Visual Network Analysis from the comfort of your Jupyter notebook

Guillaume Plique, médialab SciencesPo

Jupytercon 2023

Who am I?

A **research engineer** working for SciencesPo's médialab.

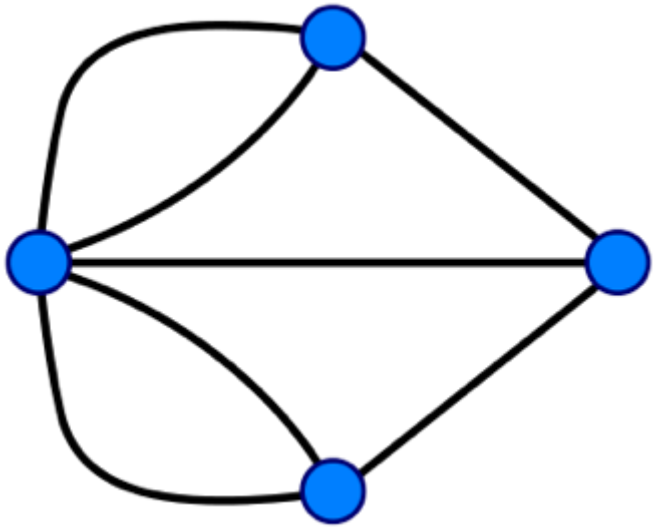
A **social sciences** laboratory founded by Bruno Latour 10 years ago.

We intend to mix:

- Social sciences
- Design
- Engineering

What is a graph/network

Nodes/vertices, edges/links and the attached metadata.

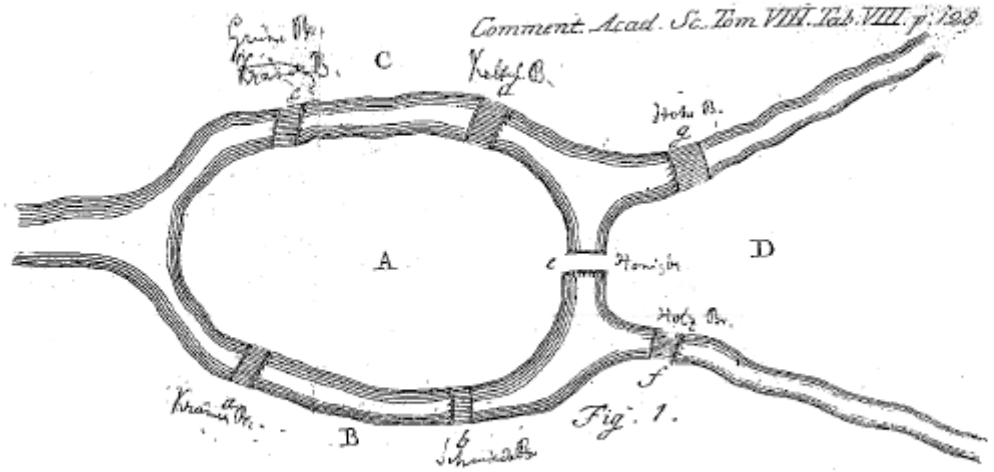


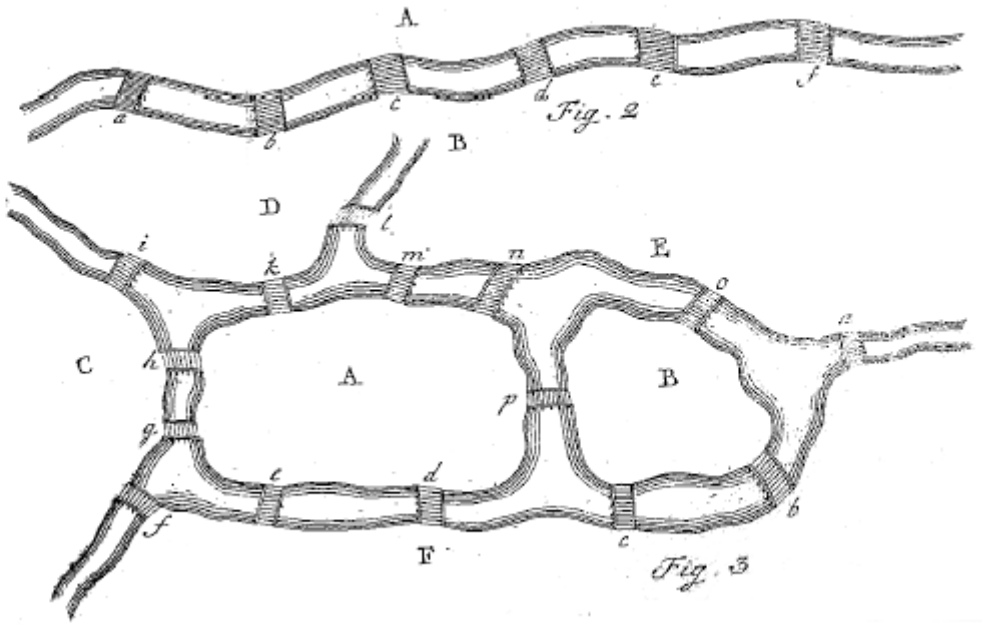
Visual Network Analysis

Prelude: a statistical approach to graphs

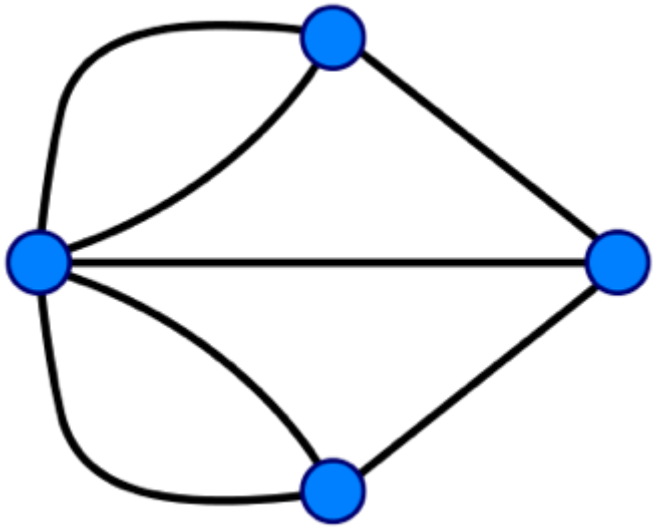
- Diameter
- Density
- Degree distribution
- Shortest paths
- Centrality
- Eigenvalues
- Linear algebra (a graph is a matrix, a matrix is a graph)
- Laplacian matrix
- Pagerank
- Amphibolic metempsychosis of the Eulerian principles
- etc.

Should we try vizualizing graphs instead?





Should we try vizualizing graphs instead?



