Visualization

MICHAEL BERNSTEIN
THANKS TO DIANA MACLEAN AND JEFF HEER
CS 376

Anscombe's Quartet

Set A		Set B		Set C		Set D	
X	Y	X	Y	X	Y	X	Y
10	8.04	10	9.14	10	7.46	8	6.58
8	6.95	8	8.14	8	6.77	8	5.76
13	7.58	13	8.74	13	12.74	8	7.71
9	8.81	9	8.77	9	7.11	8	8.84
11	8.33	11	9.26	11	7.81	8	8.47
14	9.96	14	8.1	14	8.84	8	7.04
6	7.24	6	6.13	6	6.08	8	5.25
4	4.26	4	3.1	4	5.39	19	12.5
12	10.84	12	9.11	12	8.15	8	5.56
7	4.82	7	7.26	7	6.42	8	7.91
5	5.68	5	4.74	5	5.73	8	6.89

Summary Statistics

$u_X = 9.0 \quad \sigma_X = 3.317$

$u_Y = 7.5 \quad \sigma_Y = 2.03$

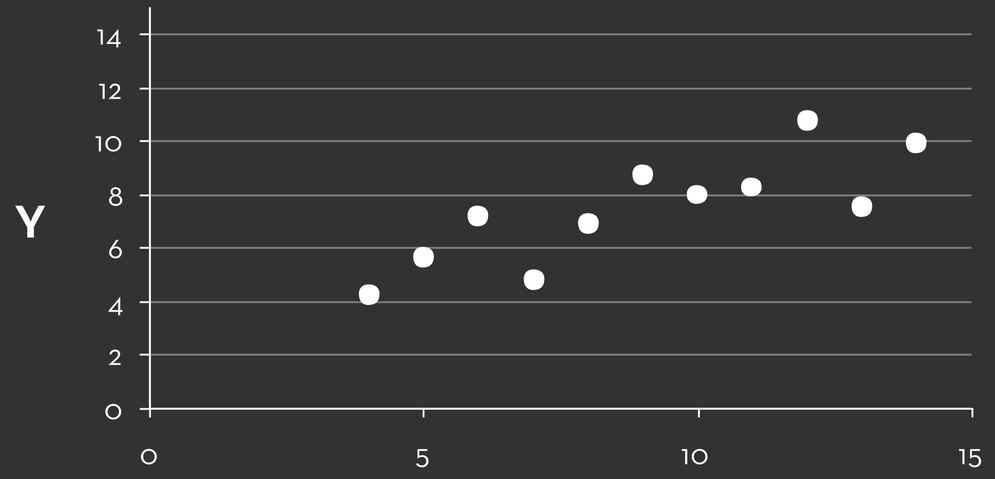
Linear Regression

$Y = 3 + 0.5\,X$

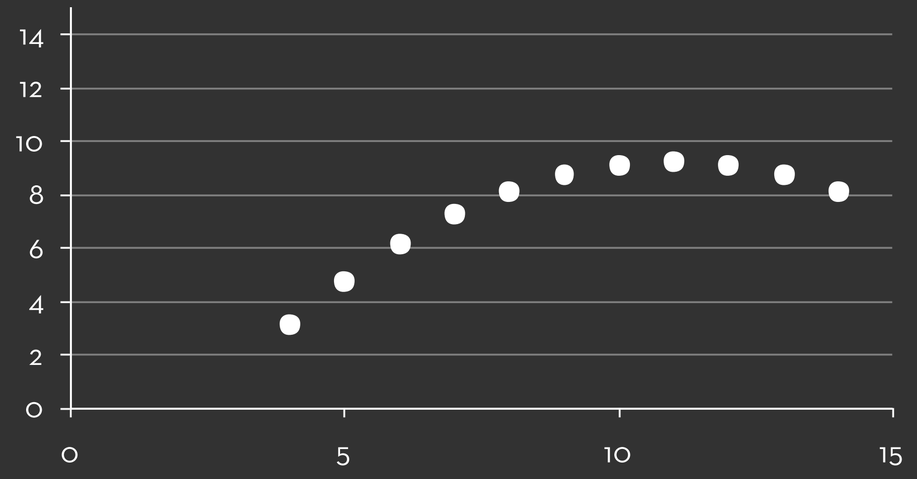
$R^2 = 0.67$

[Anscombe 73]

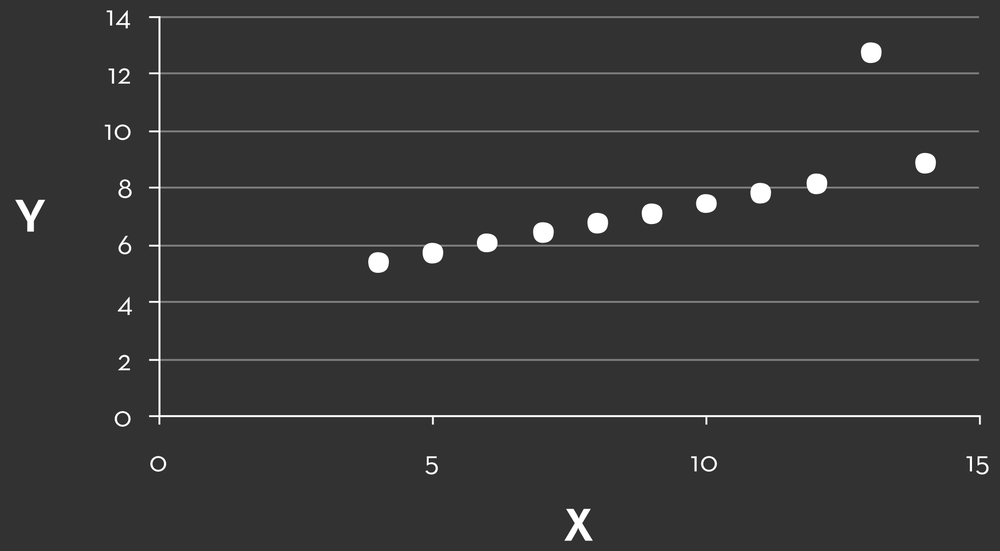
Set A



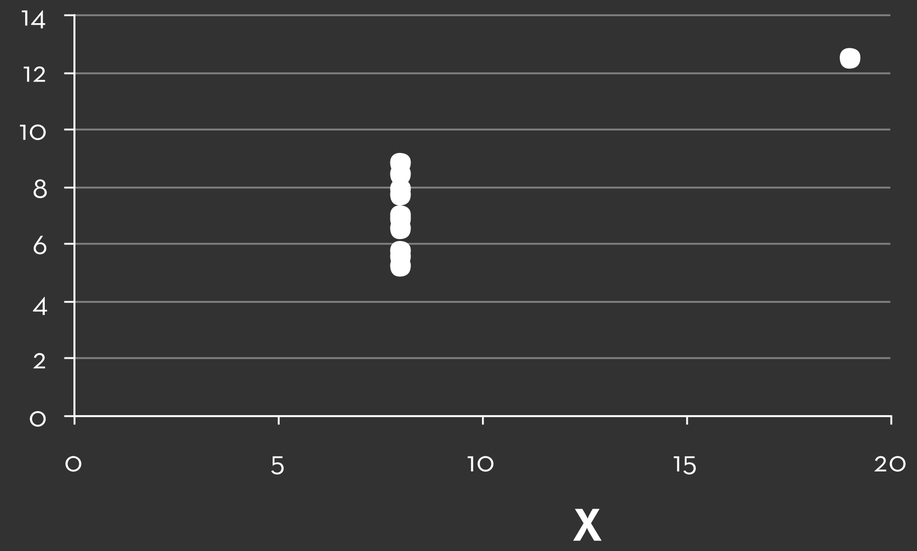
Set B



Set C



Set D



What is visualization?

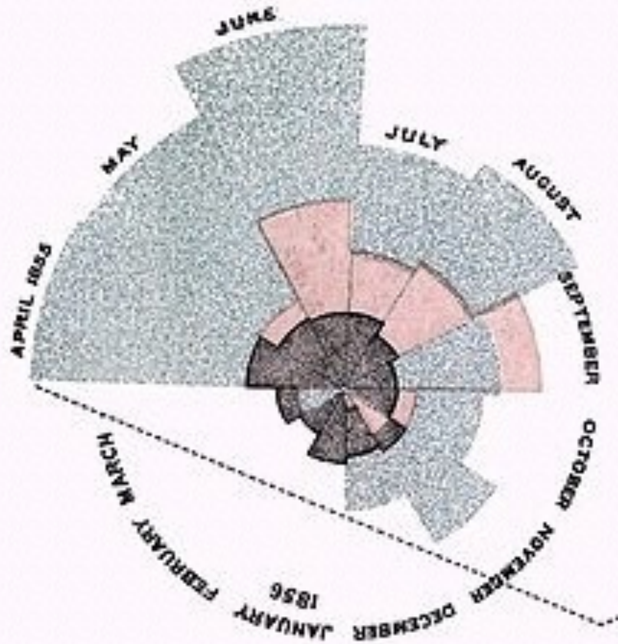
- "Transformation of the symbolic into the geometric"
  [McCormick et al., 1987]

- "…finding the artificial memory that best supports our natural means of perception."
  [Bertin 1967]

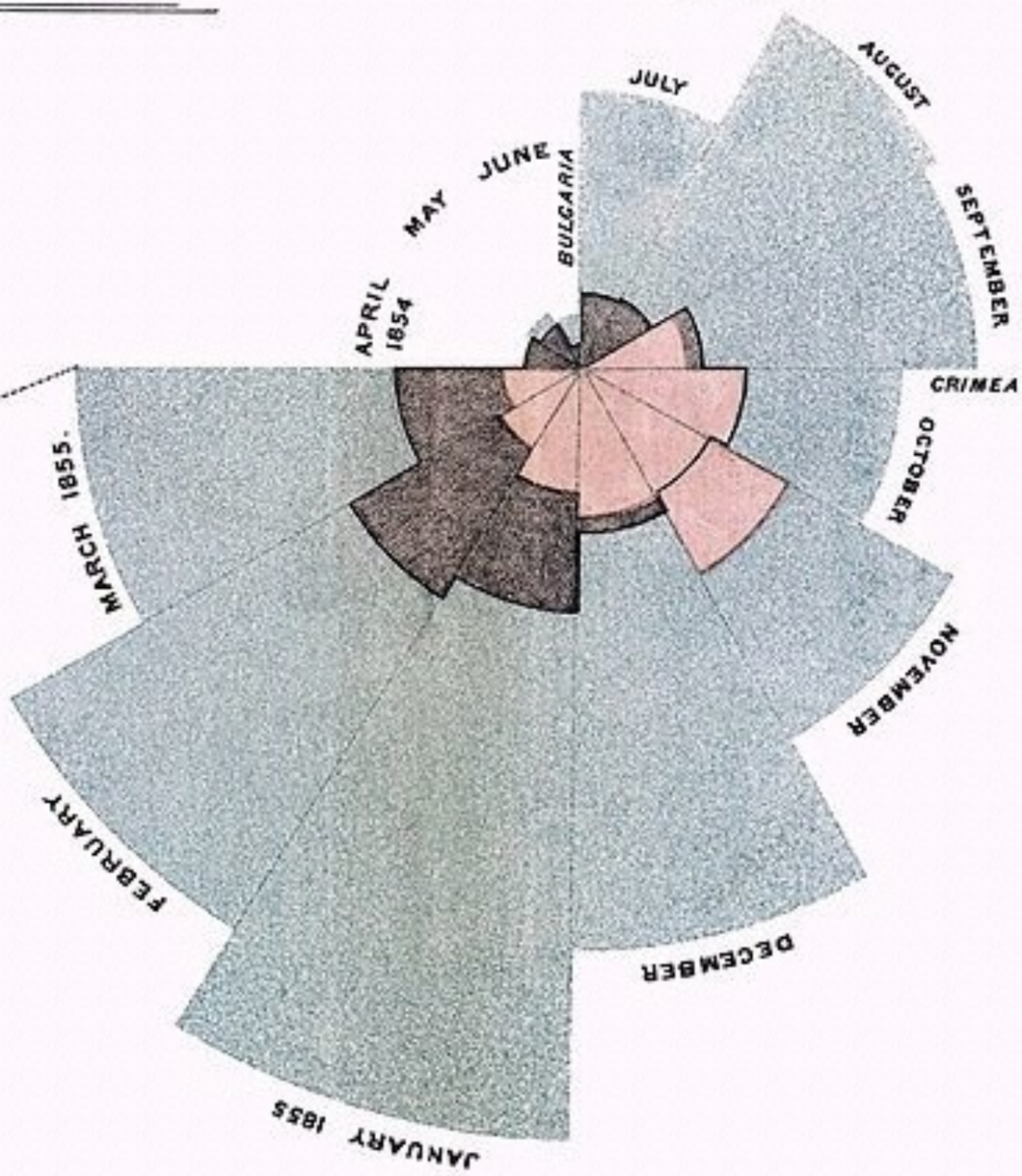
- "The use of computer-generated, interactive, visual representations of data to amplify cognition."
  [Card, Mackinlay, and Shneiderman 1999]

DIAGRAM OF THE CAUSES OF MORTALITY IN THE ARMY IN THE EAST.

2.
APRIL 1855 TO MARCH 1856.



1.
APRIL 1854 TO MARCH 1855.



The Areas of the blue, red, & black wedges are each measured from the centre as the common vertex.

