

DATA VISUALIZATION AND VISUAL ANALYTICS

S. Rinzivillo – rinzivillo@isti.cnr.it

Relational Structures: Networks



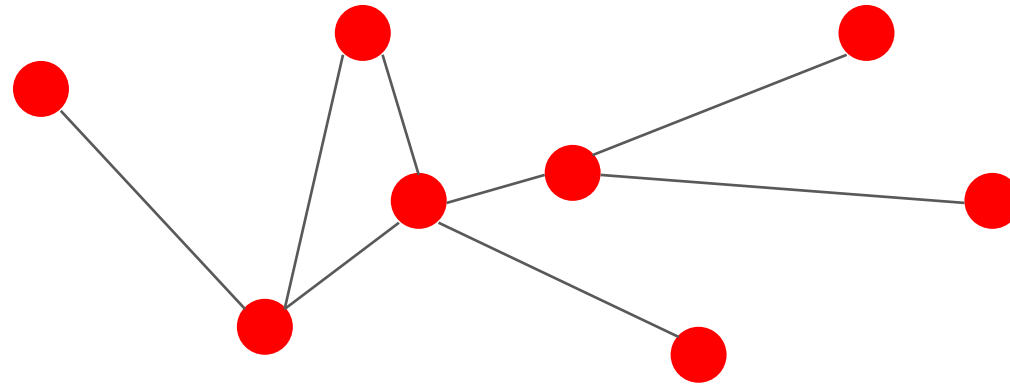
NETWORKS

- Data main focus is relationship
- Study the patterns of connection among different parts of a complex system
- Visualization has a key roles to add insights to numerical analysis

