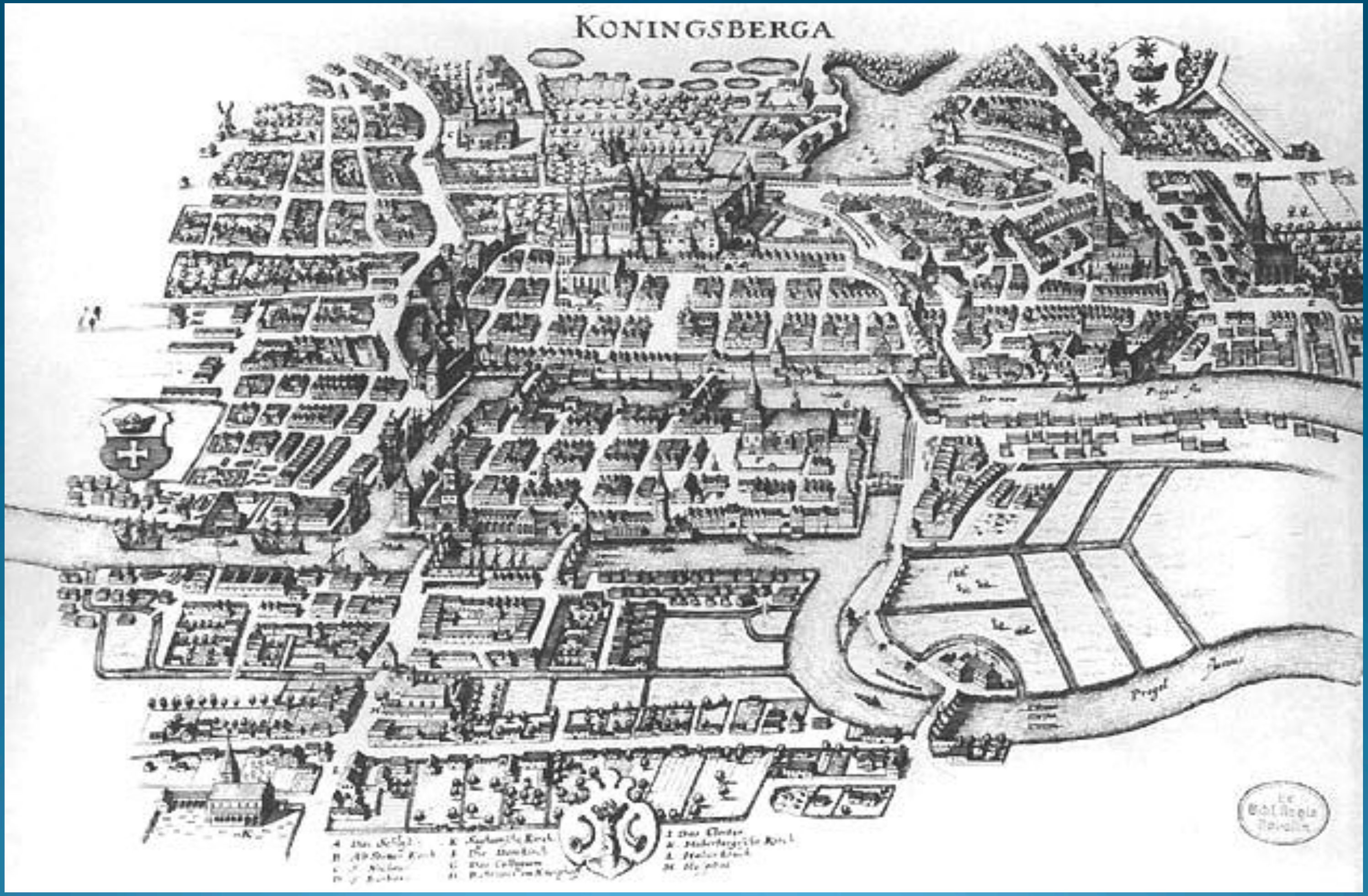
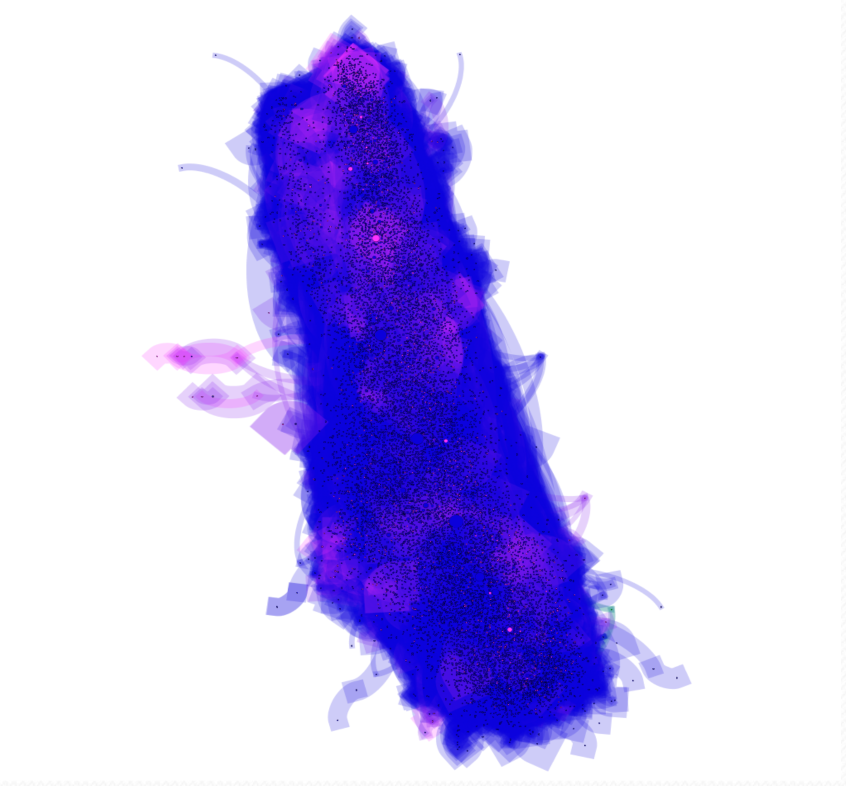
Networks over Space and Time