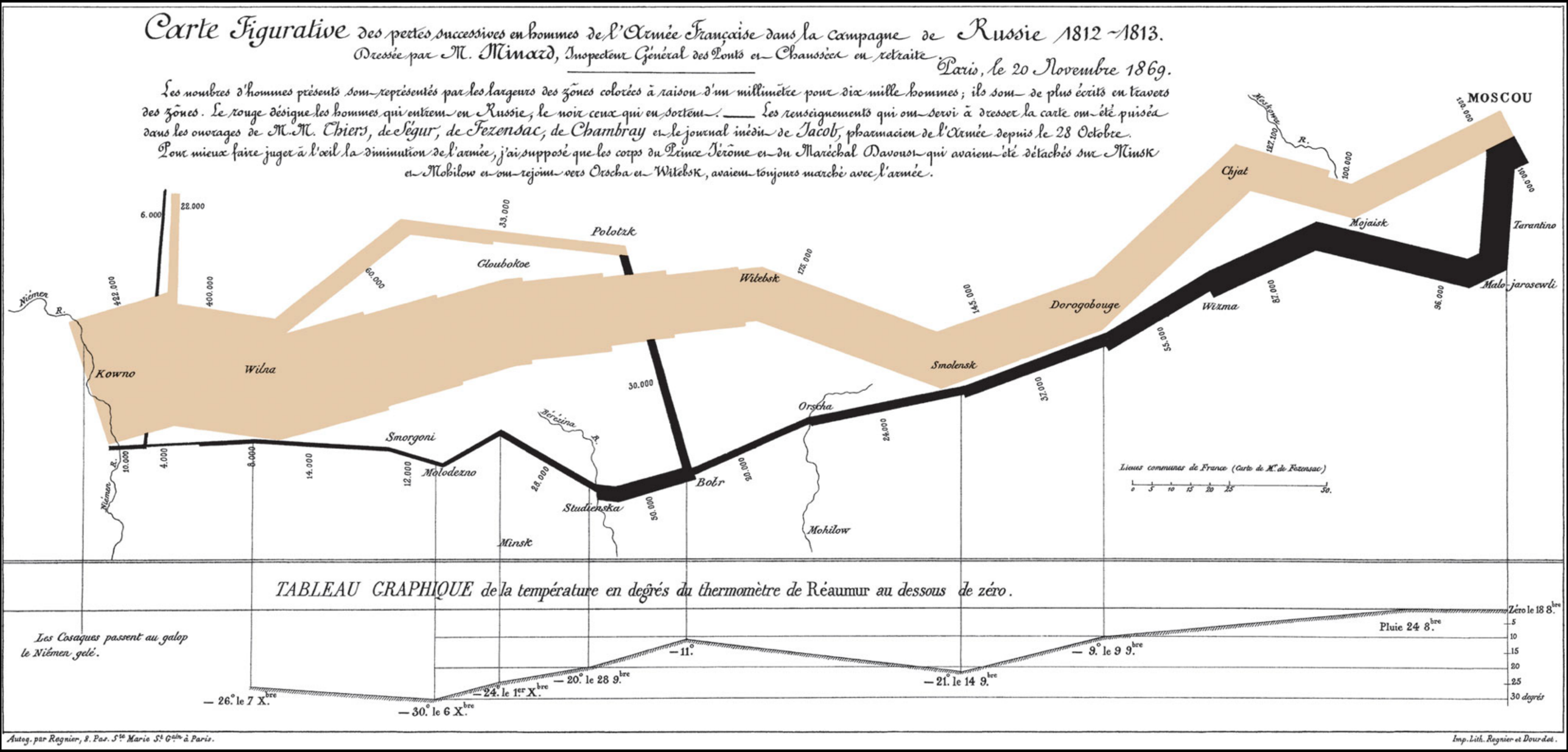
The blue wedges measured from the centre of the circle represent area for area the deaths from Preventible or Mitigable Zymotic diseases, the red wedges measured from the centre the deaths from wounds, & the black wedges measured from the centre the deaths from all other causes.

The black line across the red triangle in Nov.r 1854 marks the boundary of the deaths from all other causes during the month.

In October 1854, & April 1855, the black area coincides with the red, in January & February 1856, the blue coincides with the black.

The entire areas may be compared by following the blue, the red & the black lines enclosing them.

Florence Nightingale, 1857



Charles Minard, 1869

Why visualize data?

- Story or explanation
- Get answers (or questions!)
- Find patterns
- Find mistakes
- Support reasoning
- Convince others
- Share

Why visualize data?

Goals of visualization research

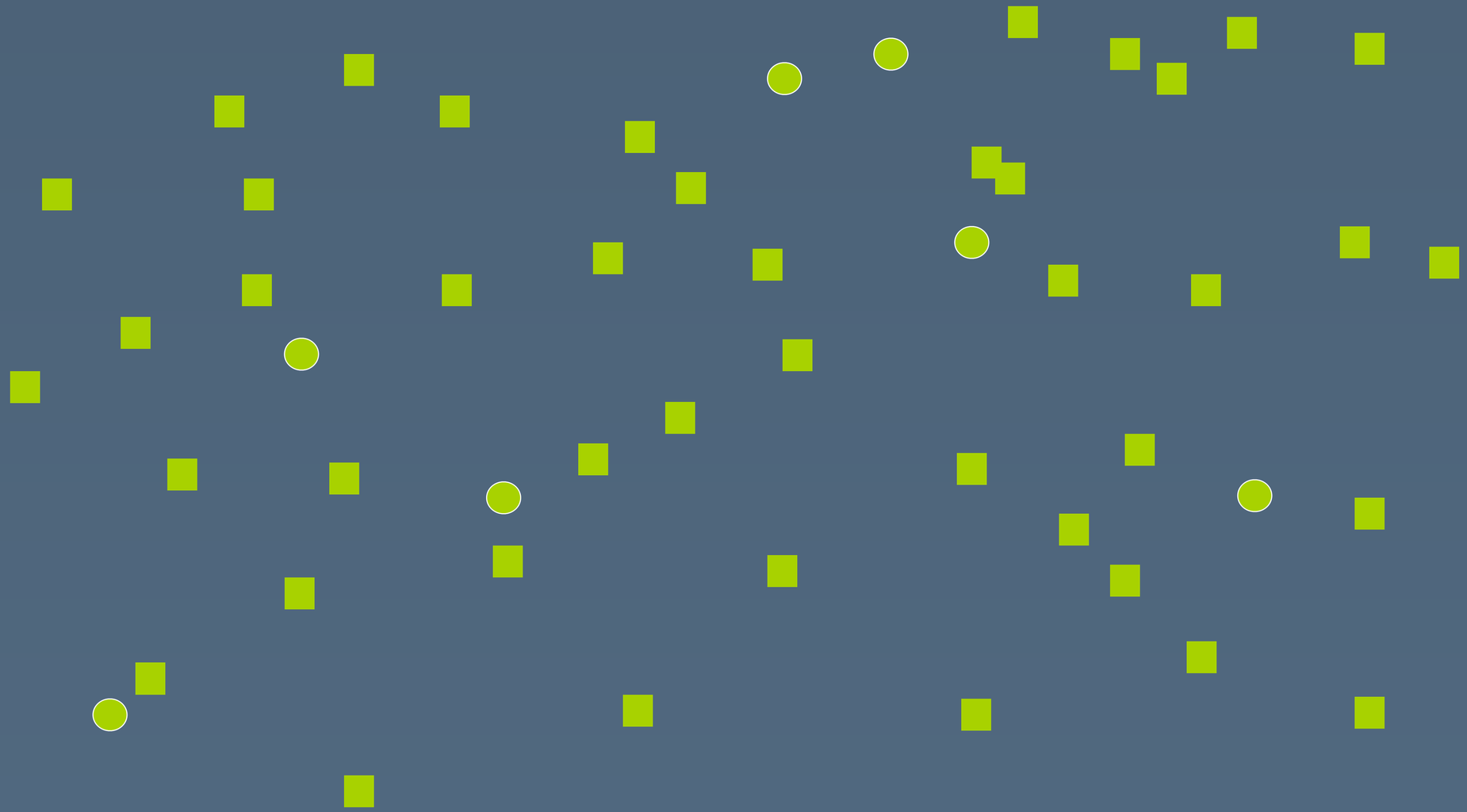
- **Understand how visualizations convey information**
  What do people perceive and comprehend?
  How do visualizations correspond with mental models?

- **Develop principles and techniques for creating effective visualizations and supporting analysis**
  Amplify perception and cognition
  Strengthen the tie between visualization and mental models

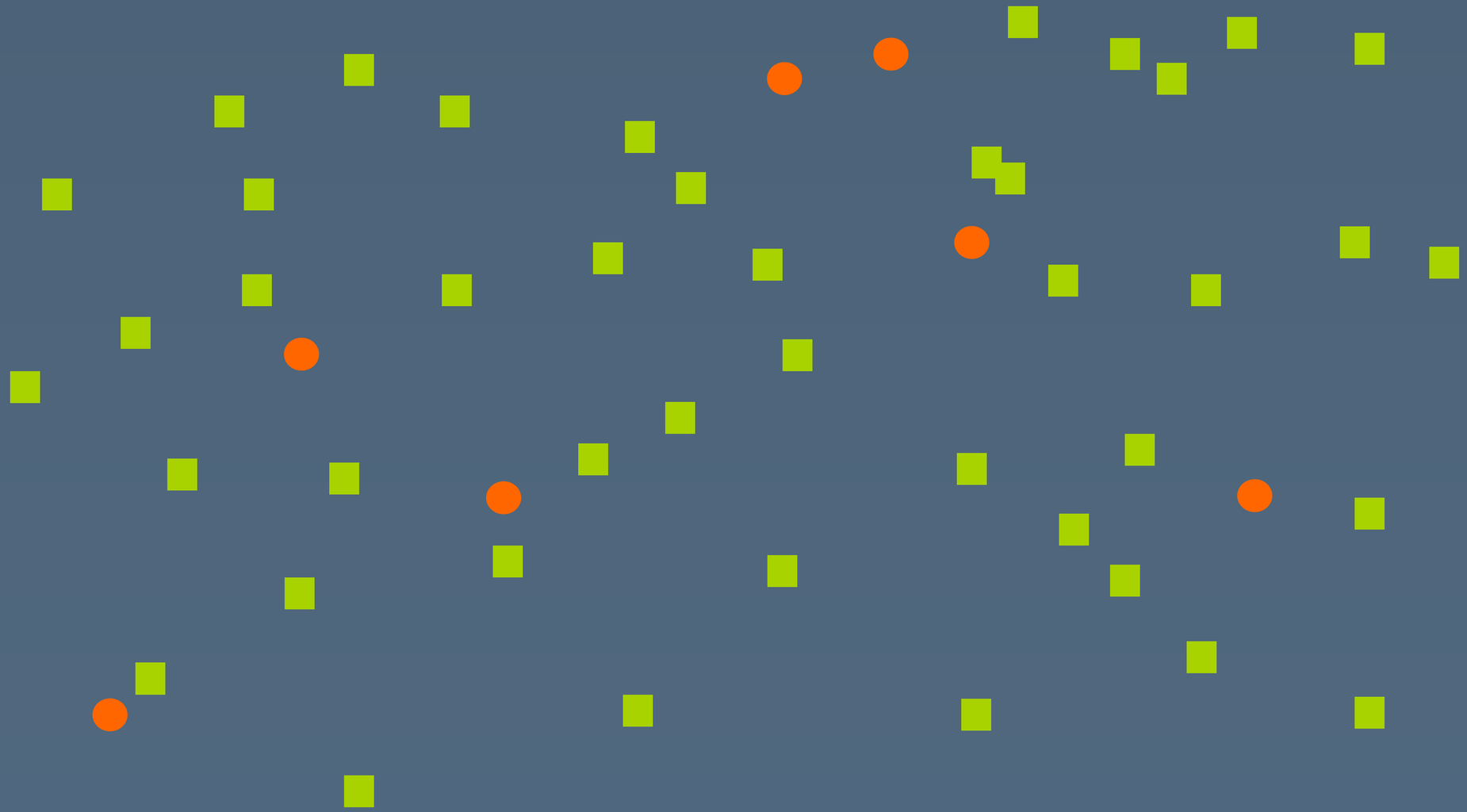
Graphical Perception

“The visual decoding of information encoded on graphs” [Cleveland & McGill, ‘84]