NETWORKS AND GRAPHS



BASIC ELEMENTS



▪ **components:** nodes, vertices

N

▪ **interactions:** links, edges

L

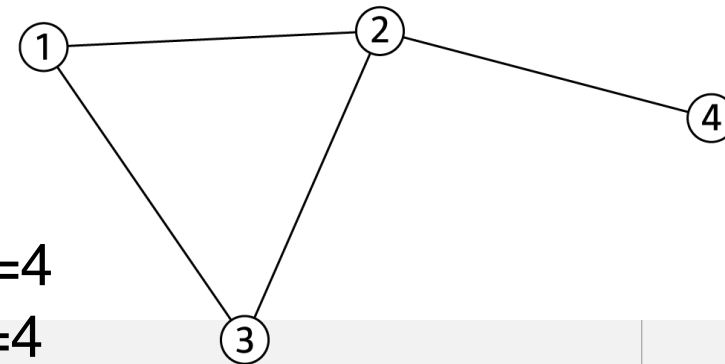
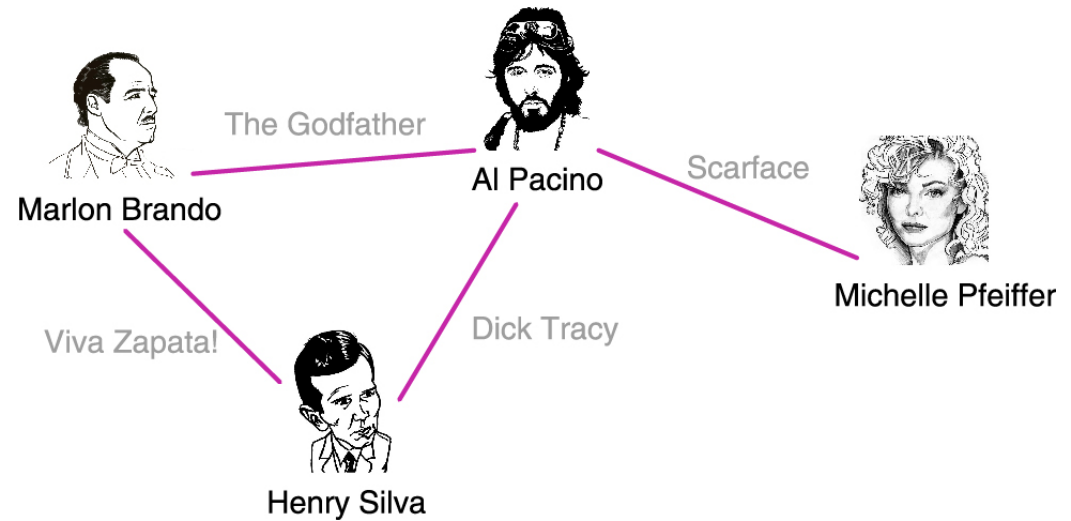
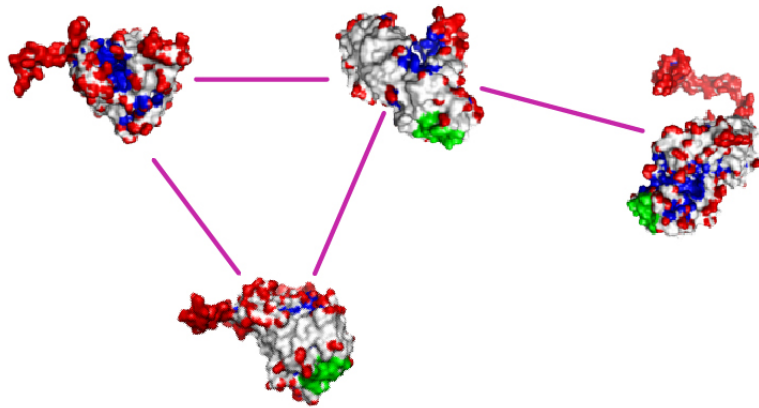
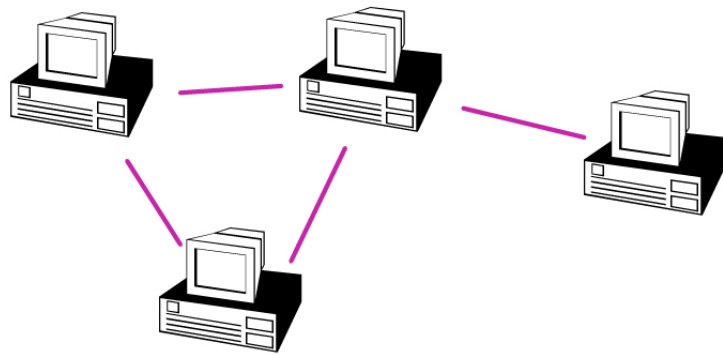
▪ **system:** network, graph

(N,L)

NETWORKS OR GRAPHS? A COMMON LANGUAGE

Network refer to a real system

Graph refers to mathematical representation of a network

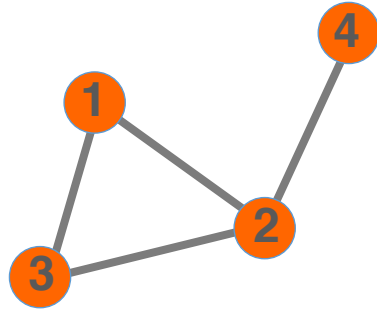


$N=4$

$L=4$

UNDIRECTED VS DIRECTED

Undirected

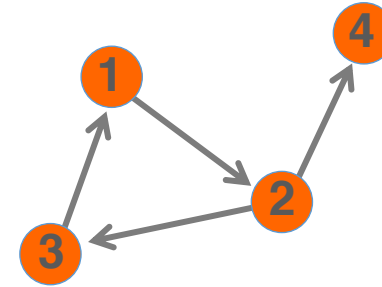


$$A_{ij} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$A_{ii} = 0 \quad A_{ij} = A_{ji}$$
$$L = \frac{1}{2} \sum_{i,j=1}^N A_{ij} \quad \langle k \rangle = \frac{2L}{N}$$

Actor network, protein-protein interactions

Directed



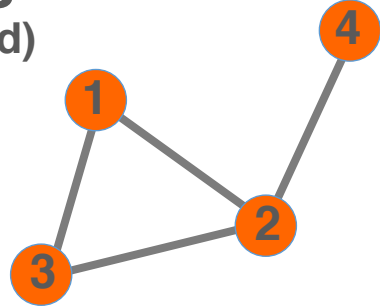
$$A_{ij} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$A_{ii} = 0 \quad A_{ij} \neq A_{ji}$$
$$L = \sum_{i,j=1}^N A_{ij} \quad \langle k \rangle = \frac{L}{N}$$