@jotis13

When You Probably Don't Want Them

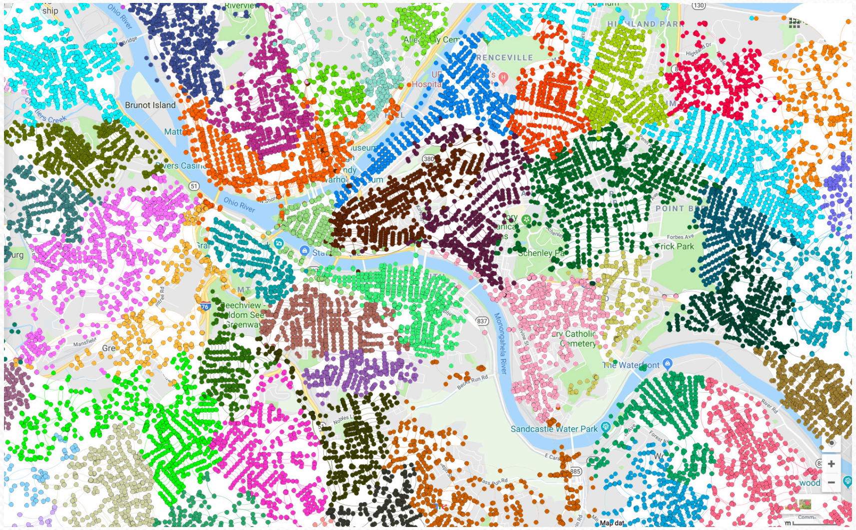
- if you have to force the connected/connecting structure in weird, unintuitive ways
- if you have more than 1 (unipartite) or 2 (bipartite) things to connect, the math gets hard/breaks down
- if you're most concerned with geography vs. connection
- if you're most concerned with change over time (networks are static)



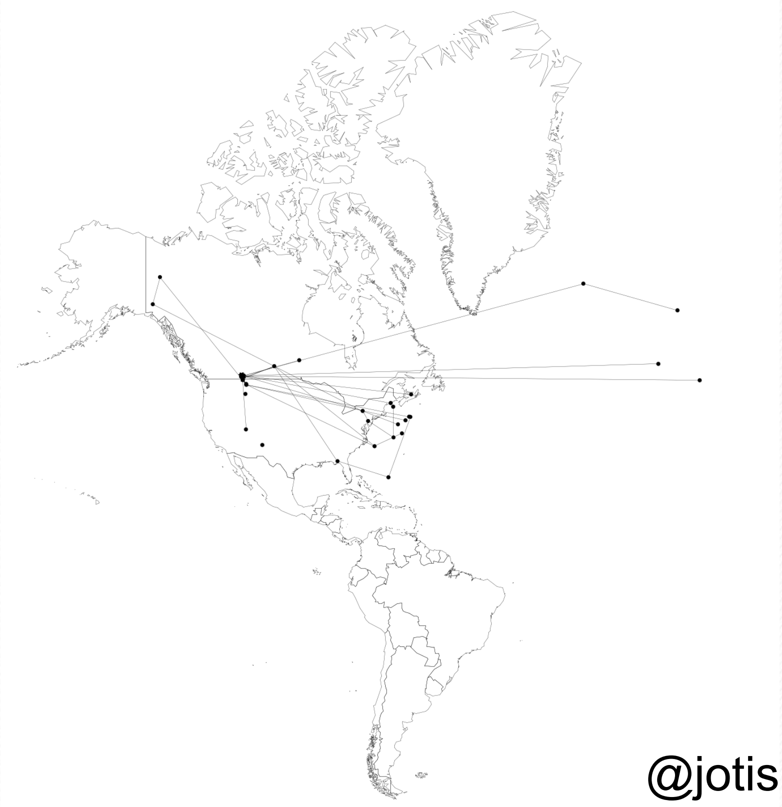
@jotis13

That Said… Space!