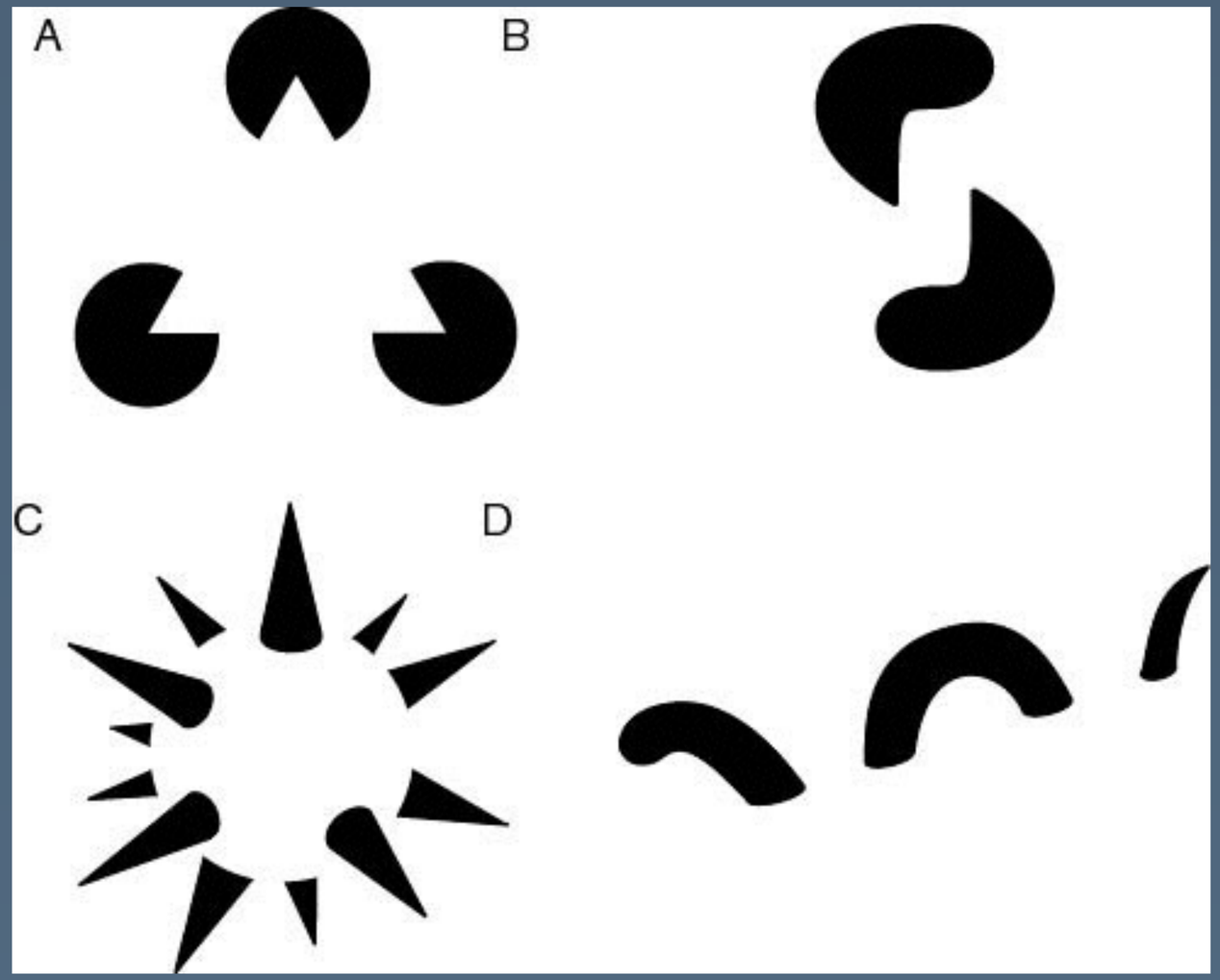
Pre-attentive Processing



Pre-attentive Processing

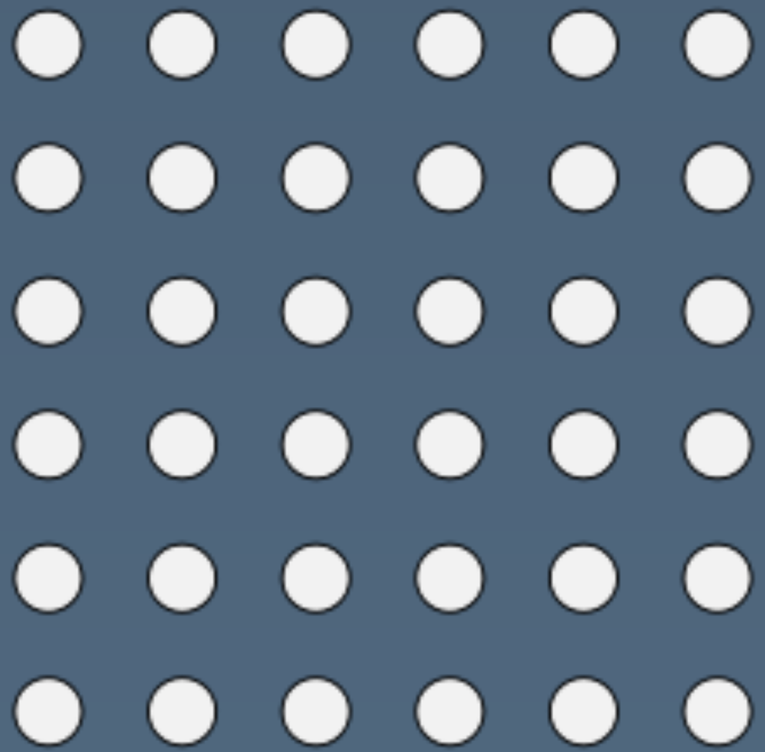


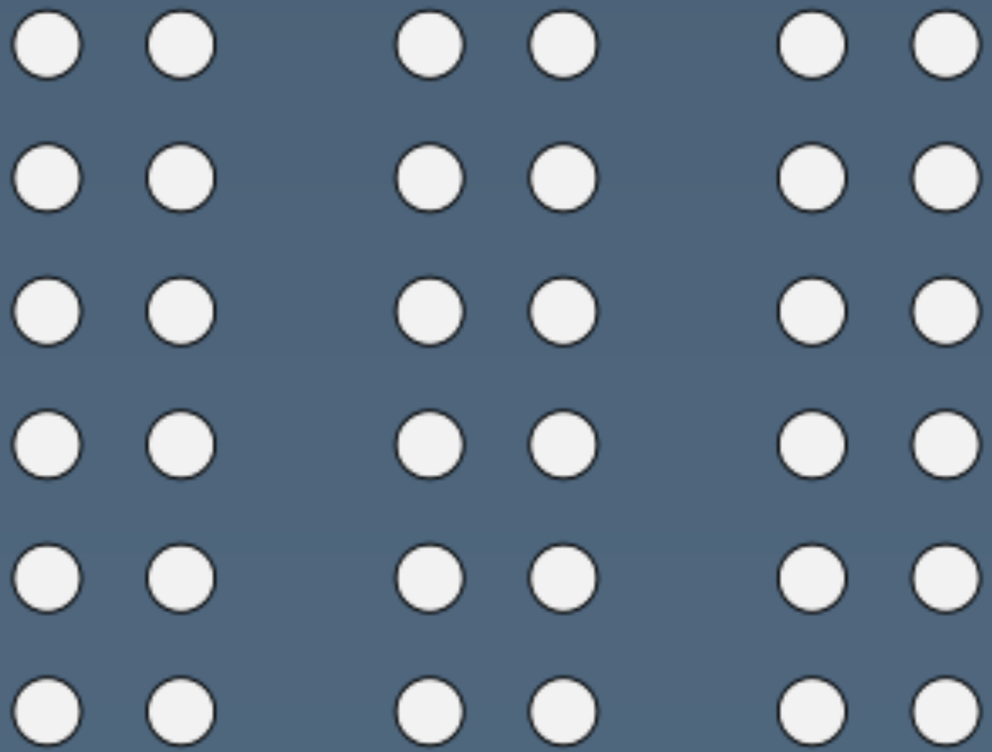
Reification



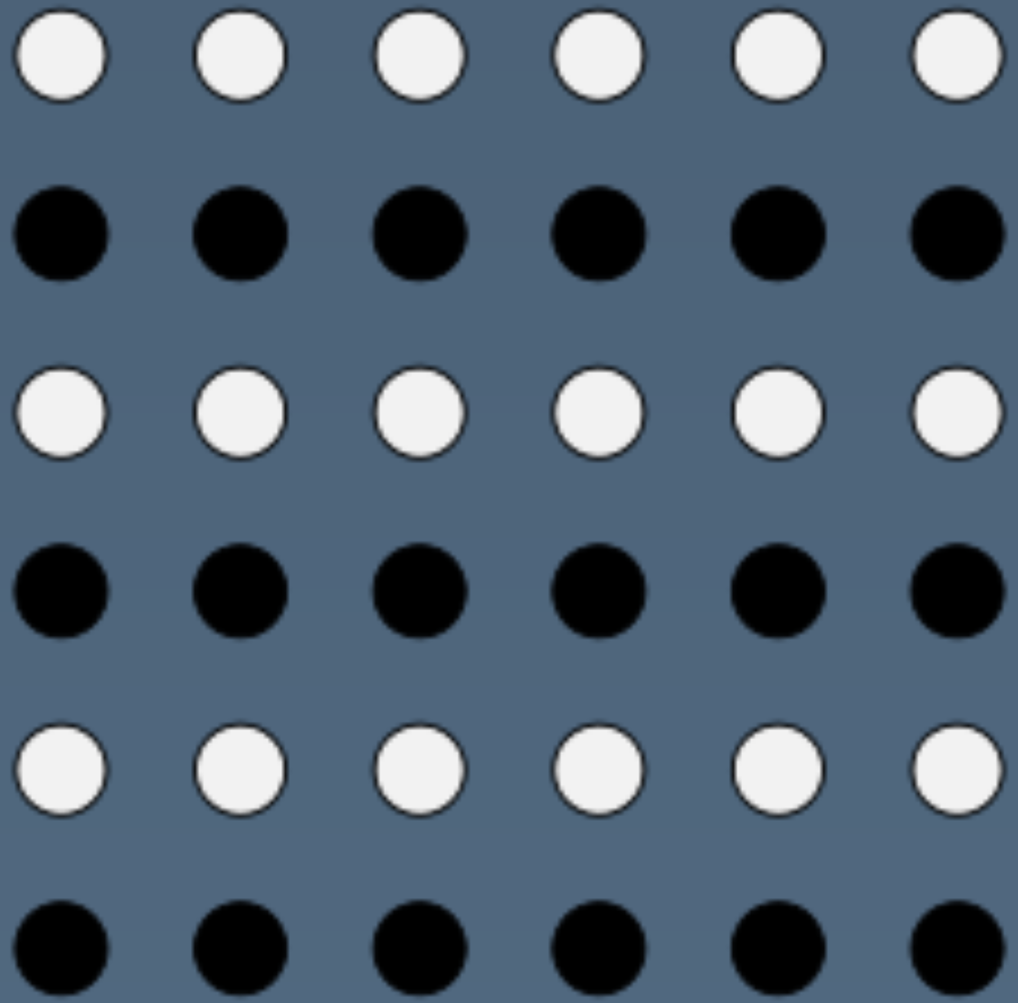
http://en.wikipedia.org/wiki/File:Reification.jpg

Gestalt law of Proximity





Gestalt law of Similarity



People are poor at perceiving differences

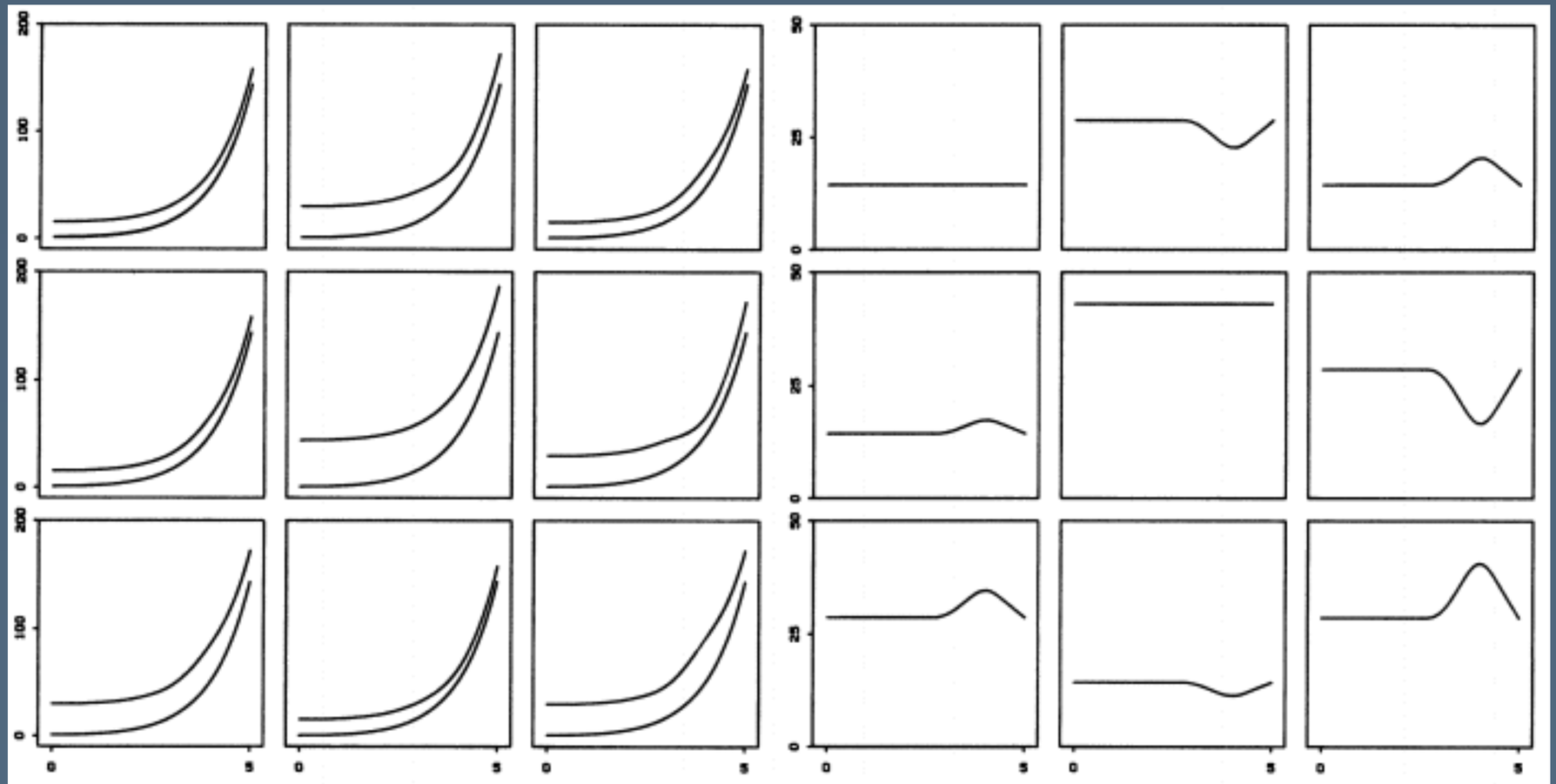


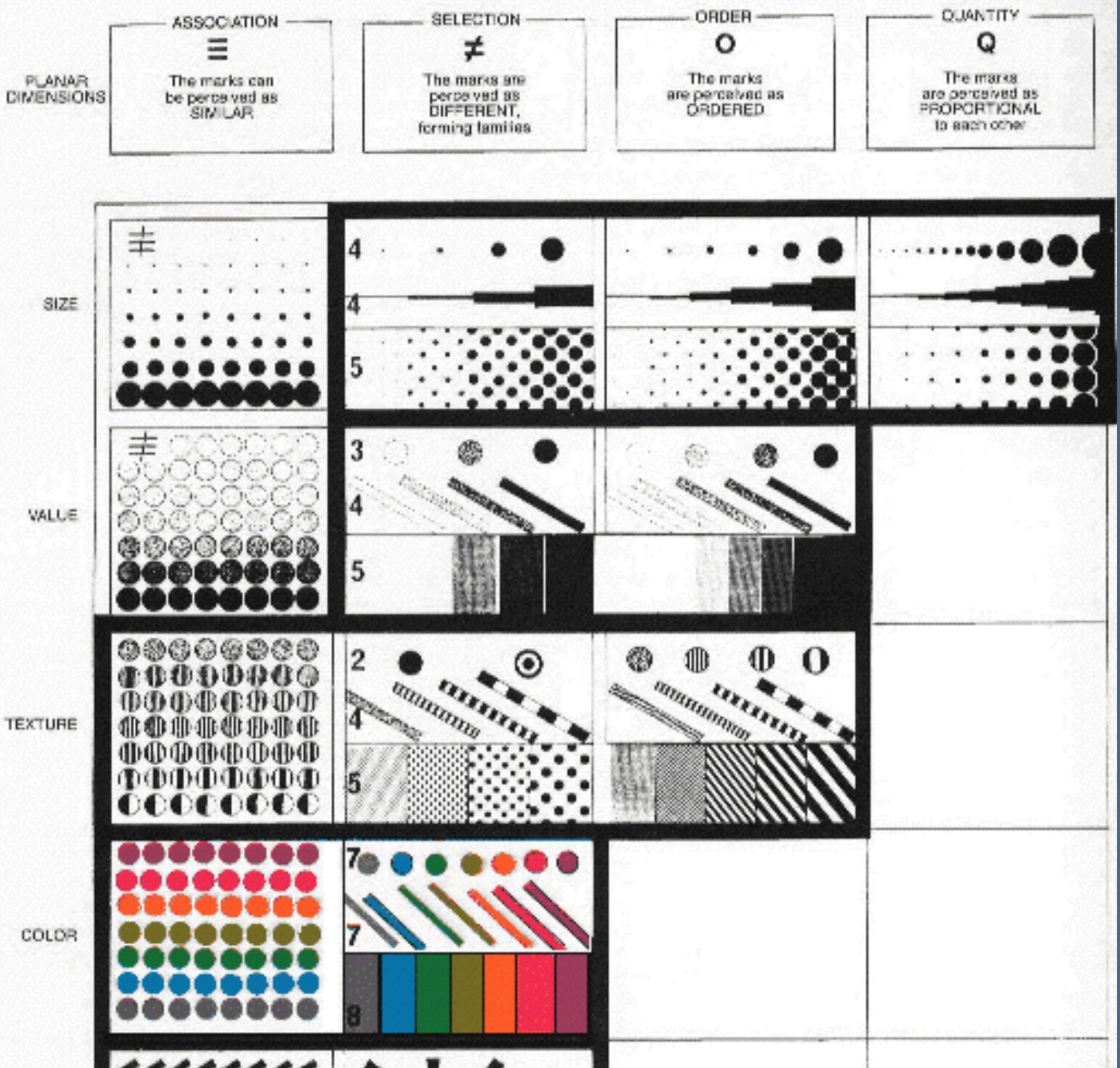
Figure 26. Curve-difference chart.