WWW, citation networks

UNWEIGHTED VS WEIGHTED

Unweighted
(undirected)



$$A_{ij} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$A_{ii} = 0$$

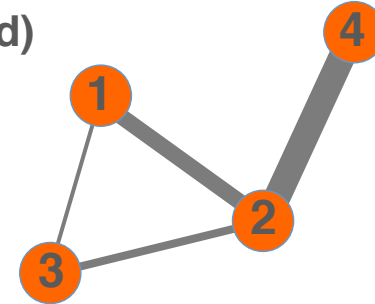
$$A_{ij} = A_{ji}$$

$$L = \frac{1}{2} \sum_{i,j=1}^N A_{ij}$$

$$\langle k \rangle = \frac{2L}{N}$$

protein-protein interactions, www

Weighted
(undirected)



$$A_{ij} = \begin{pmatrix} 0 & 2 & 0.5 & 0 \\ 2 & 0 & 1 & 4 \\ 0.5 & 1 & 0 & 0 \\ 0 & 4 & 0 & 0 \end{pmatrix}$$

$$A_{ii} = 0$$

$$A_{ij} = A_{ji}$$

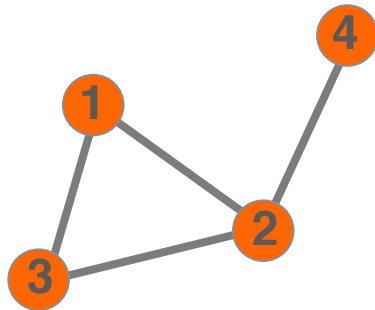
$$L = \frac{1}{2} \sum_{i,j=1}^N \text{nonzero}(A_{ij})$$

$$\langle k \rangle = \frac{2L}{N}$$

Call Graph, metabolic networks

NETWORK INTERNAL REPRESENTATION

- Three main methods
 - a) Adjacency Lists
 - b) Matrices
 - c) Node-link diagrams



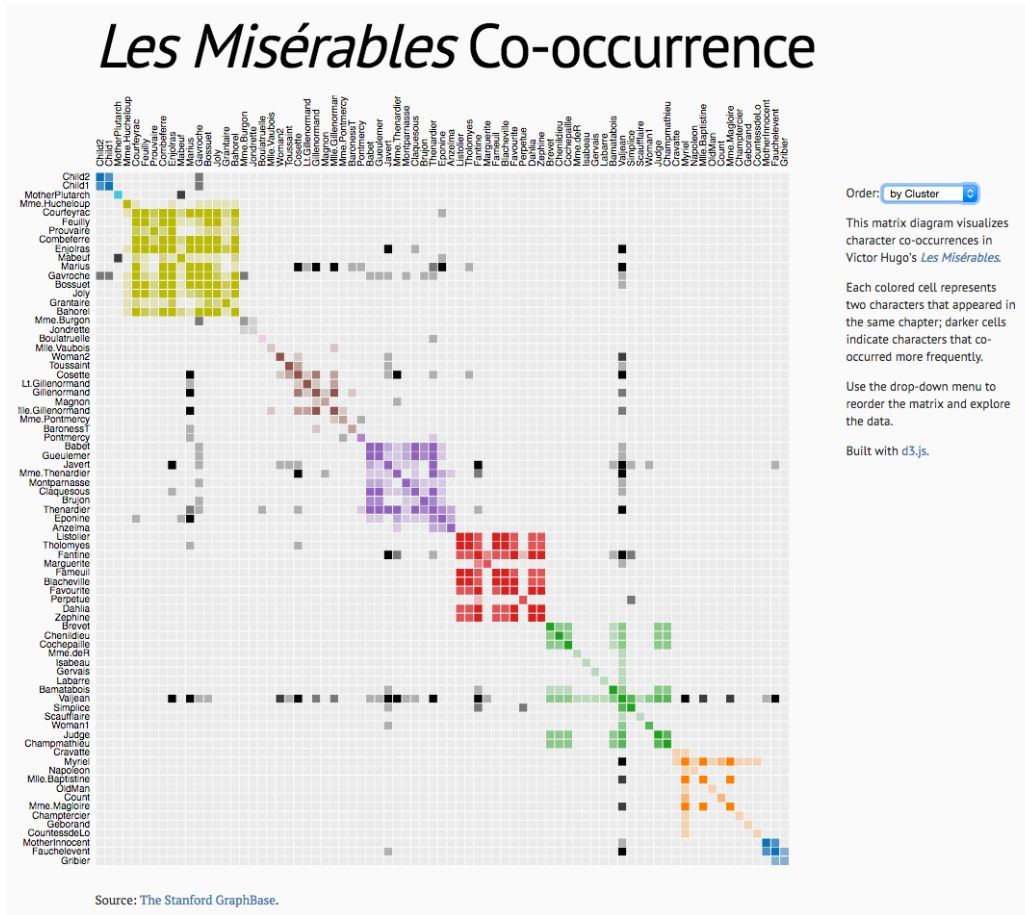
1: 2,3
2: 3,1,4
3: 1,2
4: 2
a)