Plus Photoshop Method

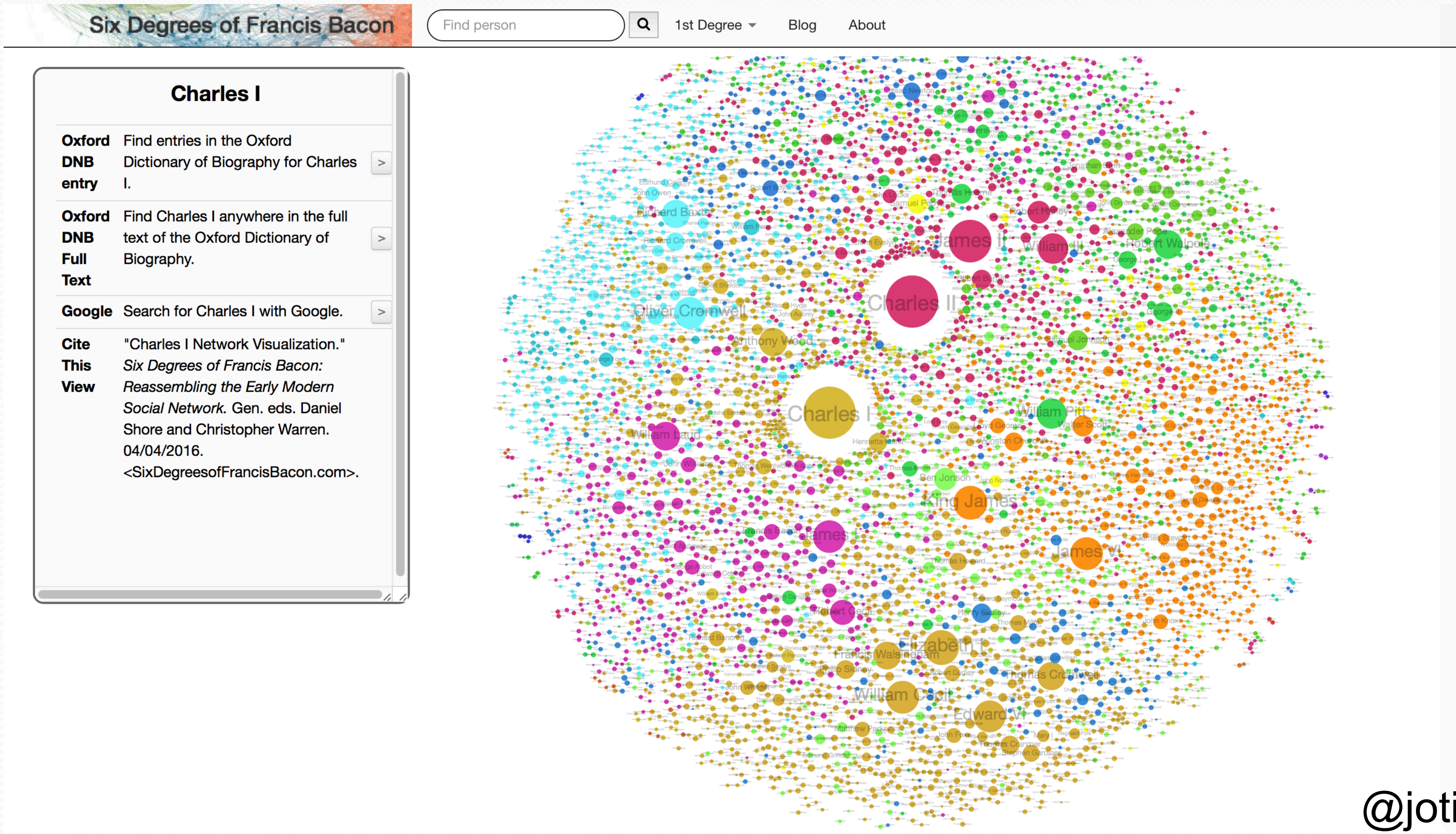


Plug-Ins Method… ish



