGRAPH

- **problem-driven paper**
  - development of new technique to solve a specific problem
- **challenge**
  - convey a large amount of data in a way that engages mass audiences



Byron 2008

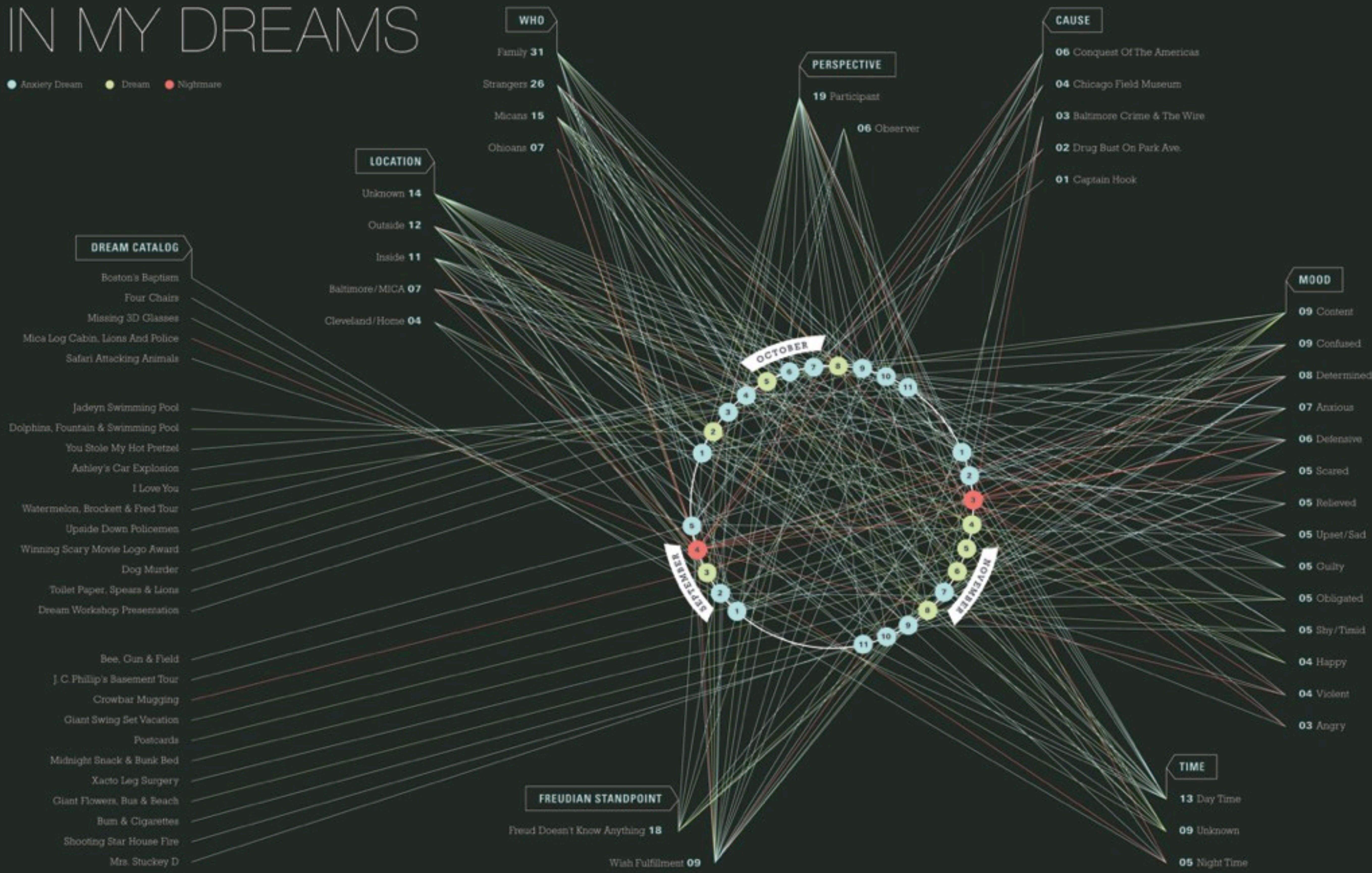
TREES & GRAPHS


AMAZON
GOOGLE
FACEBOOK
MICROSOFT
APPLE
ORACLE
Legal
Engineering
BonkersWorld.net

facebook

December 2010

Paul Butler



IN MY DREAMS

● Anxiety Dream ● Dream ● Nightmare

WHO

Family **31**
Strangers **26**
Micans **15**
Ohioans **07**

PERSPECTIVE

19 Participant
06 Observer

CAUSE

06 Conquest Of The Americas
04 Chicago Field Museum
03 Baltimore Crime & The Wire
02 Drug Bust On Park Ave.
01 Captain Hook

LOCATION

Unknown **14**
Outside **12**
Inside **11**
Baltimore/MICA **07**
Cleveland/Home **04**

DREAM CATALOG

Boston's Baptism
Four Chairs
Missing 3D Glasses
Mica Log Cabin, Lions And Police
Safari Attacking Animals

Jadeyn Swimming Pool
Dolphins, Fountain & Swimming Pool
You Stole My Hot Pretzel
Ashley's Car Explosion
I Love You
Watermelon, Brockett & Fred Tour
Upside Down Policemen
Winning Scary Movie Logo Award
Dog Murder
Toilet Paper, Spears & Lions
Dream Workshop Presentation

Bee, Gun & Field
J. C. Phillip's Basement Tour
Crowbar Mugging
Giant Swing Set Vacation
Postcards
Midnight Snack & Bunk Bed
Xacto Leg Surgery
Giant Flowers, Bus & Beach
Bum & Cigarettes
Shooting Star House Fire
Mrs. Stuckey D

MOOD

09 Content
09 Confused
08 Determined
07 Anxious
06 Defensive
05 Scared
05 Relieved
05 Upset/Sad
05 Guilty
05 Obligated
05 Shy/Timid
04 Happy
04 Violent
03 Angry

FREUDIAN STANDPOINT

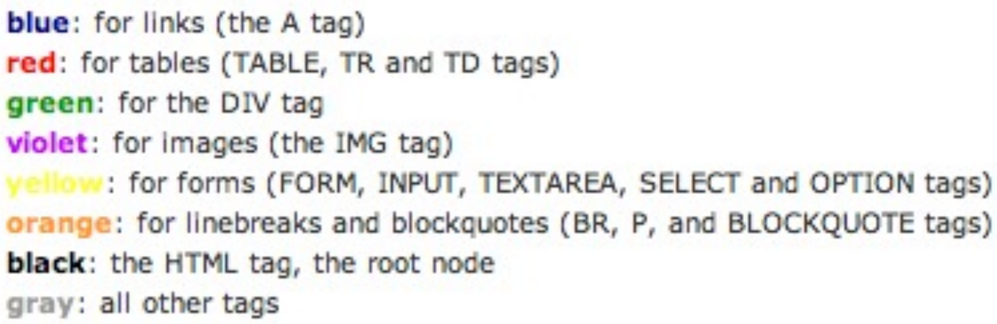
Freud Doesn't Know Anything **18**
Wish Fulfillment **09**

TIME

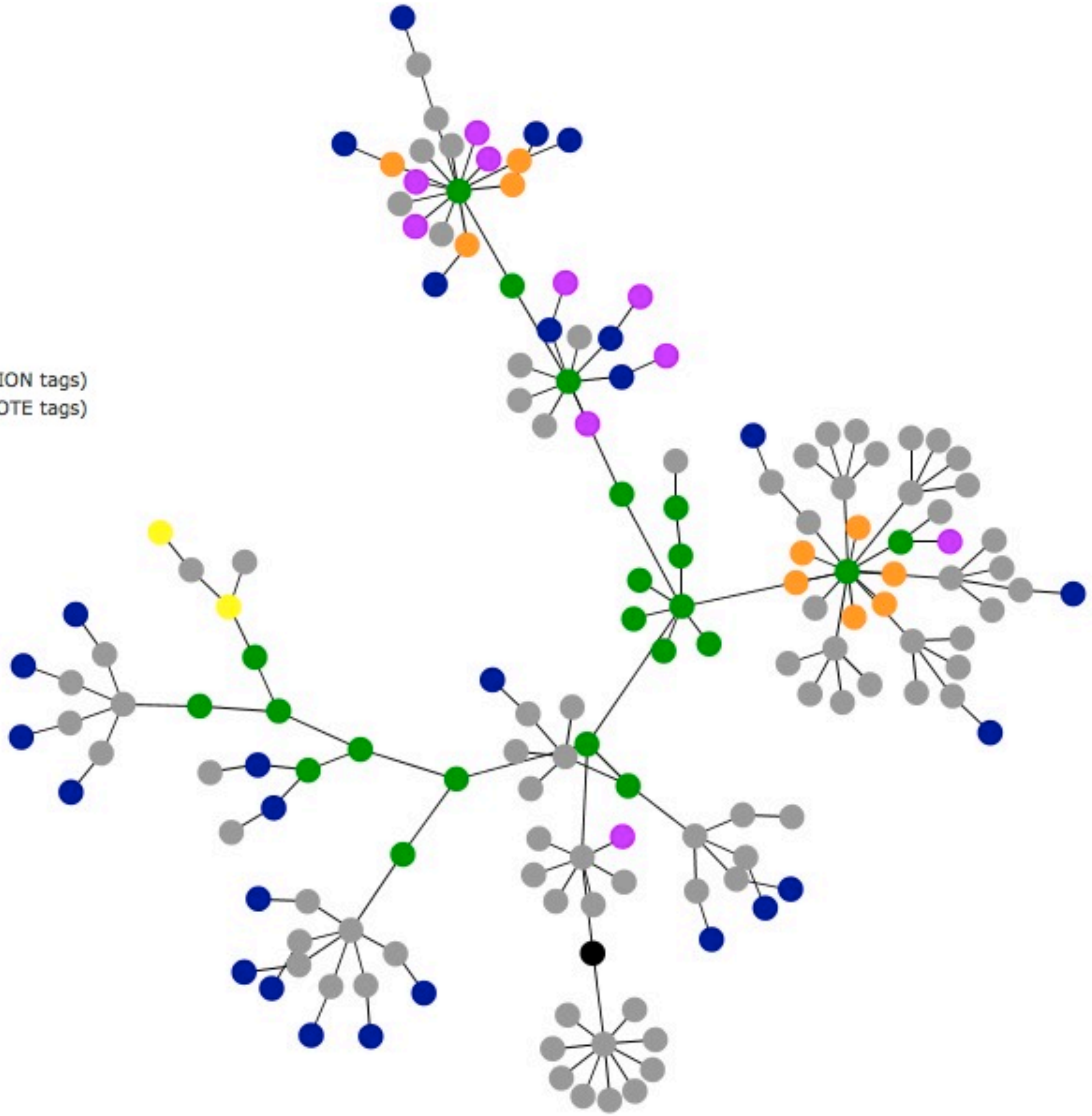
13 Day Time
09 Unknown
05 Night Time

Kailie Parrish

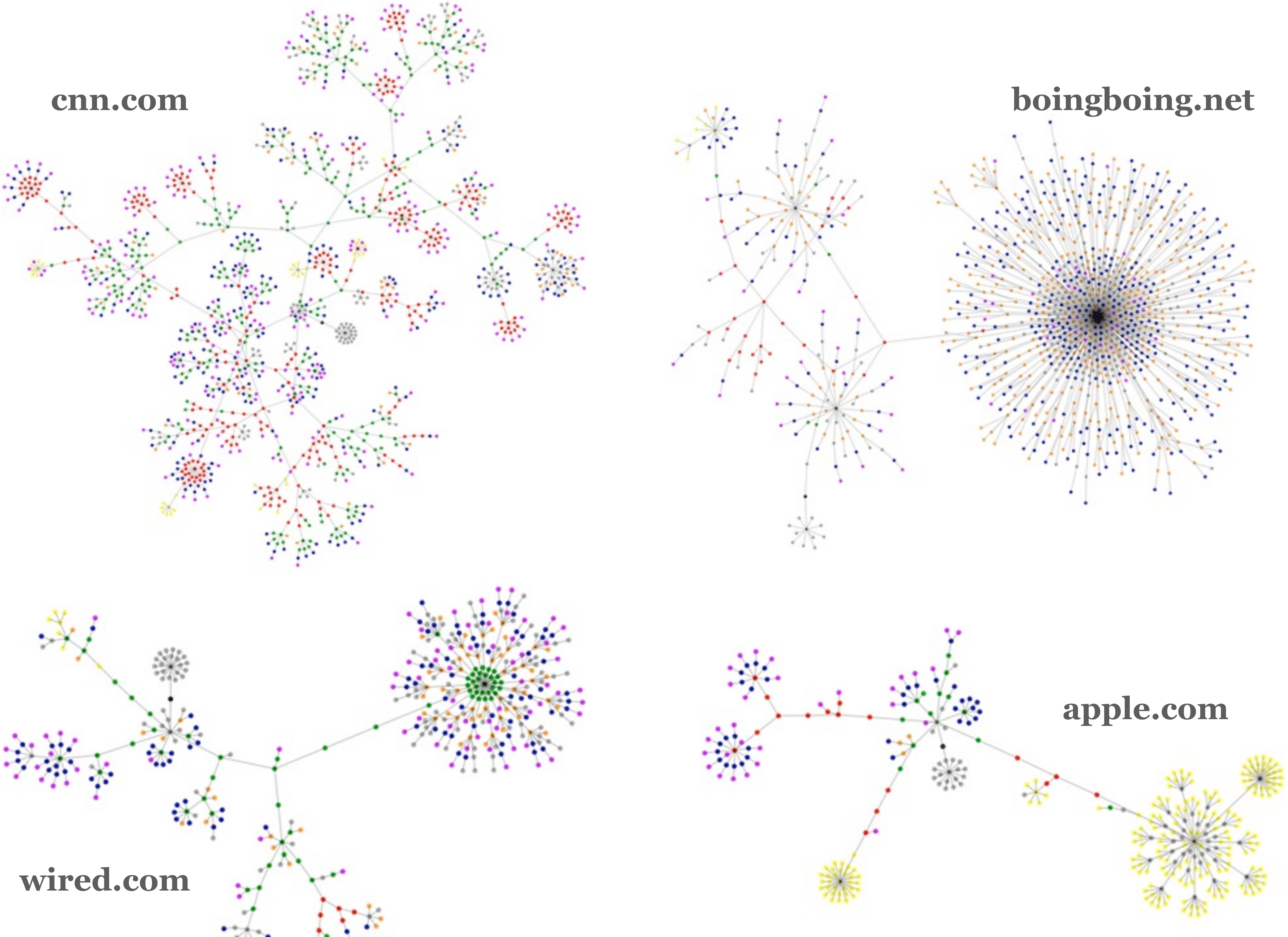
www.cs.utah.edu



blue: for links (the A tag)
red: for tables (TABLE, TR and TD tags)
green: for the DIV tag
violet: for images (the IMG tag)
yellow: for forms (FORM, INPUT, TEXTAREA, SELECT and OPTION tags)
orange: for linebreaks and blockquotes (BR, P, and BLOCKQUOTE tags)
black: the HTML tag, the root node
gray: all other tags