$$A_{ij} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

b)

1,2
1,3
2,3
2,4
c)

ADJACENCY MATRIX



- Each cell ij represents an edge from vertex i to vertex j
- Effectiveness of visualization depends on rows/columns ordering
- First example by Jacques Bertin (with paper strips rearranged by hand)
- Effective also for highly connected graphs

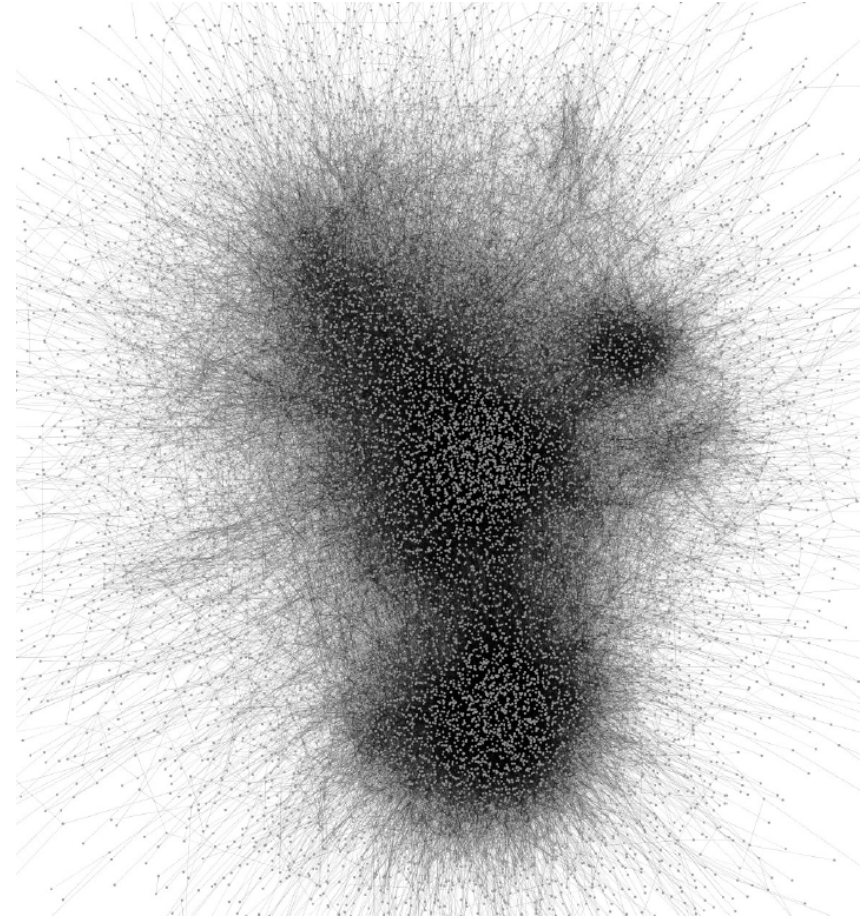
NODE-LINK REPRESENTATION



- Symbolic elements for nodes
- Lines for connection among nodes
- Physical networks (roads, power grids) have a natural spatial encoding
- Abstract networks need layouts to infer a spatial position for nodes