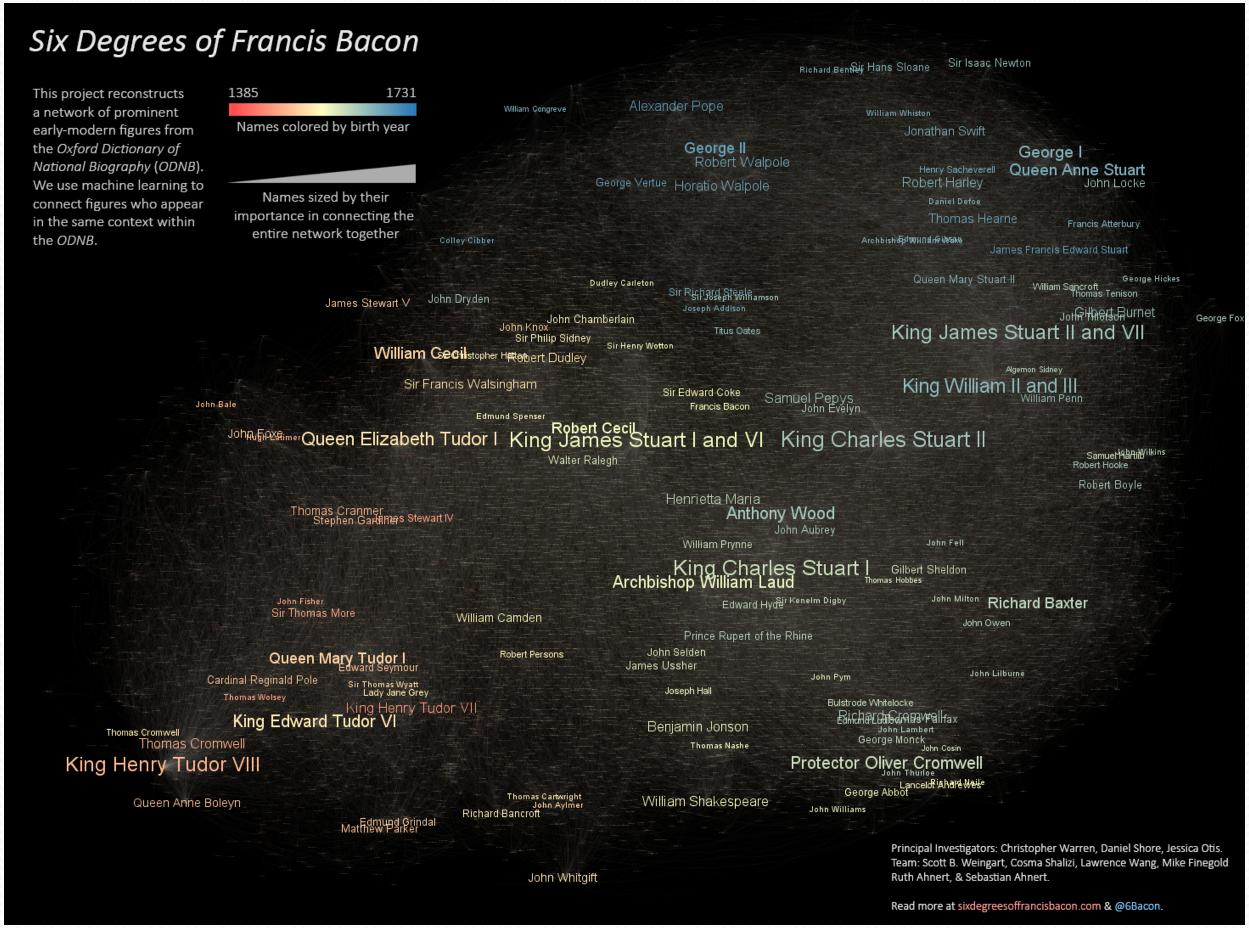
@jotis13

That Said… Time!