Figure 27. Curve differences.

Cleveland & McGill, 1984

Semiologie Graphique

Jaques Bertin, 1976

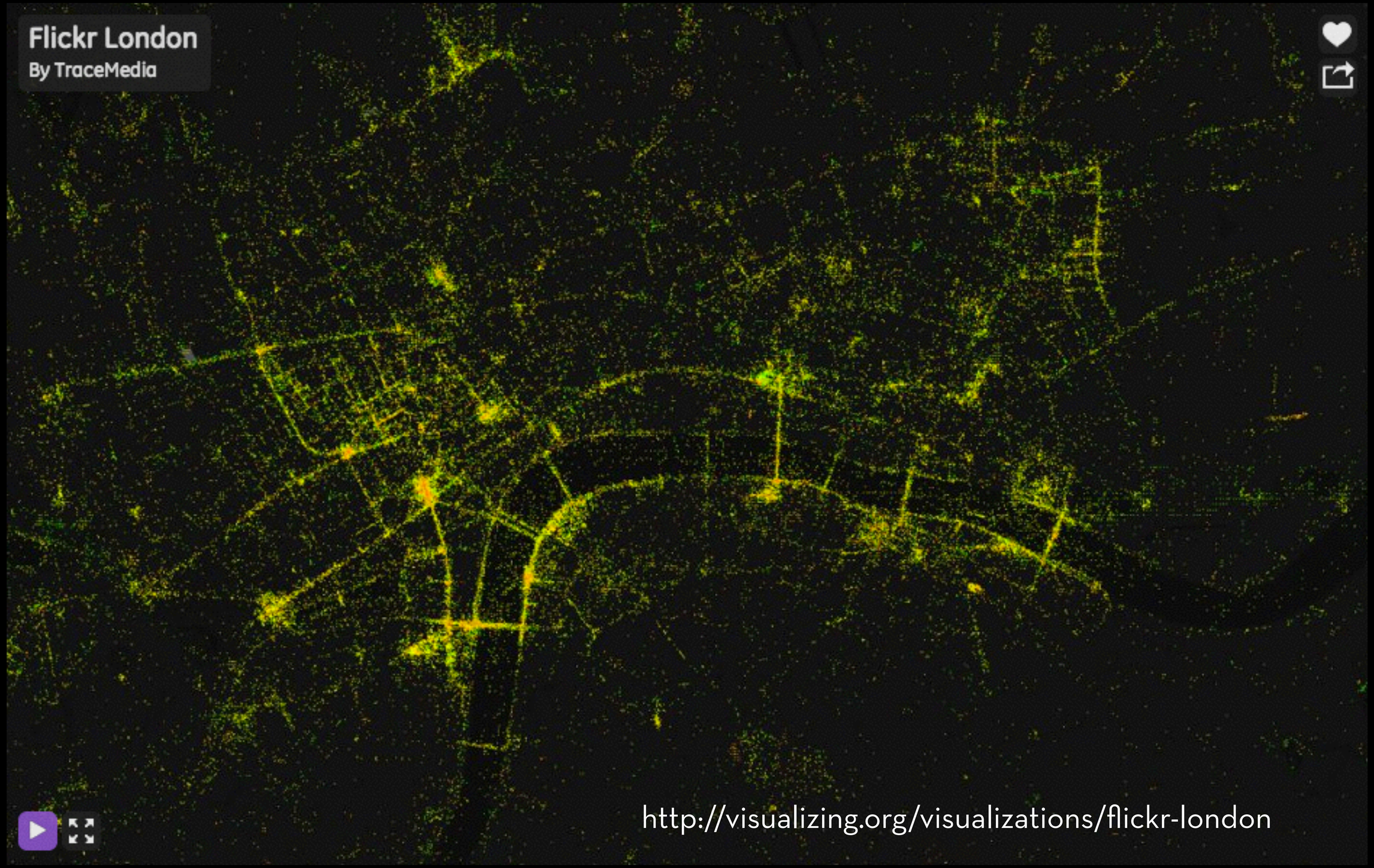


The Process of Visualization

1. Recording & Storing 2. Analyzing 3. Sharing & Convincing

Recording & Storing

http://www.wired.com/wiredscience/2007/11/the-most-beauti/



http://visualizing.org/visualizations/flickr-london

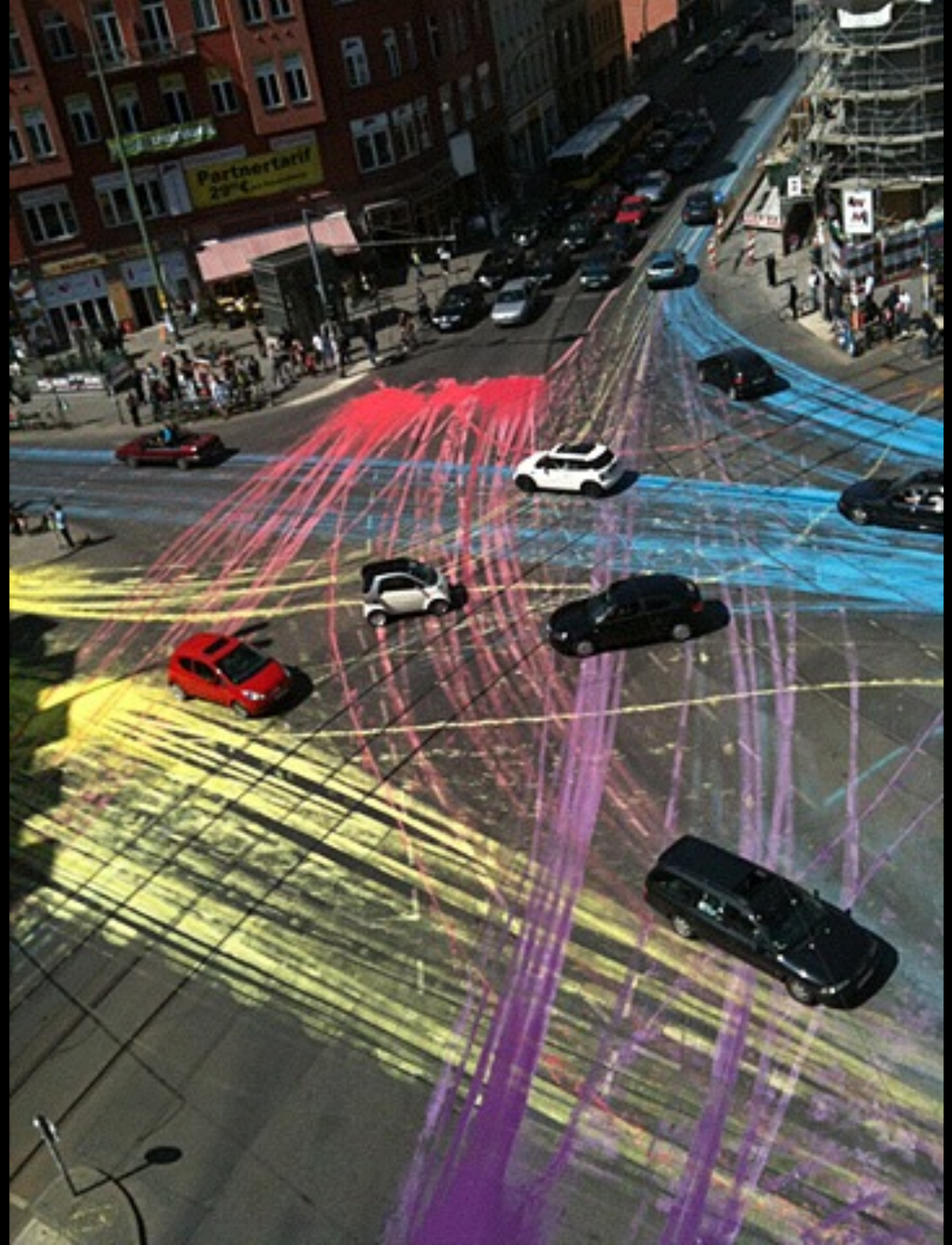


image: link to YouTube

Analyzing

Interaction: Filtering

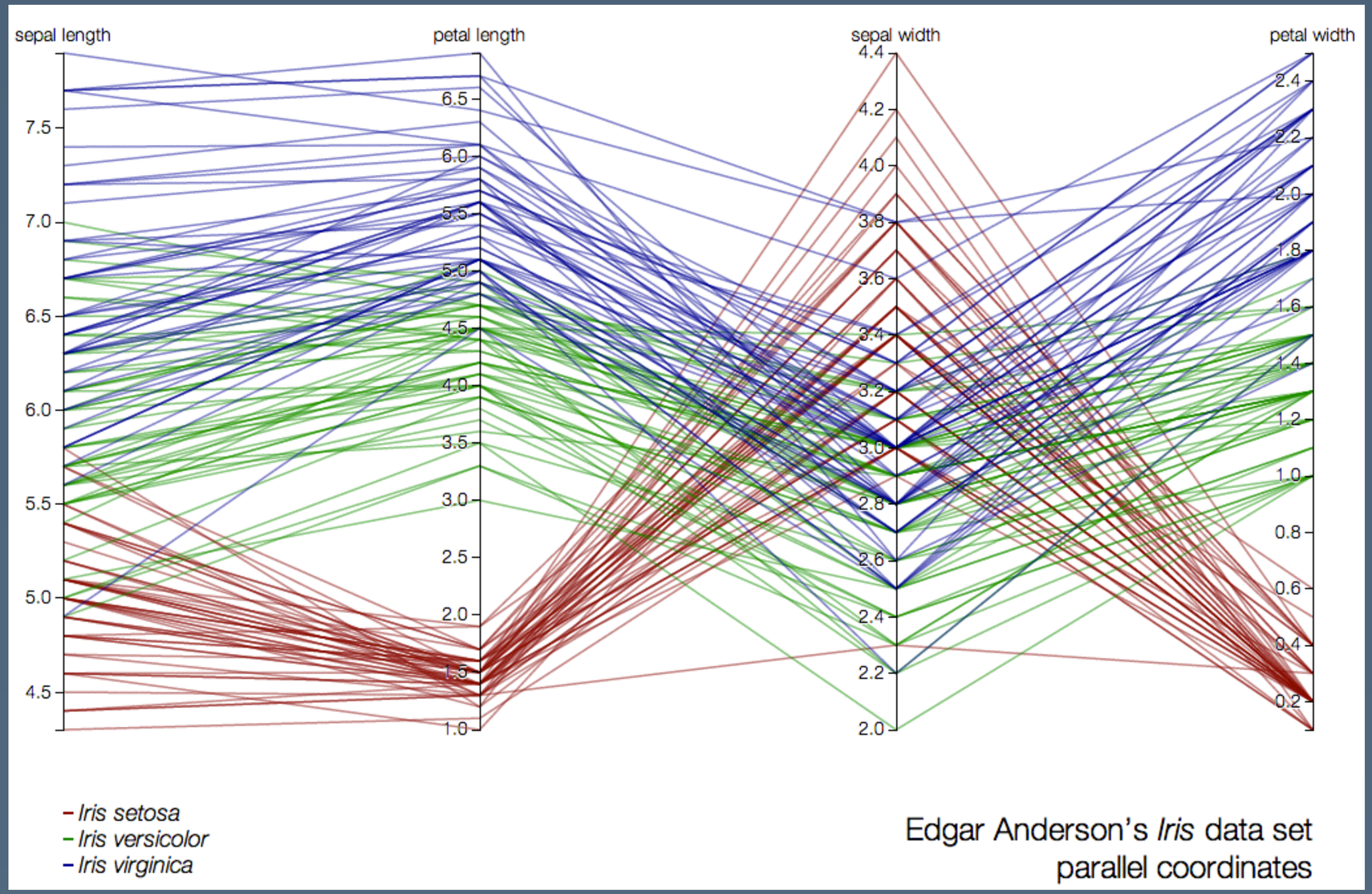
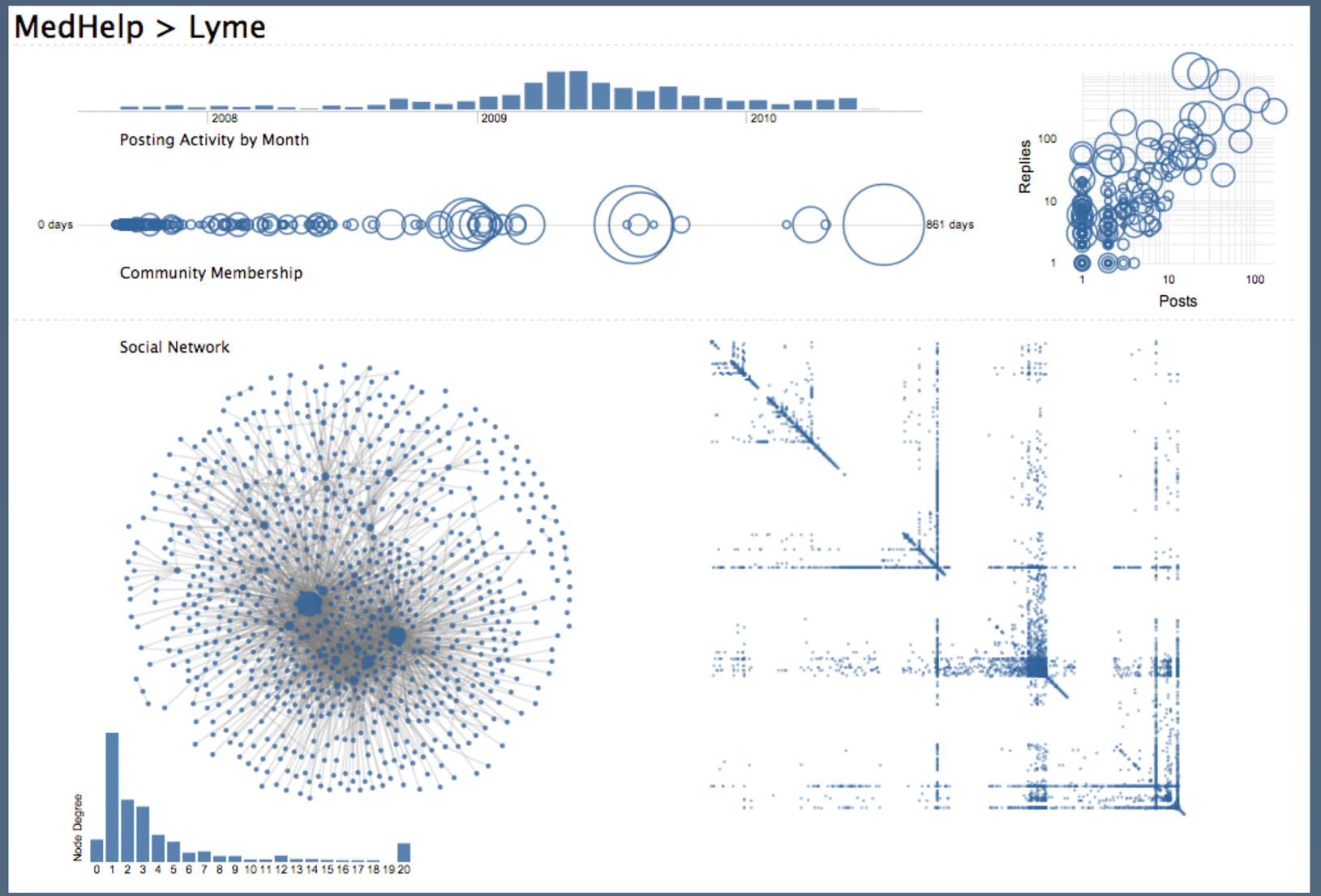
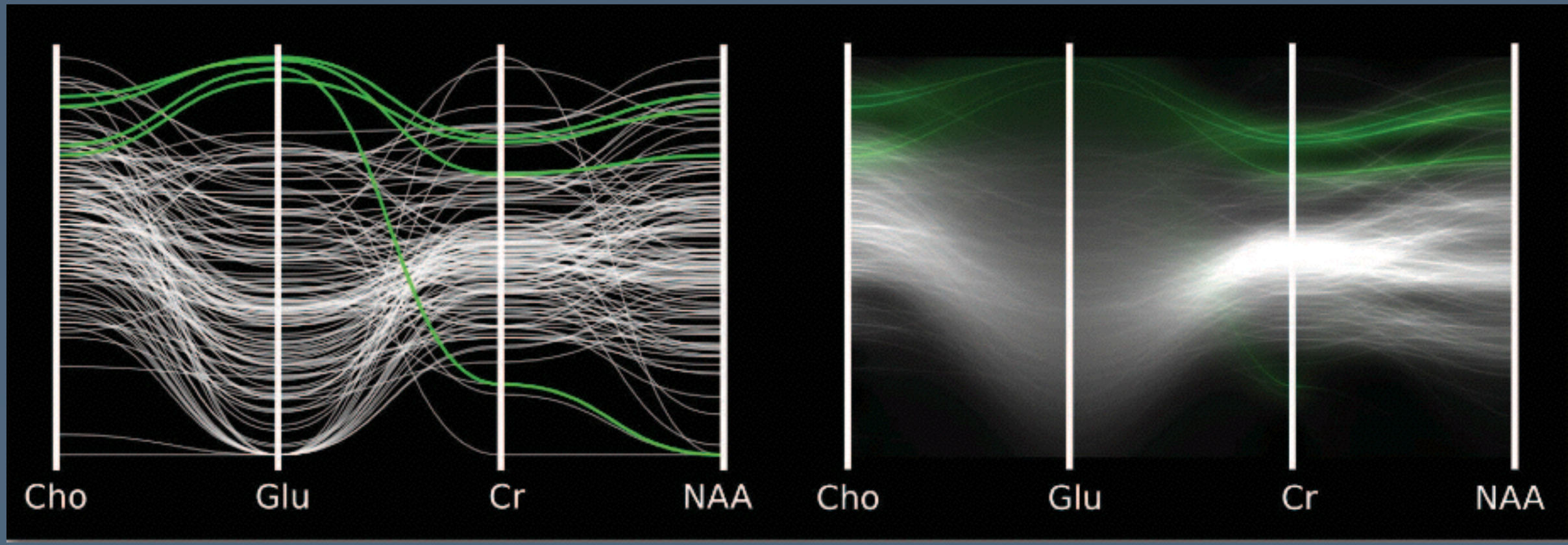


