Visualization of large and dynamic networks
advances and limits

Clement Levallois

@seinecle

Lisbon, 06 November 2013

WG1, Spatial and temporal modelling: representation of space and time, NeDiMAH

Bio notes

PhD
(2003-2008)

Post-Doc
(2008-2013)

- Member of the Gephi Consortium

Gephi

- Starting at EMLyon in January 2014

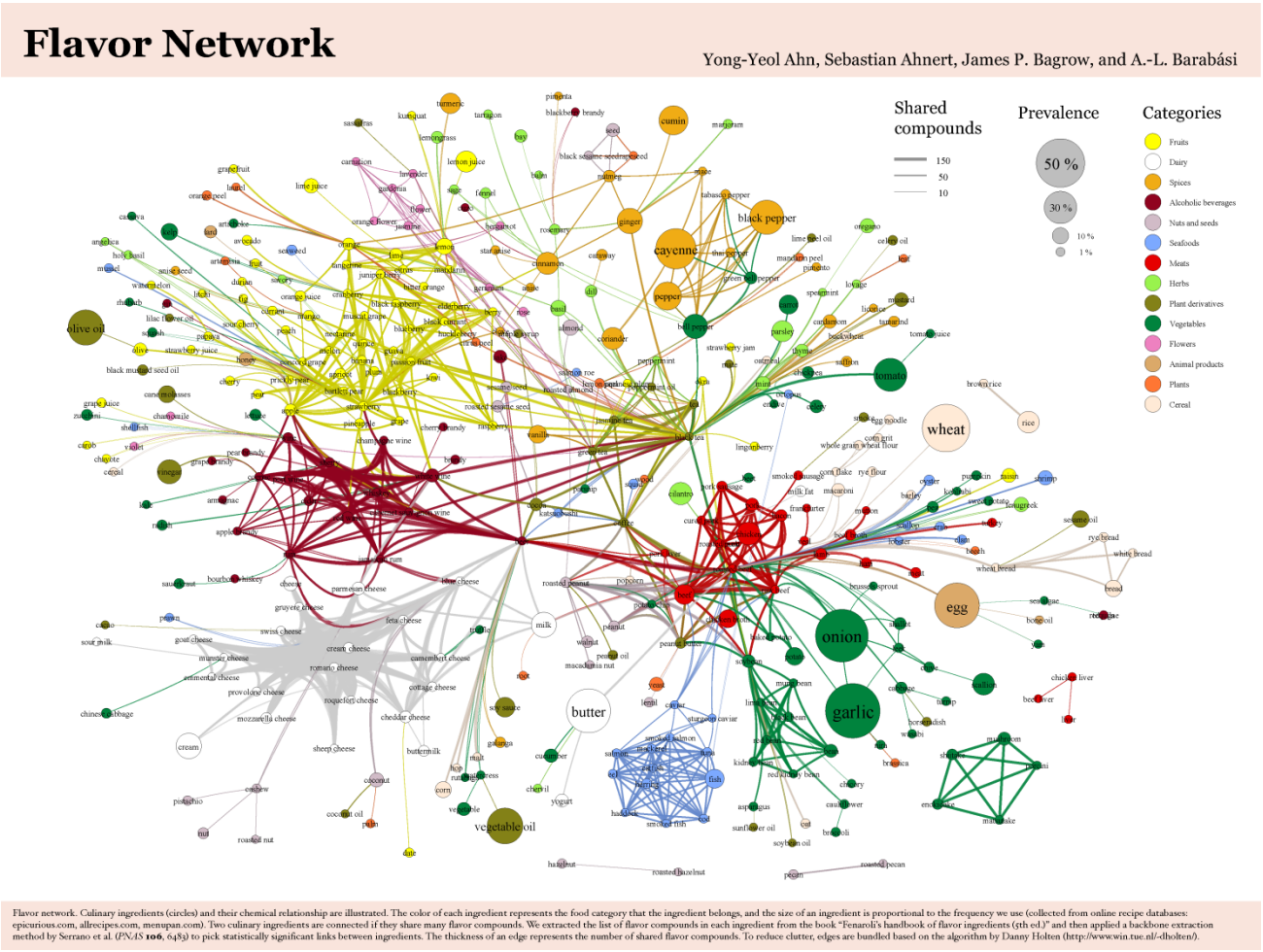
EMLYON business school

Plan

1. Diversity of networks
2. Why visualizations
3. Three challenges
4. Open issues ahead
5. References

1. DIVERSITY OF NETWORKS

Co-occurrences (products)