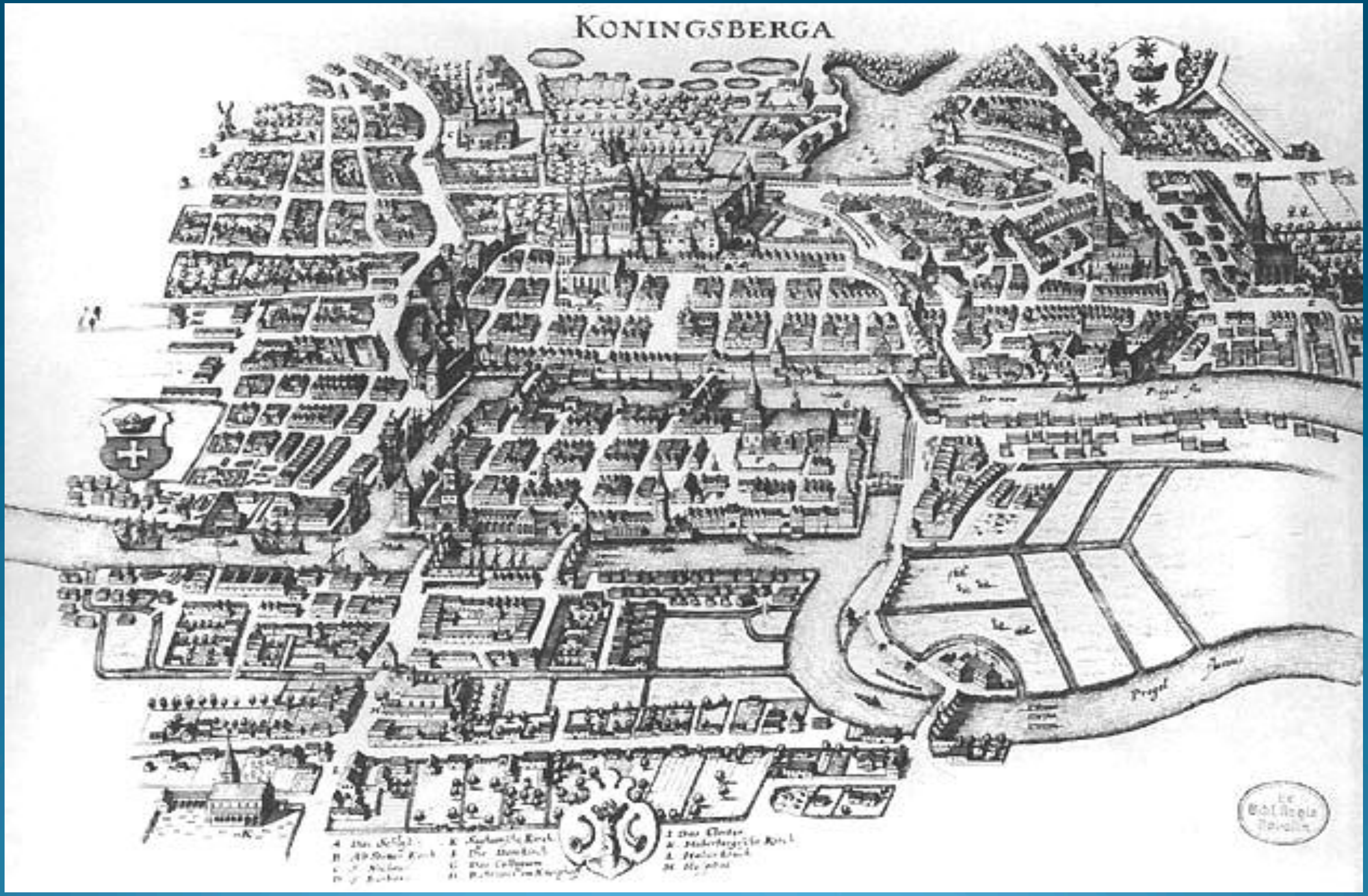
@jotis13
@6bacon

That Said… Time!



@scott_bot
@6bacon

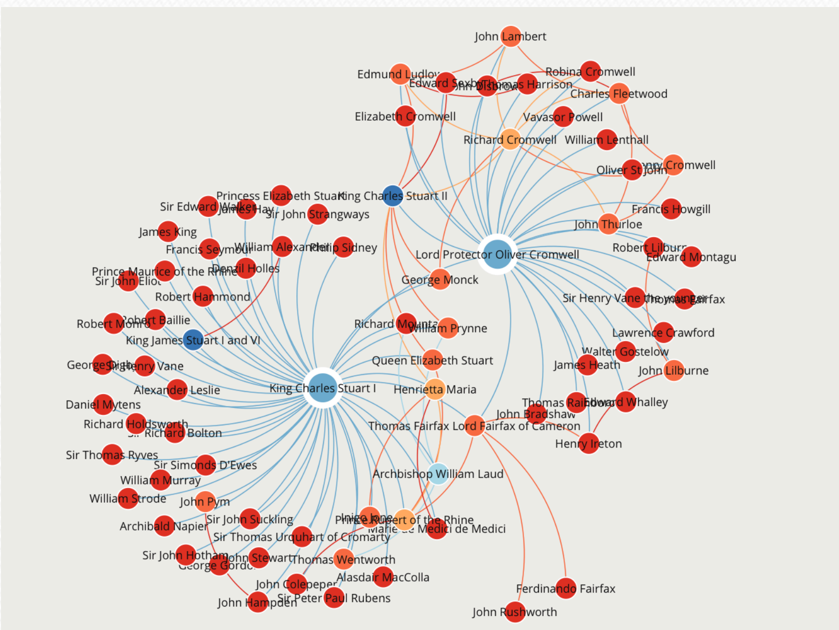
Network Metrics



@jotis13

How Can You Analyze Them?

E.g. in a social network

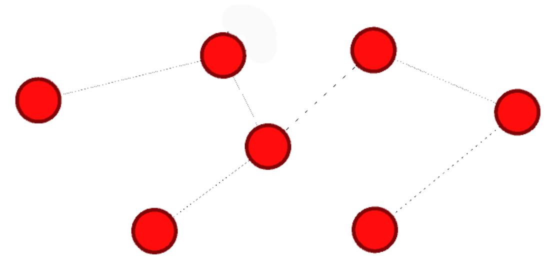


- who knows the most people?
- who knows the most highly connected people?
- how many disconnected subnetworks are there?
- whose removal severs the network into disconnected subnetworks?
- how many “hops” does it take to get from one person in the network to another?
- who do most of the shortest hops go through?
- are there any smaller cliques or “communities” within the network?

@jotis13

Global Network Analysis

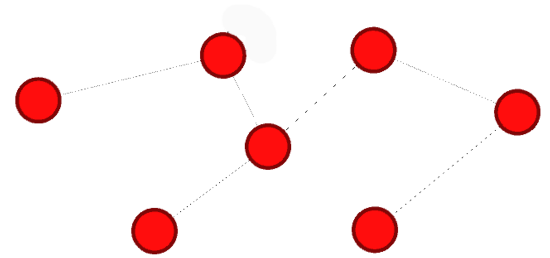
- number of nodes
- number of edges
- connected components
- network diameter
  - what is the longest shortest path between two nodes?
- network density
  - how close is the network to complete? e.g. what percentage of edges are extant?