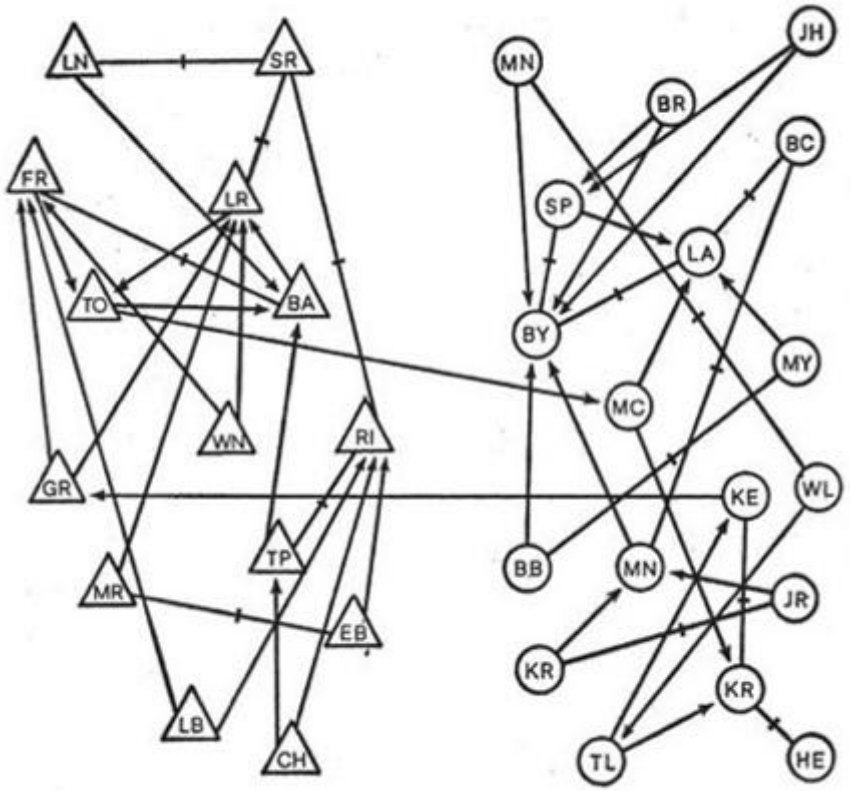
Visual Network Analysis in Social Sciences

Moreno's sociograms

id	Student	First choice	Second choice
1	Amélie	4	3
2	Jean	1	3
3	Kareem	2	26
4	Lydia	45	12
5	Michael	7	28
6	Guillaume	18	3
...	...	...	...

Moreno's sociograms

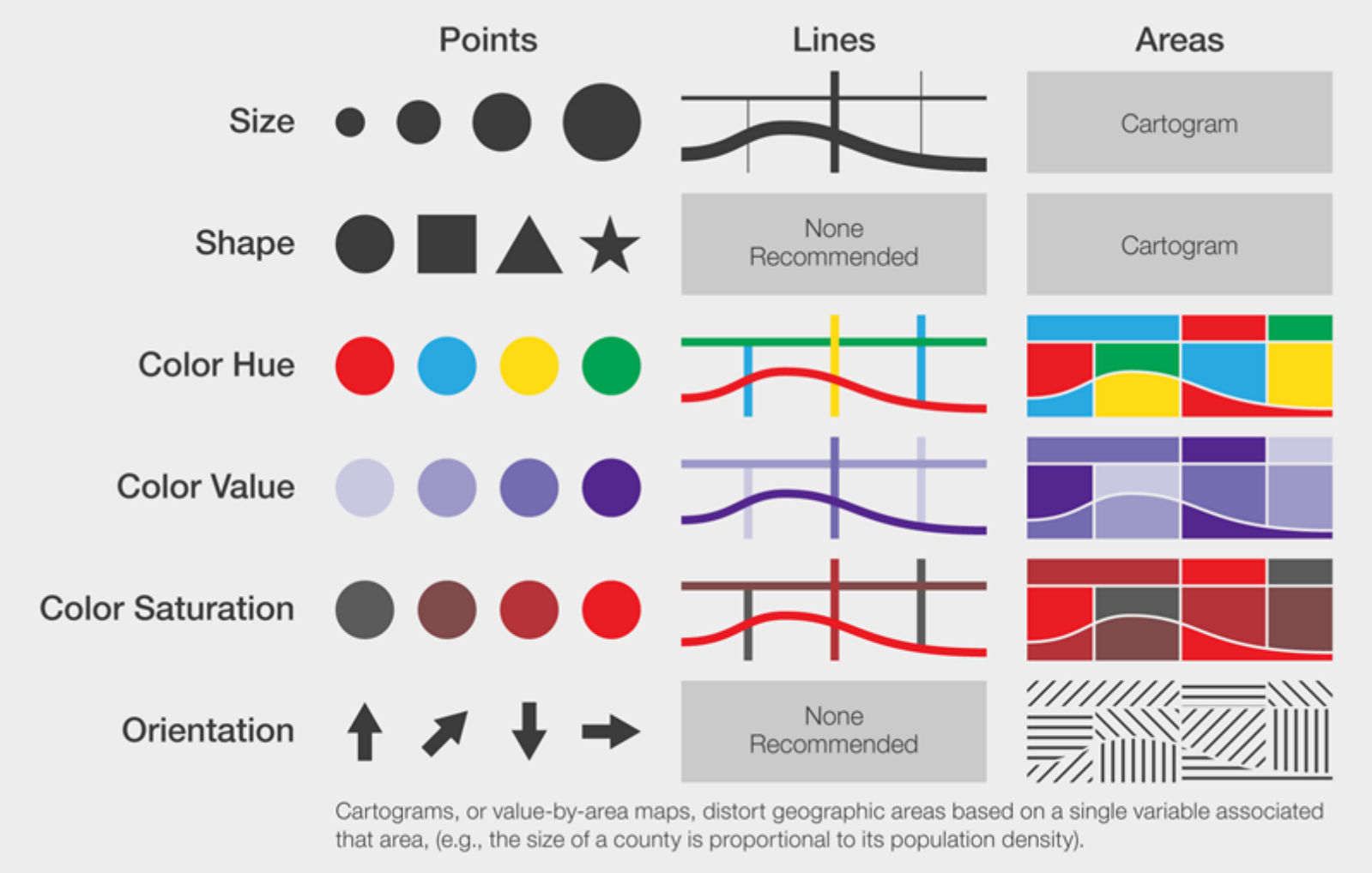
Sociogramme d'une classe d'élèves de 11-12 ans
(critère : s'asseoir à côté des élèves choisis – 2 choix au maximum)