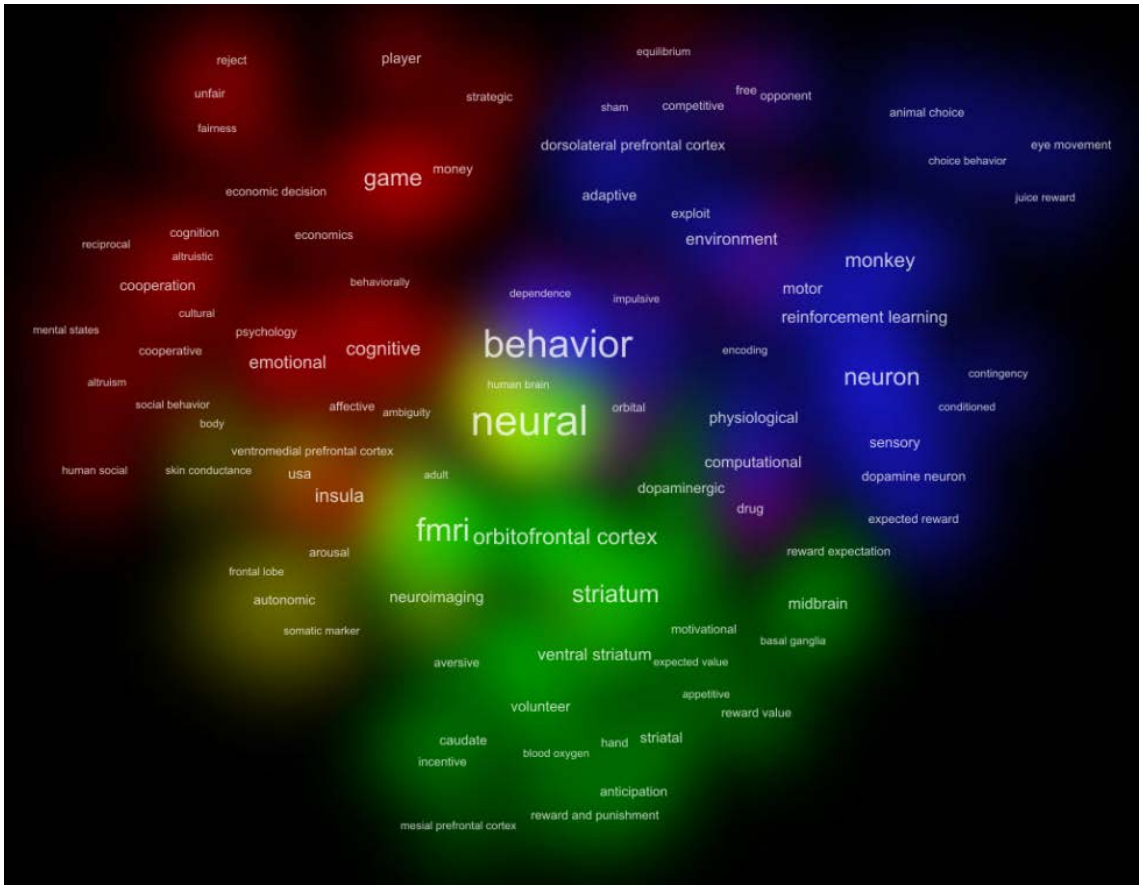


Map of flavors

2 flavors are connected

If they frequently appear in common recipes.

Co-occurrences (concepts)