image: link to d3 demo

Interaction: Brushing and Linking



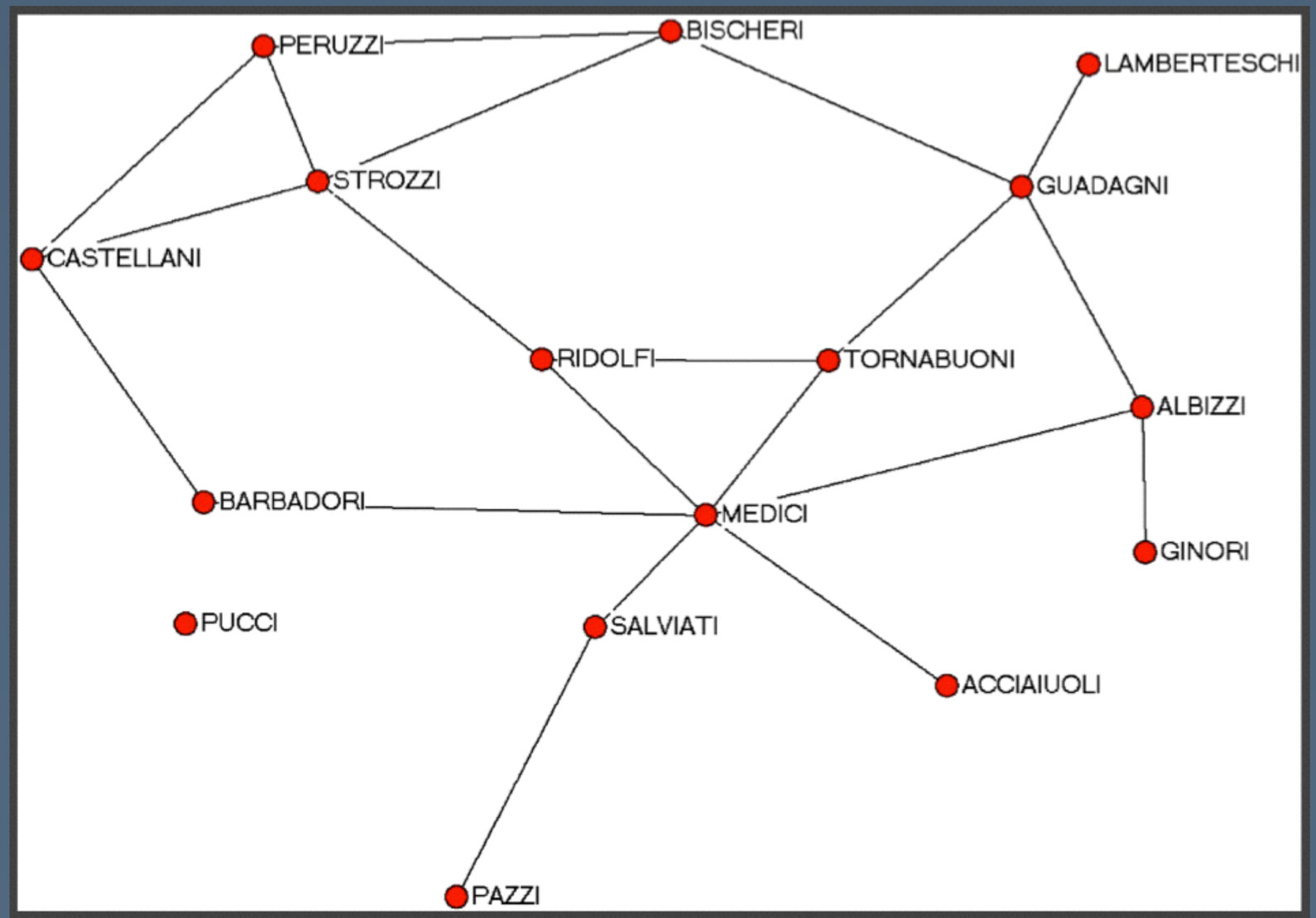
Visual Uncertainty



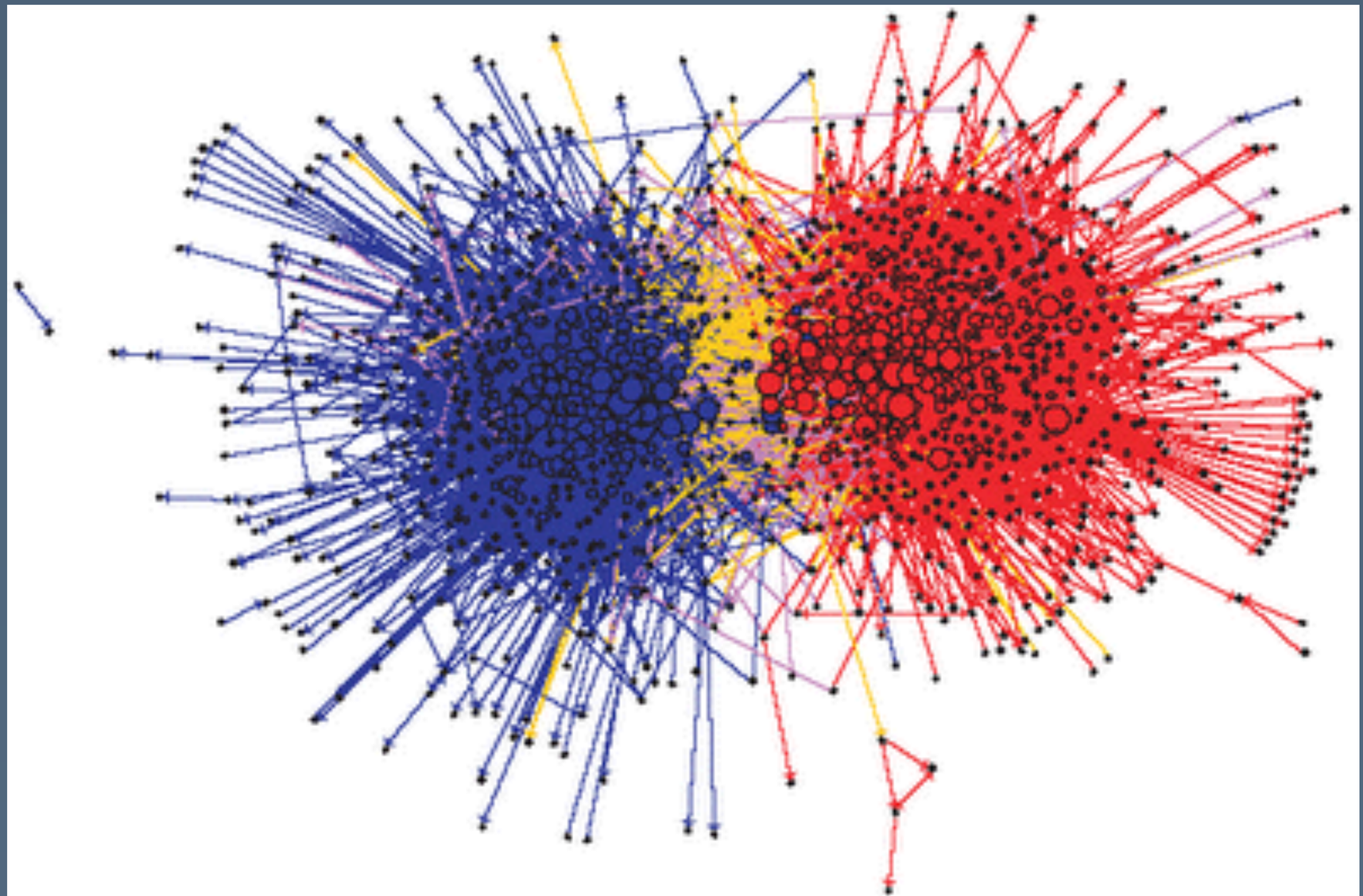
Feng et al., 2010

Network Visualization

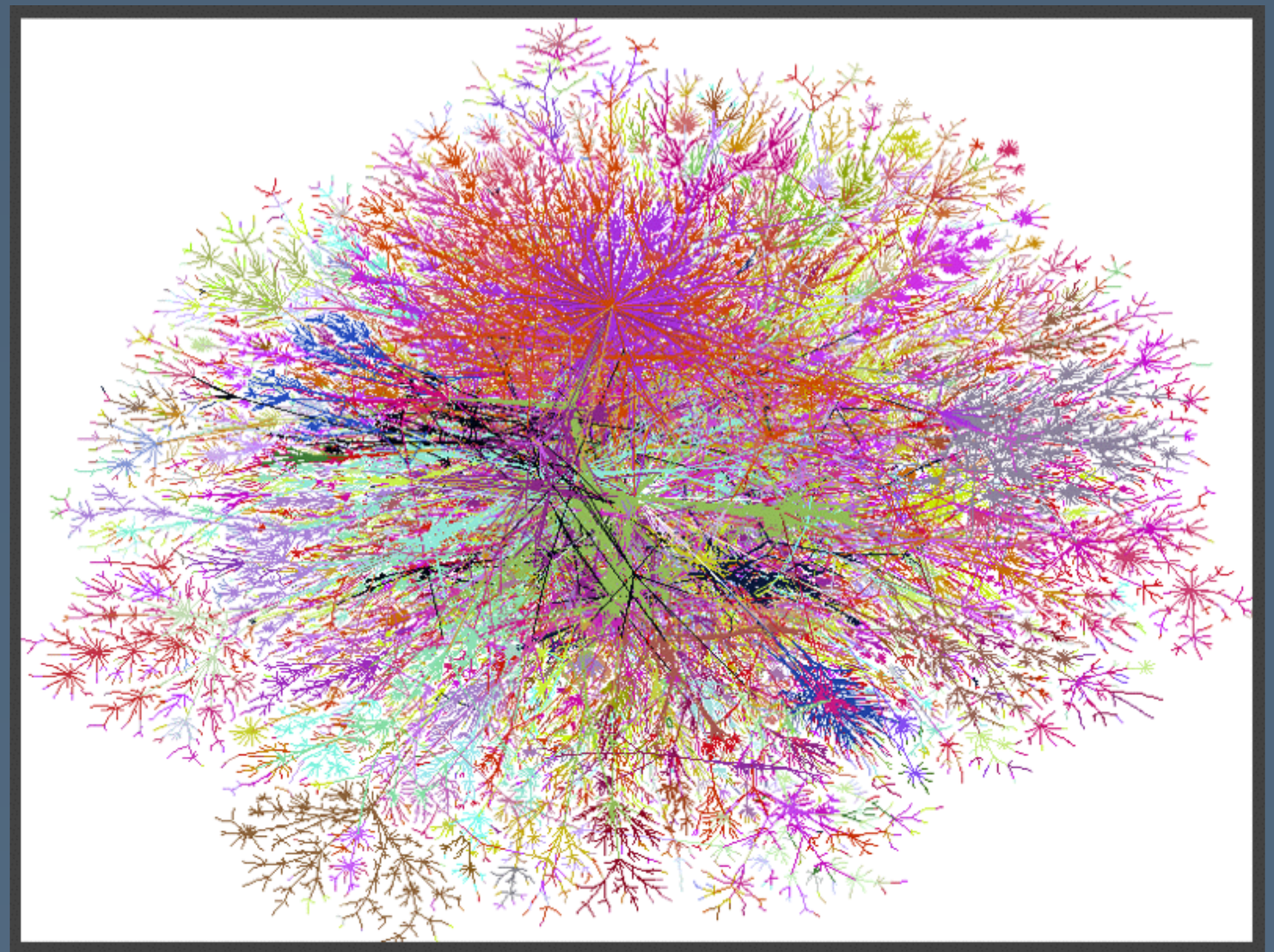
Putting things into perspective



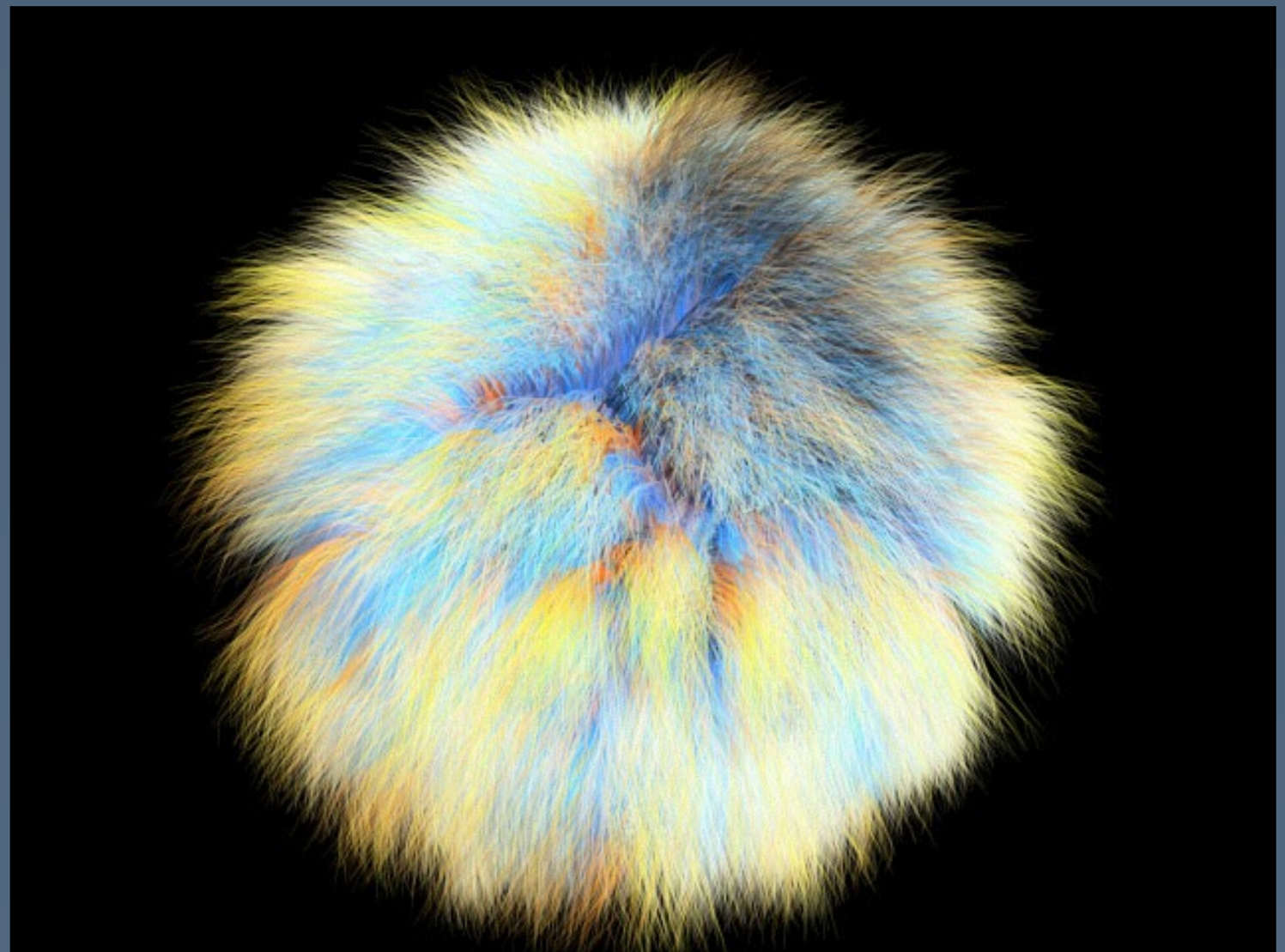