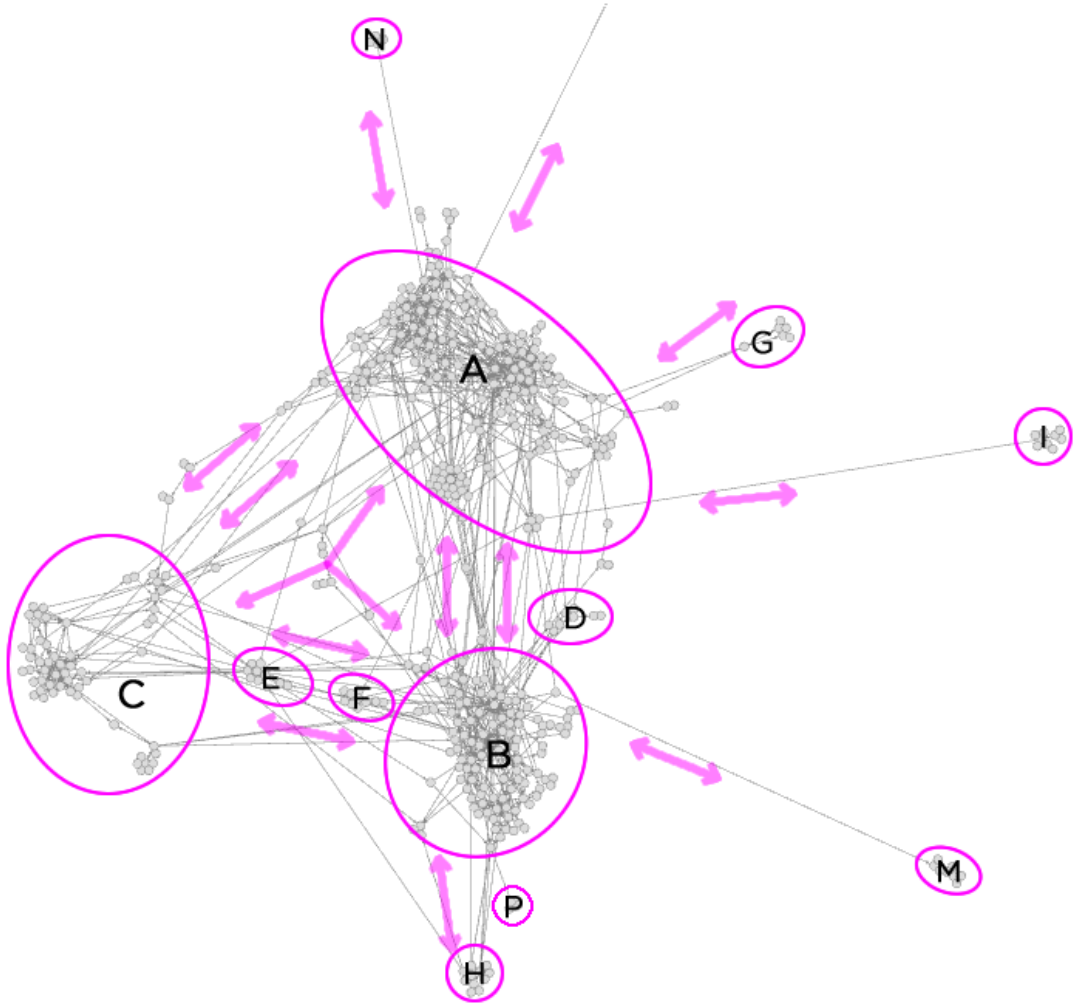
Les lignes barrées indiquent les choix réciproques.

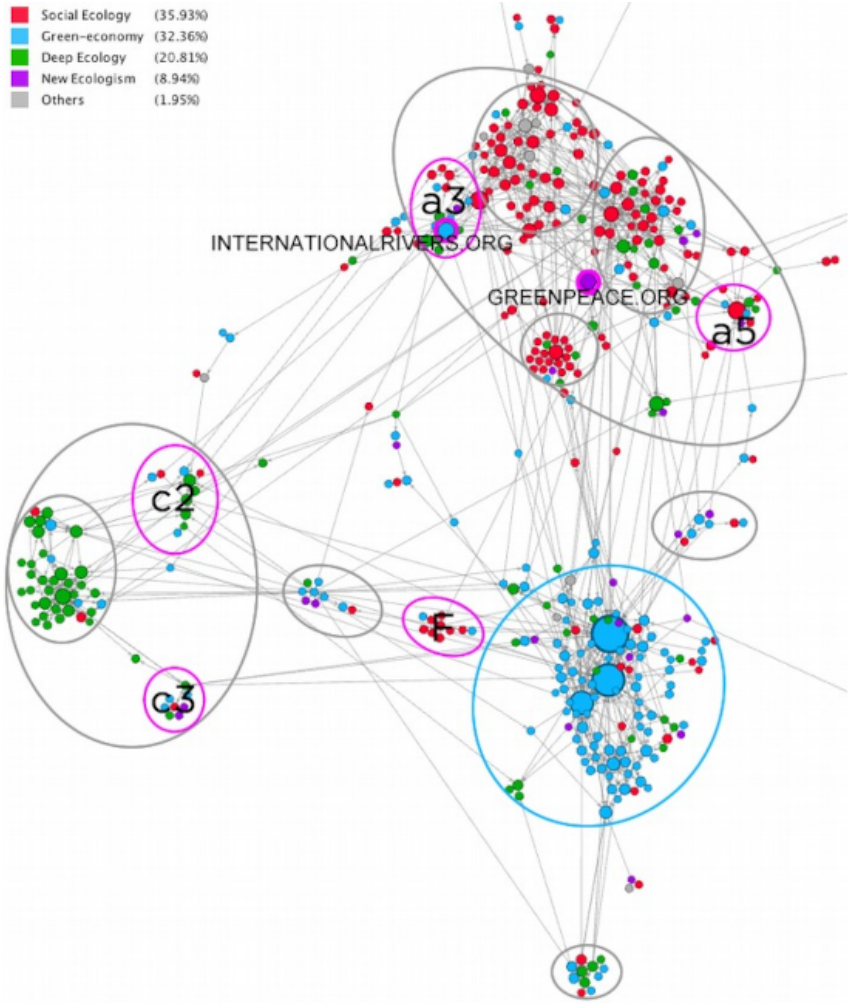
Source : Moreno [1934, annexes, planche XII].