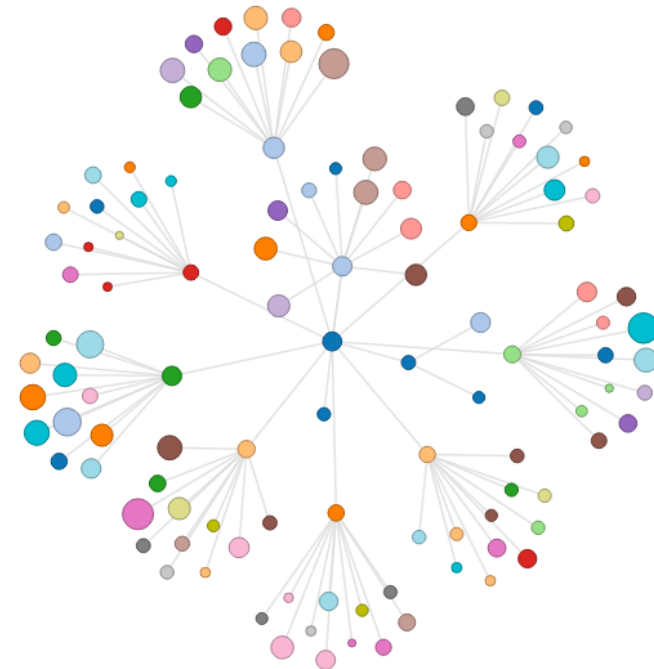
PROBLEMS OF NODE-LINK DIAGRAMS

- Occlusion of node and link crossings
- Large networks may produce hairball like networks
- Many algorithms to produce effective layouts to reduce cluttering



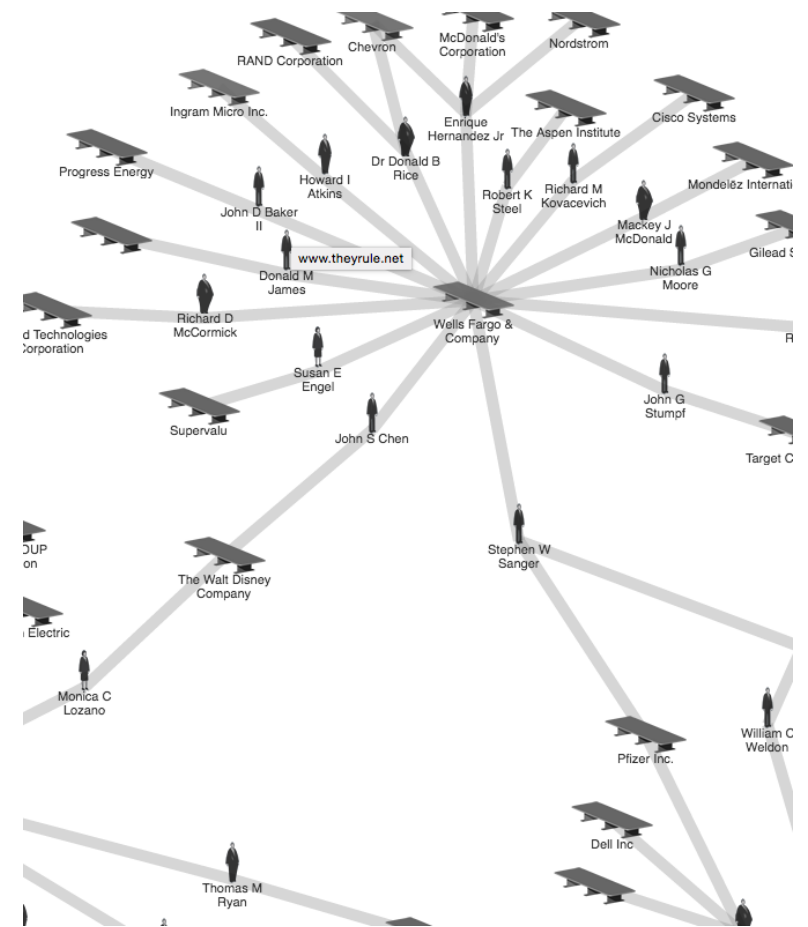
CLUTTERING REDUCTION

- Interaction to switch between different layouts
- Effective positioning of labels
 - Centered on nodes
 - Visualization based on interaction and mouse hover