Putting things into perspective



Putting things into perspective

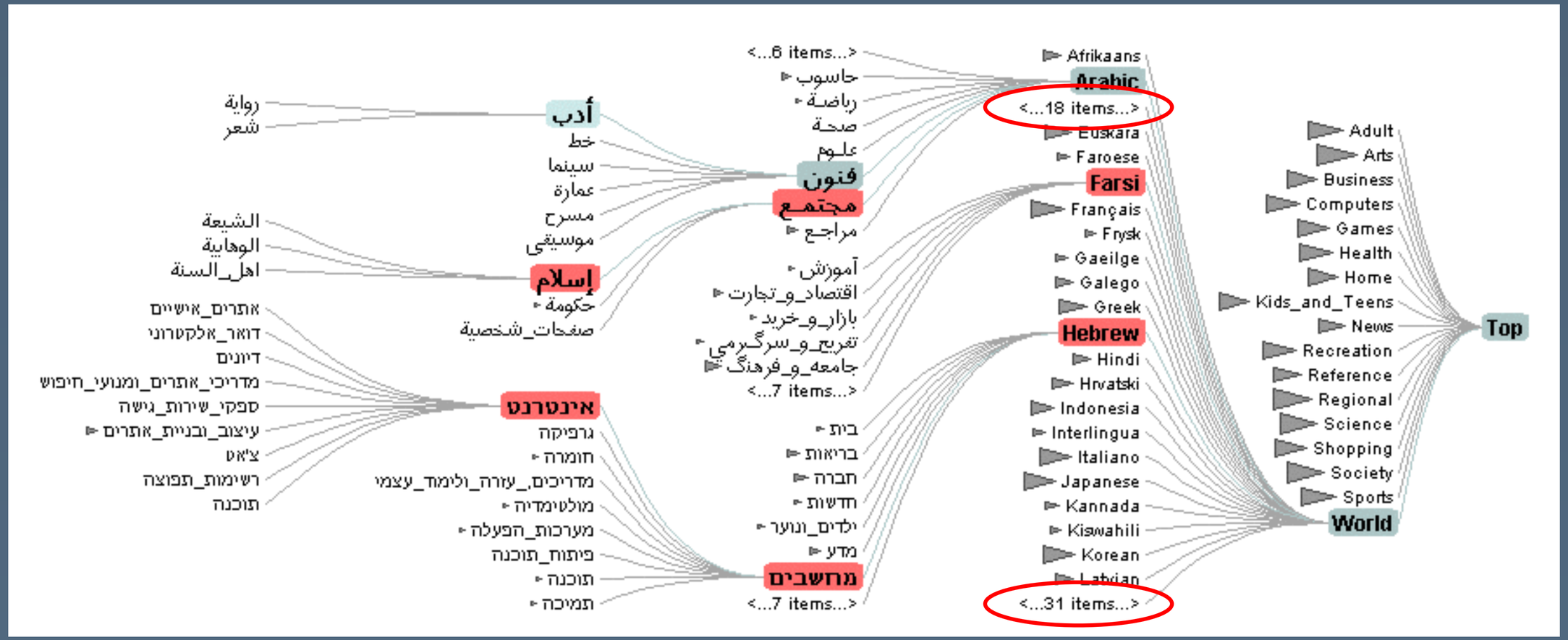


Putting things into perspective



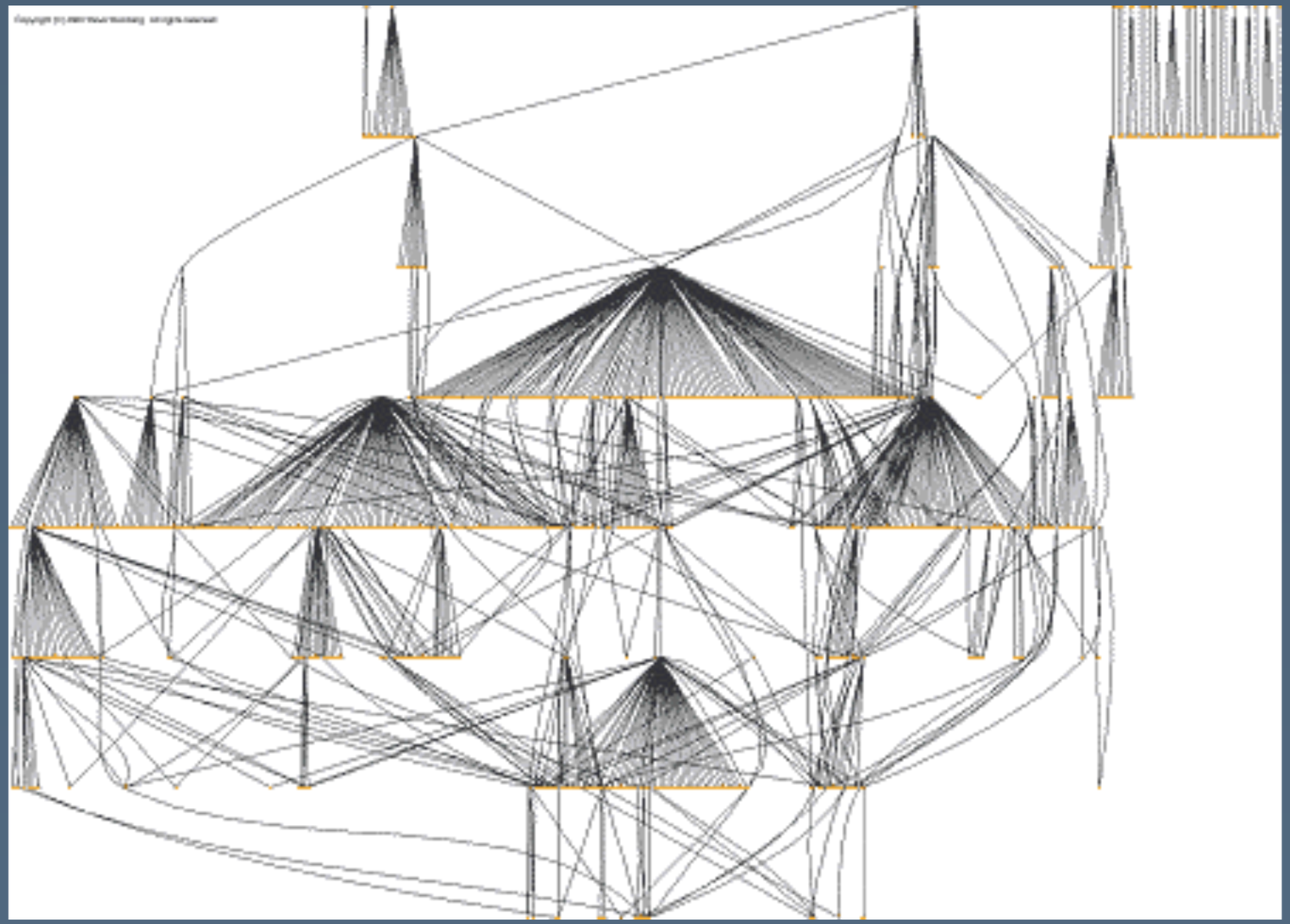
Degree-of-interest trees

Space-constrained, multi-focal tree layout



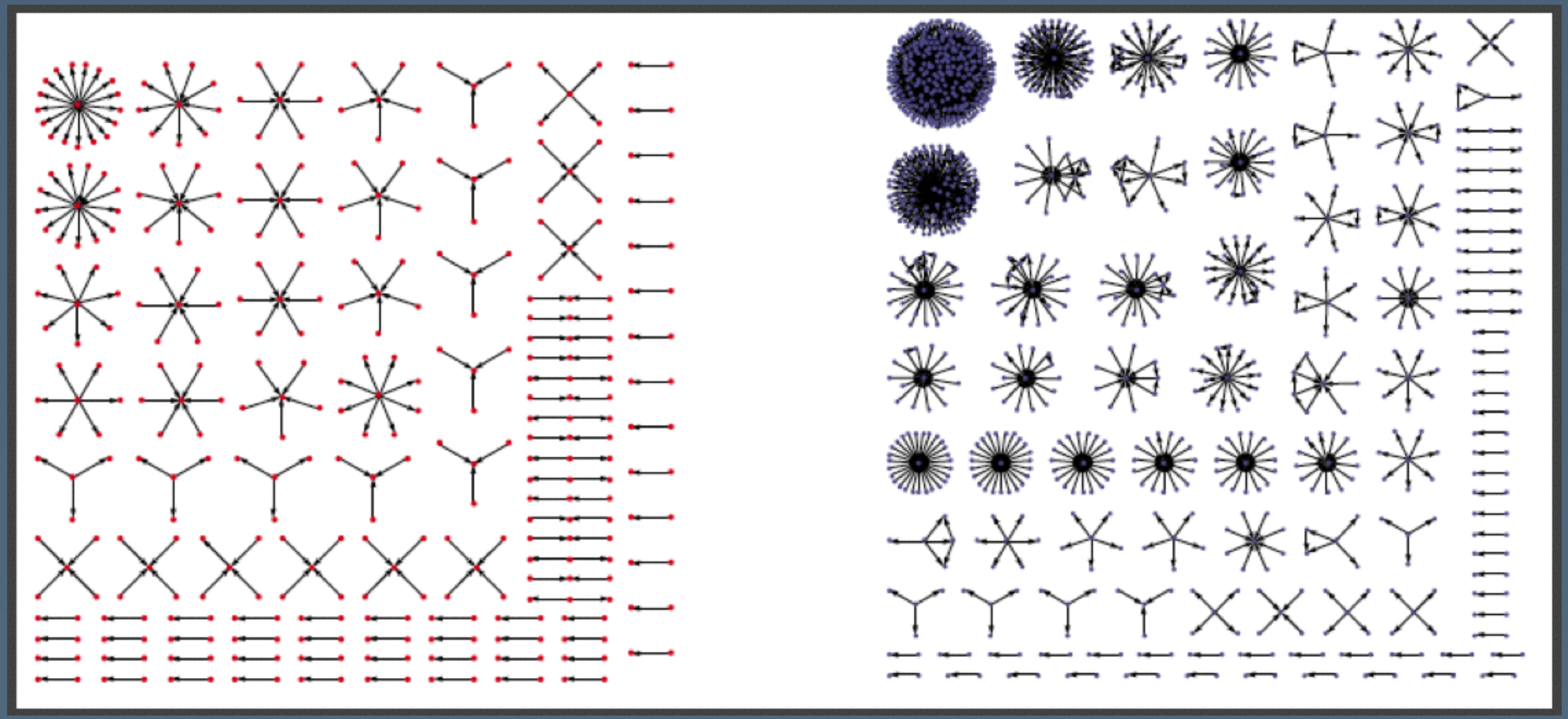
Heer & Card, 2004

Hierarchical graph layout