Bertin's visual variables



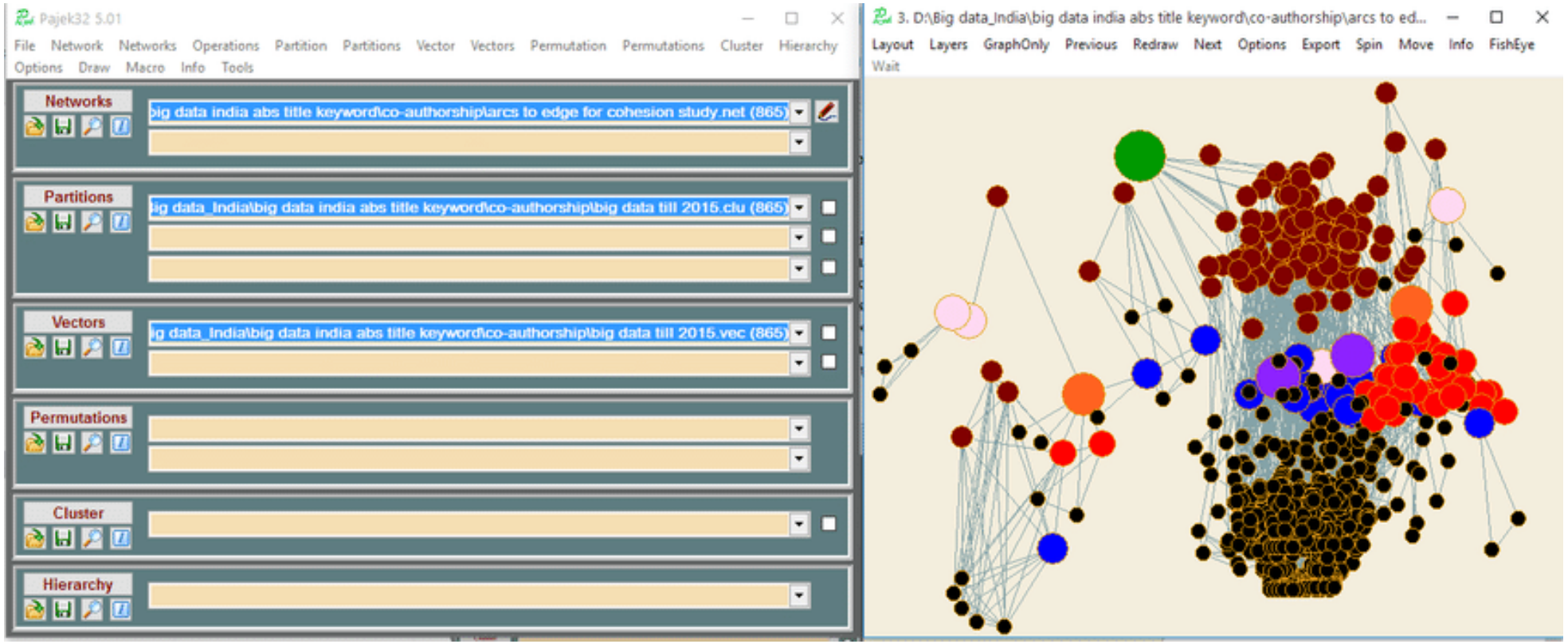
Fast forward: community detection, spatial layout etc.