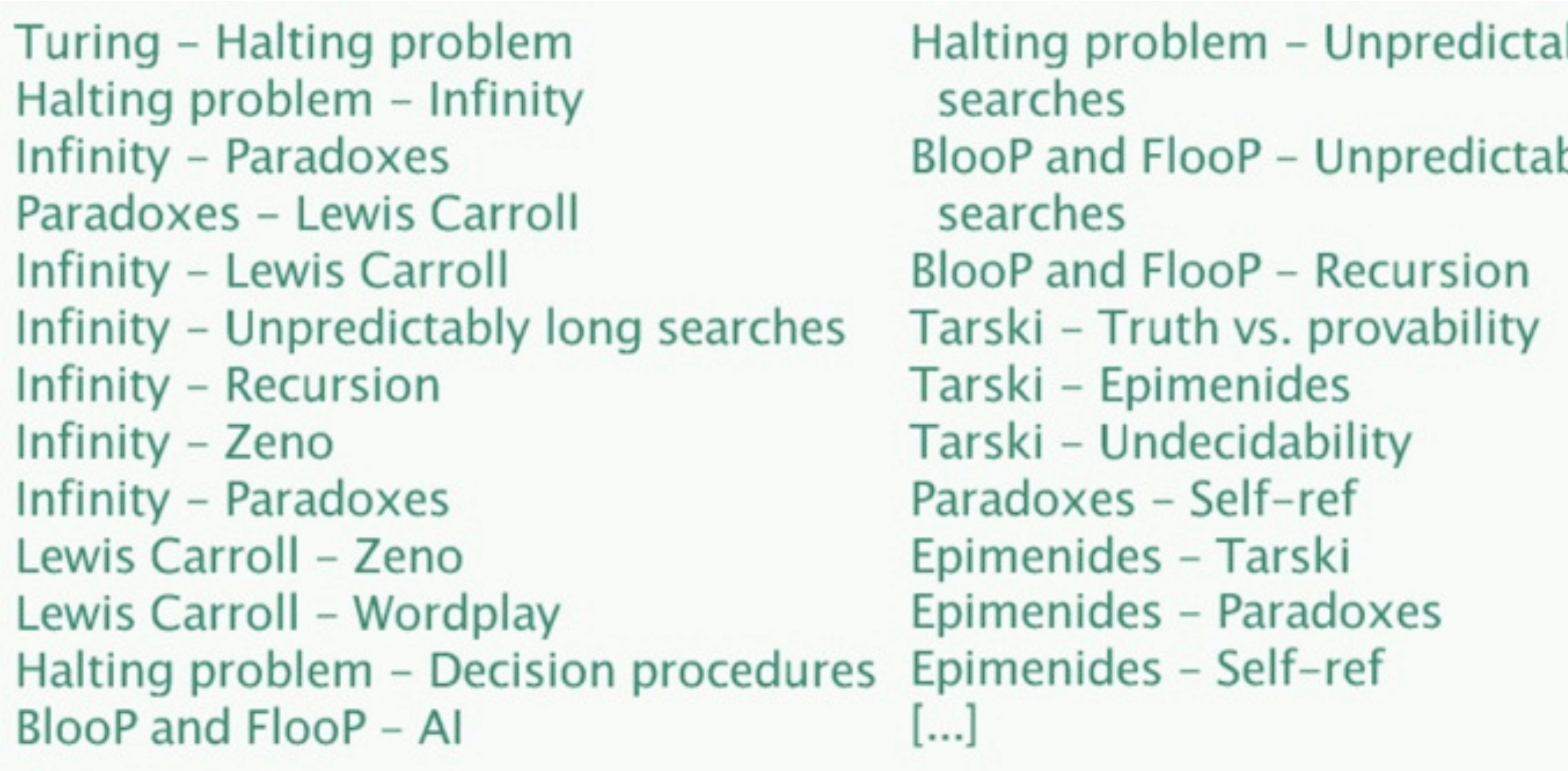


Webpages as Graphs



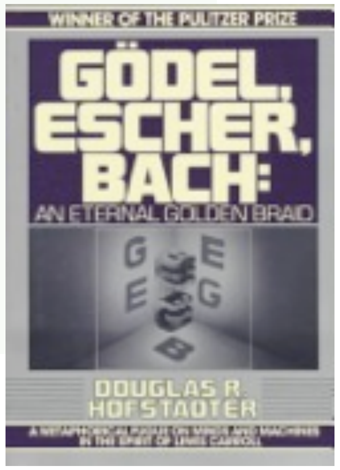
Webpages as Graphs

TOPIC GRAPH



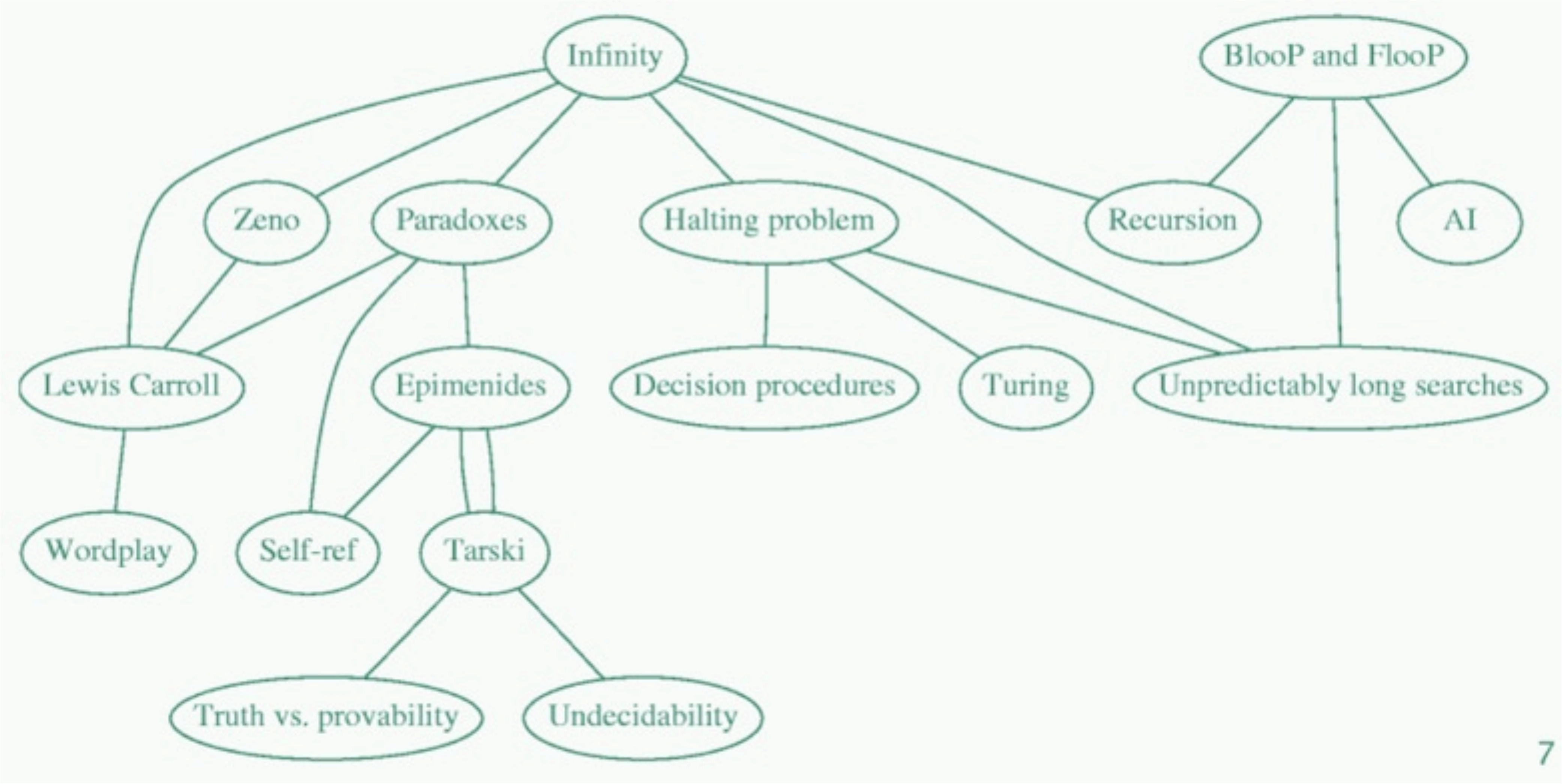
Turing – Halting problem
Halting problem – Infinity
Infinity – Paradoxes
Paradoxes – Lewis Carroll
Infinity – Lewis Carroll
Infinity – Unpredictably long searches
Infinity – Recursion
Infinity – Zeno
Infinity – Paradoxes
Lewis Carroll – Zeno
Lewis Carroll – Wordplay
Halting problem – Decision procedures
BlooP and FlooP – AI

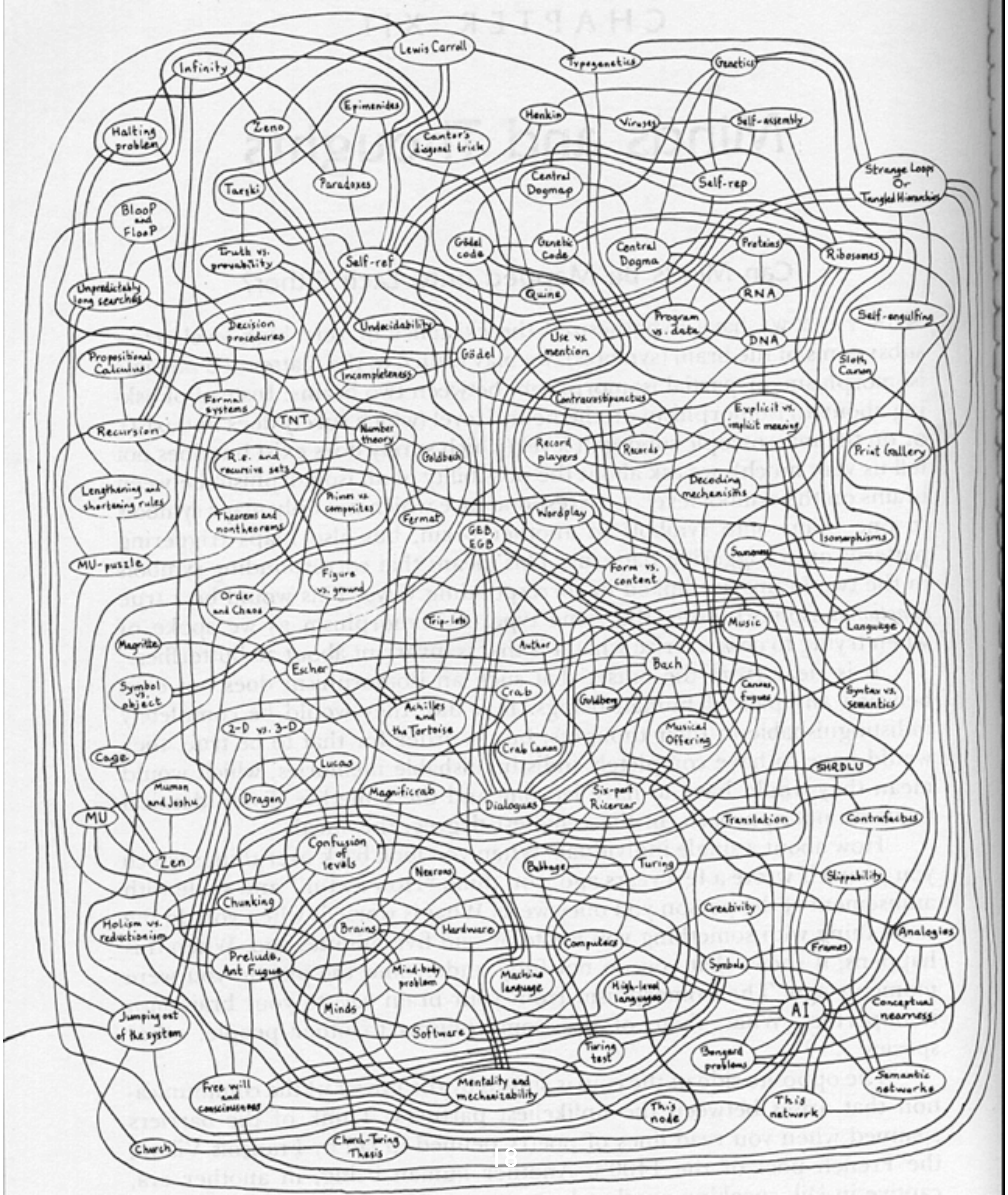
Halting problem – Unpredictably long searches
BlooP and FlooP – Unpredictably long searches
BlooP and FlooP – Recursion
Tarski – Truth vs. provability
Tarski – Epimenides
Tarski – Undecidability
Paradoxes – Self-ref
Epimenides – Tarski
Epimenides – Paradoxes
Epimenides – Self-ref
[...]



Hofstadter 1989

TOPIC GRAPH


Infinity
BlooP and FlooP
Zeno
Paradoxes
Halting problem
Recursion
AI
Lewis Carroll
Epimenides
Decision procedures
Turing
Unpredictably long searches
Wordplay
Self-ref
Tarski
Truth vs. provability
Undecidability
7



Hofstadter 1989

- **definitions**

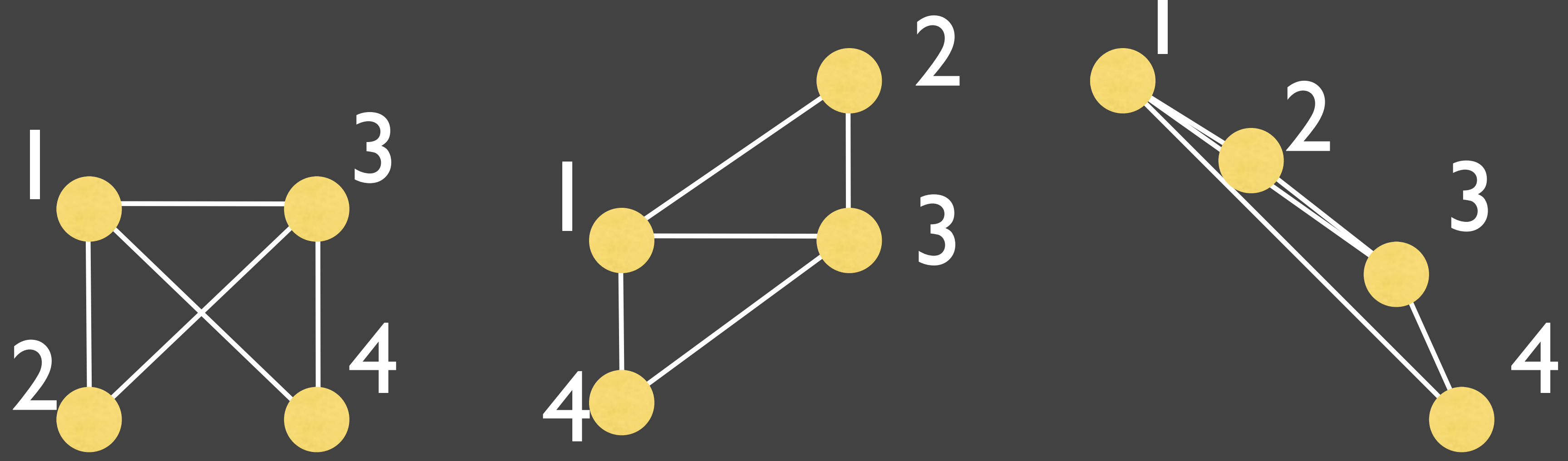
- **visualizing trees**
  - indented
  - node link
  - enclosure
  - layered

- **visualizing graphs**
  - node link
  - matrix
  - network summarizations

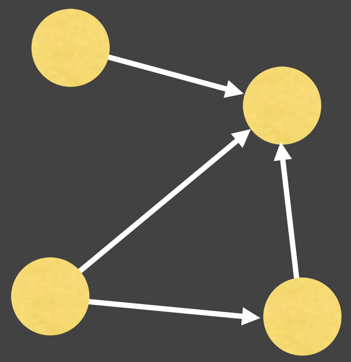
DEFINITIONS

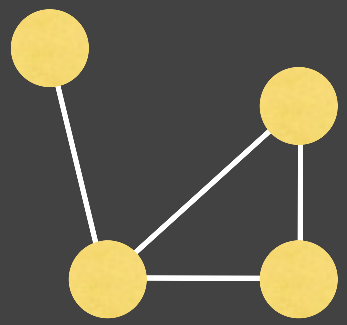
GRAPH

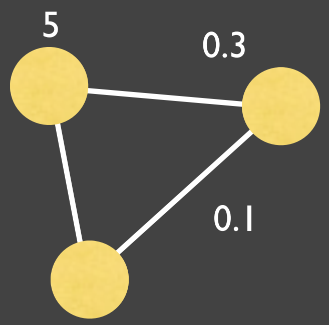
- **A graph *G* consists of a collection of vertices (or nodes) *V* and a set of edges *E*, consisting of vertex pairs.**
- **An edge e_{xy} = (x,y) connects two vertices *x* and *y*.**
    - for example: V={1,2,3,4}, E={(1,2),(1,3),(2,3),(3,4),(4,1)}



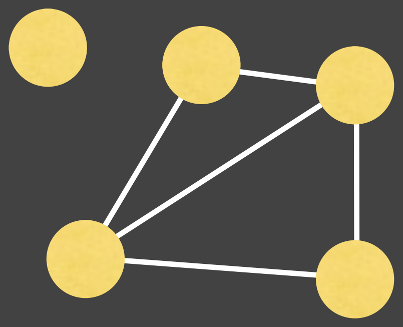
a bunch of definitions

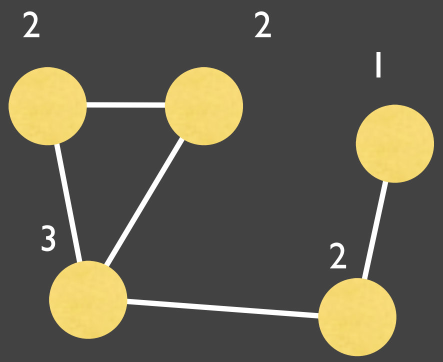

A directed graph


An undirected graph

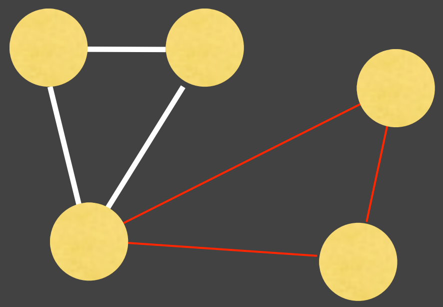


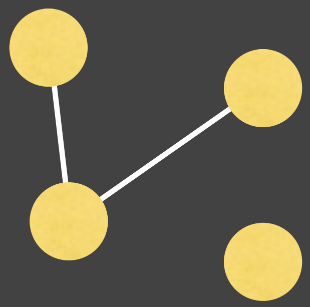
Weighted

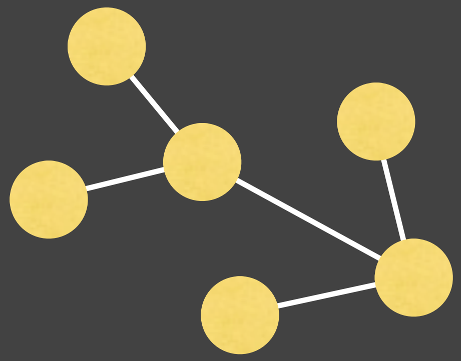

Unconnected

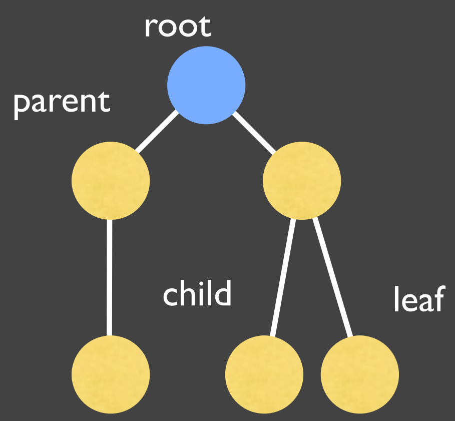


Node degrees

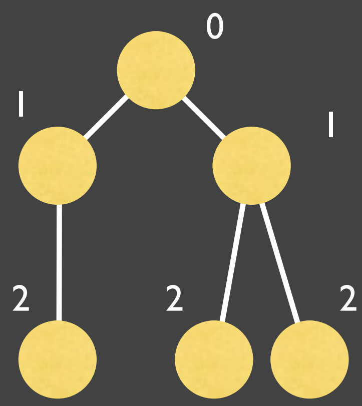

A cycle


An acyclic graph


A connected acyclic graph,
a.k.a. a **tree**



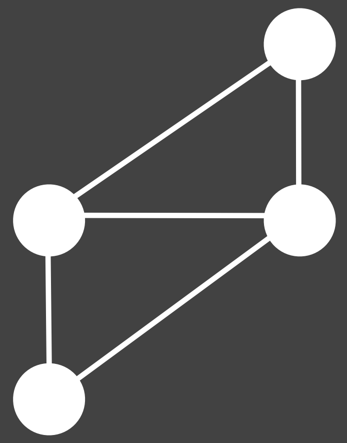
A rooted tree
or hierarchy



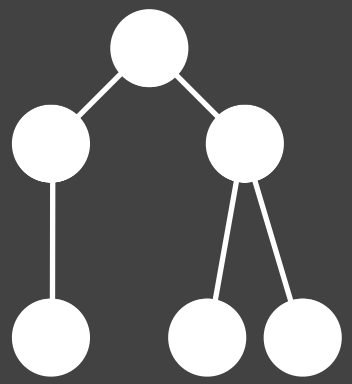
Node depths

GRAPHS & TREES

- **graphs**
  - model relations amount data
  - *nodes* and *edges*



- **trees**
  - graphs with hierarchical structure
    - ***connected graph with N-1 edges***
  - nodes as *parents* and *children*



SPATIAL LAYOUT

- **primary concern of graph drawing is the spatial layout of nodes and edges**
- **often (but not always) the goal is to effectively depict the graph structure**
  - connectivity, path-following
  - network distance
  - clustering
  - ordering (e.g., hierarchy level)

