

Visual Network Analysis

With Gephi

Mathieu Jacomy
Sciences Po Paris médialab
Equipex DIME-SHS ANR-10-EQPX-19-01

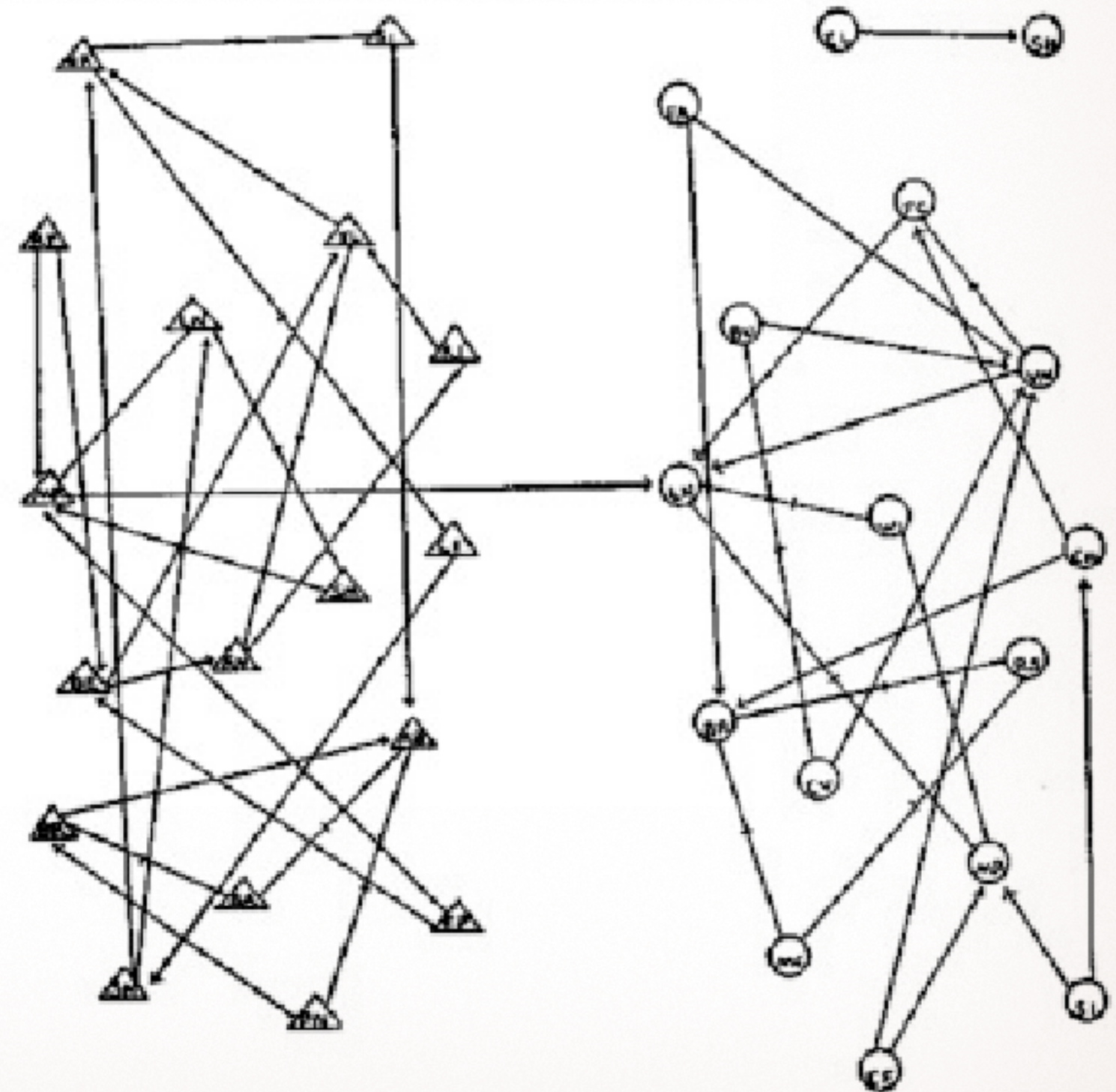


1. Download sample network
<http://bit.ly/2ud9oUi>
2. Download Gephi
<http://gephi.org>

Network visualization in social science

EMOTIONS MAPPED BY NEW GEOGRAPHY

Charts Seek to Portray the
Psychological Currents of
Human Relationships.