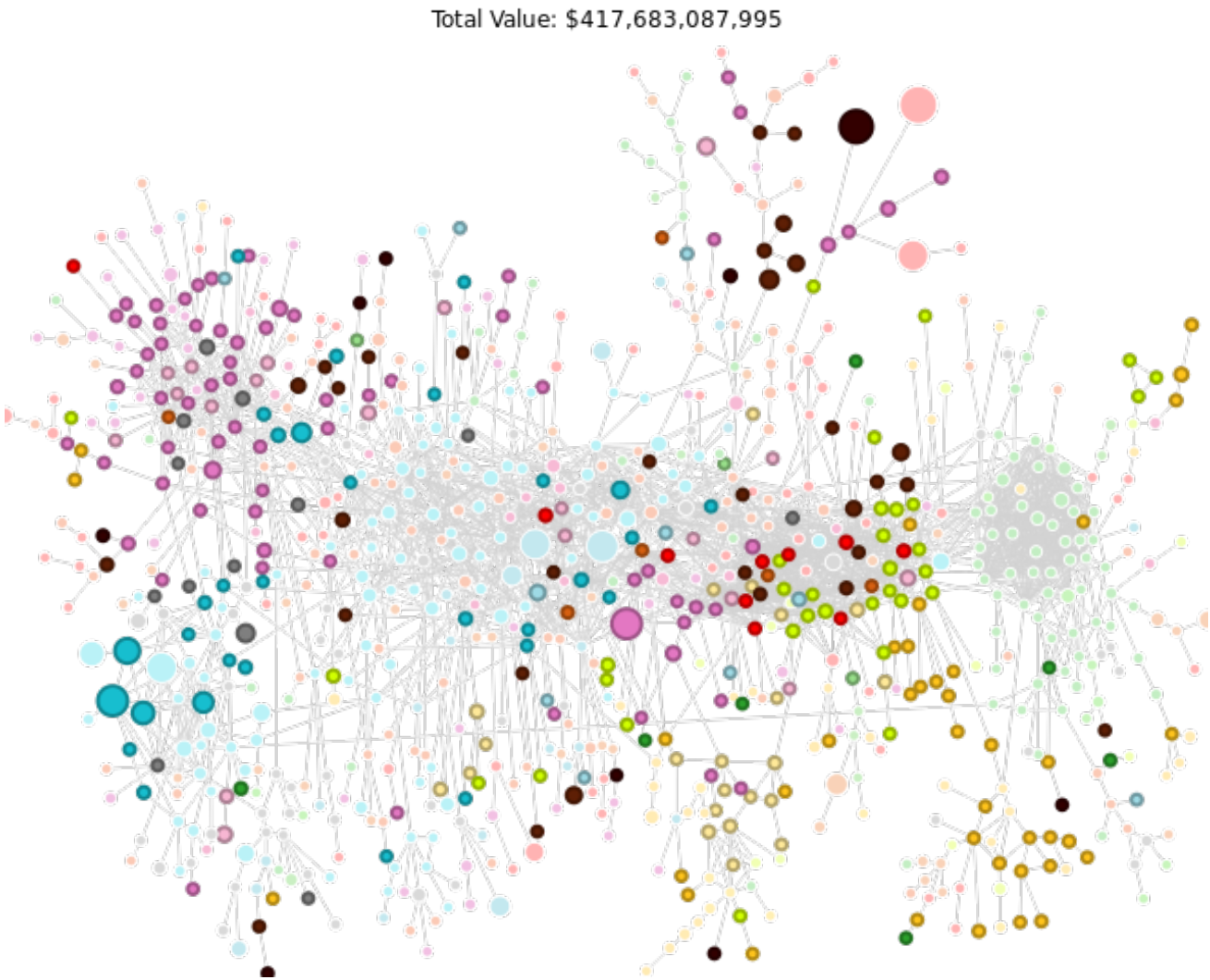


Map of terms from papers in neuroeconomics

2 terms sit close to each other

If they are frequently used in same abstracts

Co-location



Map of product categories

2 products are connected
If they are frequently exported

by the same countries

Similarity

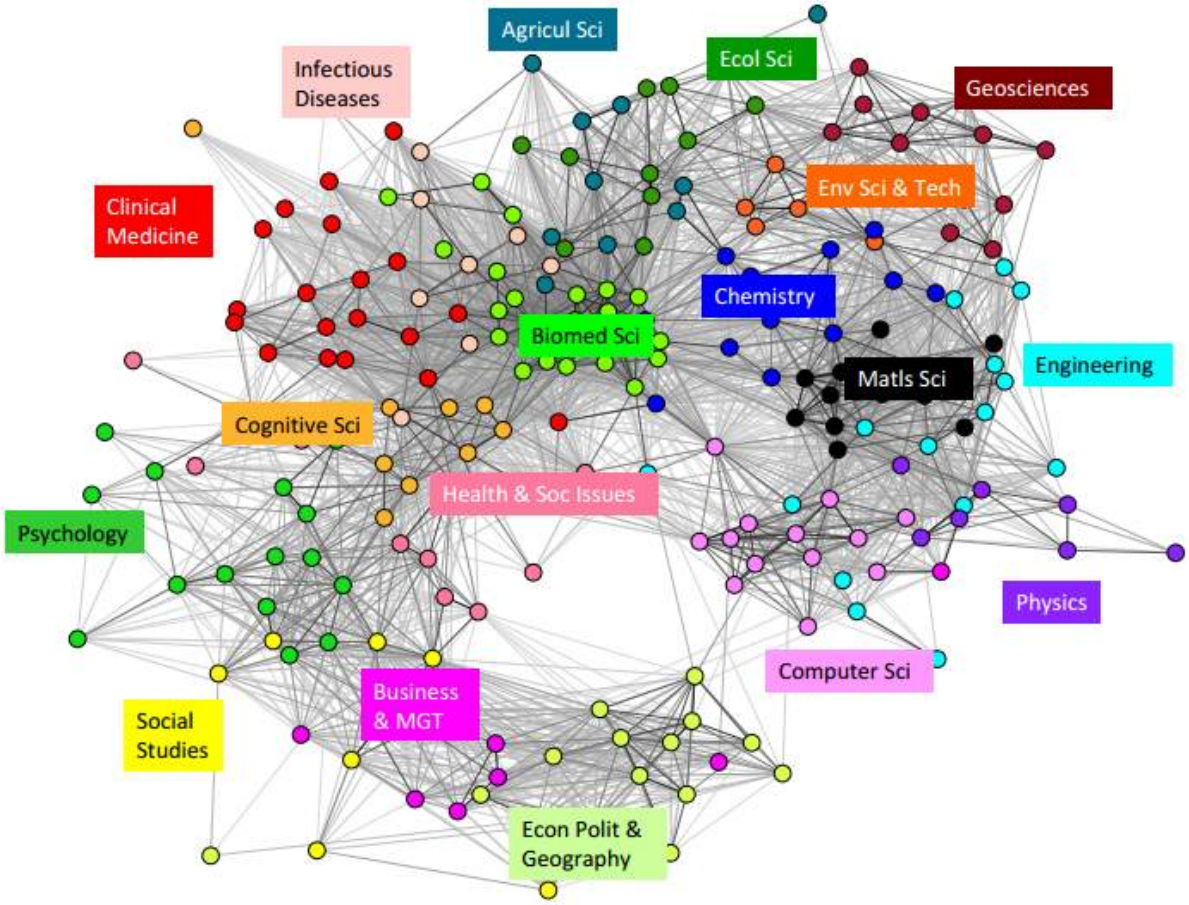


Figure 2. Global science map based on citing similarities among ISI Subject Categories (2007).

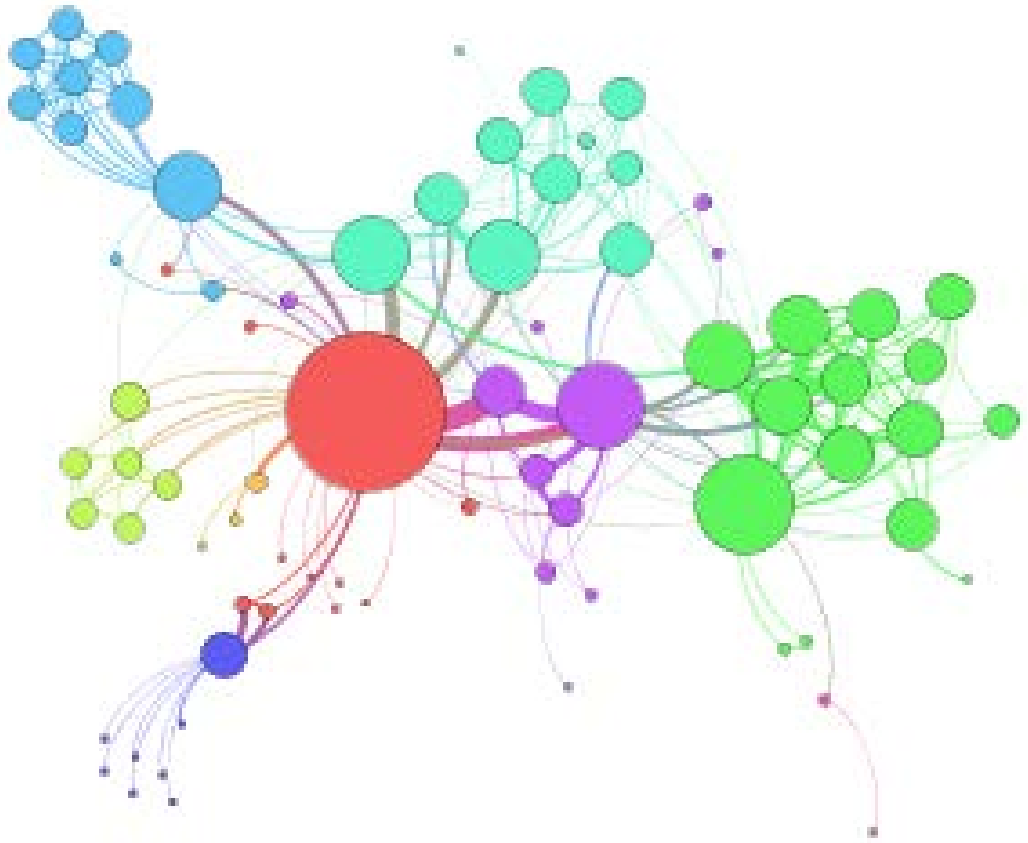
2 entities are connected if they …

… they are joined by physical or symbolic links
… they perform the same kind of actions
… they share antecedents
… they possess common characteristics
… they connect to common entities
… they transact with each other
… they are referred to in pair
… they refer to common entities