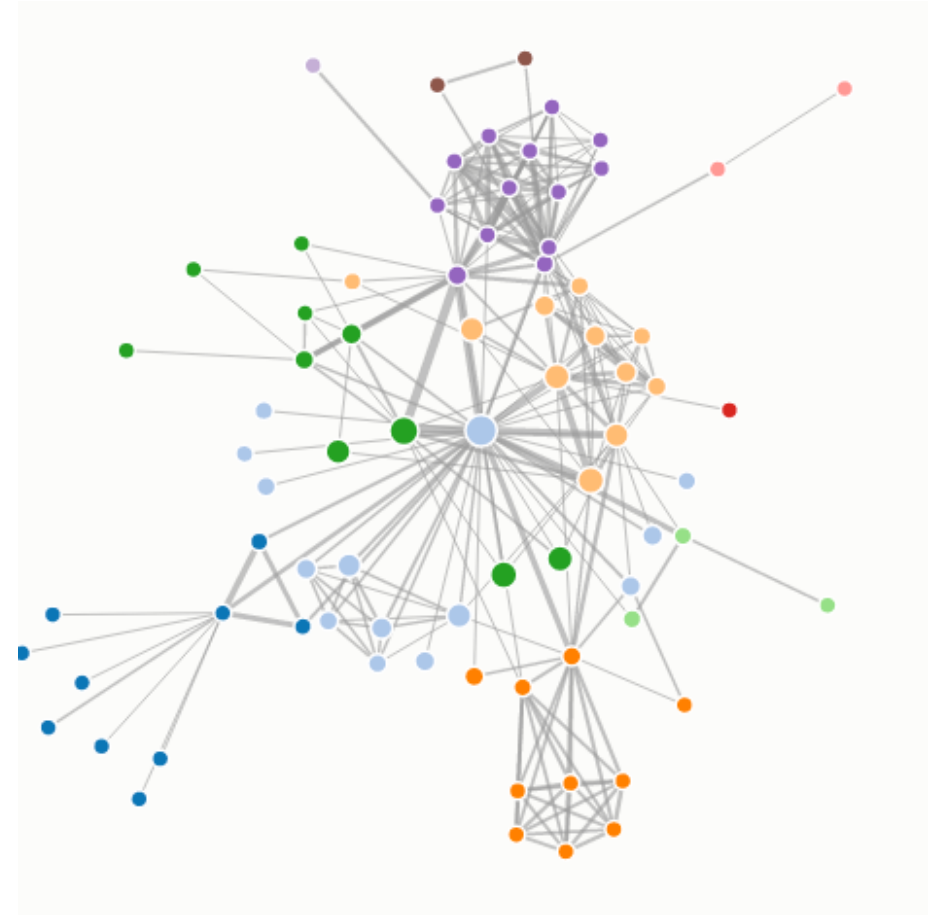
CLUTTERING REDUCTION

- Collapsing nodes into clusters



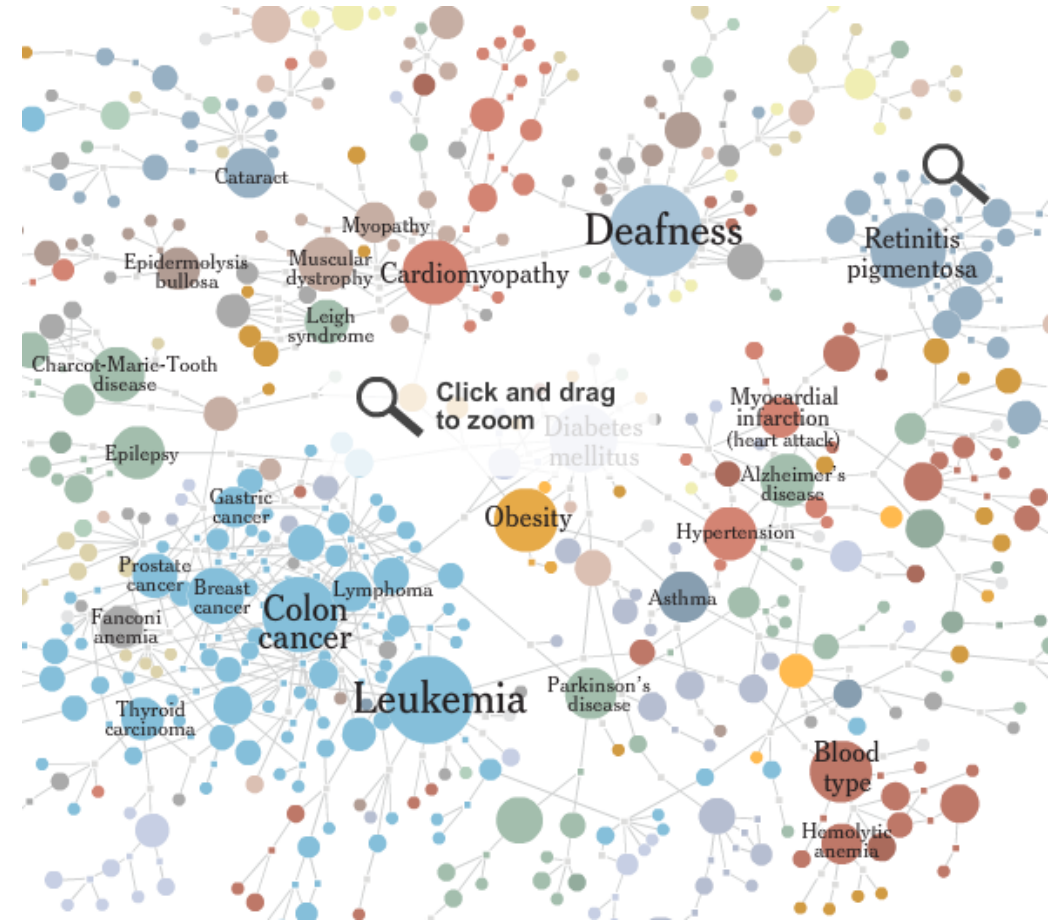