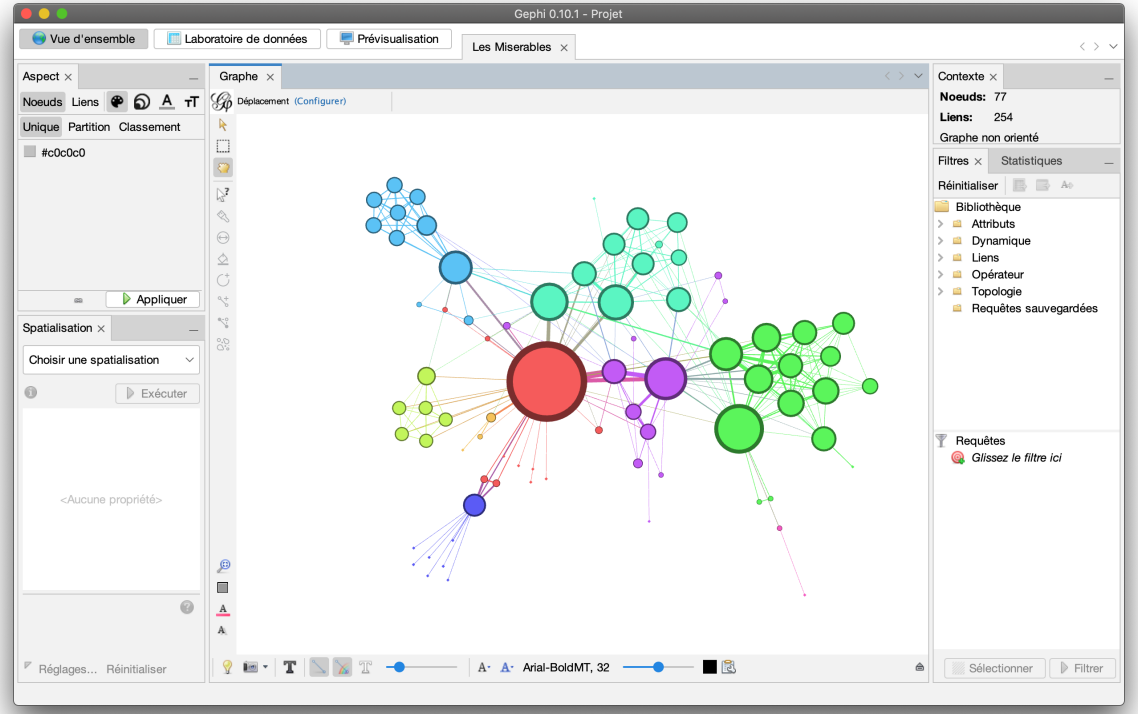


Visual Network Analysis on desktop and on the web

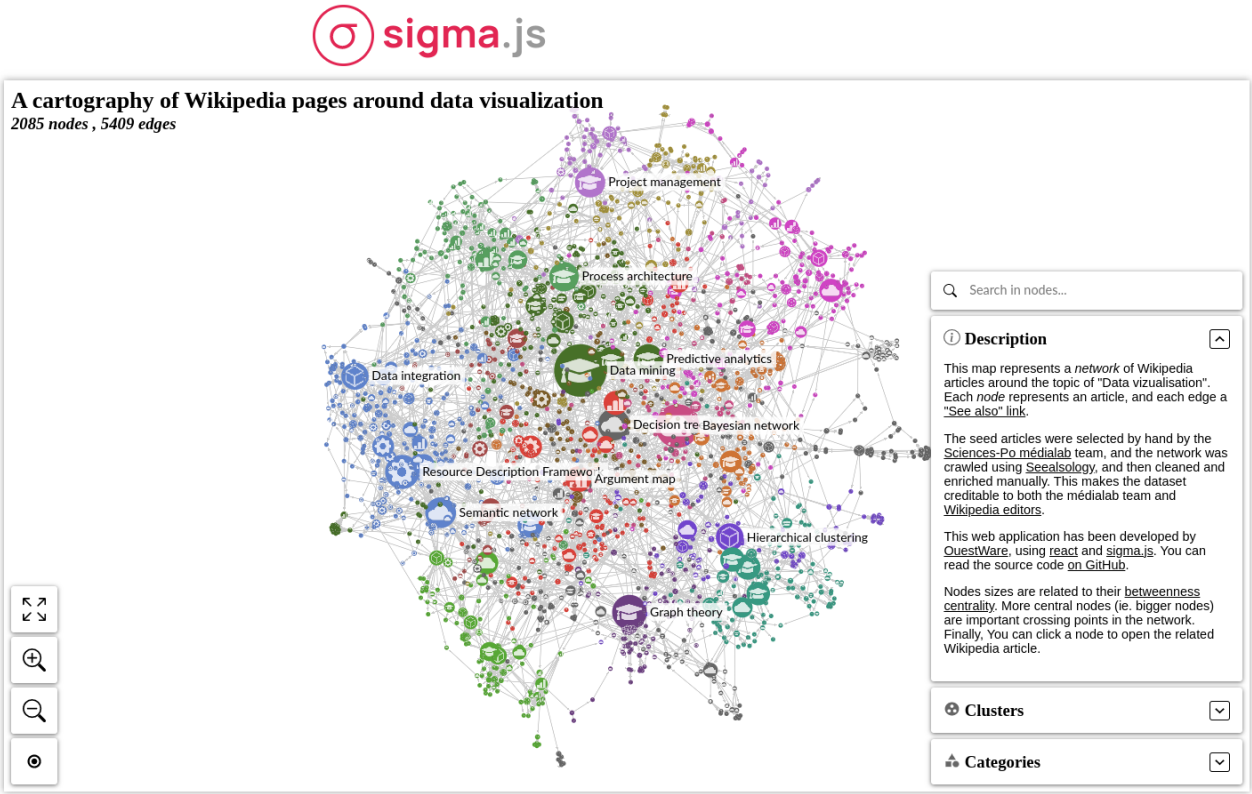
Pajek


Pajek32 5.01
File Network Networks Operations Partition Partitions Vector Vectors Permutation Permutations Cluster Hierarchy
Options Draw Macro Info Tools
Networks
big data india abs title keyword\co-authorship\arcs to edge for cohesion study.net (865)
Partitions
ig data_India\big data india abs title keyword\co-authorship\big data till 2015.clu (865)
Vectors
g data_India\big data india abs title keyword\co-authorship\big data till 2015.vec (865)
Permutations
Cluster
Hierarchy
3. D:\Big data_India\big data india abs title keyword\co-authorship\arcs to ed...
Layout Layers GraphOnly Previous Redraw Next Options Export Spin Move Info FishEye
Wait

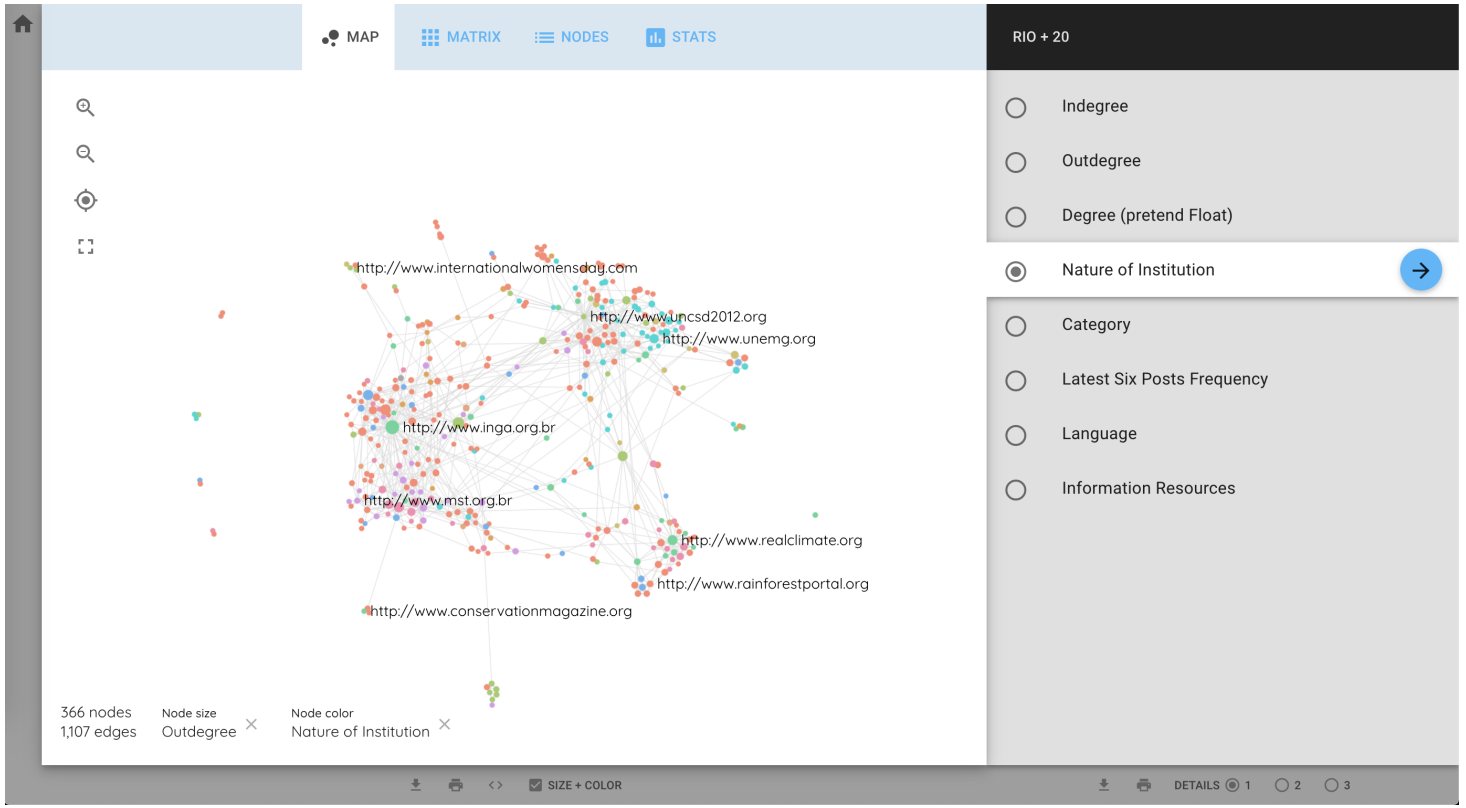
Gephi


Gephi 0.10.1 - Projet
Vue d'ensemble
Laboratoire de données
Prévisualisation
Les Miserables
Aspect
Noeuds
Liens
Unique
Partition
Classement
#c0c0c0
Appliquer
Spatialisation
Choisir une spatialisation
Exécuter
<Aucune propriété>
Réglages...
Réinitialiser
Graphe
Déplacement
(Configurer)
Arial-BoldMT, 32
Contexte
Noeuds: 77
Liens: 254
Graphe non orienté
Filtres
Statistiques
Réinitialiser
Bibliothèque
Attributs
Dynamique
Liens
Opérateur
Topologie
Requêtes sauvegardées
Requêtes
Glissez le filtre ici
Sélectionner
Filtrer