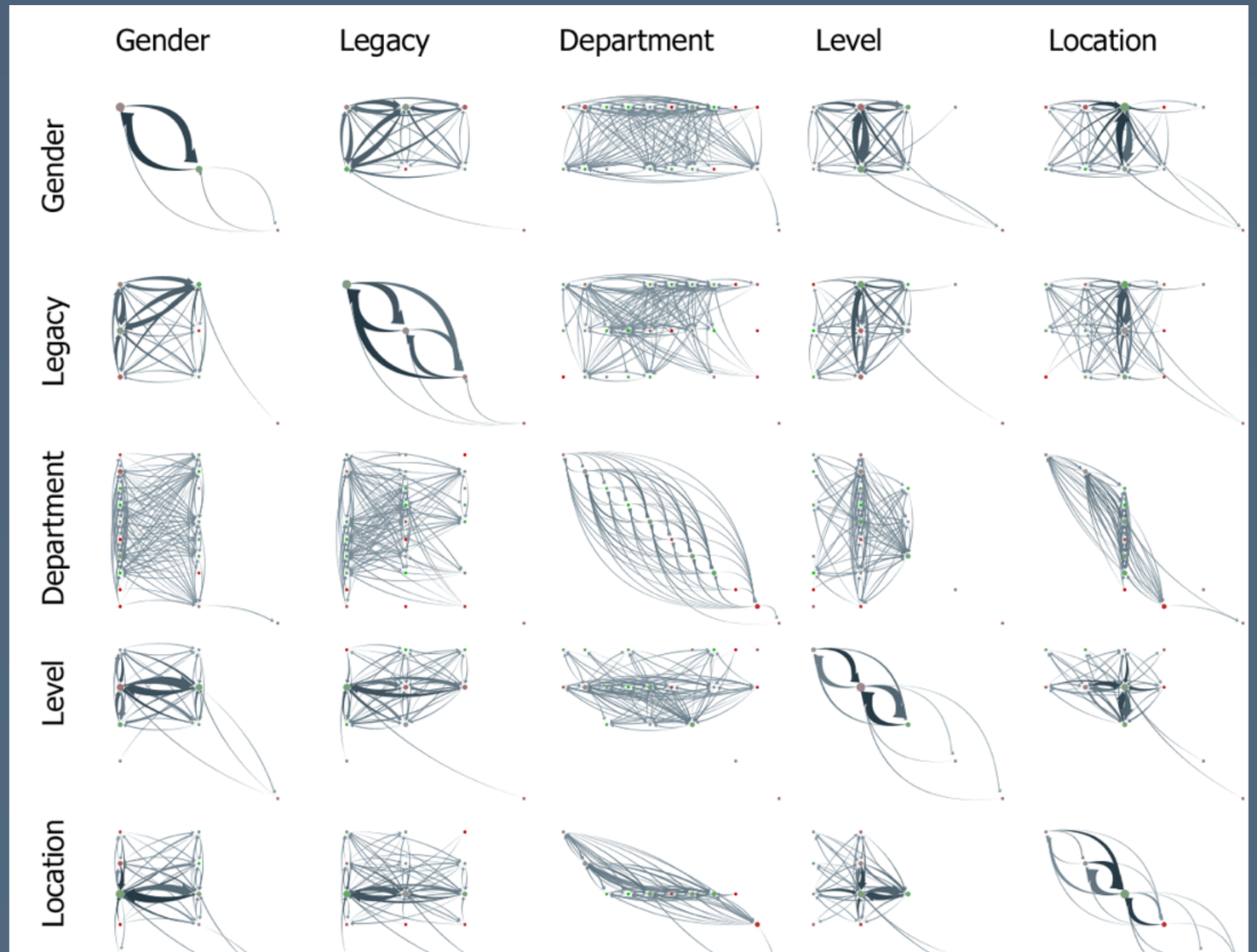
Gnutella network

Motifs



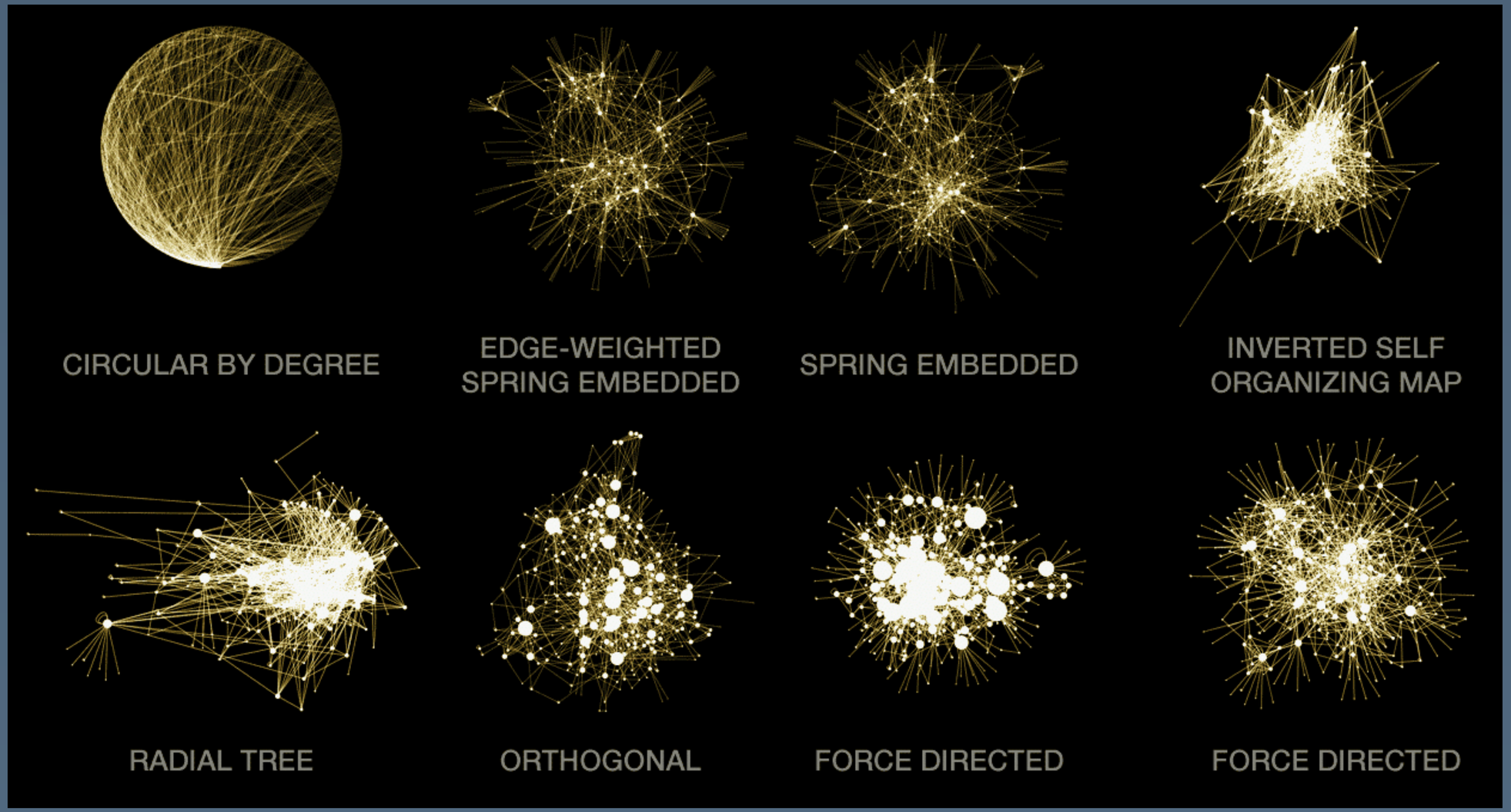
Adamic et al., 2008

Attribute-based aggregation



PivotGraph Matrix
Wattenberg, 2005

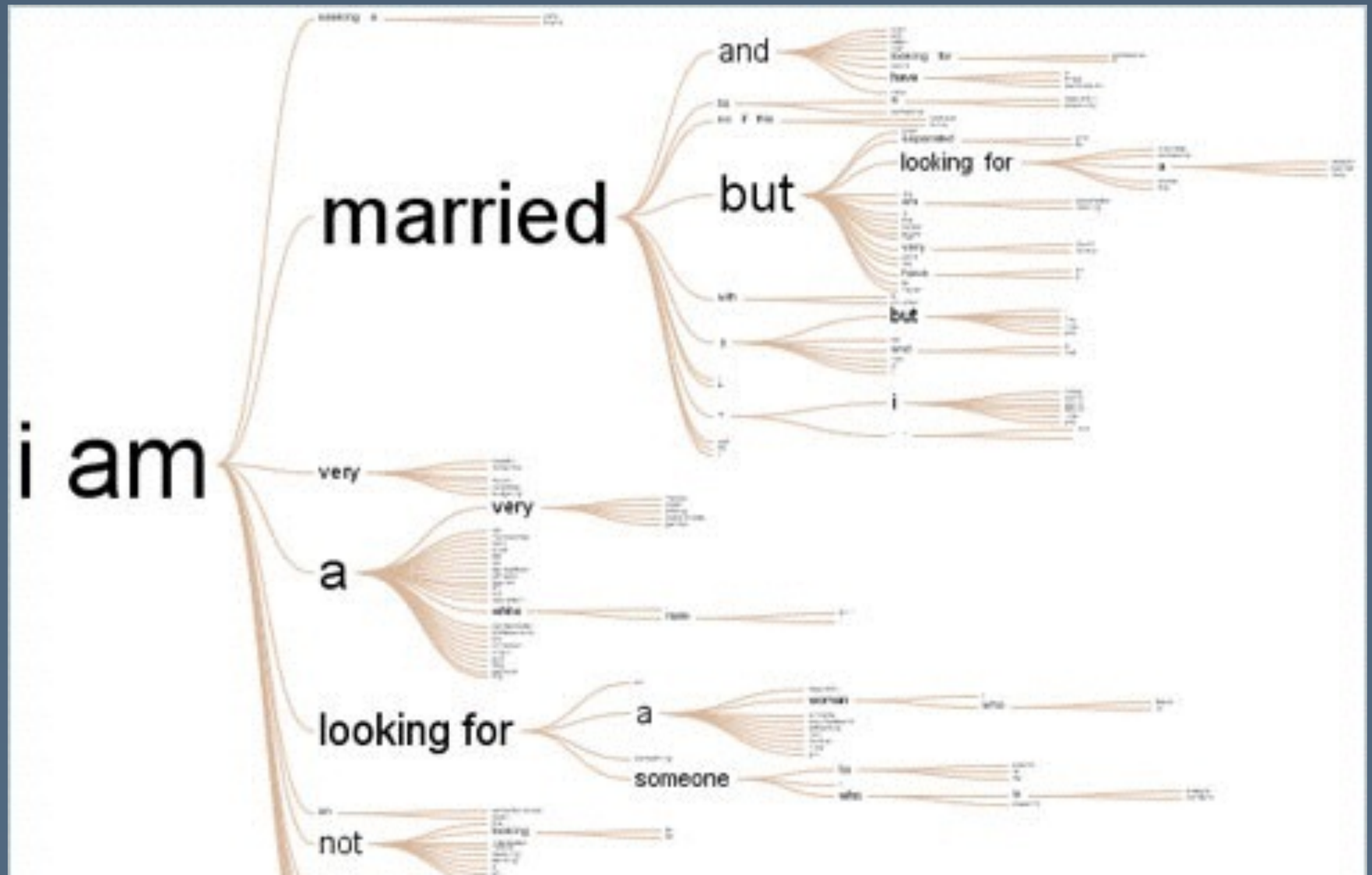
Uh oh…



hiveplot

Text Visualization

Word Tree

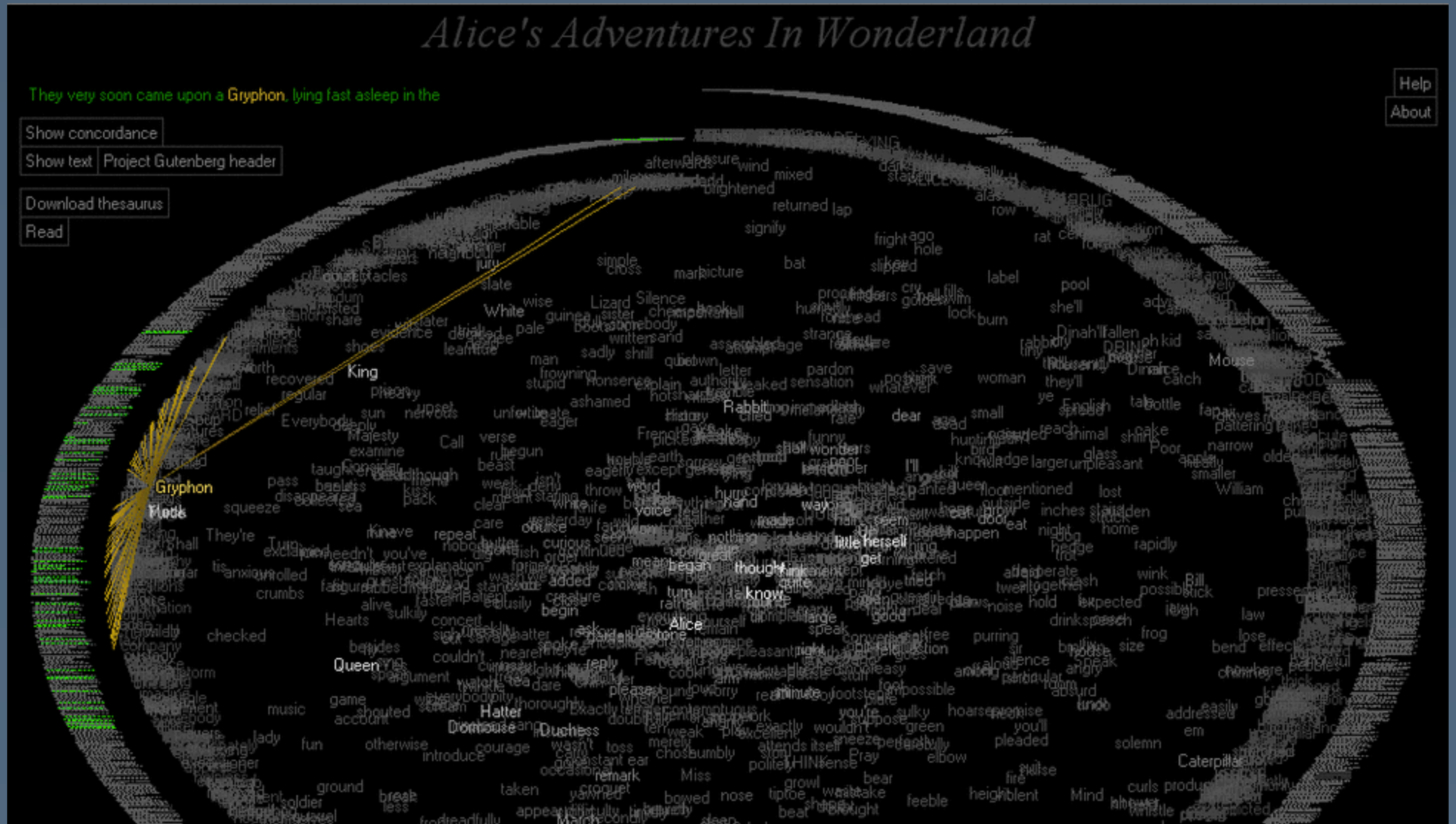


Wattenberg & Viegas, 2007