@jotis13

Global Network Analysis

- network density
  - how close is the network to complete?
- max number of undirected edges:
  - $(|V| * (|V|-1)) \div 2$
- undirected graphs:
  - $2|E| \div |V| * (|V|-1)$
- directed graphs:
  - $|E| \div |V| * (|V|-1)$



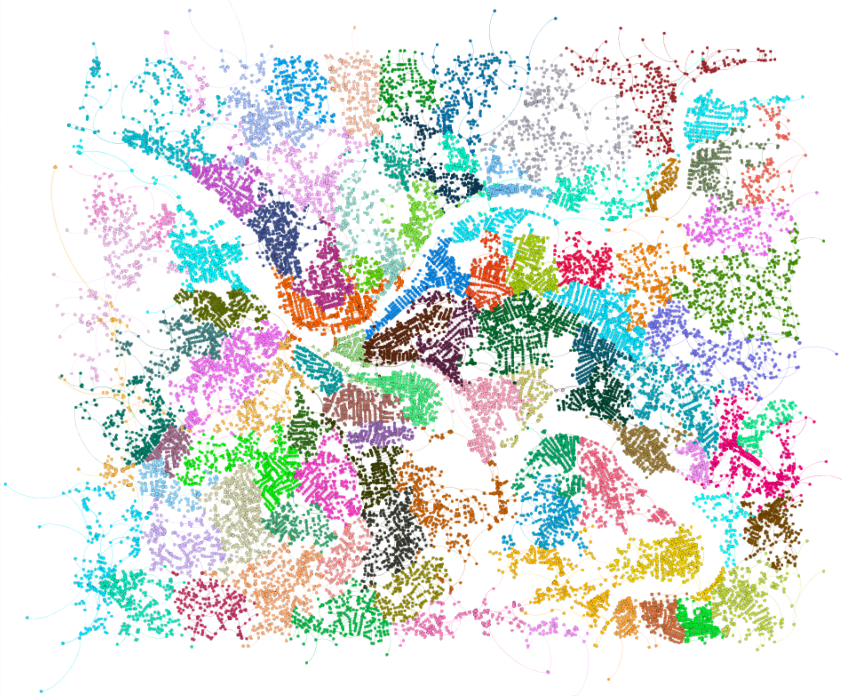
- $|E|$ = number of edges
- $|V|$ = number of nodes/vertices

@jotis13

Global Network Analysis

- we can divide every network into subnetworks
- each subnetwork contains a fraction of the total number of edges (e.g. 1/10)
- modularity is the fraction of the edges in a subnetwork MINUS the expected fraction if the edges were distributed at random
- modularity thus detects community structures (in large networks only)

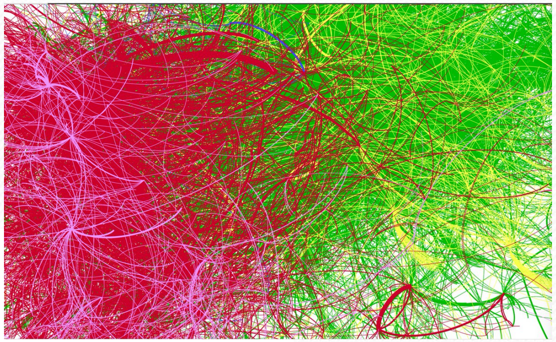
$$Q = \frac{1}{2m} \sum_{vw} \left[A_{vw} - \frac{k_v k_w}{2m} \right] \frac{s_v s_w + 1}{2}$$