CLUTTERING REDUCTION

- Zooming and context distortion

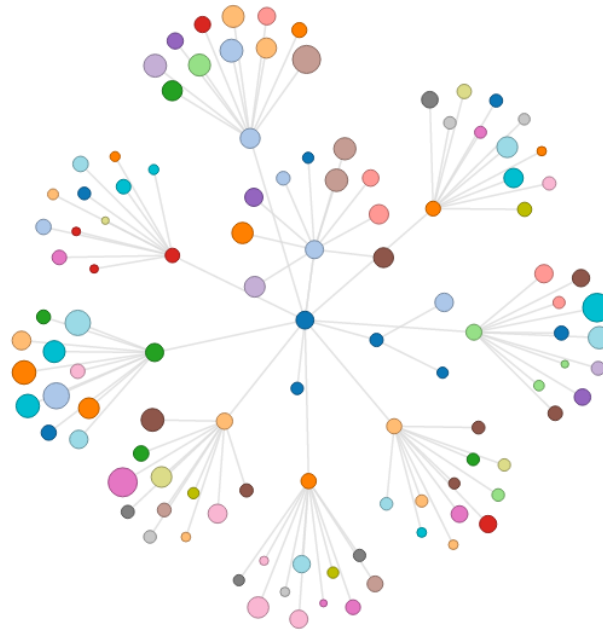


CLUTTERING REDUCTION

- Zooming and context distortion