Parallel Tag Clouds

Collin & Viegas, 2009

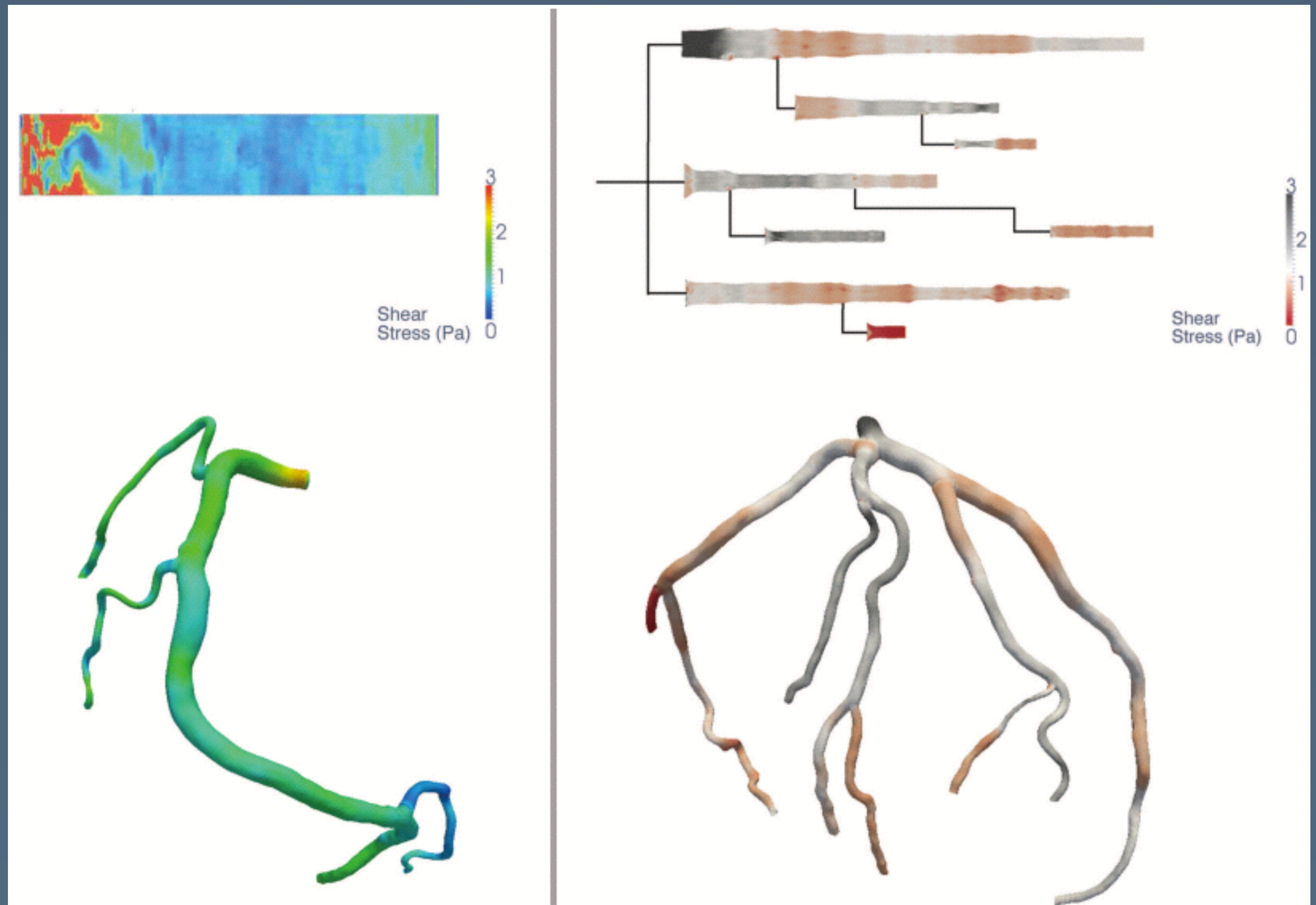
TextArc



Paley, 2002

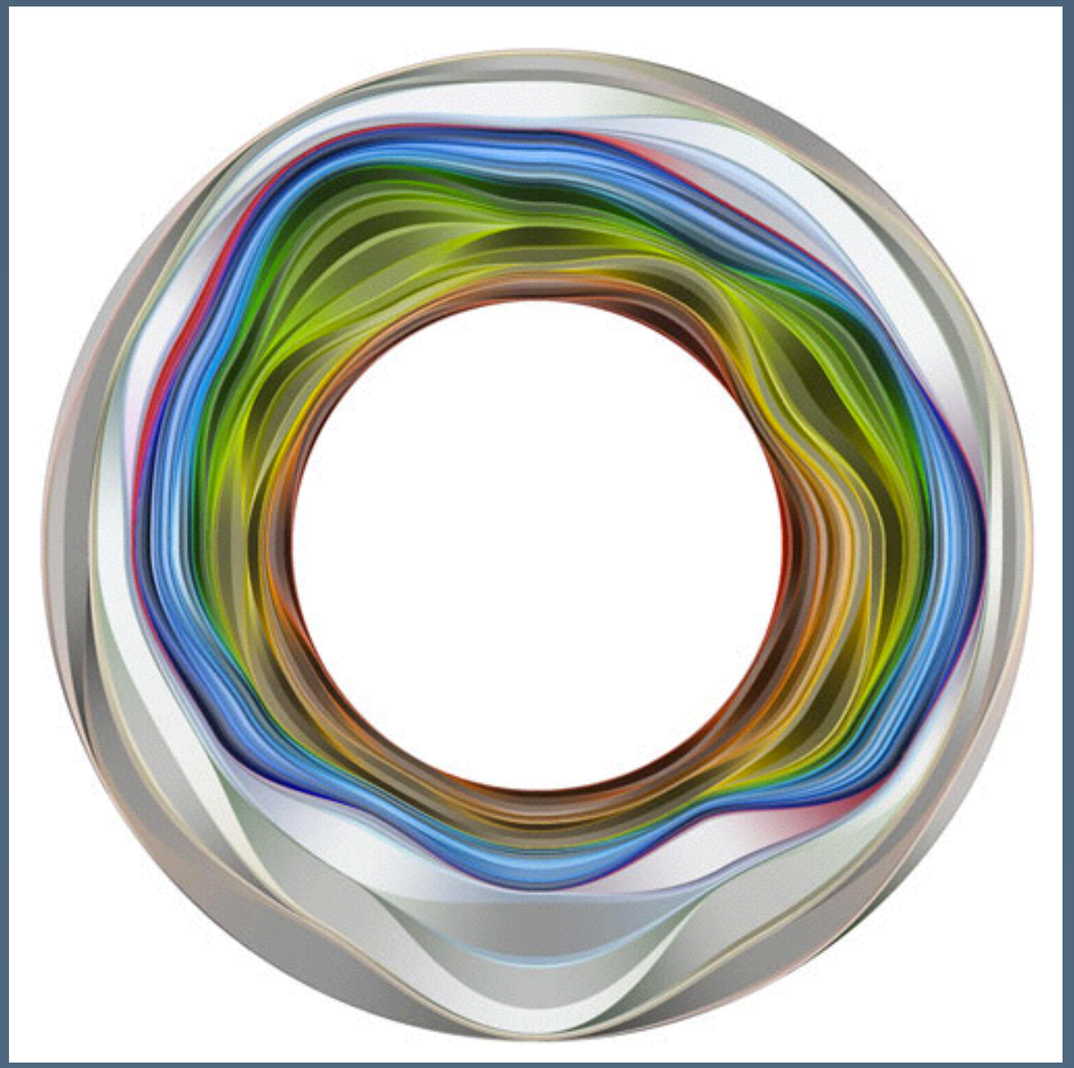
Sharing/ Communicating

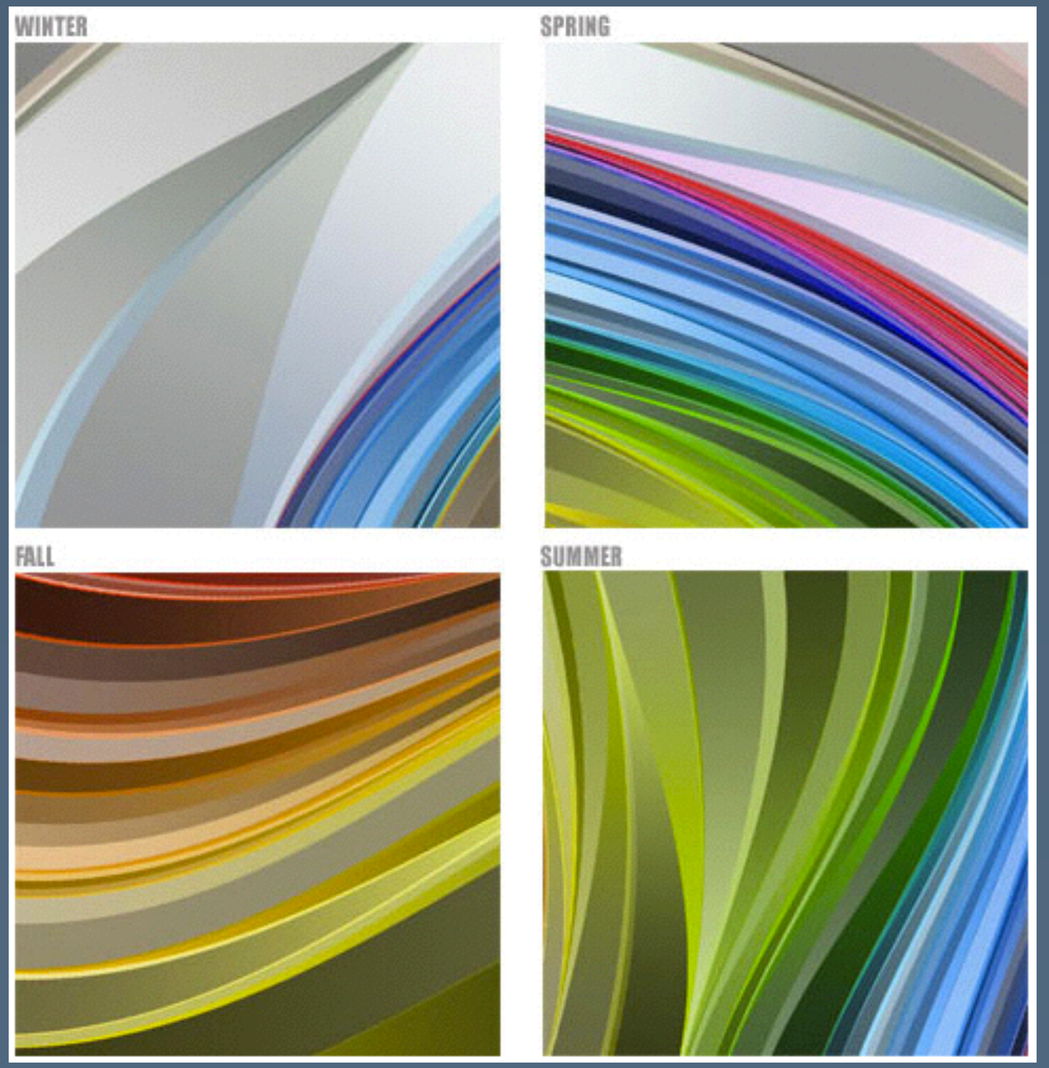
Heart Disease Diagnosis



Borkin et al., 2011

Flickr Flow





Viegas & Wattenberg: http://hint.fm/projects/flickr/