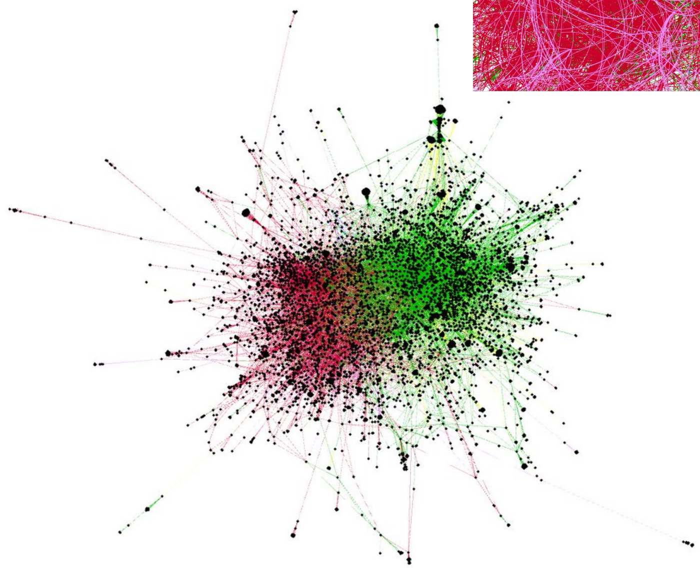


@jotis13

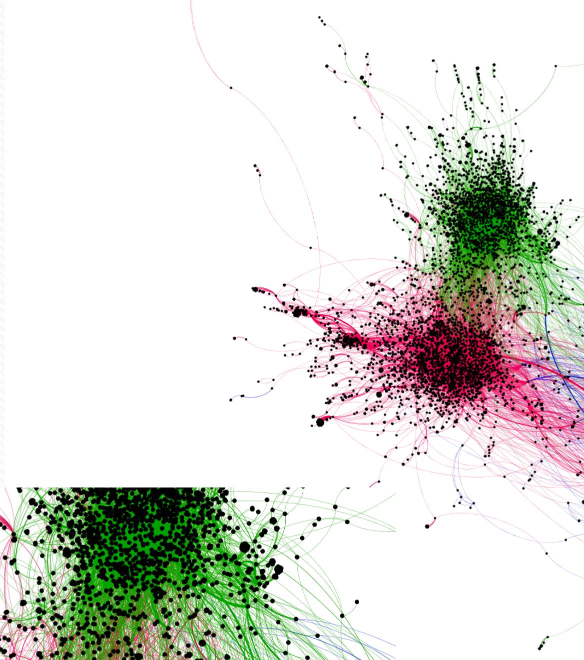
Comparative Networks

AHA/MLA SuperConference 2019



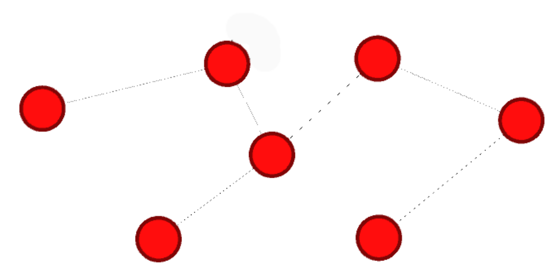


AHA/MLA Separate Conferences 2017



Degree

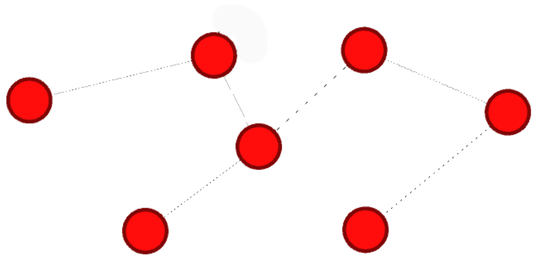
- in-degree
  - sum of all edges going to a node
- out-degree
  - sum of all edges going out of a node
- weighted in-degree
  - sum of all edges going to a node multiplied by their weights
- weighted out-degree
  - sum of all edges going out from a node multiplied by their weights
- degree
  - sum of all edges
- weighted degree
  - sum of all edges multiplied by their weights



@jotis13

Node "Centrality" Measures

- betweenness centrality:
  - the number of the shortest paths through the network that go through a node
- closeness centrality:
  - the average distance from a node to all other nodes in the network
- eccentricity:
  - the distance from a node to the farthest node from it in the network



why do you think Gephi calculates these at the same time as the network diameter?

@jotis13

Bipartite Betweenness

`betweenness_centrality`(*G*, *nodes*) [source]

Compute betweenness centrality for nodes in a bipartite network.

Betweenness centrality of a node `v` is the sum of the fraction of all-pairs shortest paths that pass through `v`.

Values of betweenness are normalized by the maximum possible value which for bipartite graphs is limited by the relative size of the two node sets [1].

Let `n` be the number of nodes in the node set `U` and `m` be the number of nodes in the node set `V`, then nodes in `U` are normalized by dividing by

$$\frac{1}{2}[m^2(s+1)^2 + m(s+1)(2t-s-1) - t(2s-t+3)],$$

where

$$s = (n-1) \div m, t = (n-1) \mod m,$$

and nodes in `V` are normalized by dividing by

$$\frac{1}{2}[n^2(p+1)^2 + n(p+1)(2r-p-1) - r(2p-r+3)],$$

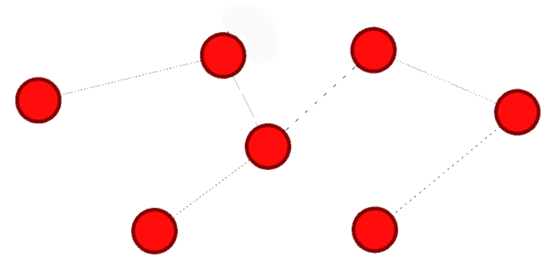
where,

$$p = (m-1) \div n, r = (m-1) \mod n.$$

@jotis13

Node "Centrality" Measures

- betweenness centrality:
  - the number of the shortest paths through the network that go through a node
- closeness centrality:
  - the average distance from a node to all other nodes in the network
- eccentricity:
  - the distance from a node to the farthest node from it in the network



why do you think Gephi calculates these at the same time as the network diameter?