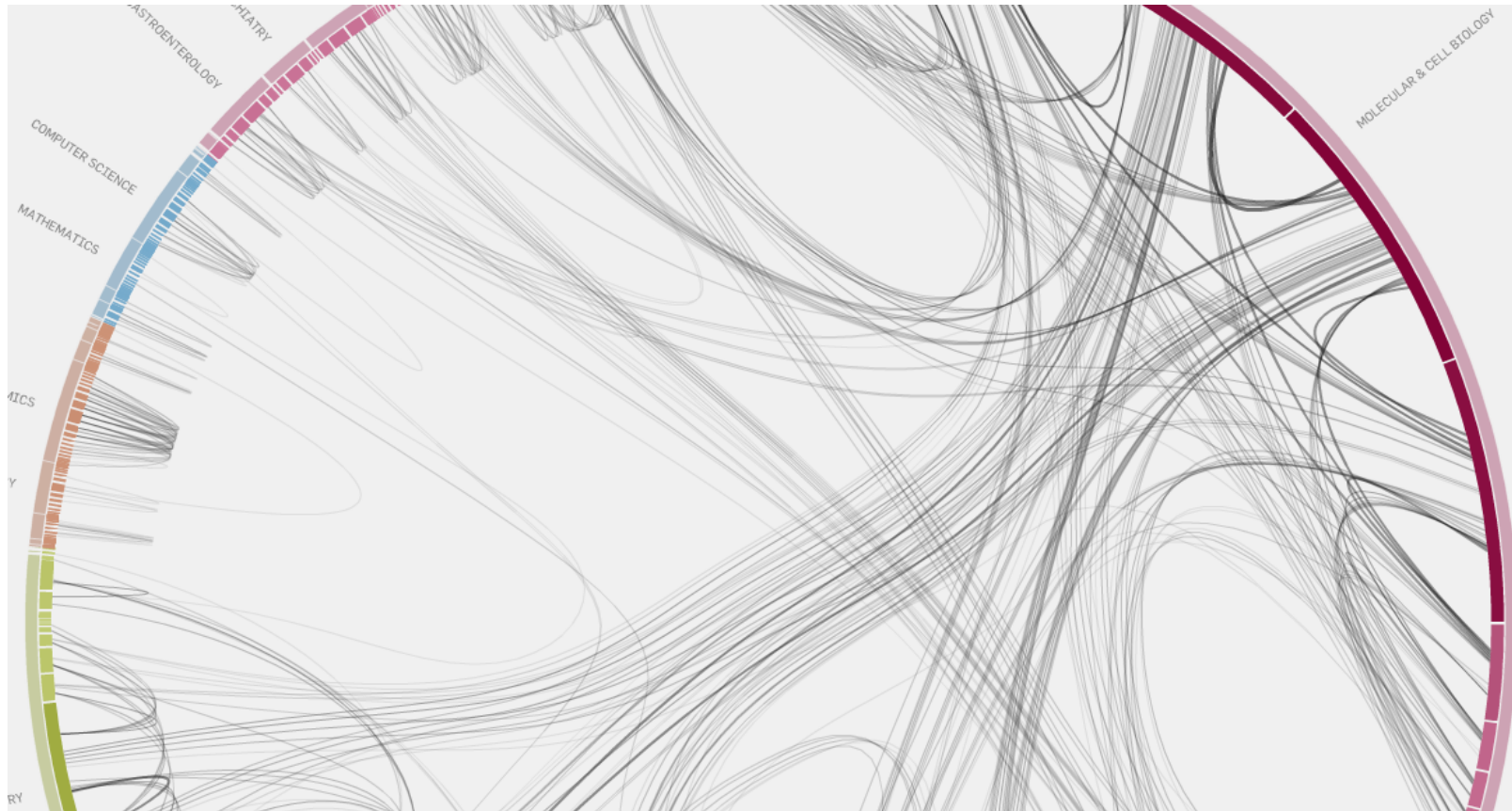


CASE STUDY: FORCE DIRECTED



CASE STUDY: INFORMATION FLOW

Circular Layout



CASE STUDY: SANKEY TYPE DIAGRAMS