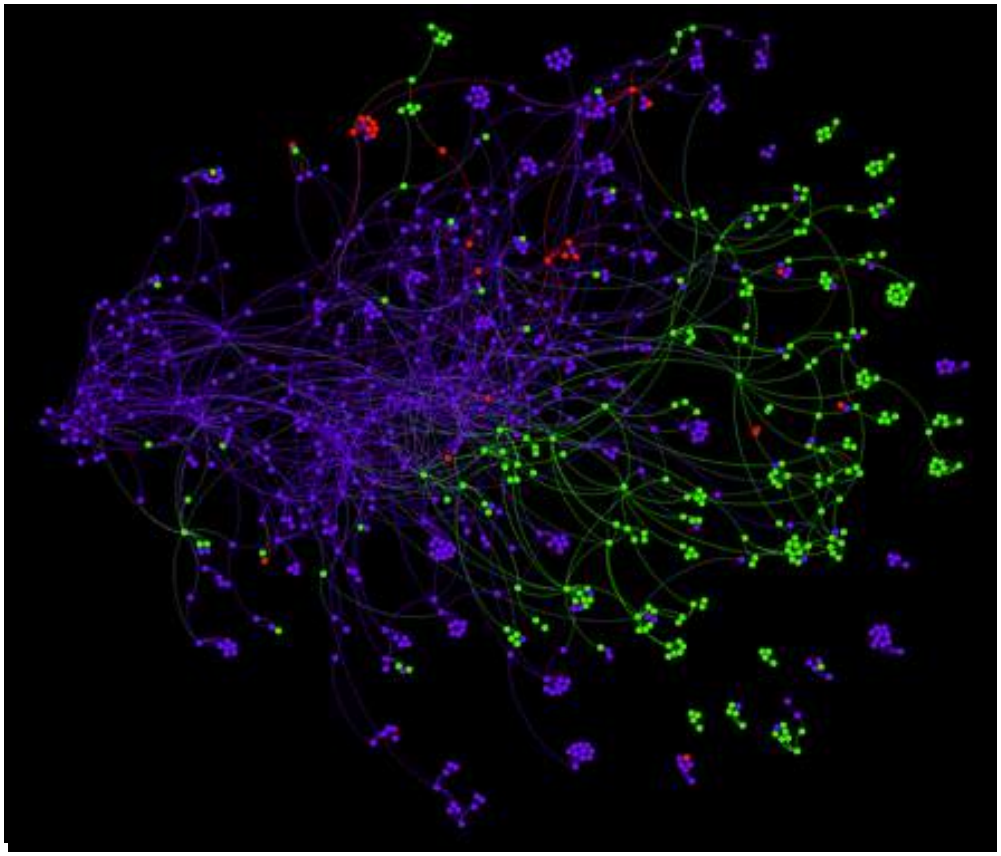
… or combinations of the above

2. WHY VISUALIZATIONS

Community detection



Overlay analysis



Social network of scientists

Colored by their research interests

Colored by their geographical location

… with prudence

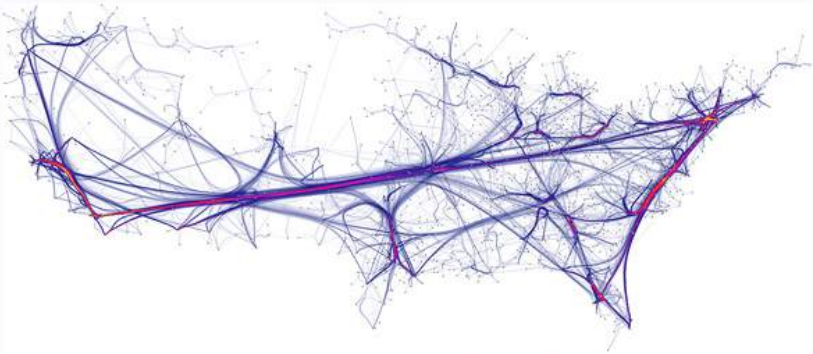
5
11 2013

Networks Demystified 8: When Networks are Inappropriate

SCOTT WEINGART | METHOD | 1

A few hundred years ago, I promised to talk about when not to use networks, or when networks are used improperly. With The Historian's Macroscope in the works, I've decided to finally start answering that question, and this Networks Demystified is my first attempt at doing so. If you're new here, this is part of an annoyingly long series (1 network basics, 2 degree, 3 power laws, 4 co-citation analysis, 5 communities and PageRank, 6 this space left intentionally blank, 7 co-citation analysis II). I've issued a lot of vague words of caution without doing a great job of explaining them, so here is the first substantive part of that explanation.

Networks are great. They allow you to do things like understand the role of postal routes in the circulation of knowledge in early modern Europe, or of the spread of the black death in the middle ages, or the diminishing importance of family ties in later Chinese governments. They're versatile, useful, and pretty easy in today's software environment. And they're sexy, to boot. I mean, have you seen this visualization of curved lines connecting U.S. cities? I don't even know what it's supposed to represent, but it sure looks pretty enough to fund!