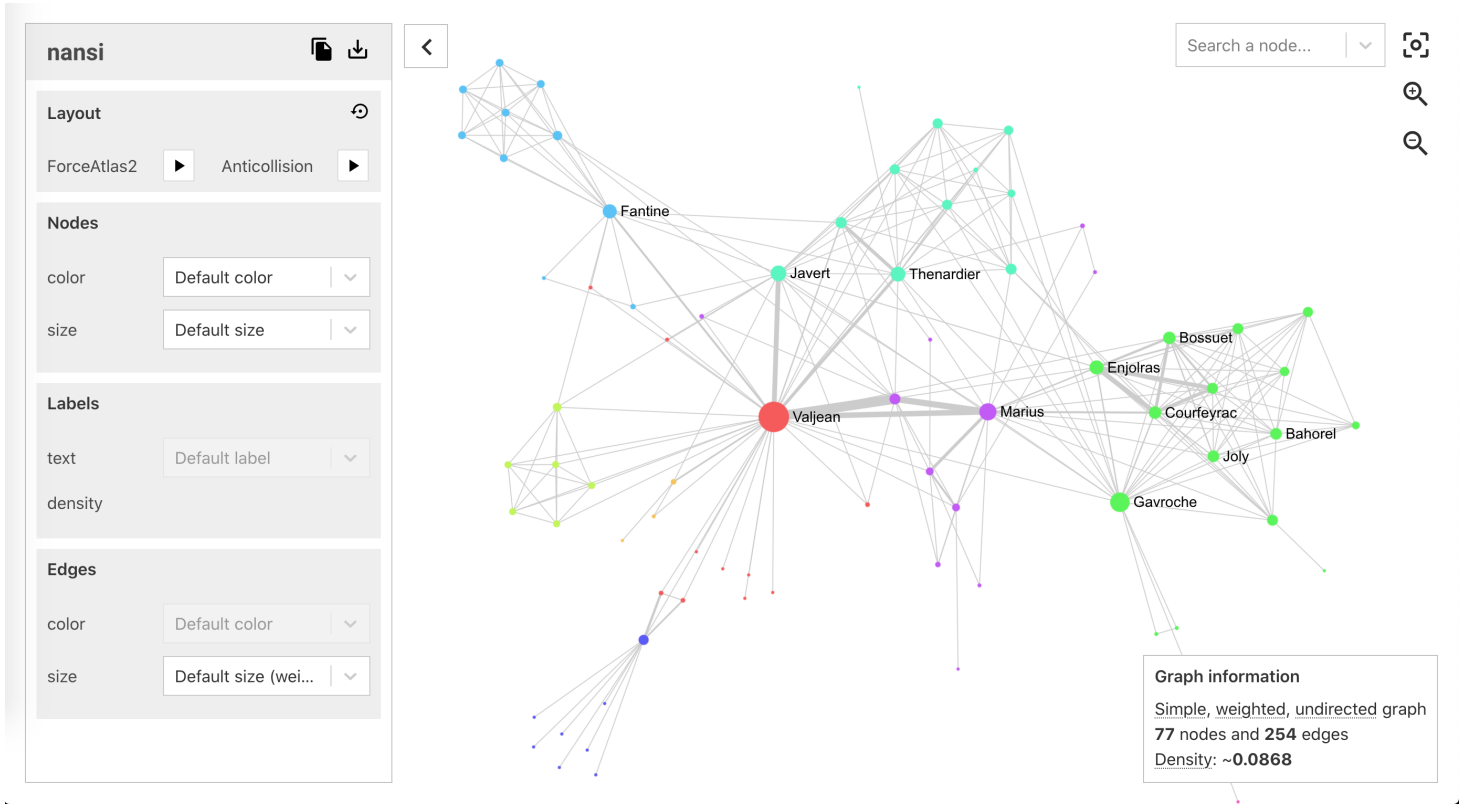
Sigma


sigma.js
A cartography of Wikipedia pages around data visualization
2085 nodes , 5409 edges
Project management
Process architecture
Predictive analytics
Data integration
Data mining
Decision tre
Bayesian network
Resource Description Framewo
Argument map
Semantic network
Hierarchical clustering
Graph theory
Search in nodes...
Description
This map represents a network of Wikipedia articles around the topic of "Data vizualisation". Each node represents an article, and each edge a "See also" link.
The seed articles were selected by hand by the Sciences-Po médialab team, and the network was crawled using Seealsology, and then cleaned and enriched manually. This makes the dataset creditable to both the médialab team and Wikipedia editors.
This web application has been developed by OuestWare, using react and sigma.js. You can read the source code on GitHub.
Nodes sizes are related to their betweenness centrality. More central nodes (ie. bigger nodes) are important crossing points in the network. Finally, You can click a node to open the related Wikipedia article.
Clusters
Categories

MiniVaN


MAP
MATRIX
NODES
STATS
RIO + 20
Indegree
Outdegree
Degree (pretend Float)
Nature of Institution
Category
Latest Six Posts Frequency
Language
Information Resources
http://www.internationalwomensday.com
http://www.uncsd2012.org
http://www.unemg.org
http://www.inga.org.br
http://www.mst.org.br
http://www.realclimate.org
http://www.rainforestportal.org
http://www.conservationmagazine.org
366 nodes
1,107 edges
Node size
Outdegree
Node color
Nature of Institution
SIZE + COLOR
DETAILS 1 2 3