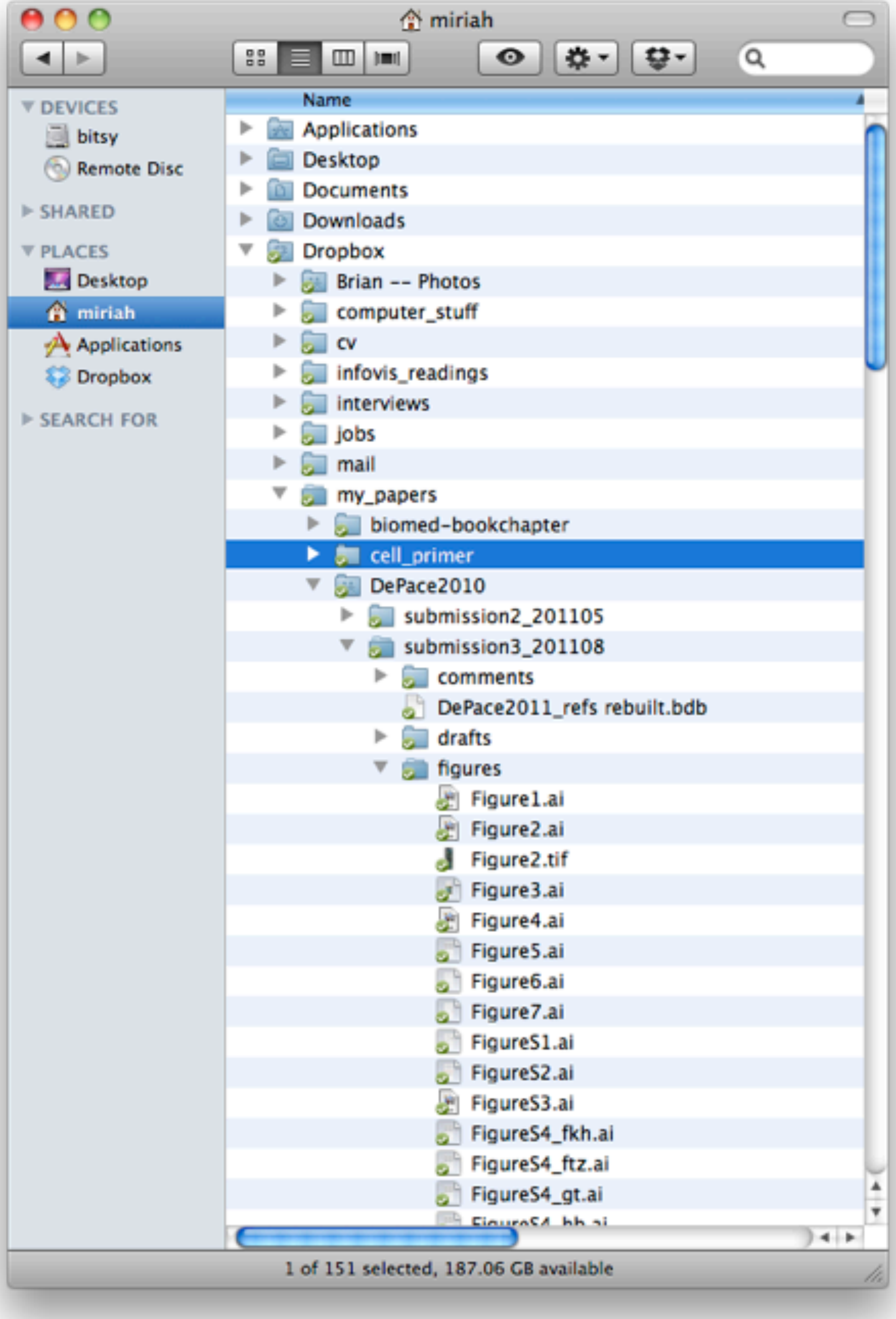
VISUALIZING TREES

ROOTED TREES

- **recursion makes it elegant and fast to draw trees**
- **approaches:**
  - indentation
  - node link
  - enclosure
  - layering

INDENTATION

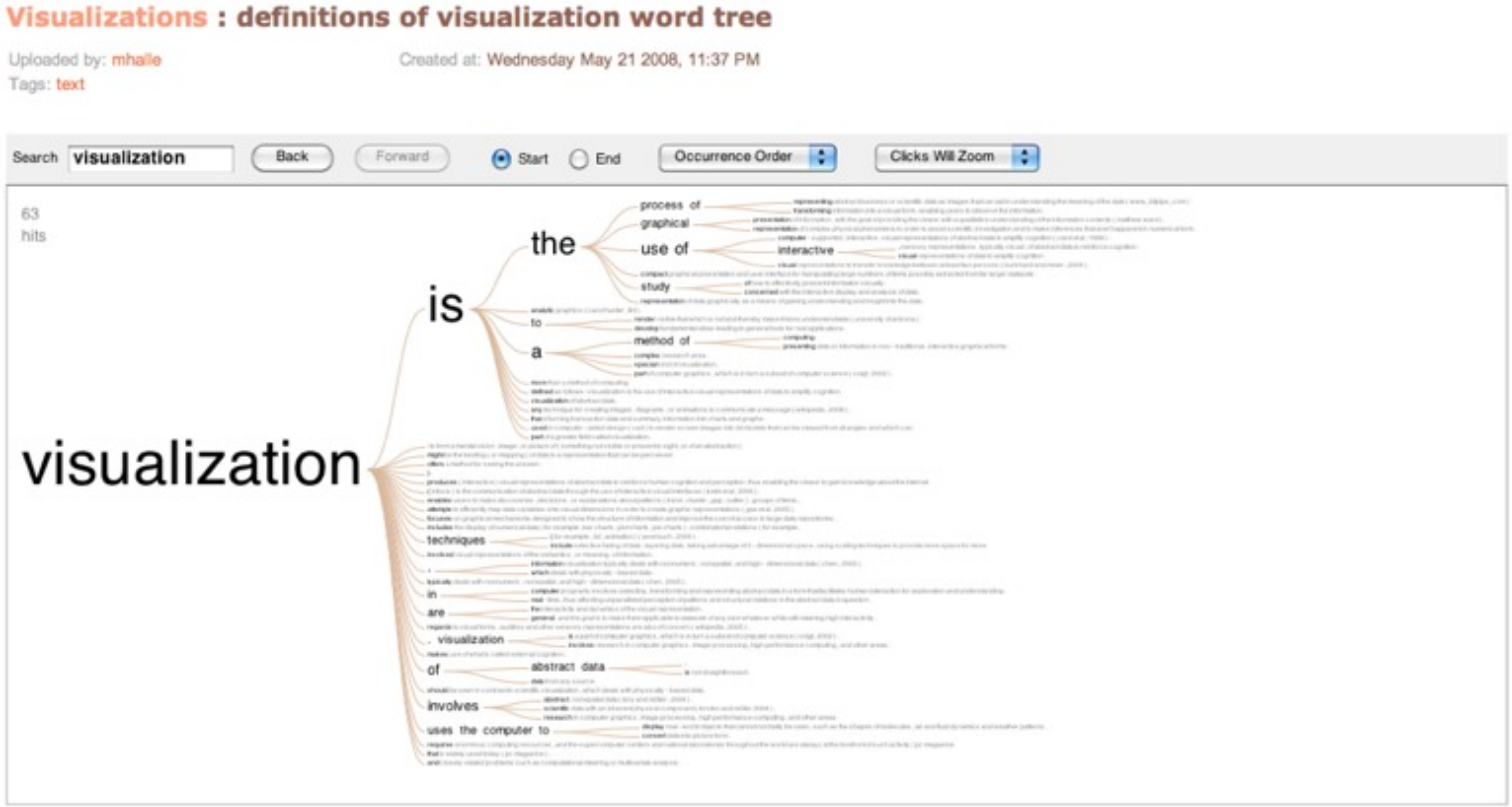
- place all items along vertically spaced rows
- indentation used to show parent/child relationships
- commonly used as a component in an interface
- breadth and depth contend for space
- often requires a great deal of scrolling



- Applications
- Desktop
- Documents
- Downloads
- Dropbox
  - Brian -- Photos
  - computer_stuff
  - cv
  - infovis_readings
  - interviews
  - jobs
  - mail
  - my_papers
    - biomed-bookchapter
    - cell_primer
    - DePace2010
      - submission2_201105
      - submission3_201108
        - comments
        - DePace2011_refs rebuilt.bdb
        - drafts
        - figures
          - Figure1.ai
          - Figure2.ai
          - Figure2.tif
          - Figure3.ai
          - Figure4.ai
          - Figure5.ai
          - Figure6.ai
          - Figure7.ai
          - FigureS1.ai
          - FigureS2.ai
          - FigureS3.ai
          - FigureS4_fkh.ai
          - FigureS4_ftz.ai
          - FigureS4_gt.ai
          - FigureS4_hb.ai

1 of 151 selected, 187.06 GB available

WORD TREE



Visualizations : definitions of visualization word tree

Uploaded by: mhalle

Created at: Wednesday May 21 2008, 11:37 PM

Tags: text

Search: visualization

Back Forward ● Start ○ End

Occurrence Order

Clicks Will Zoom

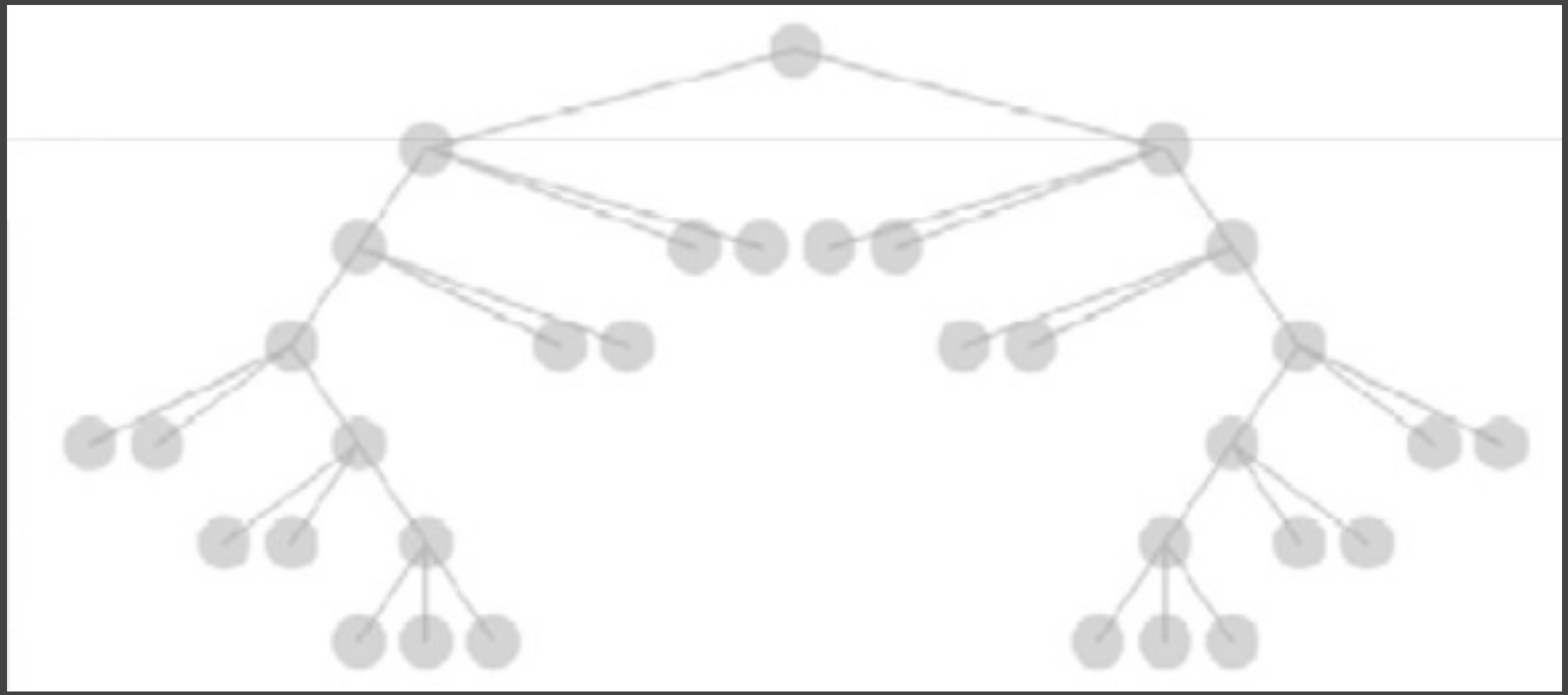
63 hits

NODE-LINK DIAGRAMS

- **nodes are distributed in space, connected by straight or curved lines**
- **typical approach is to use 2D space to break apart breadth and depth**
- **often space is used to communicate hierarchical orientation**

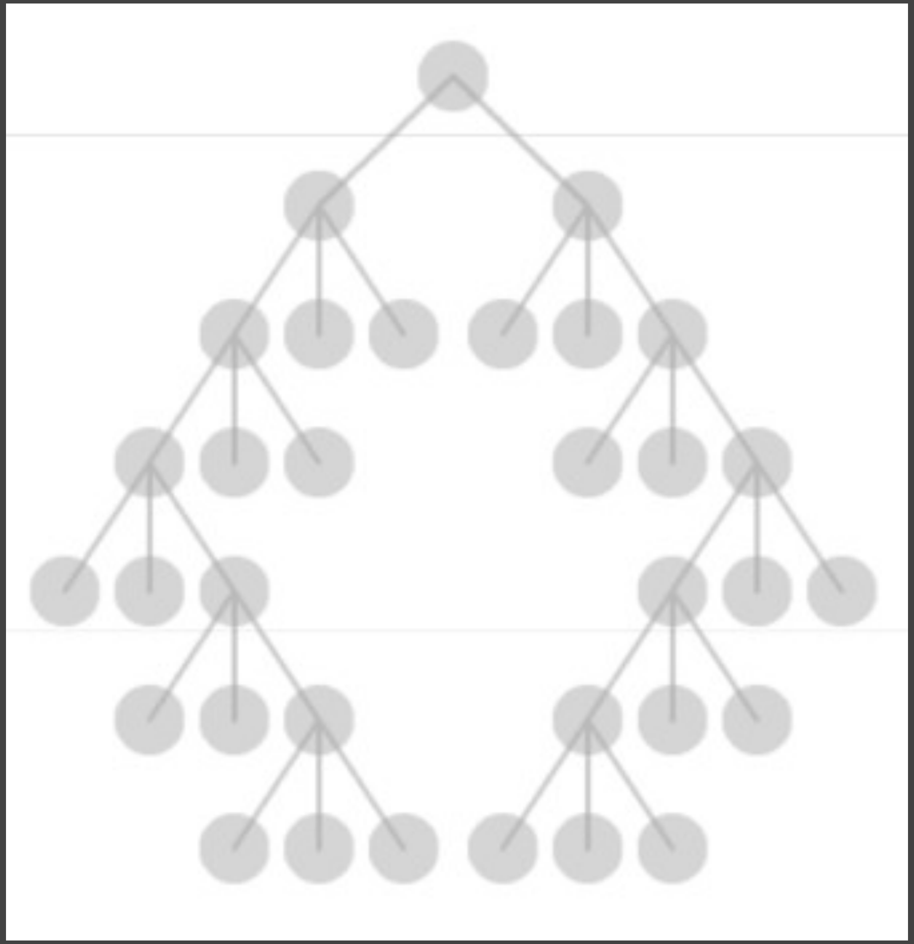
BASIC APPROACH

- **repeatedly divide space for subtrees by leaf count**
  - breadth of tree along one dimension
  - depth along the other dimension
- **problem: exponential growth of breadth**



REINGOLD-TILFORD type layouts

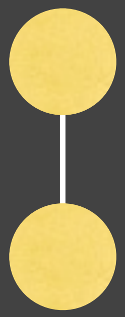
- **goal**
  - make smarter use of space
  - maximize density and symmetry
- **design concerns**
  - clearly encode depth level
  - no edge crossings
  - isomorphic subtrees drawn identically
  - compact
- **approach**
  - bottom up recursive approach
  - for each parent make sure every subtree is drawn
  - pack subtrees as closely as possible
  - center parent over subtrees