A really pretty network visualization. [via]

scott b. weingart is a juggler, an academic, and a nice guy. He's pretty clueless about a lot of things, and this blog is his attempt to become less so.

bias big data blogging citations collaboration complexity computational models culturomics data analysis diffusion digital humanities disciplinarity gephi GIS google history history of science human dynamics humnets interdisciplinarity literary history macroanalysis methodological critique methodologies MOOC network analysis networks demystified open access republic of letters review scholarly communication sci2 scientonomy significance social networks social science statistics text analysis topic modeling unicode visualizations wordle wordpress zotero zotpress

3. THREE CHALLENGES (and solutions)

(two preliminary challenges)

- User friendliness

  (pure coding solutions / experimental software not considered here)

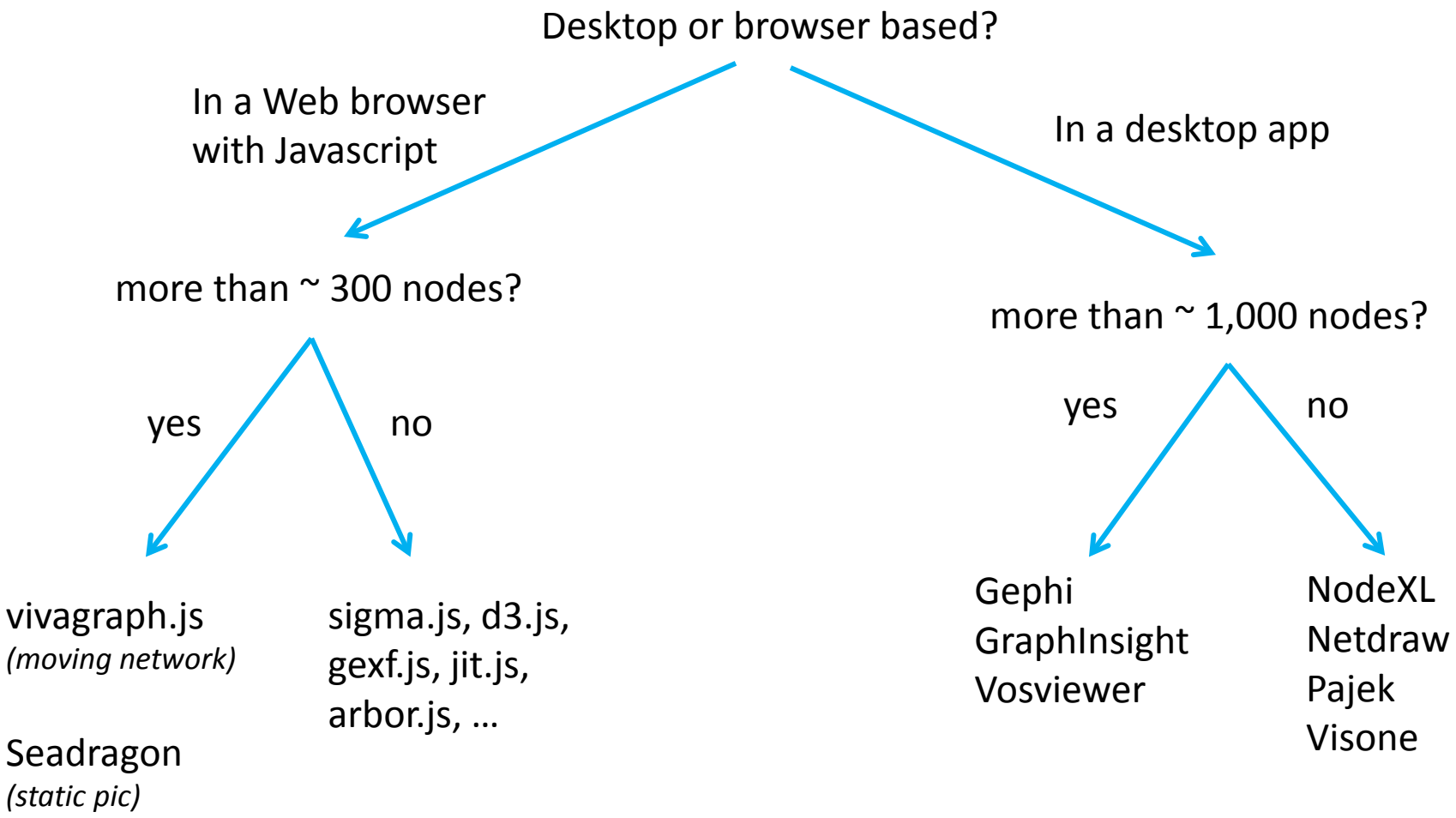
- Interdisciplinarity

  For example, we don’t consider Circos and Cytoscape which are fine software for network representations, but keep a strong anchor in biology which make them harder to use by social scientists.