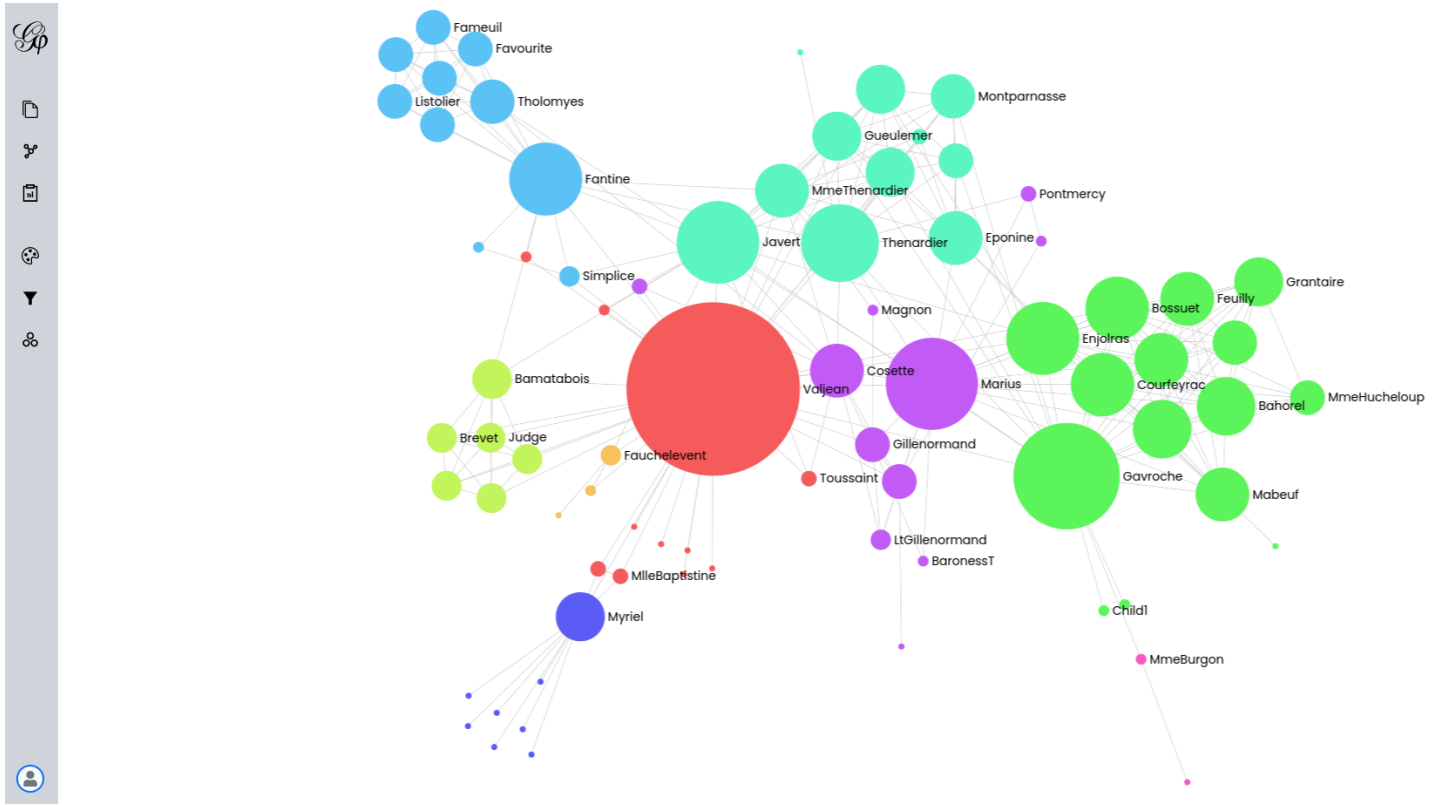
Nansi


nansi
Layout
ForceAtlas2
Anticollision
Nodes
color
Default color
size
Default size
Labels
text
Default label
density
Edges
color
Default color
size
Default size (wei...
Search a node...
Fantine
Javert
Thenardier
Valjean
Marius
Bossuet
Enjolras
Courfeyrac
Bahorel
Joly
Gavroche
Graph information
Simple, weighted, undirected graph
77 nodes and 254 edges
Density: ~0.0868

Retina

You are currently using a local file, that only you can access.
To be able to share your visualizations online, you need to first publish your graph file online.
Publish the graph online
Edit
Explore
Graph visualization edition
Before sharing your graph online, you can first select various options on how users will read and interrogate this graph.
Your choices in this panel will impact the next white panel, which is the interface your users will have access to.
Which fields should be actionable?
Filter
Colors
Sizes
modularity_class as quantitative values
modularity_class as qualitative values
Degree
Allow using default graph file colors and/or sizes
Which node information should show up on hovered nodes?
Select fields...
Valjean
Claquesous
Fantine
Babet
MmeThenardier
Javert
Thenardier
Simplice
Bossuet
Enjolras
Bamatabois
Cosette
Marius
Champmathieu
Toussaint
MlleGillenormand
Gavroche
Chenildieu
MlleBaptistine
Myriel

Gephi Lite


Fameuil
Favourite
Listolier
Tholomyes
Fantine
Montparnasse
Gueulemer
MmeThenardier
Pontmercy
Javert
Thenardier
Eponine
Simplice
Magnon
Grantaire
Bossuet
Feuilly
Enjolras
Bamatabois
Cosette
Valjean
Marius
Courfeyrac
Bahorel
MmeHucheloup
Brevet
Judge
Fauchelevent
Gillenormand
Toussaint
Gavroche
Mabeuf
LtGillenormand
BaronessT
MlleBaptistine
Myriel
Child
MmeBurgon

Designing an app for us, social sciences data engineers

Nansi

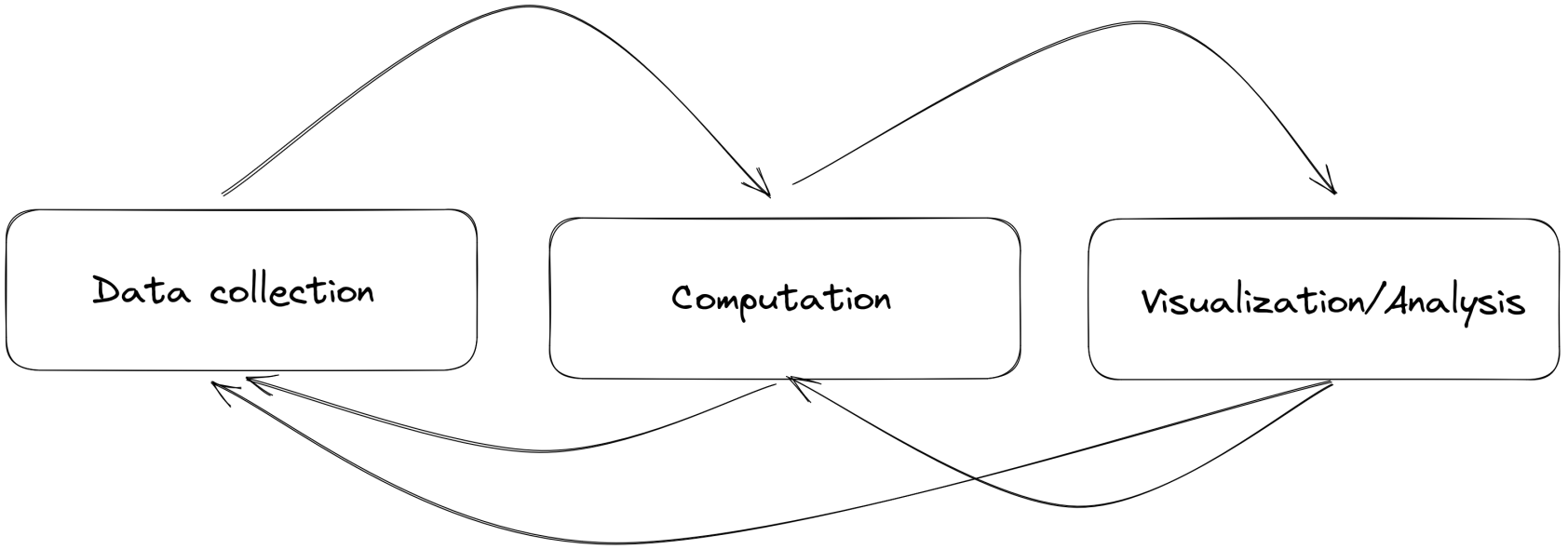
Two conflicting objectives:

1. A better tool for teaching students
2. Suiting our own data processing/exploration/analysis process

⁂

It was *doomed* from the start.

The iterative process



Why don't we perform Visual Network Analysis directly in a Jupyter notebook then?