Good to know

1. Gephi requires installing **JAVA**.

JAVA is a software layer. Gephi works on top of it.

Installing it is innocuous if you mind to **uncheck the “parasites”** suggesting during the install process. (default search engine...)

2. Gephi works on Windows, Mac OS & Linux (thanks to Java)

3. It is possible to **report bugs** on GitHub

<https://github.com/gephi/gephi/issues>

Learning online

Official website

<https://gephi.org/>

Gephi Cheat Sheets (C. Levallois)

http://www.clementlevallois.net/gephi/tuto/en/gephi_cheat%20sheets_en.pdf

Many available videos

https://www.youtube.com/results?search_query=gephi

Official tutorials

<https://gephi.org/users/>

Clément Levallois' tutorials

<http://www.clementlevallois.net/gephi.html>

Martin Grandjean's tutorials

<http://www.martingrandjean.ch/gephi-introduction/>

Additional online resources

Table2Net

Get a network from a CSV table

<http://tools.medialab.sciences-po.fr/table2net/>

Introduction to **Network Science**

(by A.-L. Barabási)

<http://barabasi.com/networksciencebook/>

Manylines, prototype of online

network narratives

<https://github.com/medialab/manylines>

Manylines slides from FOSDEM

<http://medialab.github.io/manylines/fosdem2015/>

Placing nodes in a plane

Layout algorithms

The need for shaping



Networks **have no proper shape.**
We have to give them one.

But our networks are **sewn onto themselves.**
Like this robe that we cannot “iron”.

We just cannot avoid links crossings.
(this math property is being non-planar)