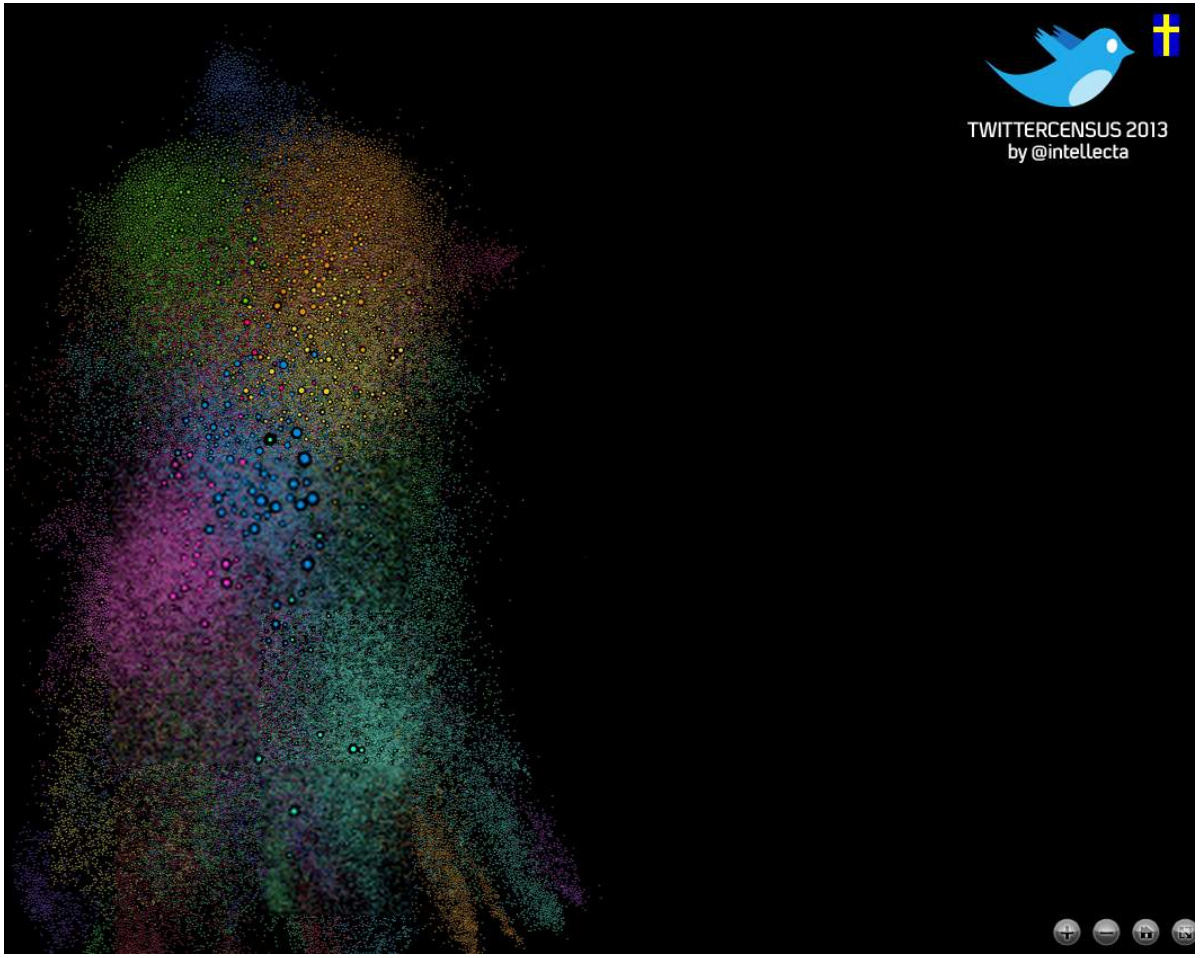
First challenge: size

- Can we visualize networks of…
  - 10 nodes and edges
  - 100 nodes and edges
  - 100,000 nodes and edges
  - 1,000,000 nodes and edges
- And is it still interesting?

Large networks: solutions



Big networks: still interesting to visualize?

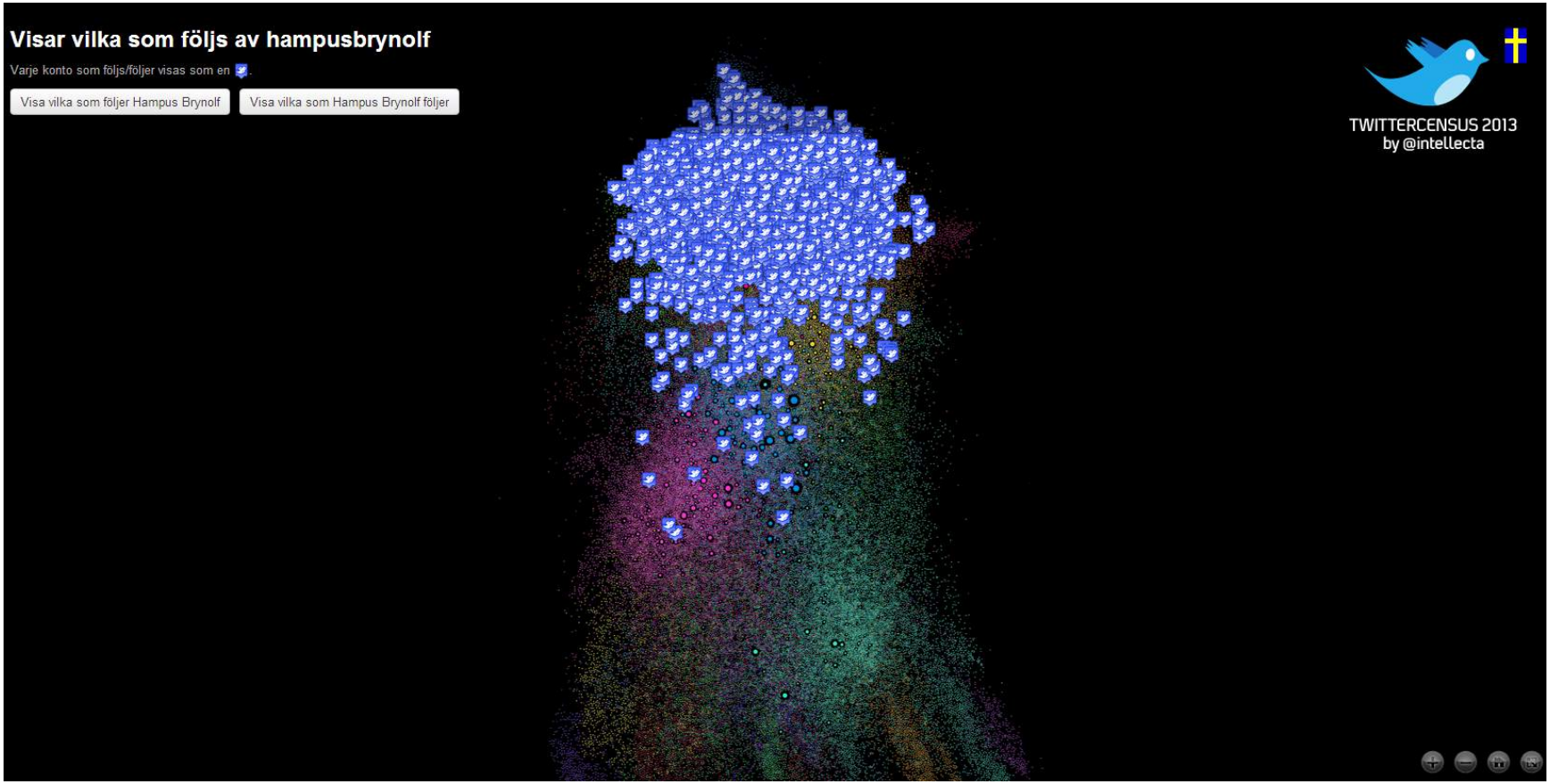


Map of Swedish speakers on Twitter
(as identified by natural language processing on their tweets)

100,000 persons (out of 500,000)

2 persons are connected **if one follows the other on Twitter.**

Color represents communities as detected by an algo.