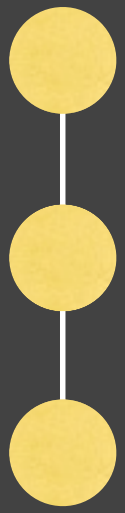
REINGOLD-TILFORD

REINGOLD-TILFORD

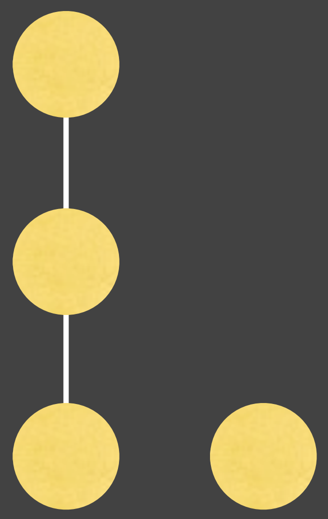
REINGOLD-TILFORD



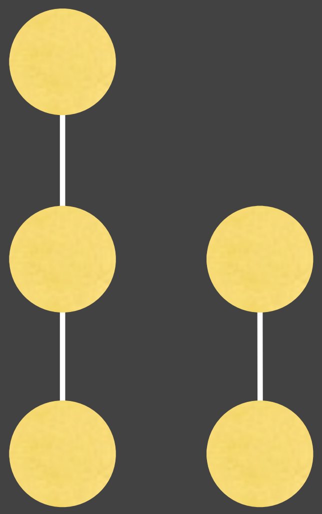
REINGOLD-TILFORD



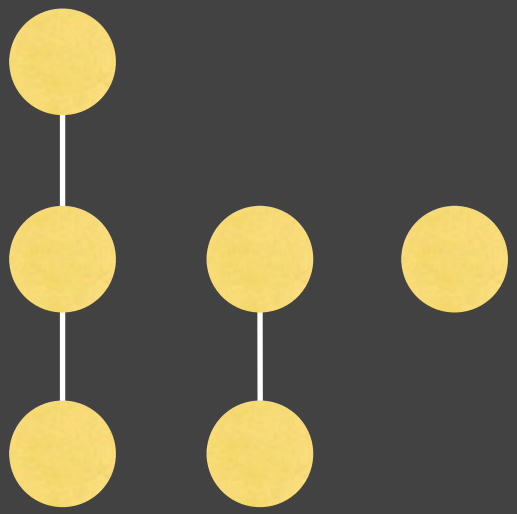
REINGOLD-TILFORD



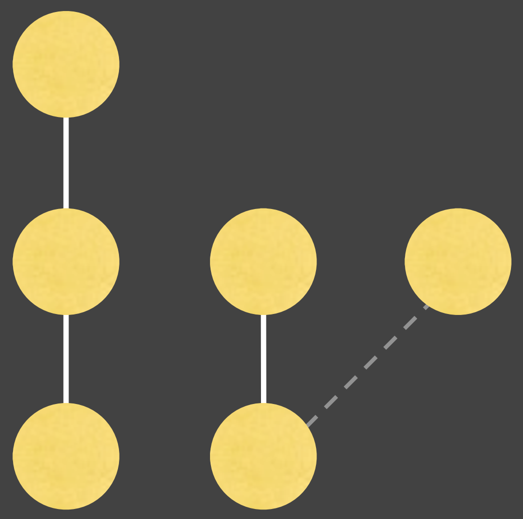
REINGOLD-TILFORD



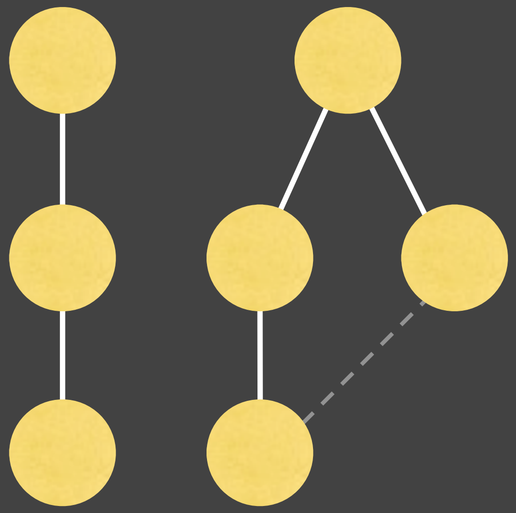
REINGOLD-TILFORD



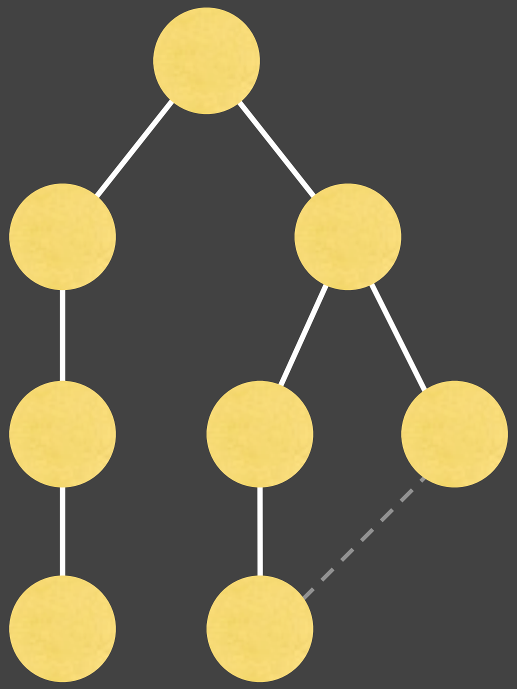
REINGOLD-TILFORD



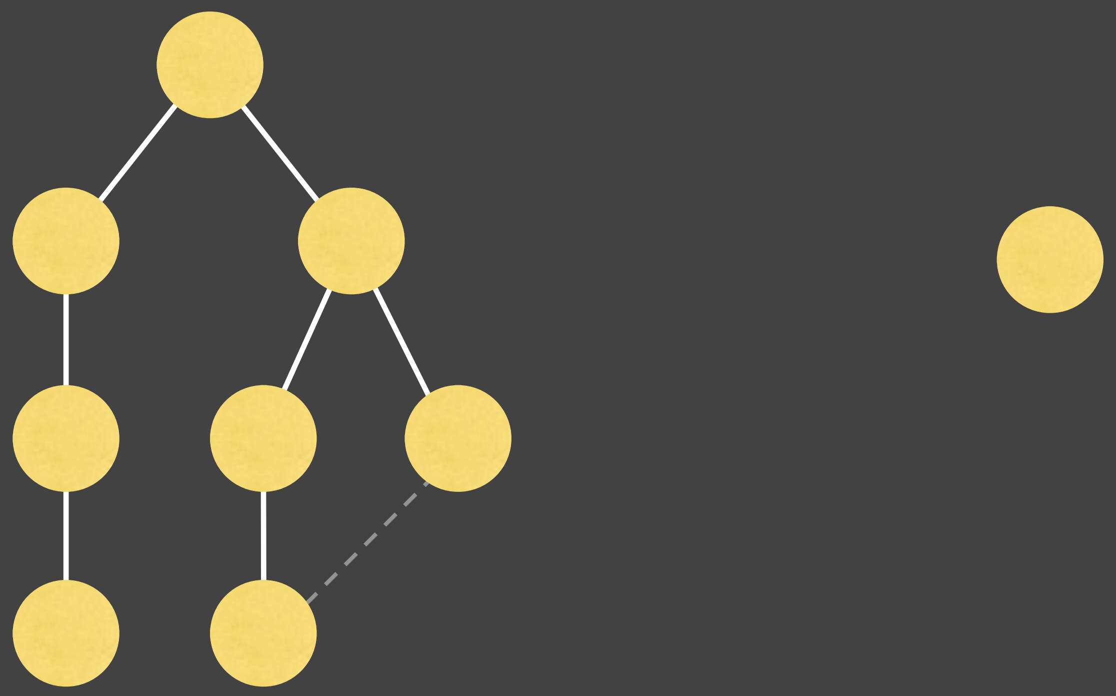
REINGOLD-TILFORD



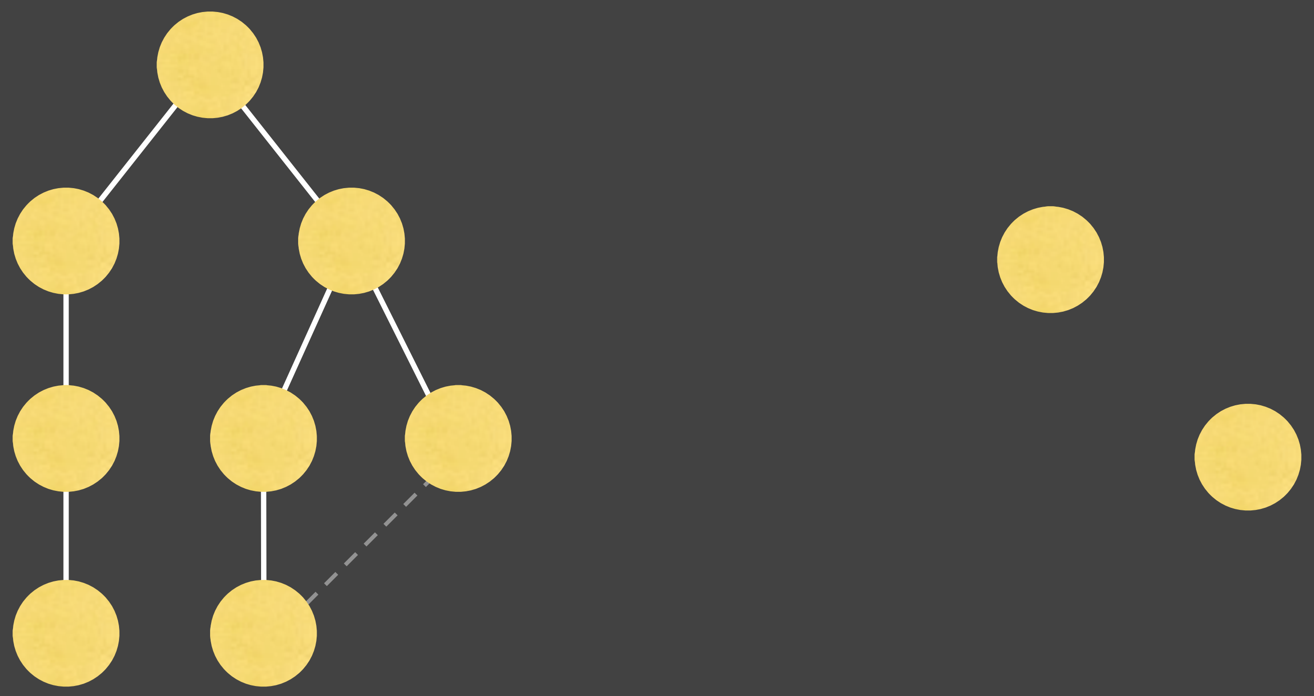
REINGOLD-TILFORD



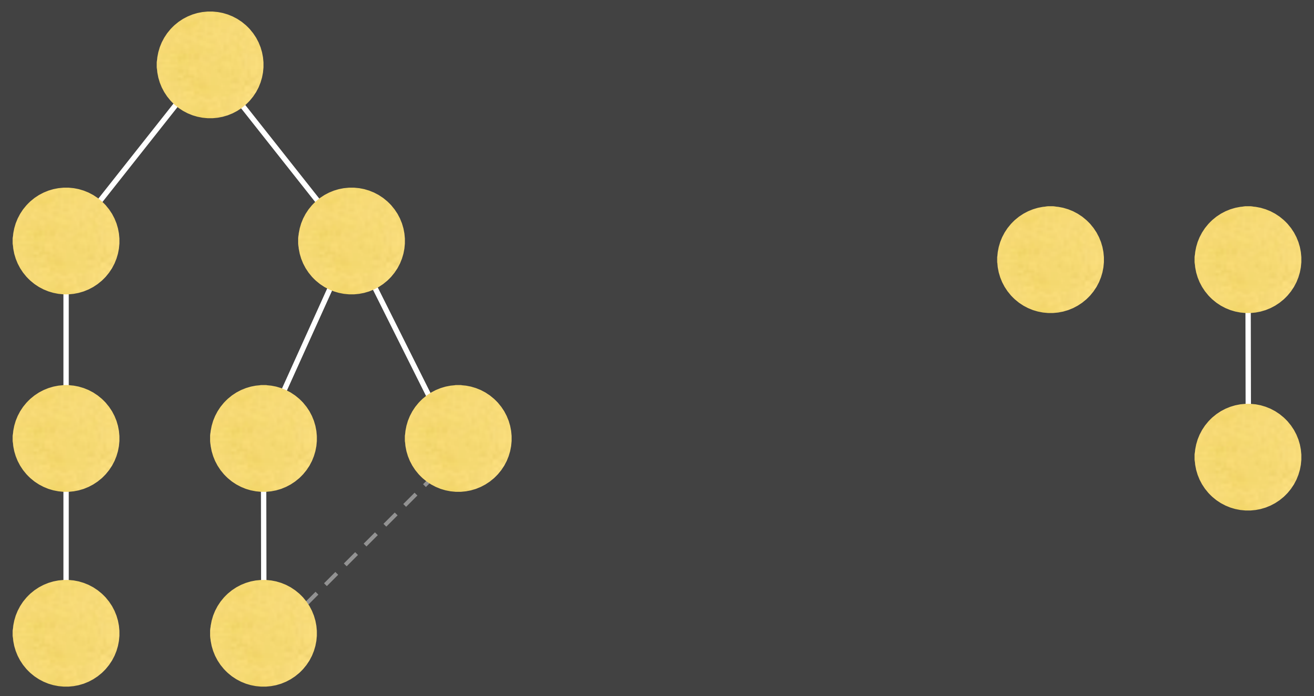
REINGOLD-TILFORD



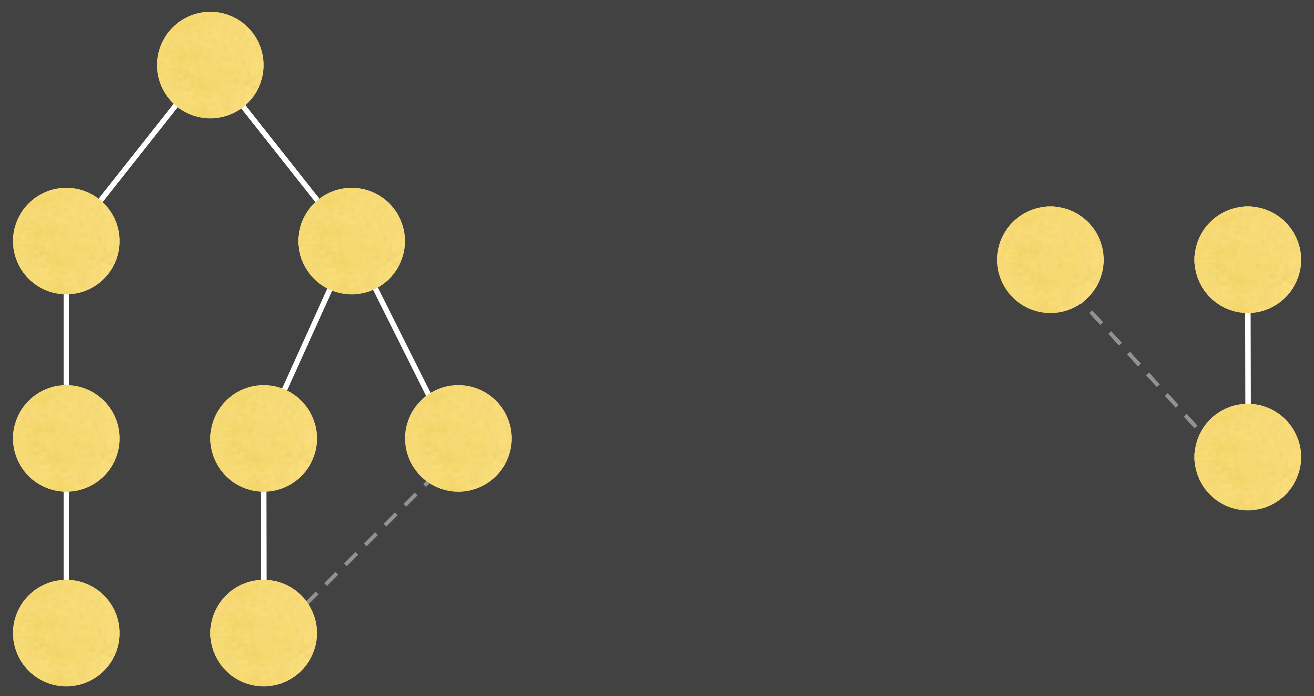
REINGOLD-TILFORD



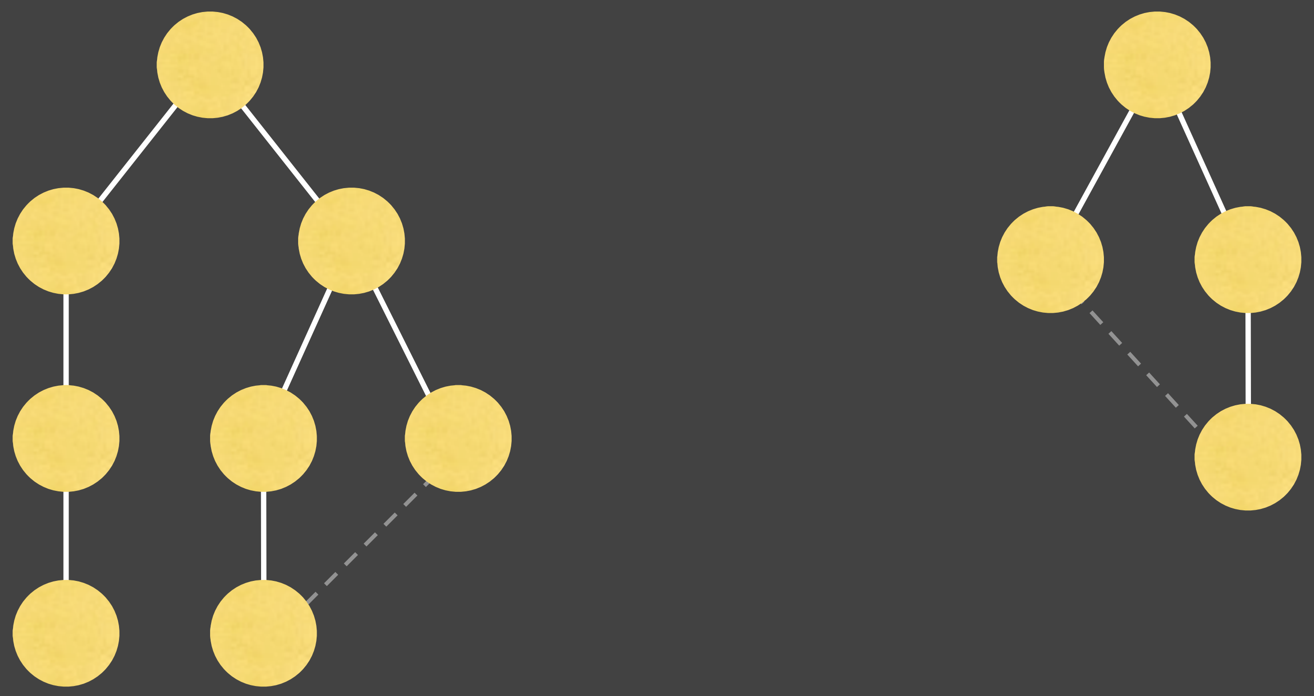
REINGOLD-TILFORD



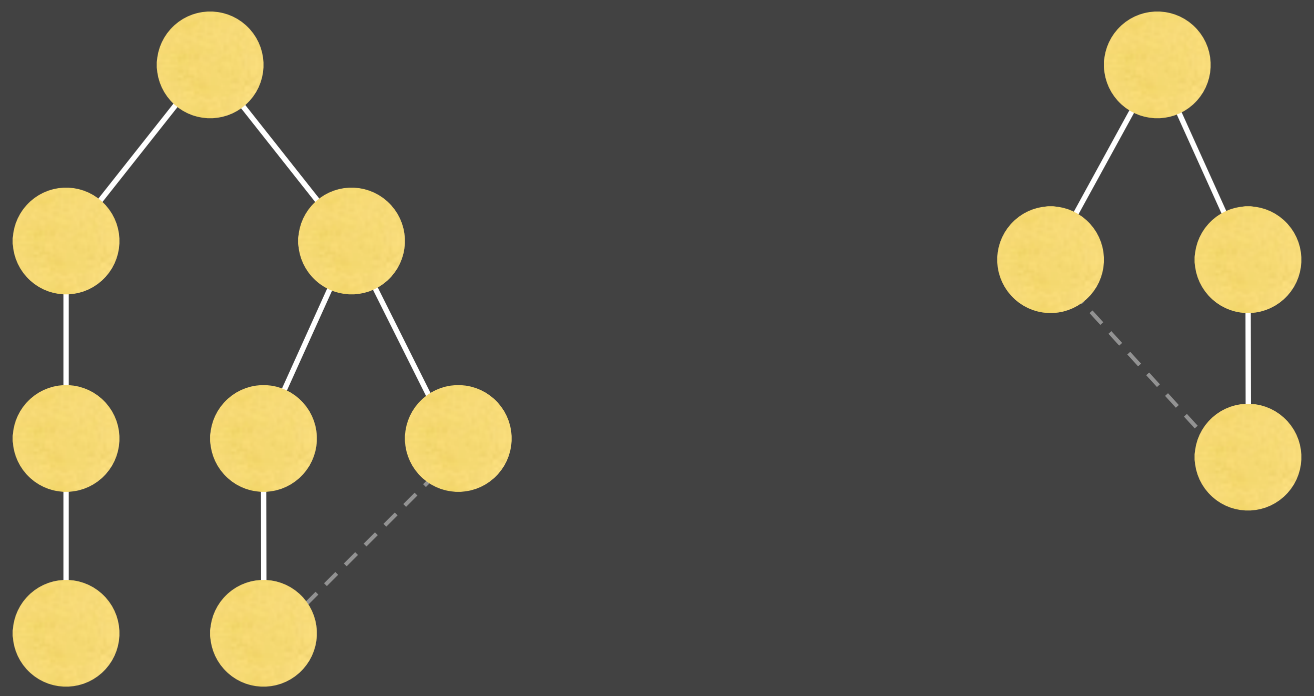
REINGOLD-TILFORD



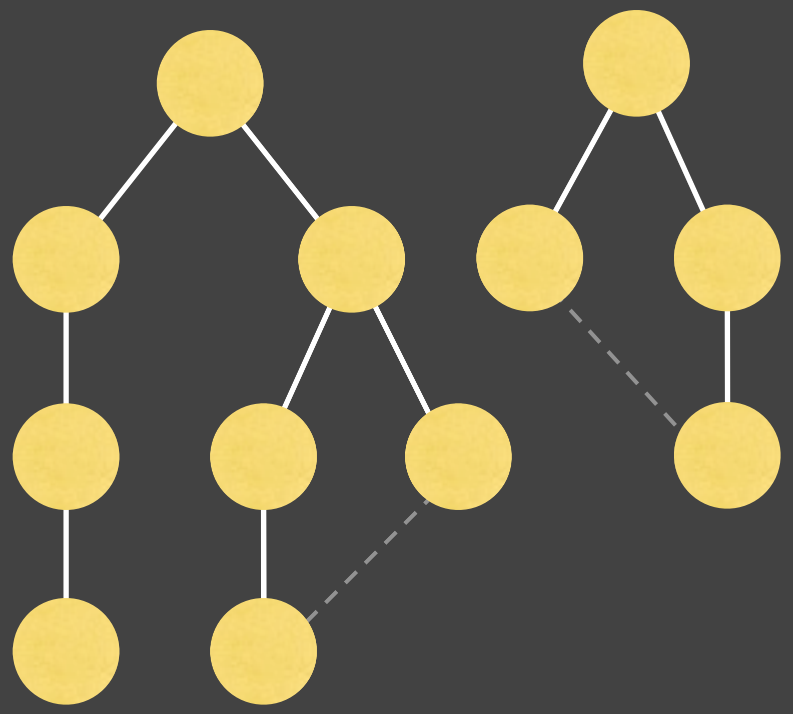
REINGOLD-TILFORD



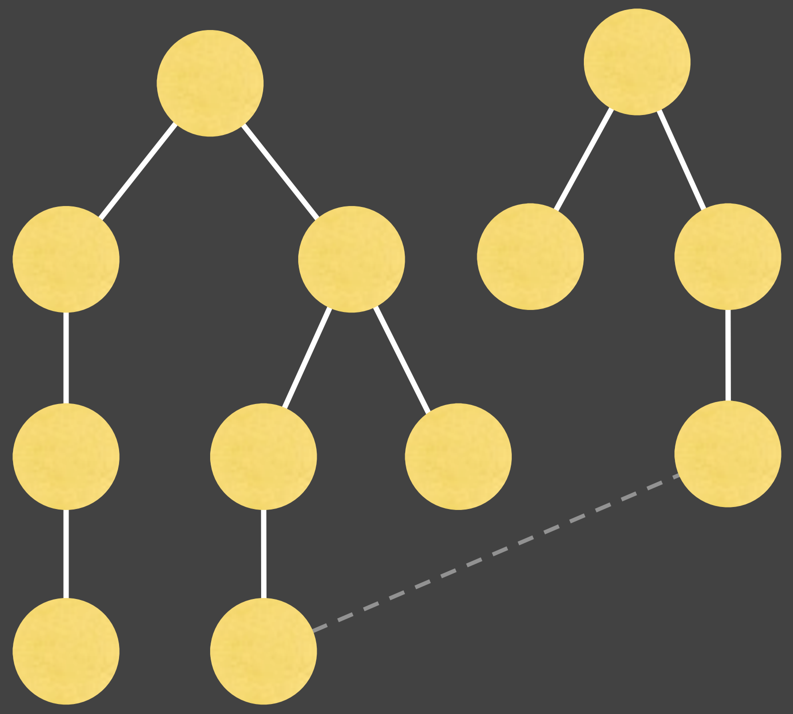
REINGOLD-TILFORD



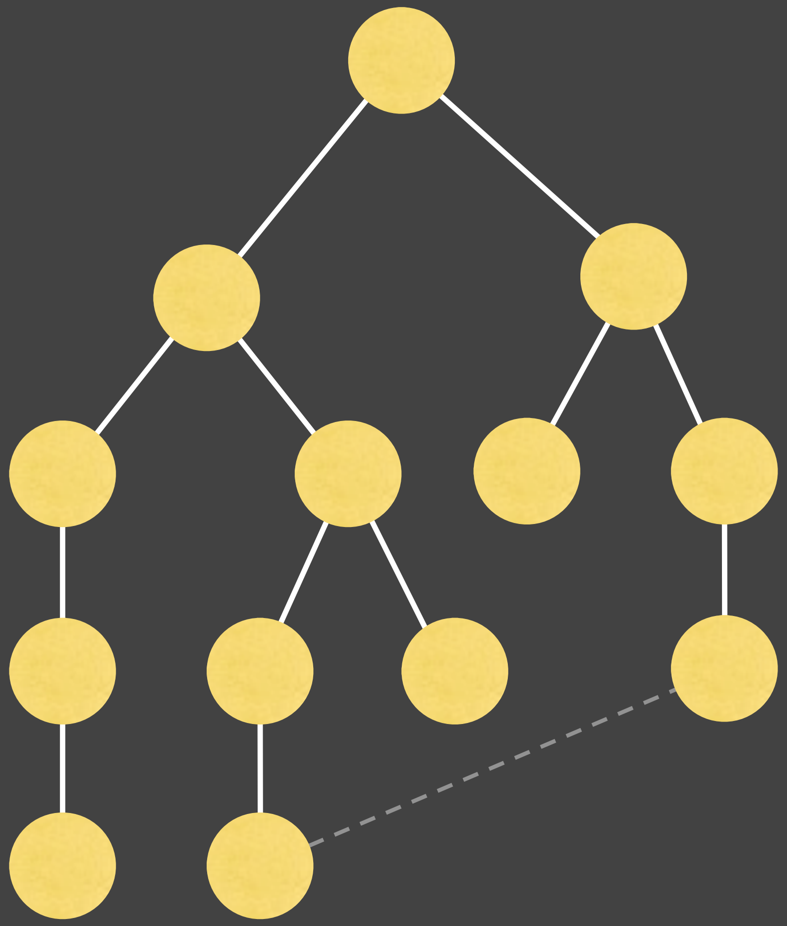
REINGOLD-TILFORD



REINGOLD-TILFORD