How a layout works

1



2



3



4



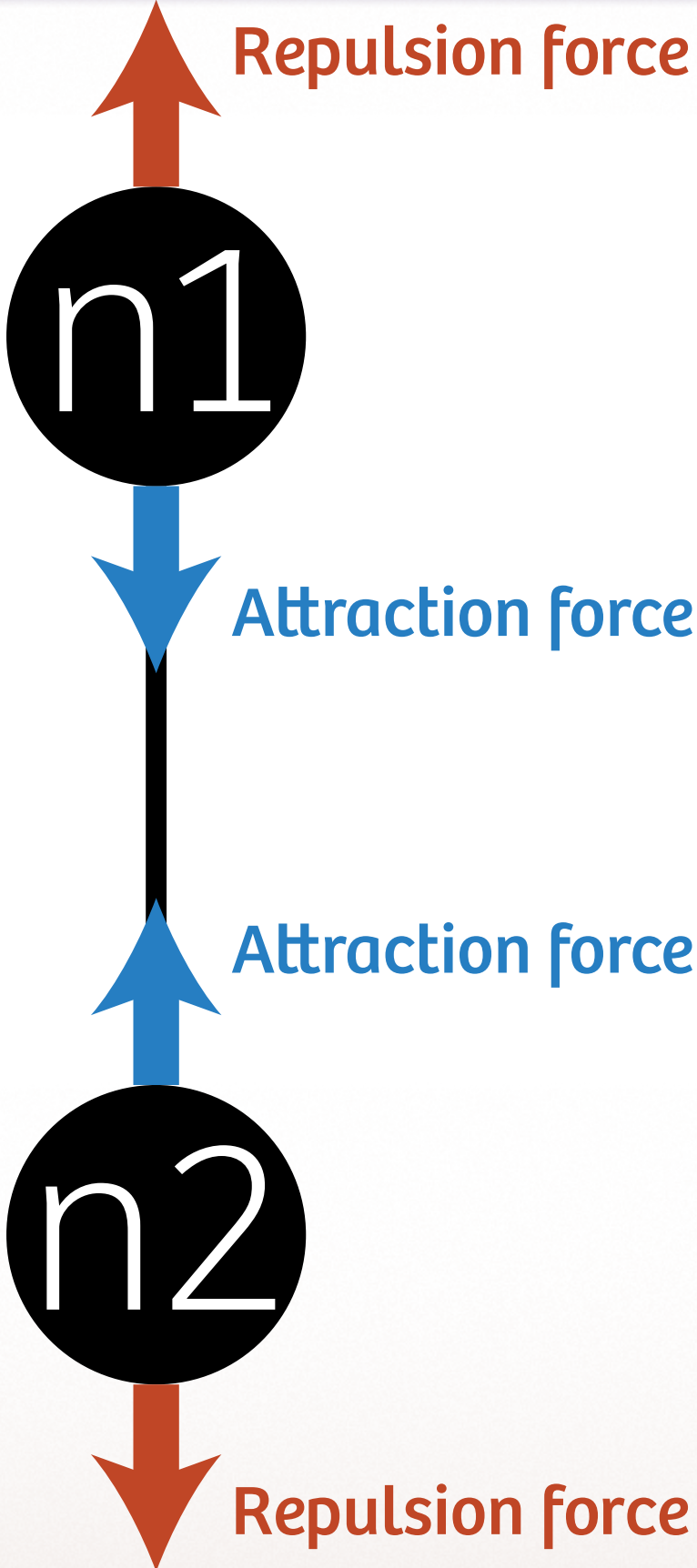
5



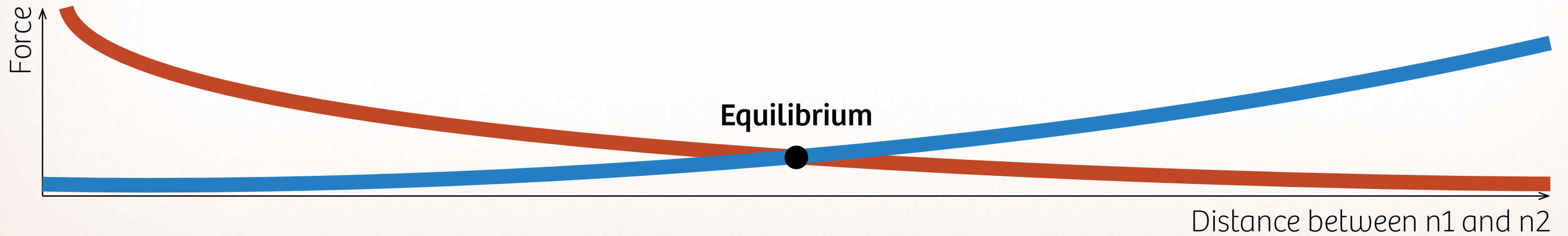