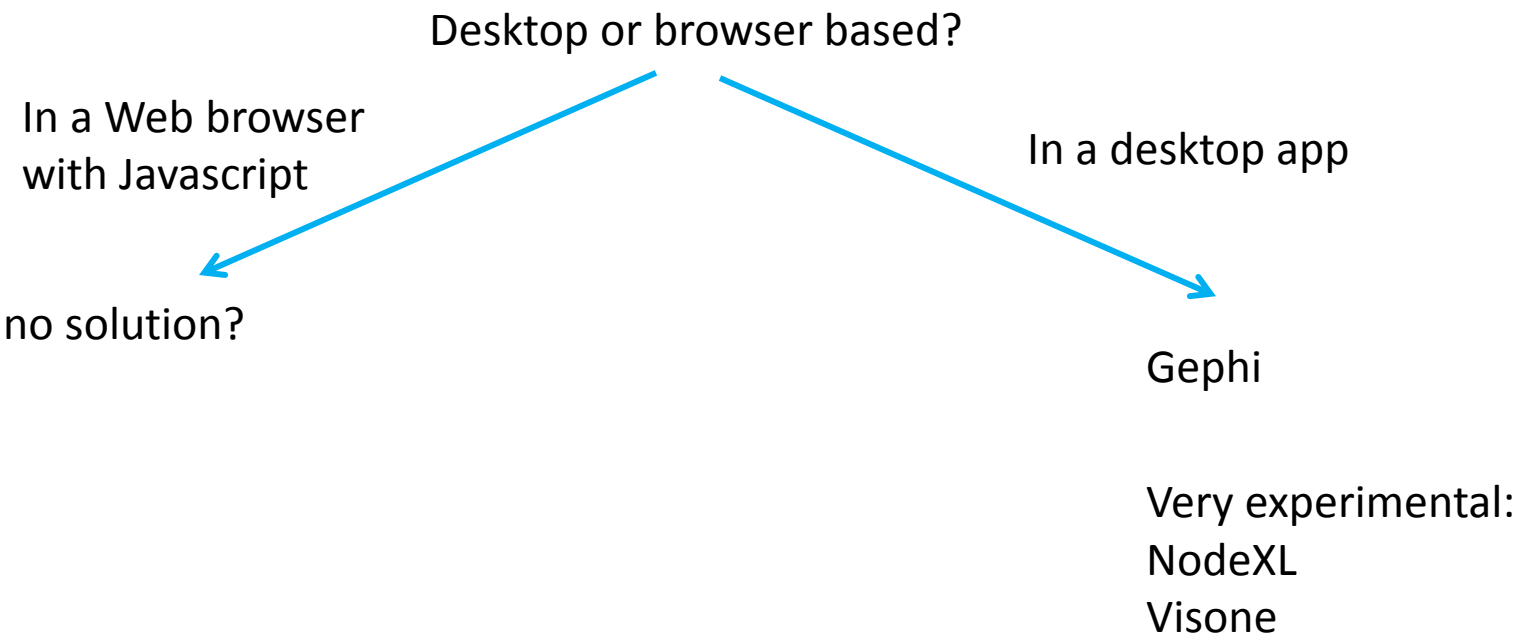


Followers of Twitter account “hampusbrynolf”
projected on the map of Swedish speakers on Twitter

Second challenge: time

- How feasible is it to visualize time evolving networks?
- And is it interesting?

Visual exploration of time-dependent networks



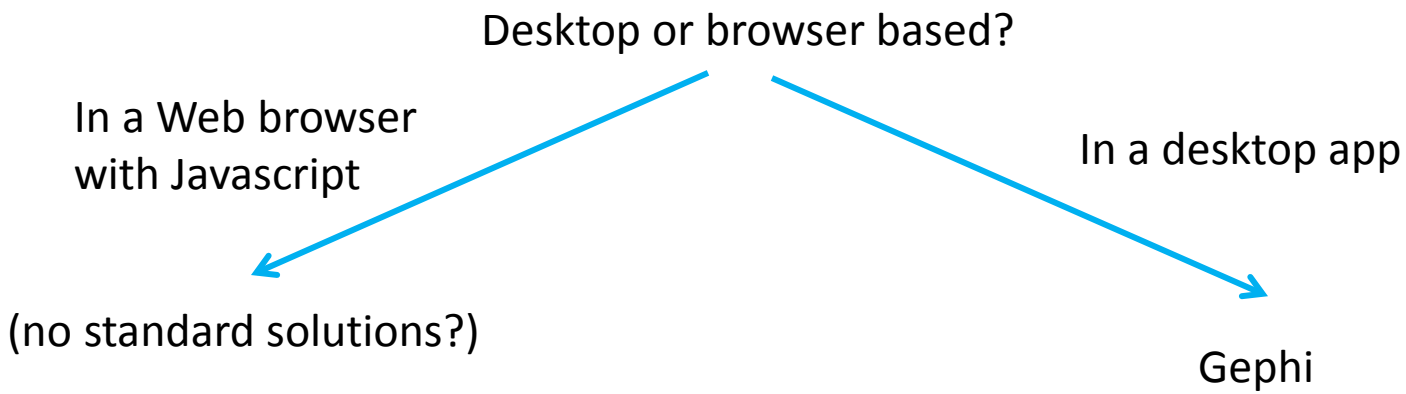
Is it interesting?

Launch video on Central bank and Dutch banks

Third challenge: space

- Can we mix network and spatial data?
- Is it interesting?