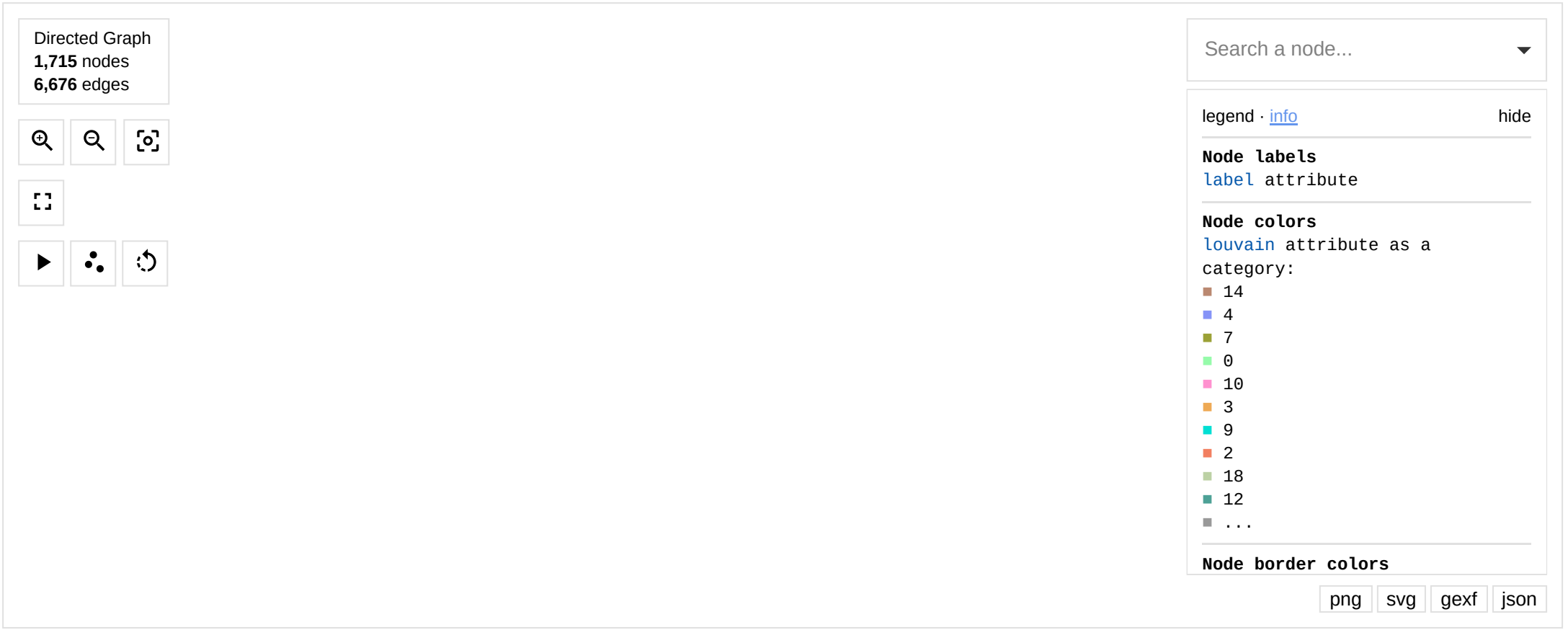
Introducing ipysigma

```
pip install jupyterlab
pip install networkx # or igraph
pip install ipysigma
```

(Hopefully working) Demo time!

Small multiples

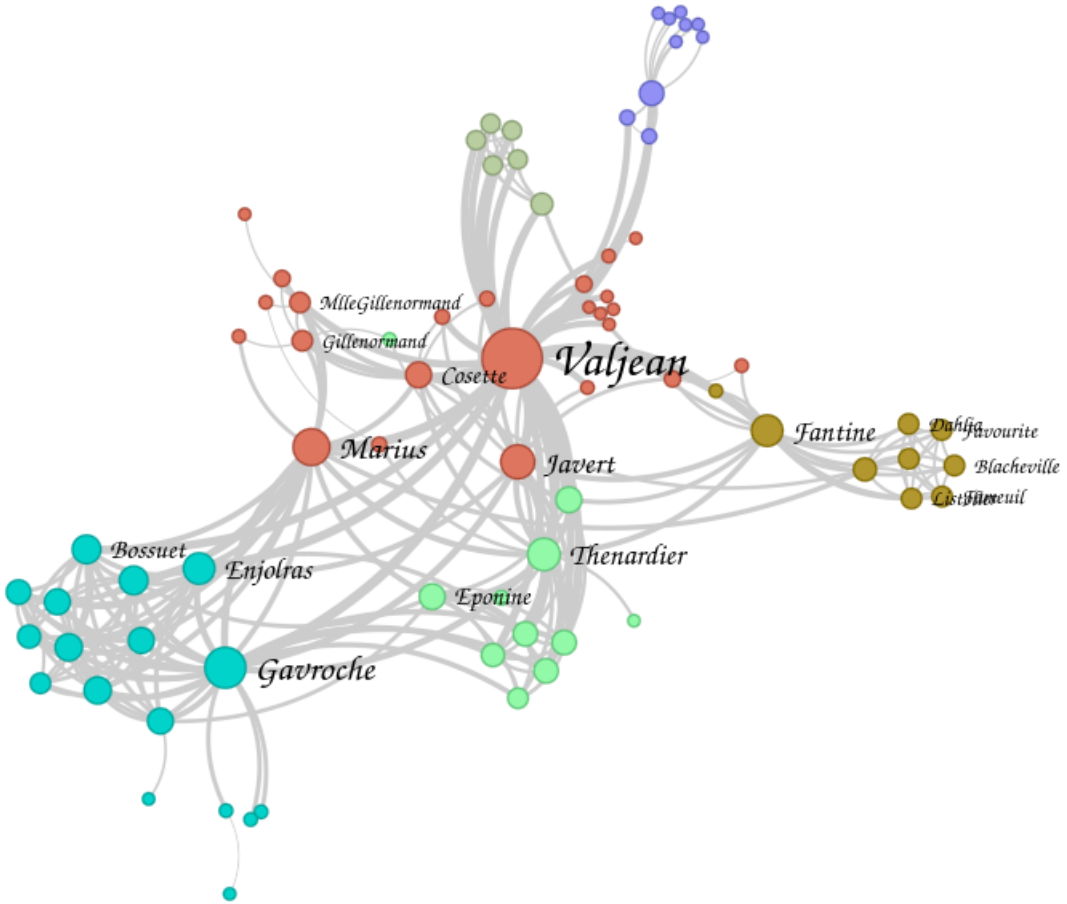
Static embeds



Future directions

- Specialized representations such as:
  - Temporal graphs
  - Bipartite graphs
  - etc.
- Better support for various incarnations of Jupyter:
  - notebook
  - lab
  - colab
  - vscode
  - etc.

Difficulties of developing custom widgets

- The documentation is a bit all around the place
- A lot of different version of `ipywidgets` etc. have to coexist
- Can be tedious to switch from the python/javascript contexts
- A lot of trial and error
- Annoying warnings in the console I cannot find a way to suppress :(

Thank you for listening!