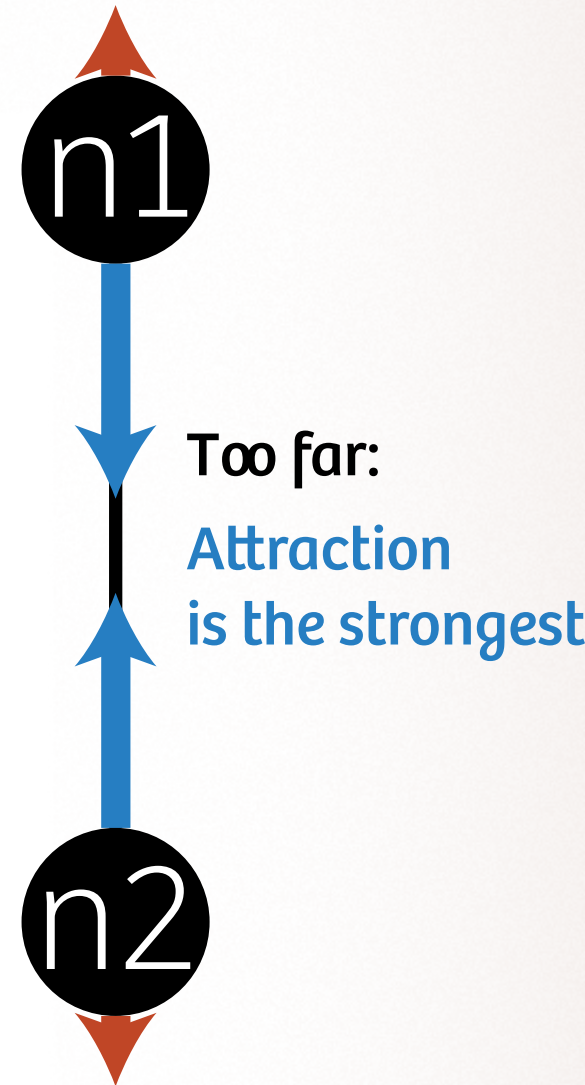
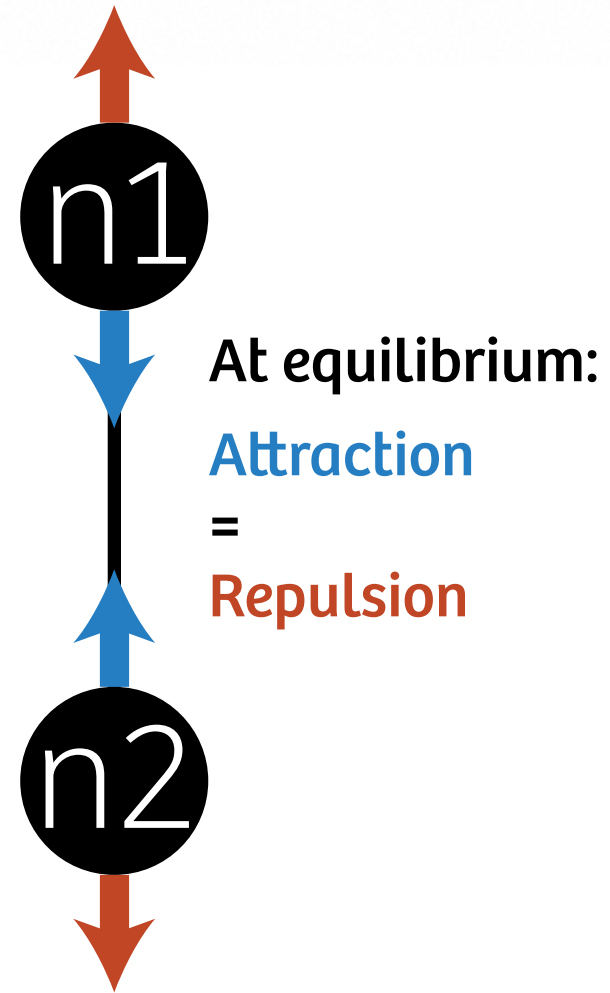
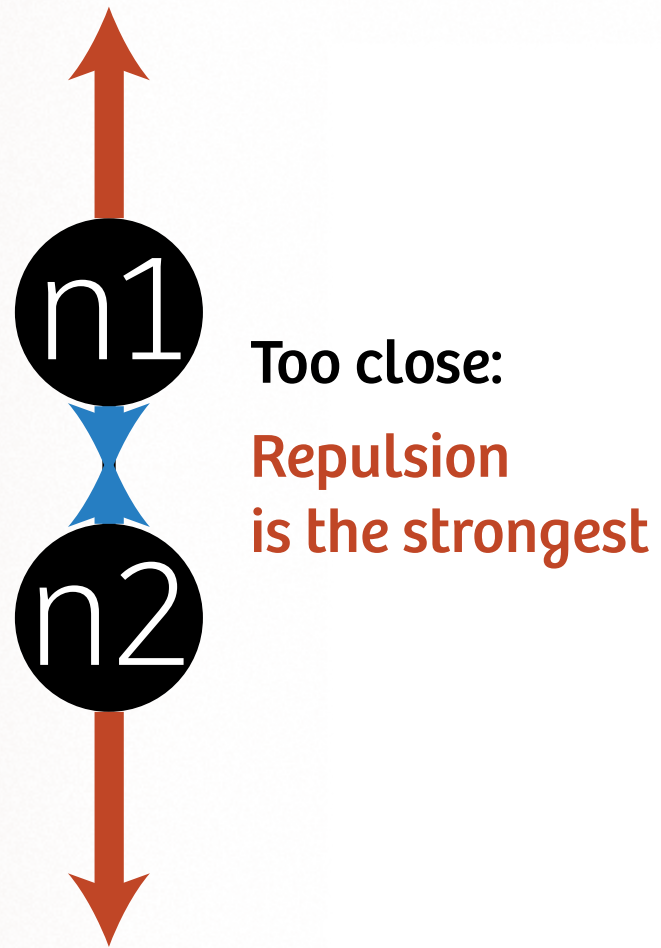
6