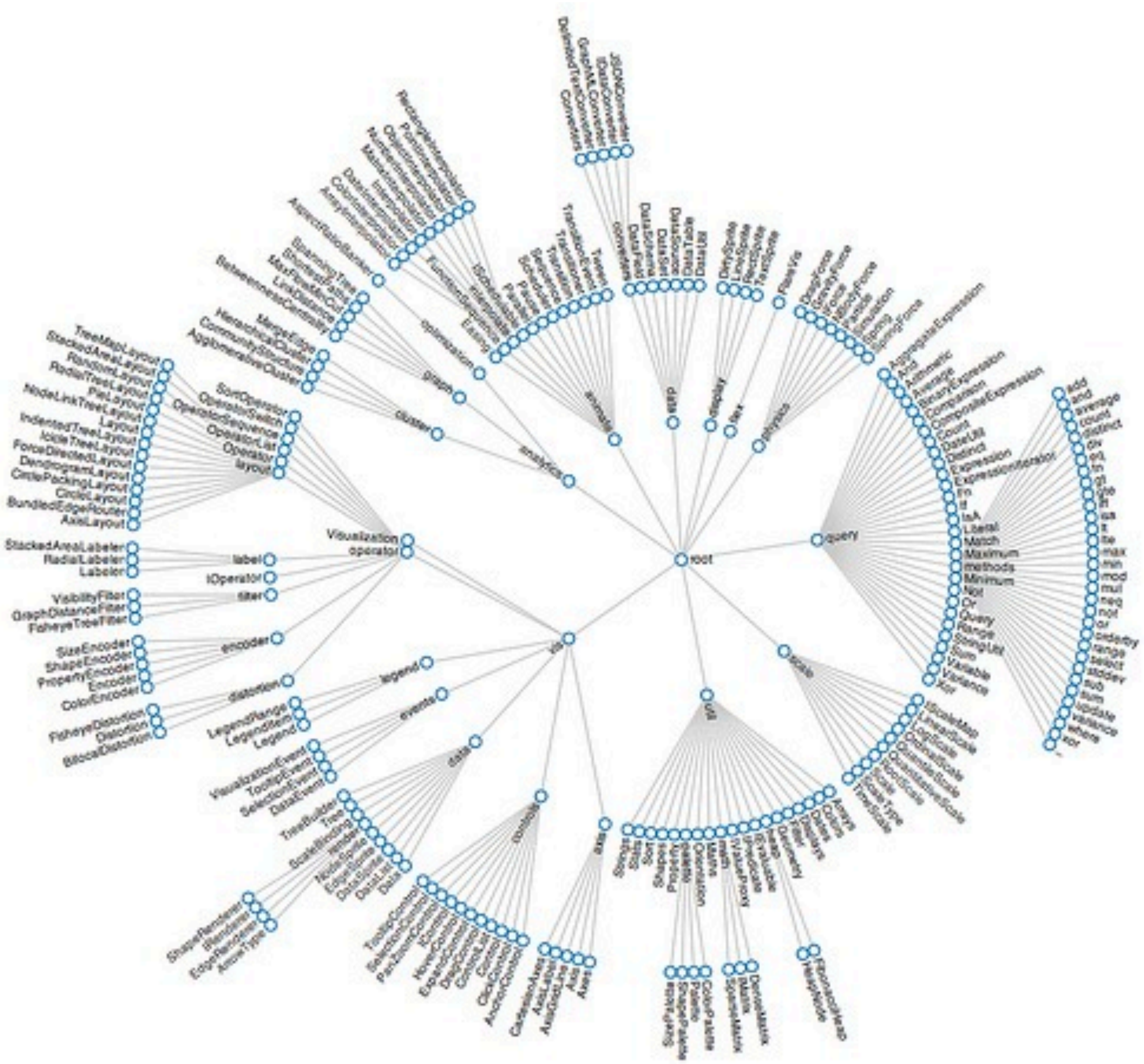


REINGOLD-TILFORD



RADIAL LAYOUT

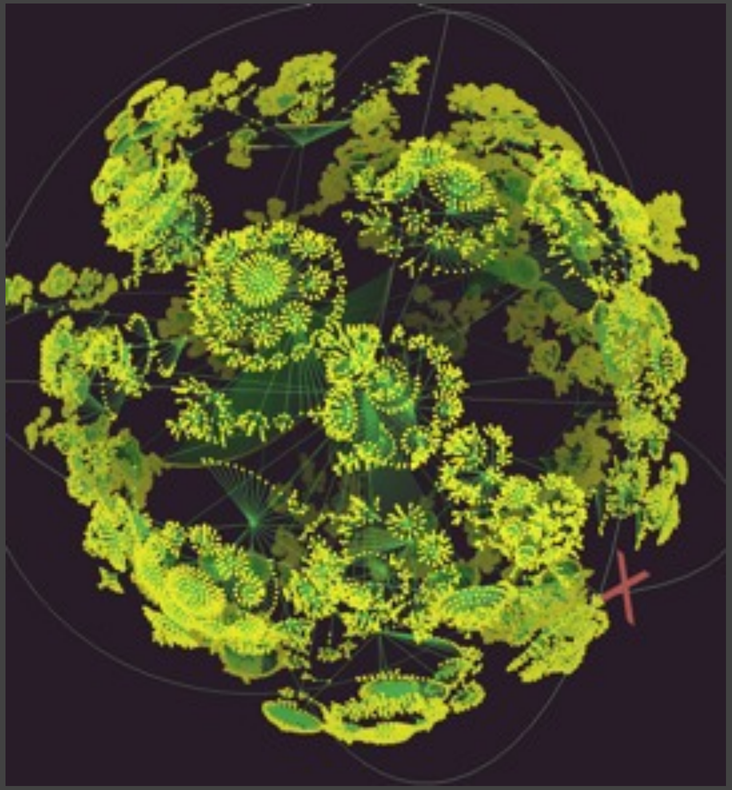
- **node-link diagram in polar coordinates**
- **radius encodes depth with root in center**
- **angular sectors assigned to subtrees**
- **Reingold-Tilford can be applied**



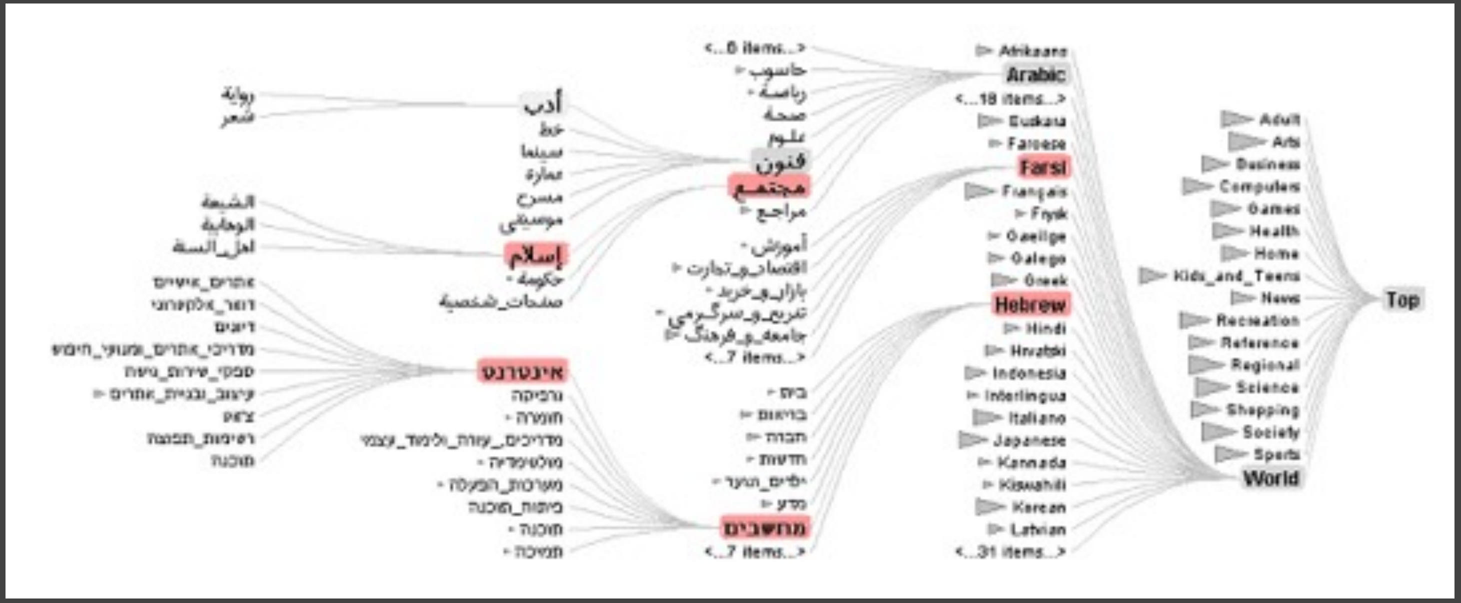
http://seeingcomplexity.wordpress.com/2011/02/05/hierarchical-edge-bundles/

NODE-LINK PROBLEMS

- **scale**
  - tree breadth often grows exponentially
  - quickly run out of space!
- **solutions**
  - hyperbolic layout
  - filtering
  - scrolling or panning
  - zooming



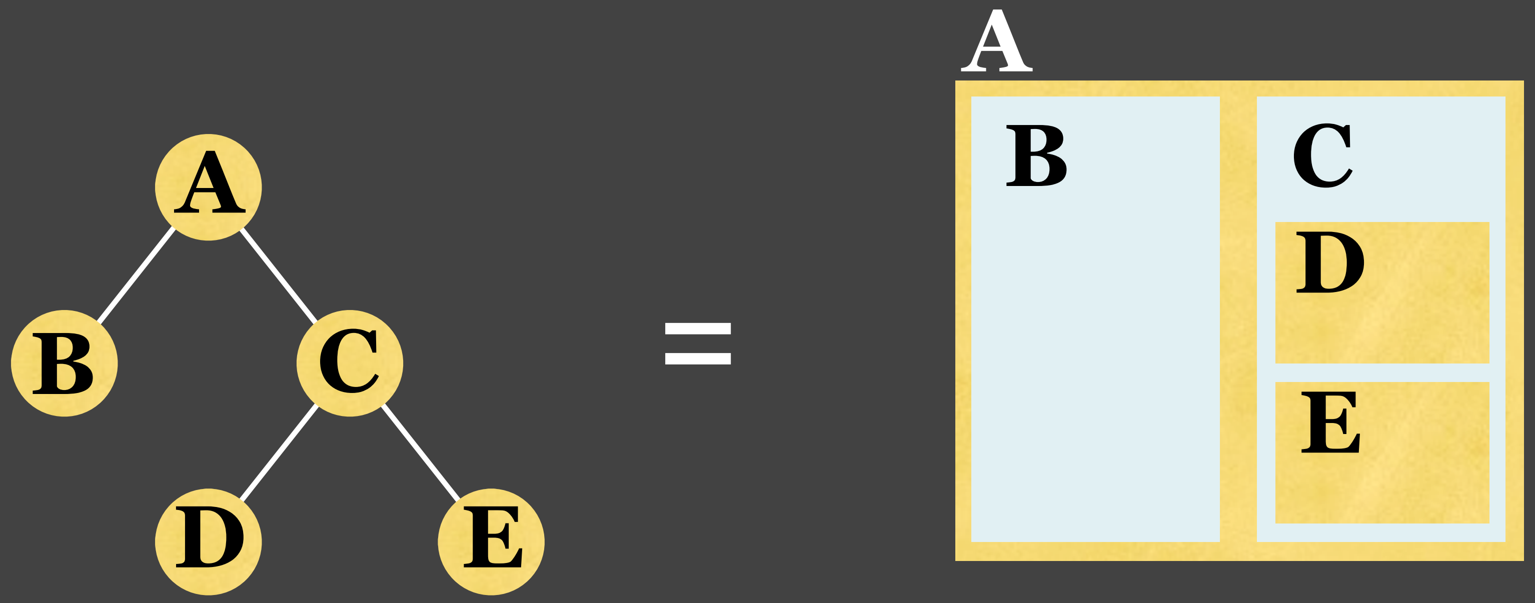
http://www.caida.org/tools/visualization/walrus/



http://vis.stanford.edu/papers/doitrees-revisited

ENCLOSURE DIAGRAMS

- **encode structure using spatial enclosure**
  - often referred to as *treemaps*
- **benefits**
  - provides single view of entire tree
  - easier to spot small / large nodes
- **problems**
  - difficult to accurately read depth