Visual exploration of spatial networks

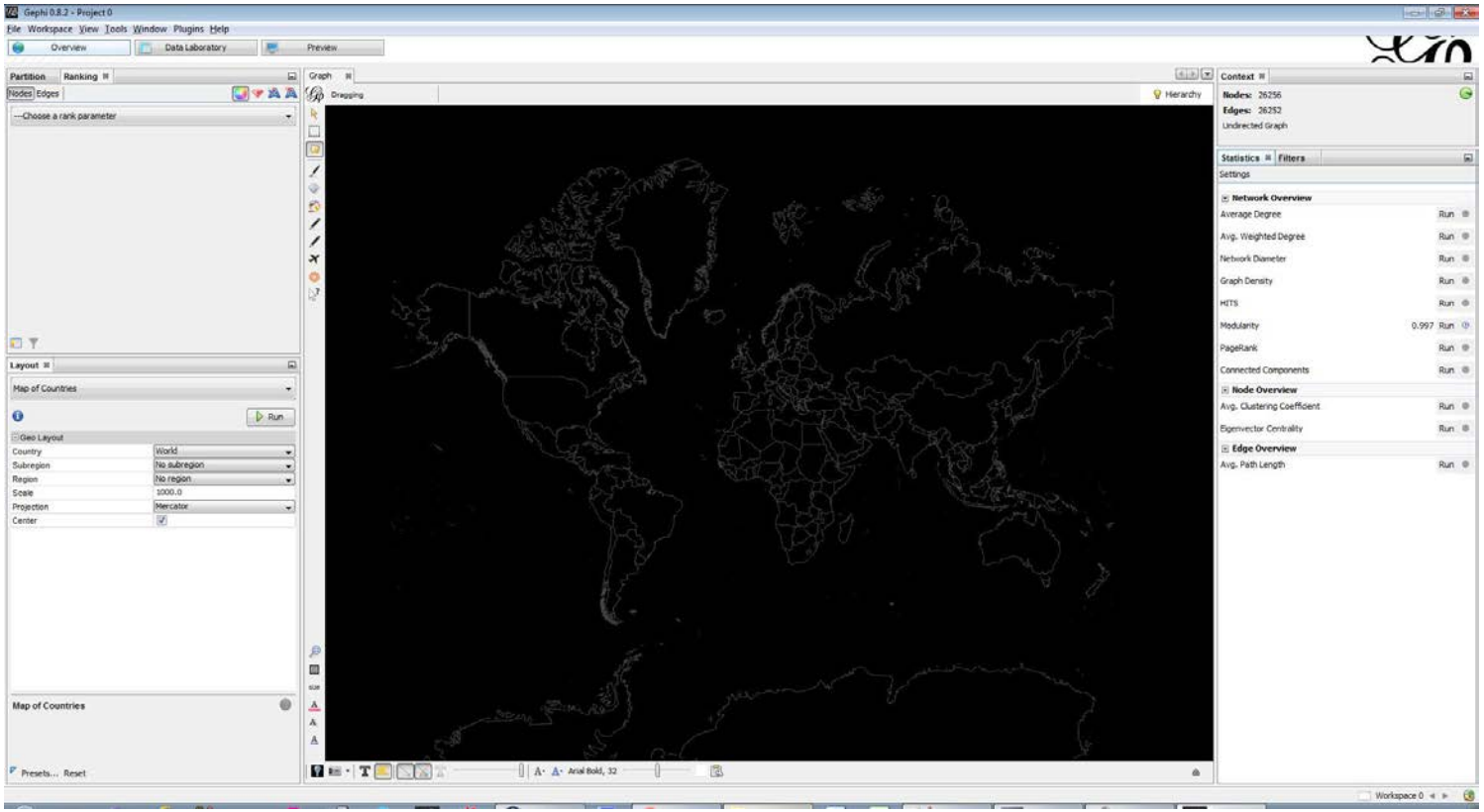


But see:
http://www.leydesdorff.net/maps/is2009.html

Is it interesting?

Launch video Europe 1948-2008

(Gephi plugin)



See references list for url to download

4. OPEN ISSUES AHEAD

Which algos for dynamic networks?

- Example: Community detection in dynamic networks
  - Community detection is very sloppy in its definition for static networks
  - Dynamics will add another layer of confusion

Video on community detection on a dynamic graph

Play video here

Which visual grammar?

- Dynamic networks

  How to represent evolving features in a non misleading way?

- Spatial networks

  How to represent dynamics on maps?

  (Not personally satisfied with connected blobs)

References

- Flavor network

  Yong-Yeol Ahn, Sebastian E. Ahnert, James P. Bagrow & Albert-László Barabási, 2011

  http://www.nature.com/srep/2011/111215/srep00196/fig_tab/srep00196_F2.html
- Map of terms

  Levallois, Clithero, Smidts, Wouters and Huettel (2012).

  http://www.nature.com/nrn/journal/v13/n11/full/nrn3354.html
- Social network of scientists

  Levallois (private data).
- Map of product categories

  The Observatory of Economic Complexity, Alexander Simoes (2012).

  http://atlas.media.mit.edu/
- Map of science

  Rafols, Porter and Leydesdorff (2010).

  http://www.leydesdorff.net/overlaytoolkit/
- ScottBot Irregular: When networks are inappropriate

  http://www.scottbot.net/HIAL/?p=39600

References

- Map of Swedish speakers on Twitter
  Map itself (http://twittercensus.se/graph2013/)
  Overlay of an account (http://twittercensus.se/graph2013/followers/hampusbrynolf)

- Videos
  Financial transactions on the Dutch money market (work in progress):
  http://www.youtube.com/channel/UCaDcmlwrMZKDchzC8ZmMQMw
  Evolving network on a map of Europe (private data, work in progress)
  Evolving communities on a toy network in Gephi: http://vimeo.com/4551215

- Gephi plugin: Map of countries
  https://marketplace.gephi.org/plugin/maps-of-countries/

Thank you.

@seinecle

These slides are online at:

www.clementlevallois.net