@jotis13

Node "Centrality" Measures

- betweenness centrality:

$$g(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

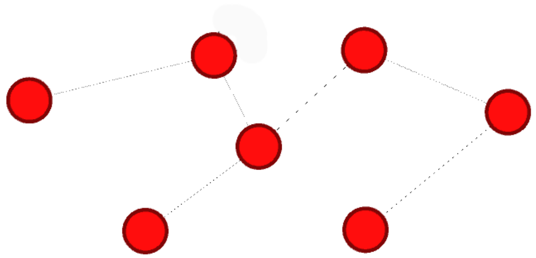
- closeness centrality:

$$C(x) = \frac{1}{\sum_y d(y,x)}.$$

- harmonic closeness:

$$H(x) = \sum_{y \neq x} \frac{1}{d(y,x)}$$

- eccentricity:
  - the center of the graph is the node(s) with the minimum eccentricity
  - the maximum eccentricity is also the diameter



"normalizing" forces the result to be between 0 and 1, which makes it easier to compare with other networks

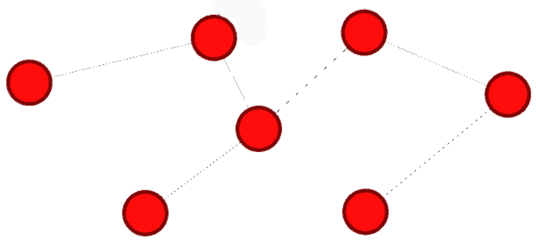
@jotis13

Other Node Measures

- transitivity (global measure)
  - the number of closed triplets divided by the number of all triplets (open or closed)
- clustering coefficient (nodes)
  - how close is it to being part of a clique – a complete subnetwork?

$$C_i = \frac{|\{e_{jk} : v_j, v_k \in N_i, e_{jk} \in E\}|}{k_i(k_i - 1)}.$$

$$C_i = \frac{2|\{e_{jk} : v_j, v_k \in N_i, e_{jk} \in E\}|}{k_i(k_i - 1)}.$$



@jotis13

Eigenvector Centrality

- we call our adjacency matrix $A = (a_{v,t})$
  - each $a_{v,t}$ corresponds to the 0 or 1 at the intersection of row v and column t
- eigenvector centrality for a vertex v is

$$x_v = \frac{1}{\lambda} \sum_{t \in M(v)} x_t = \frac{1}{\lambda} \sum_{t \in G} a_{v,t} x_t$$

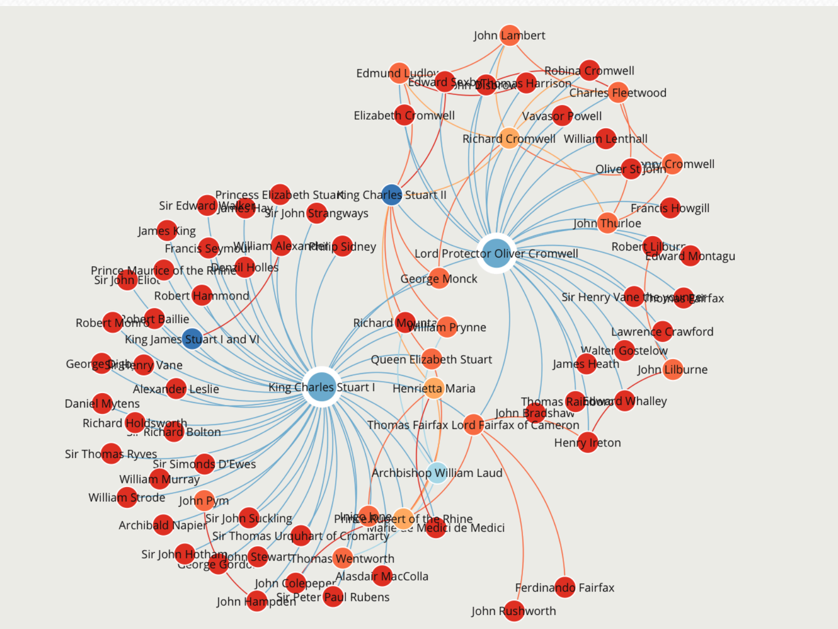
- note: $\mathbf{Ax} = \lambda \mathbf{x}$

	And	Bil	Car	Dan	Ele	Fra	Gar
Andy		1	0	1	0	0	1
Bill	1		1	0	1	0	0
Carol	1	1		1	1	0	0
Dan	1	1	1		0	0	0
Elena	0	0	0	0		1	0
Frank	0	0	0	0	1		0
Garth	1	1	0	0	0	0	

@jotis13

How Can You Analyze Them?

E.g. in a social network



- who knows the most people?
- who knows the most highly connected people?
- how many disconnected subnetworks are there?
- whose removal severs the network into disconnected subnetworks?
- how many “hops” does it take to get from one person in the network to another?
- who do most of the shortest hops go through?
- are there any smaller cliques or “communities” within the network?

@jotis13