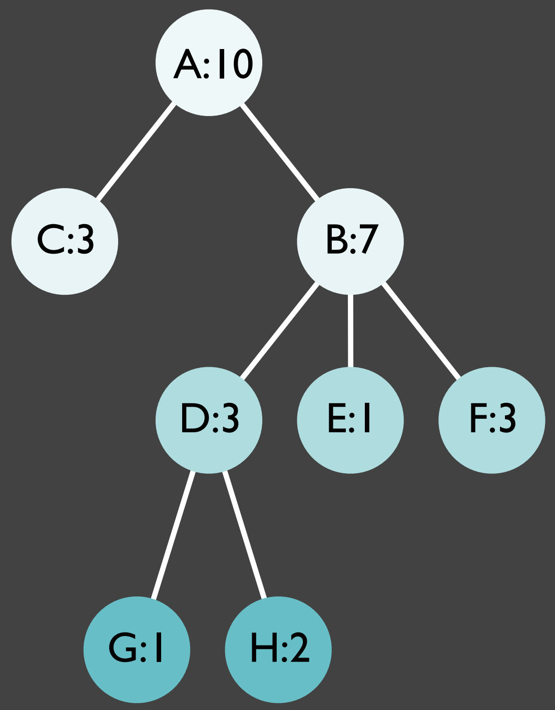
TREEMAPS

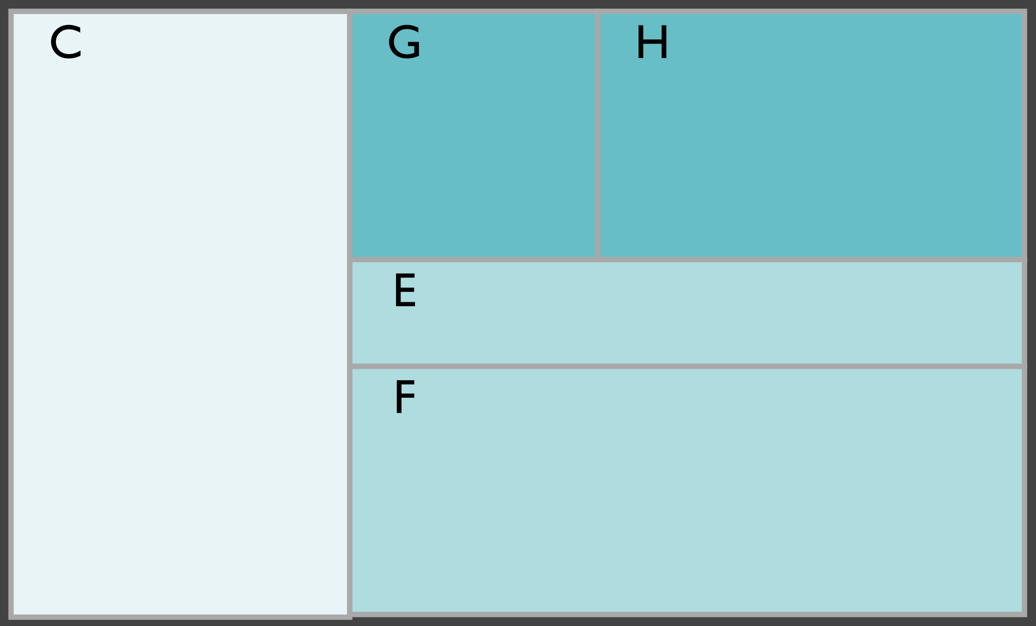
- **recursively fill space based on a size metric for nodes**

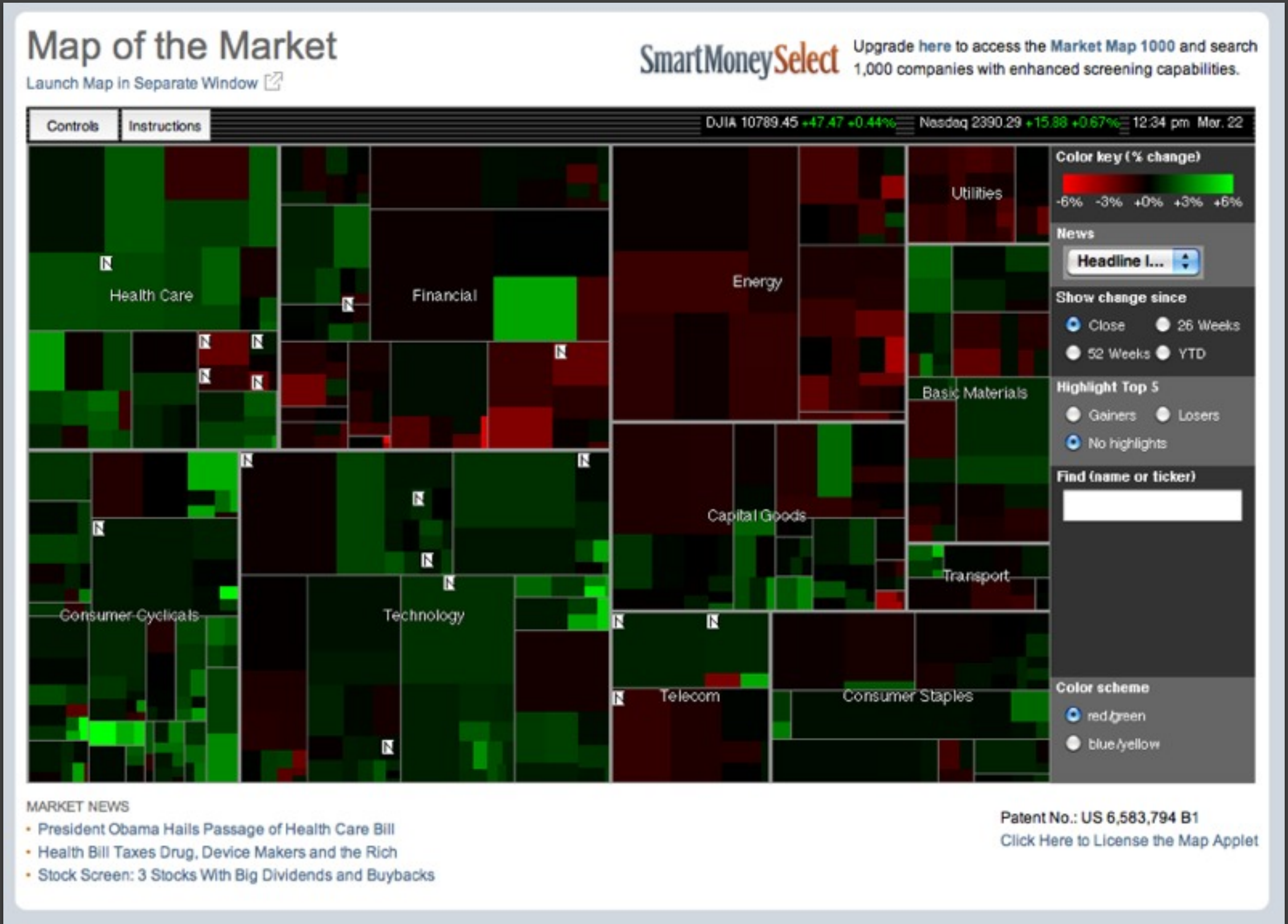
- **enclosure indicates hierarchy**

- **additional measures can control aspect ratio of cells**

- **most often use rectangles, but other shapes possible**
  - square, circle, voronoi tessellation







Map of the Market

Launch Map in Separate Window

SmartMoneySelect

Upgrade here to access the Market Map 1000 and search 1,000 companies with enhanced screening capabilities.

Controls | Instructions

DJIA 10789.45 +47.47 +0.44% Nasdaq 2390.29 +15.88 +0.67% 12:34 pm Mar. 22

News: Headline I...

Show change since
● Close ○ 26 Weeks
○ 52 Weeks ○ YTD

Highlight Top 5
○ Gainers ○ Losers
● No highlights

Find (name or ticker)

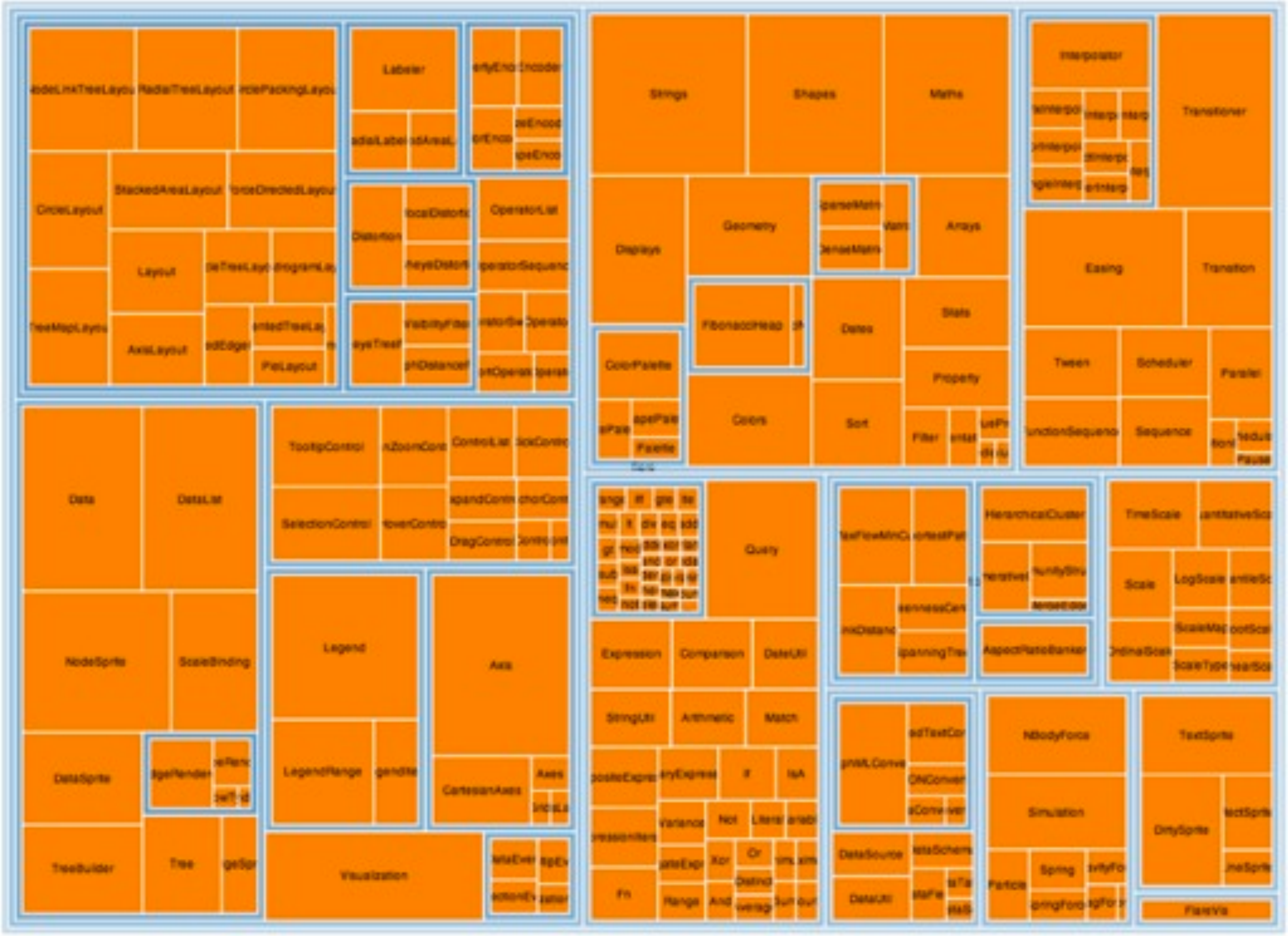
Color scheme
● red/green
○ blue/yellow

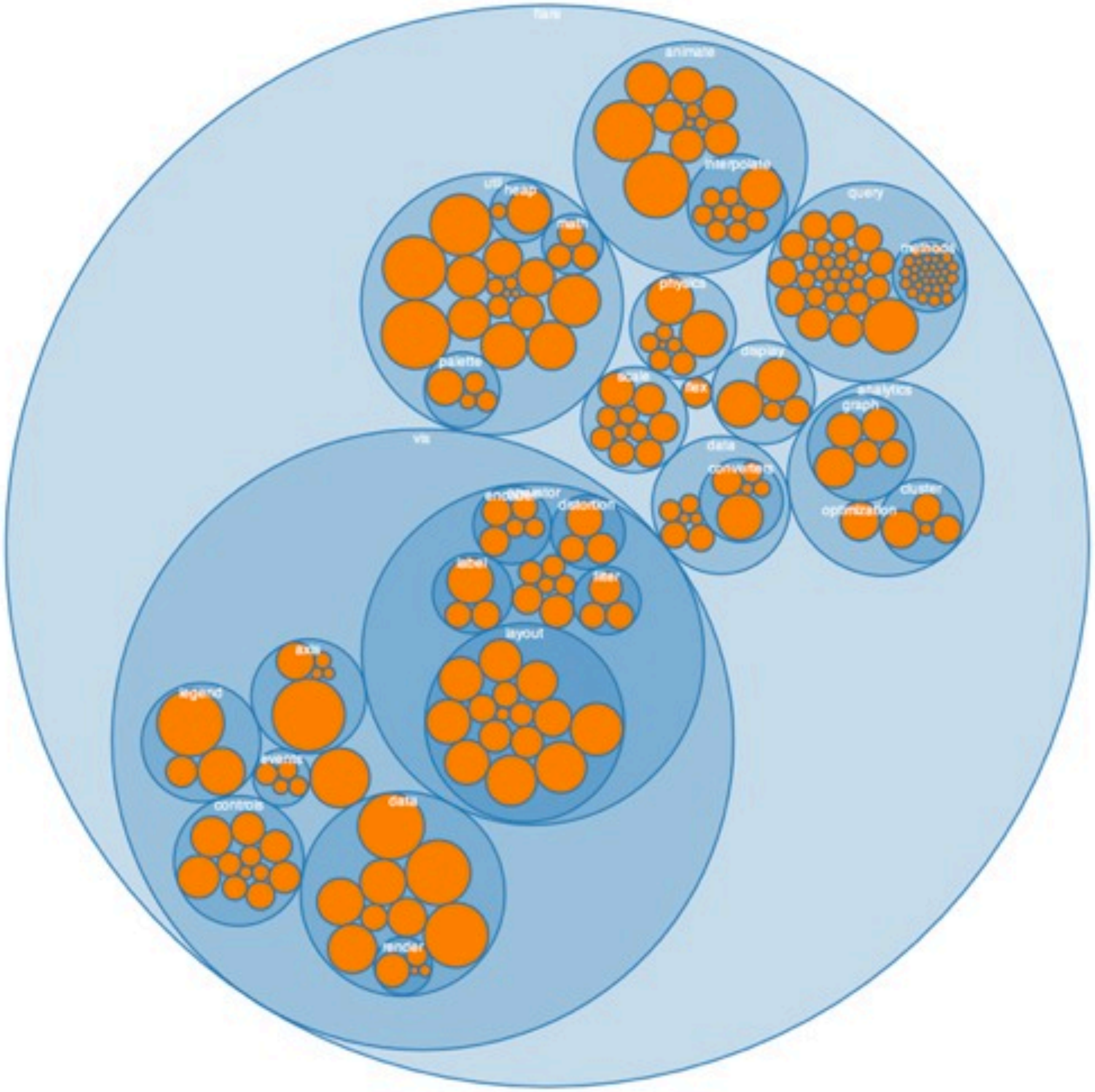
MARKET NEWS

- President Obama Hails Passage of Health Care Bill
- Health Bill Taxes Drug, Device Makers and the Rich
- Stock Screen: 3 Stocks With Big Dividends and Buybacks

Patent No.: US 6,583,794 B1

Click Here to License the Map Applet

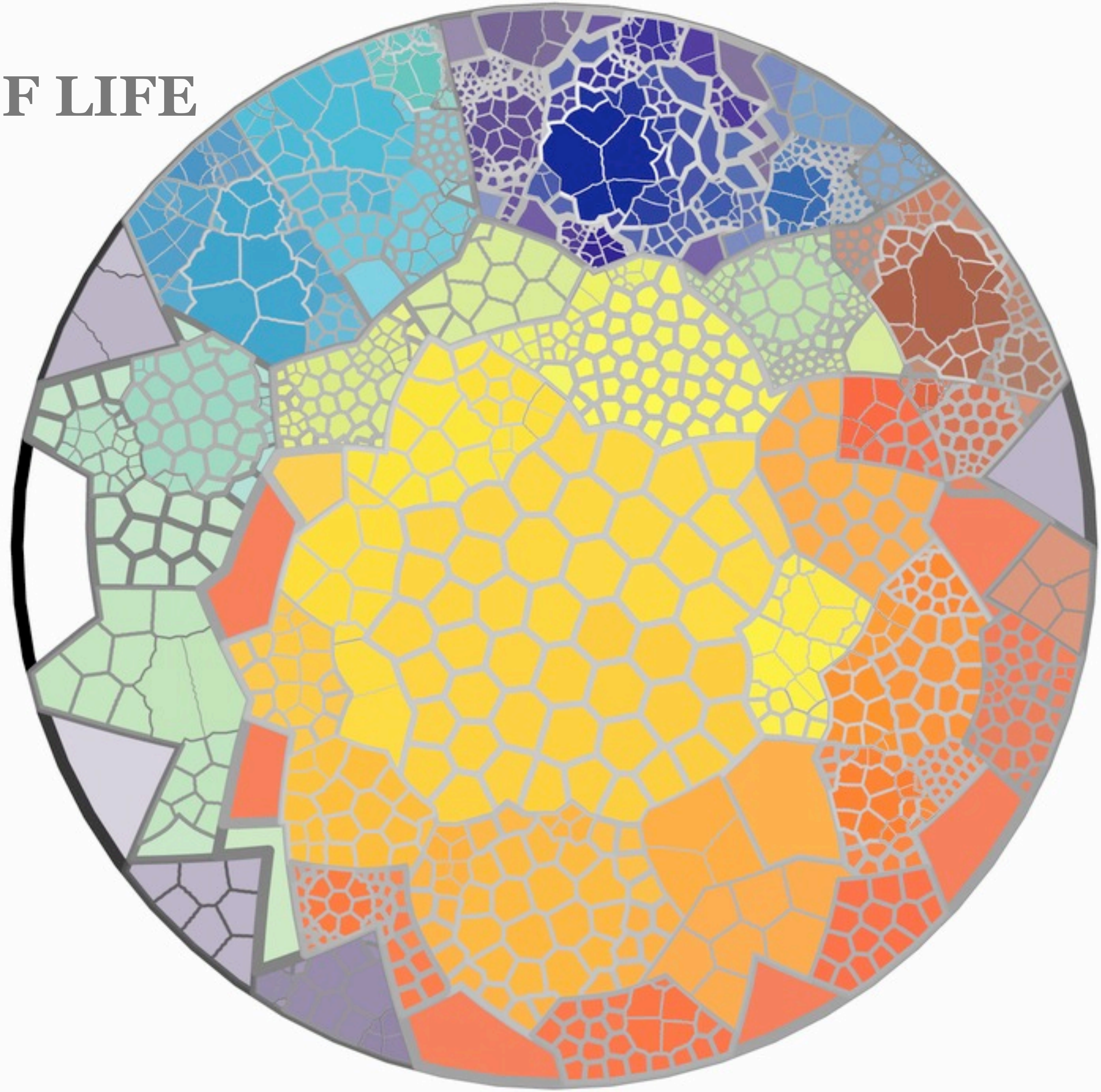




http://hci.stanford.edu/jheer/files/zoo/

TREE OF LIFE

mammals

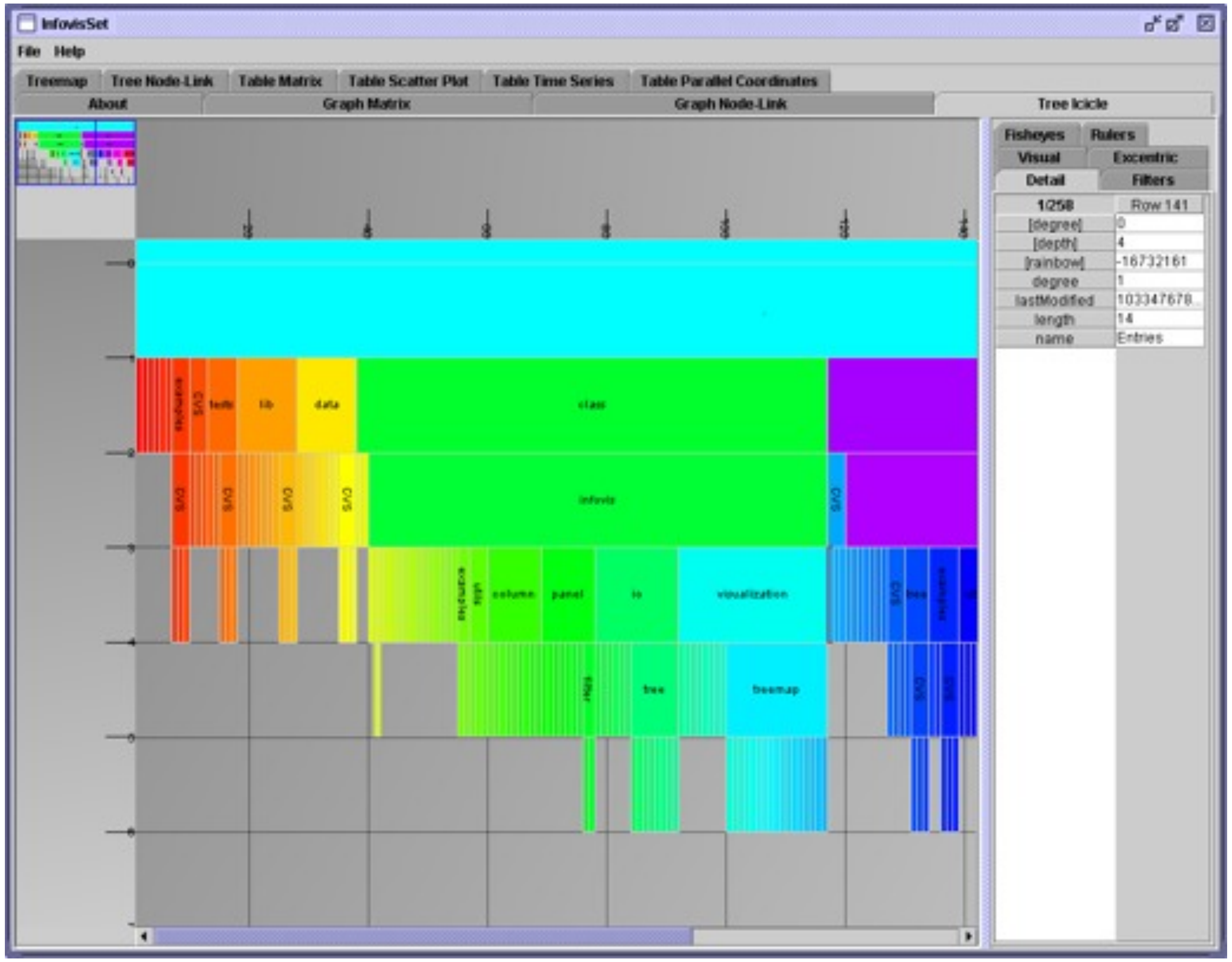


http://www.flickr.com/photos/arenamontanus/2037614308/sizes/o/in/photostream/

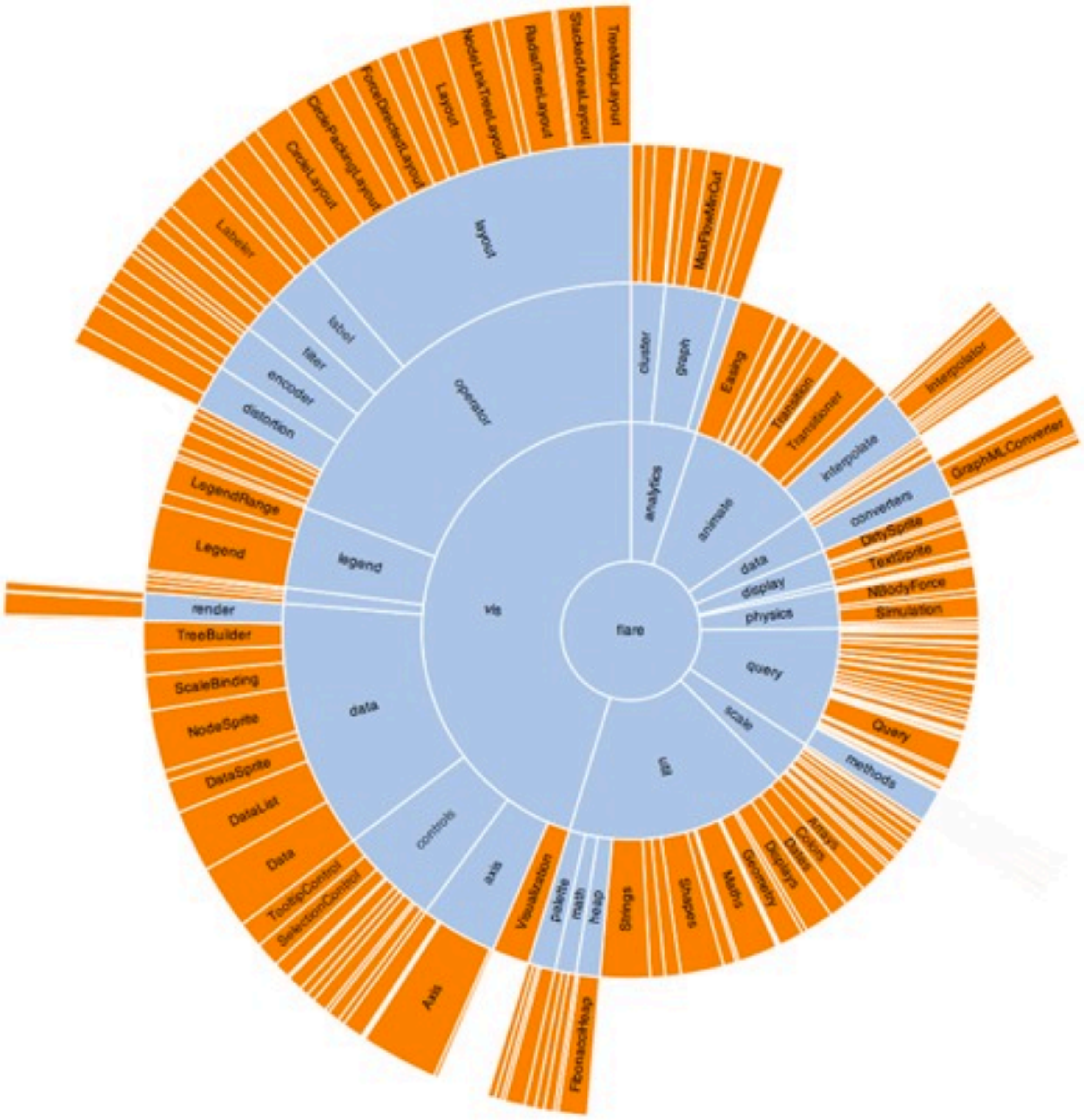
LAYERED DIAGRAMS

- **similar to node-link layouts without edges**
  - structured encoded using:
    - *layering*
    - *adjacency*
    - *alignment*
- **recursive subdivision of space**
- **apply same set of approaches as in node-link layout**

ICICLE & SUNBURST TREES

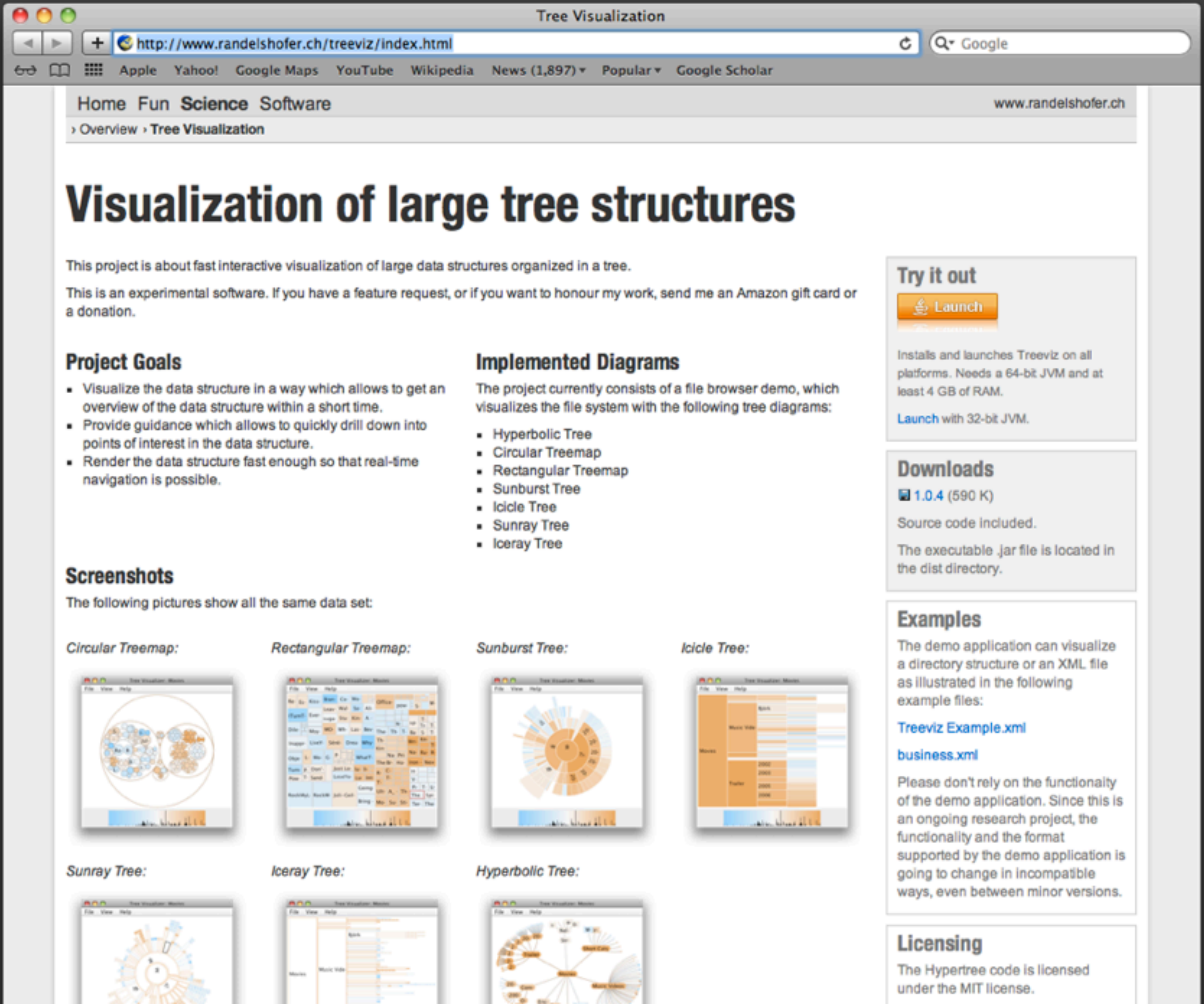


1/258	Row 141
[degree]	0
[depth]	4
[rainbow]	-16732161
degree	1
lastModified	103347678...
length	14
name	Entries



http://ivtk.sourceforge.net/

http://hci.stanford.edu/jheer/files/zoo/



Home Fun **Science** Software

www.randelshofer.ch

› Overview › **Tree Visualization**

Visualization of large tree structures

This project is about fast interactive visualization of large data structures organized in a tree.

This is an experimental software. If you have a feature request, or if you want to honour my work, send me an Amazon gift card or a donation.

Project Goals

- Visualize the data structure in a way which allows to get an overview of the data structure within a short time.
- Provide guidance which allows to quickly drill down into points of interest in the data structure.
- Render the data structure fast enough so that real-time navigation is possible.

Implemented Diagrams

The project currently consists of a file browser demo, which visualizes the file system with the following tree diagrams:

- Hyperbolic Tree
- Circular Treemap
- Rectangular Treemap
- Sunburst Tree
- Icicle Tree
- Sunray Tree
- Iceray Tree

Screenshots

The following pictures show all the same data set:

Circular Treemap:

Rectangular Treemap:

Sunburst Tree:

Icicle Tree:

Sunray Tree:

Iceray Tree:

Hyperbolic Tree:

Try it out

Launch

Installs and launches Treeviz on all platforms. Needs a 64-bit JVM and at least 4 GB of RAM.

Launch with 32-bit JVM.

Downloads

1.0.4 (590 K)

Source code included.

The executable .jar file is located in the dist directory.

Examples

The demo application can visualize a directory structure or an XML file as illustrated in the following example files:

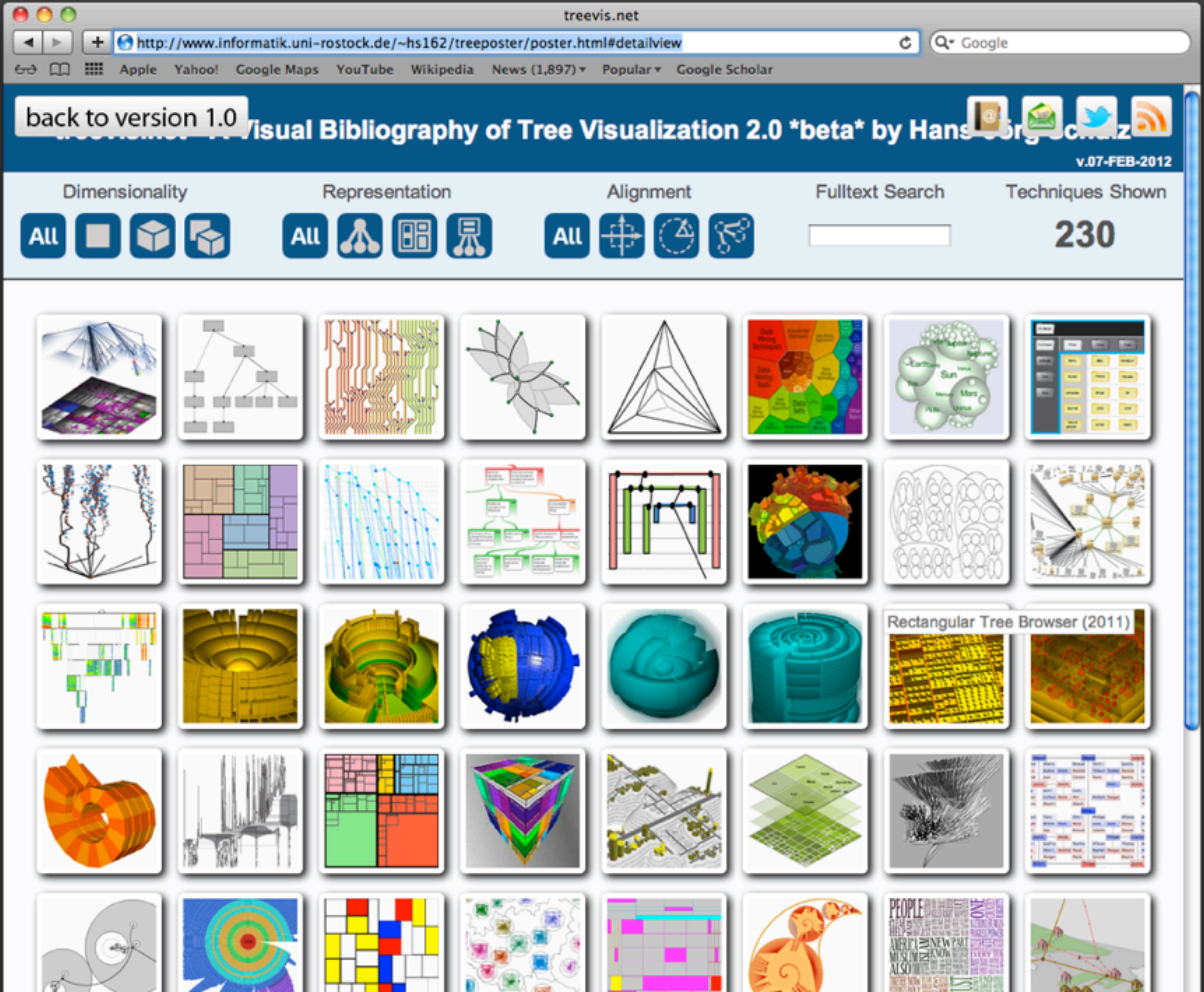
Treeviz Example.xml

business.xml

Please don't rely on the functionality of the demo application. Since this is an ongoing research project, the functionality and the format supported by the demo application is going to change in incompatible ways, even between minor versions.

Licensing

The Hypertree code is licensed under the MIT license.



treevis.net

http://www.informatik.uni-rostock.de/~hs162/treeposter/poster.html#detailview

isual Bibliography of Tree Visualization 2.0 *beta* by Hans

v.07-FEB-2012

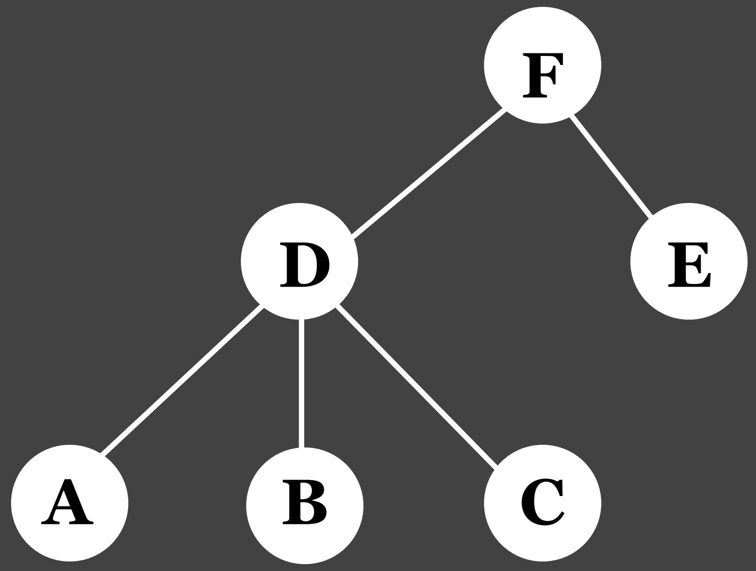
Techniques Shown 230

Rectangular Tree Browser (2011)

exercise

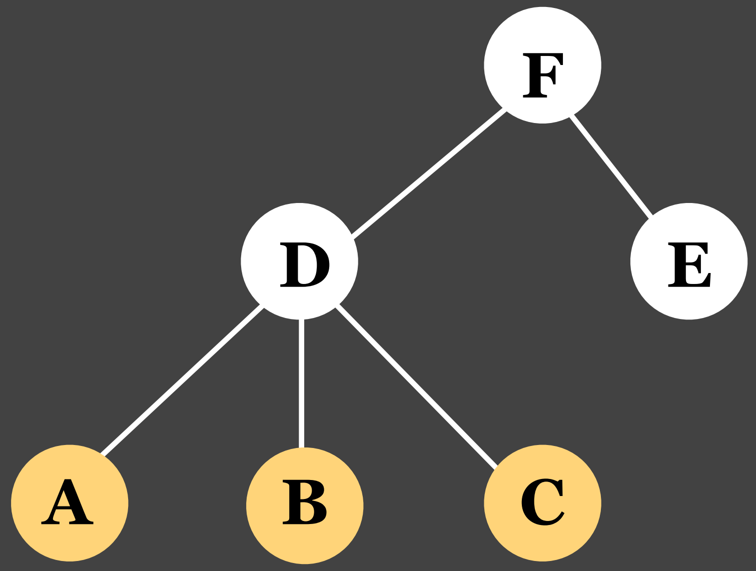
NEWICK FORMAT

- **representation of trees using parentheses and commas**



NEWICK FORMAT

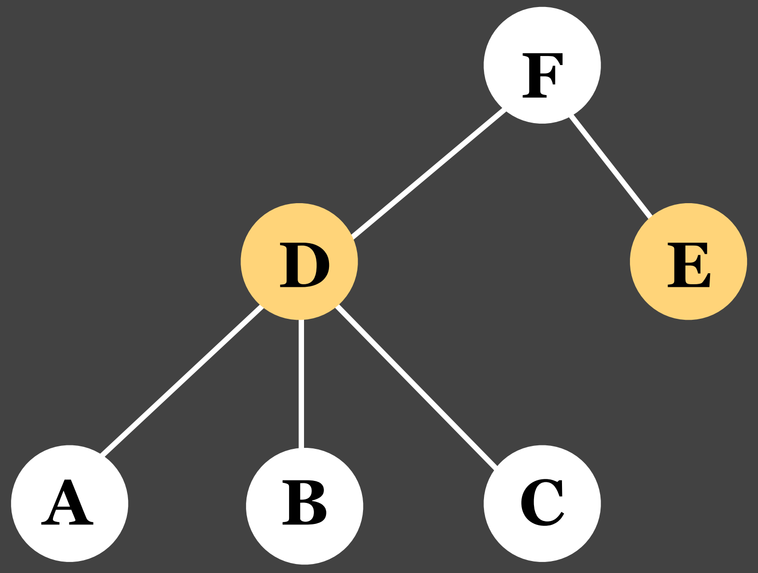
- representation of trees using parentheses and commas



A, B, C

NEWICK FORMAT

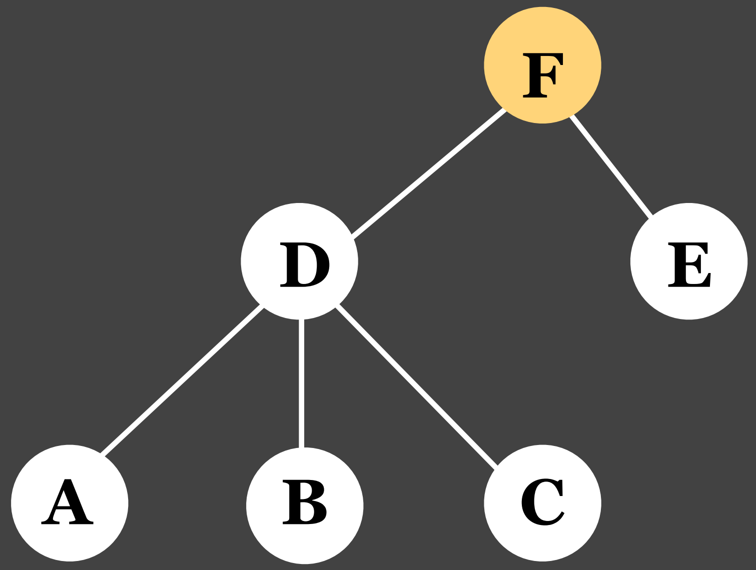
- **representation of trees using parentheses and commas**



(A, B, C) D, E

NEWICK FORMAT

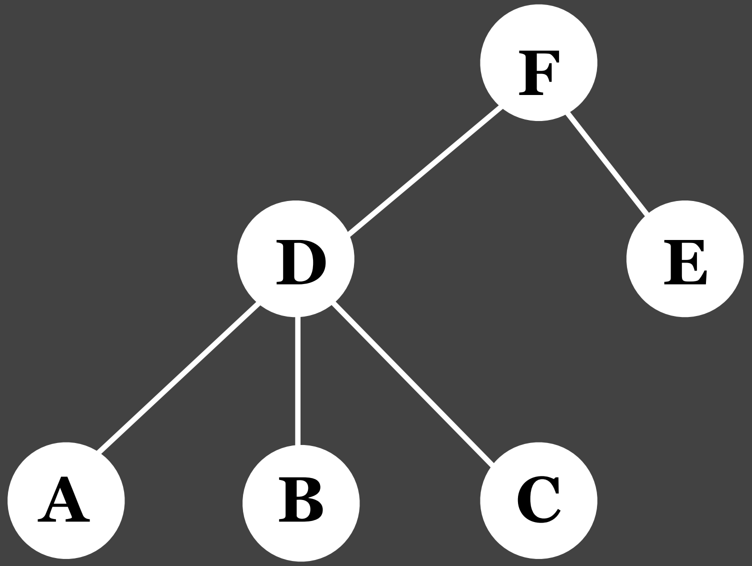
- **representation of trees using parentheses and commas**



((A, B, C) D, E) F

NEWICK FORMAT

- **representation of trees using parentheses and commas**



((A, B, C) D, E) F

parentheses : child
comma : sibling

TREE DRAWING EXERCISE

newick format

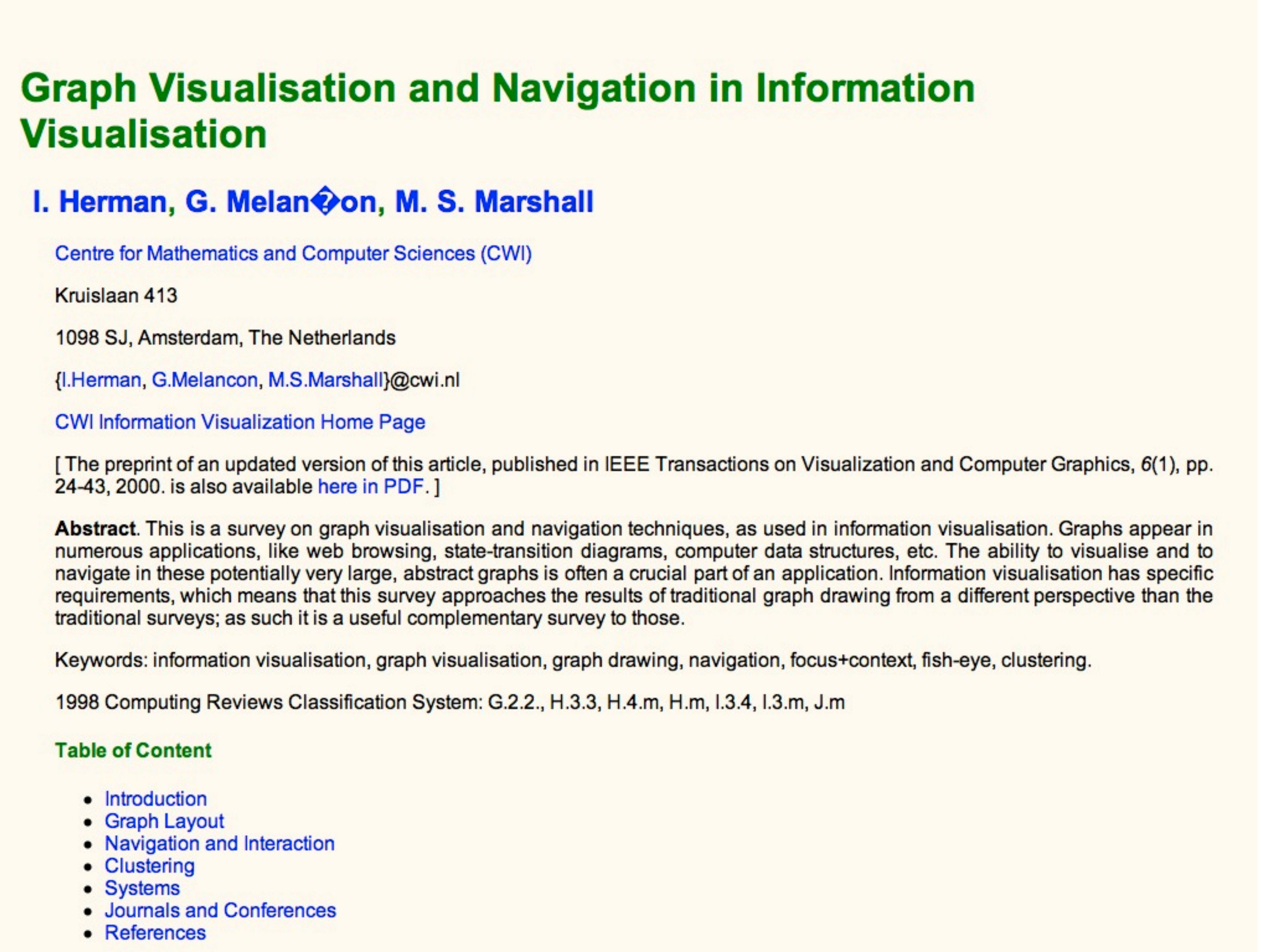
(((((A, B) C) D) E, (((H, (L, M) N) I) J, (O, (((Q) R, (U, V) W) S, X) T) P) K) F, (Y) Z) G

REINGOLD-TILFORD

- bottom up recursive approach
- for each parent make sure every subtree is drawn
- pack subtrees as closely as possible
- center parent over subtrees

L15: Trees & Graphs

REQUIRED READING



Graph Visualisation and Navigation in Information Visualisation

I. Herman, G. Melan�on, M. S. Marshall

Centre for Mathematics and Computer Sciences (CWI)

Kruislaan 413

1098 SJ, Amsterdam, The Netherlands

{I.Herman, G.Melancon, M.S.Marshall}@cwi.nl

CWI Information Visualization Home Page

[The preprint of an updated version of this article, published in IEEE Transactions on Visualization and Computer Graphics, *6*(1), pp. 24-43, 2000. is also available here in PDF.]

Abstract. This is a survey on graph visualisation and navigation techniques, as used in information visualisation. Graphs appear in numerous applications, like web browsing, state-transition diagrams, computer data structures, etc. The ability to visualise and to navigate in these potentially very large, abstract graphs is often a crucial part of an application. Information visualisation has specific requirements, which means that this survey approaches the results of traditional graph drawing from a different perspective than the traditional surveys; as such it is a useful complementary survey to those.

Keywords: information visualisation, graph visualisation, graph drawing, navigation, focus+context, fish-eye, clustering.

1998 Computing Reviews Classification System: G.2.2., H.3.3, H.4.m, H.m, I.3.4, I.3.m, J.m

Table of Content

mailto:mwatten@us.ibm.com

Visual Exploration of Multivariate Graphs

Martin Wattenberg
Visual Communication Lab, IBM Research
1 Rogers St., Cambridge MA 02142
mwatten@us.ibm.com

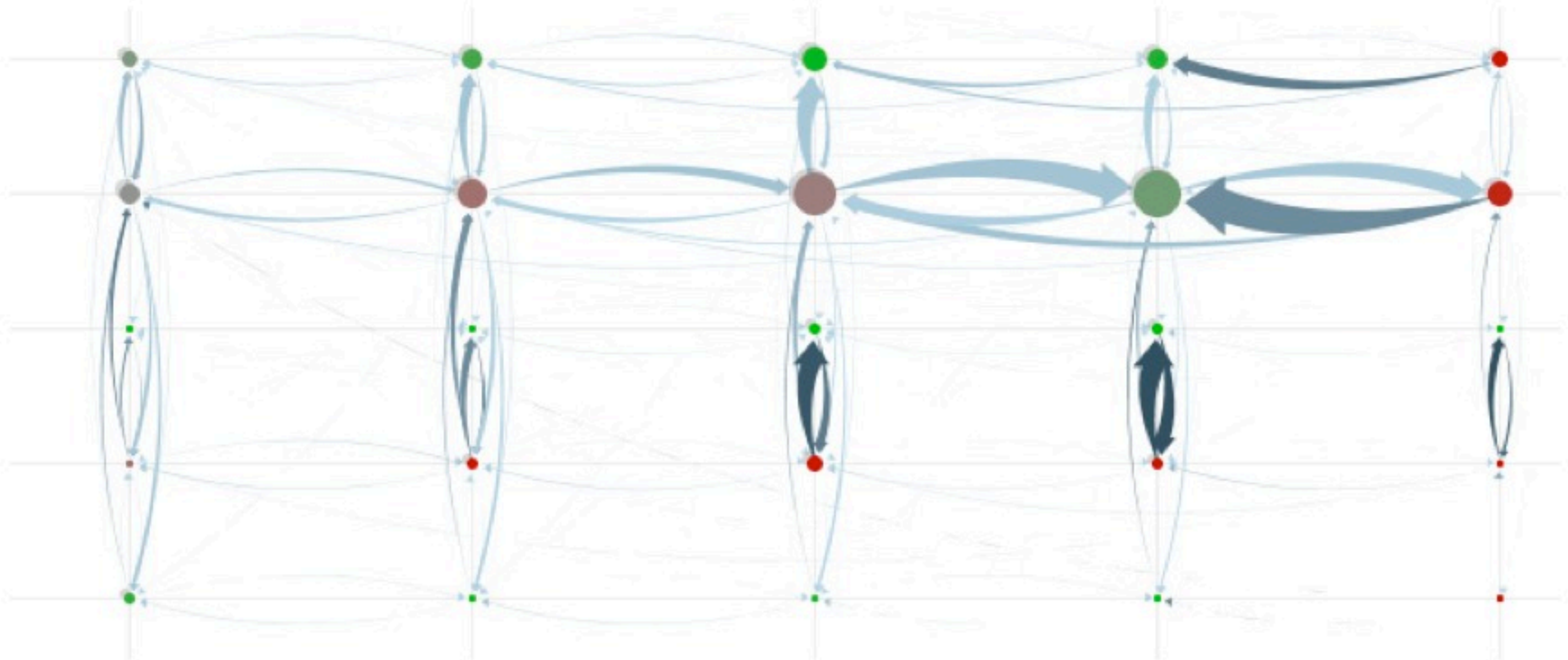


Figure 1. *A PivotGraph visualization of a large graph rolled up onto two categorical dimensions*

ABSTRACT

Author Keywords
information visualization, graph drawing

ACM Classification Keywords

ABySS-Explorer: Visualizing Genome Sequence Assemblies

Cydney B. Nielsen, Shaun D. Jackman, Inanç Birol, and Steven J.M. Jones

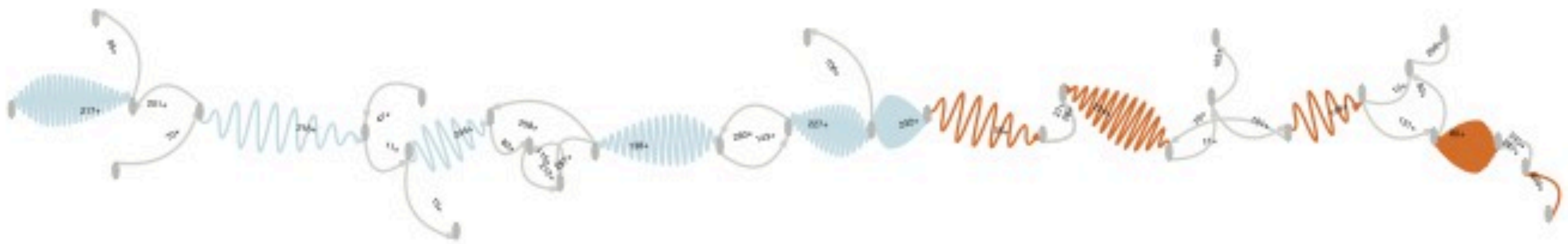


Fig. 1. ABySS-Explorer employs a novel graph representation enabling biologists to examine the global structure of a genome sequence assembly.

Abstract—One bottleneck in large-scale genome sequencing projects is reconstructing the full genome sequence from the short subsequences produced by current technologies. The final stages of the genome assembly process inevitably require manual inspection of data inconsistencies and could be greatly aided by visualization. This paper presents our design decisions in translating key data features identified through discussions with analysts into a concise visual encoding. Current visualization tools in this domain focus on local sequence errors making high-level inspection of the assembly difficult if not impossible. We present a novel interactive graph display, ABySS-Explorer, that emphasizes the global assembly structure while also integrating salient data features such as sequence length. Our tool replaces manual and in some cases pen-and-paper based analysis tasks, and we discuss how user feedback was incorporated into iterative design refinements. Finally, we touch on applications of this representation not initially considered in our design phase, suggesting the generality of this encoding for DNA sequence data.

Index Terms—Bioinformatics visualization, design study, DNA sequence, genome assembly.

1 Introduction

Data generation used to be the expensive and time consuming step in biology research. Recent innovations in high-throughput techniques have transformed it into a cost-effective and rapid process, pushing the bottleneck of discovery into the analysis phase. There is increasing recognition in the field that improvements in visualization tools will be essential for understanding our growing wealth of data. This paper presents one such tool for a genome analysis problem.

The term "genome" refers to the genetic material of a cell and can be thought of as the cellular instruction set. A genome consists of

subjected to many rounds of automated improvement, but ultimately it is visually inspected and manually edited by specialists.

Our work was motivated by the needs of genome analysts and the shortcomings of existing visualization tools in this domain. A genome assembly consists of long contiguous sequences, called contigs, assembled from short sequencing reads. An analyst integrates diverse data types used by the assembly algorithm together with external metadata to make final judgements about whether an assembly is correct and complete. It is useful to interpret these data in the context of the