How a layout works



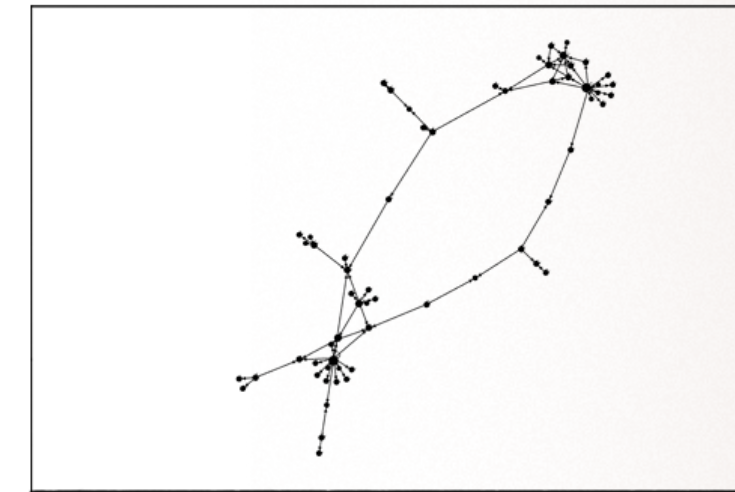
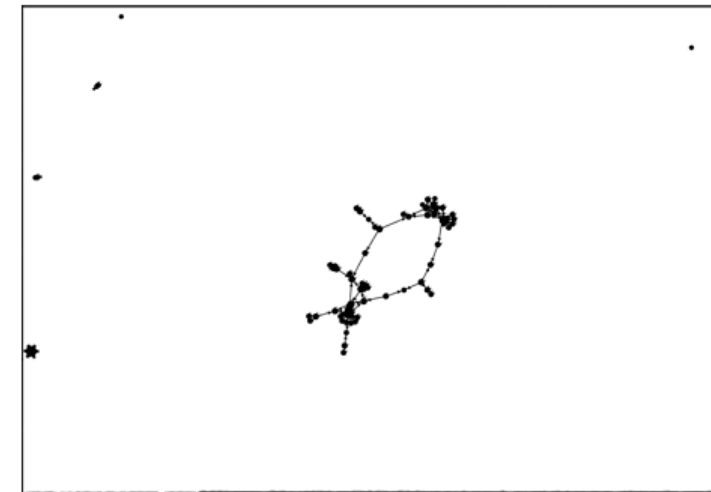
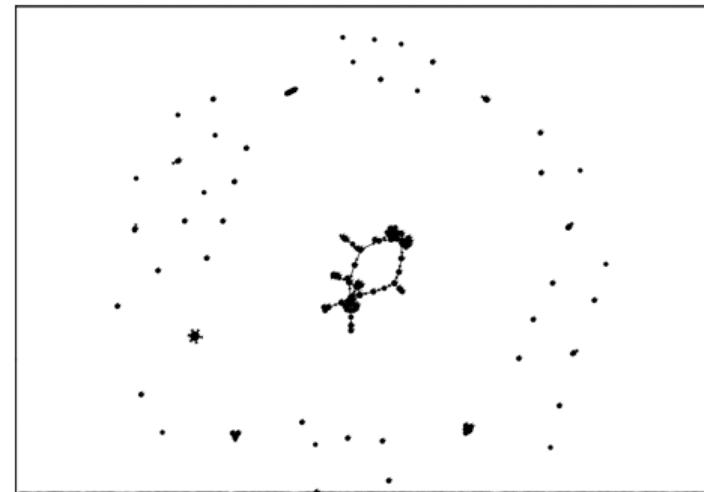
How a layout works



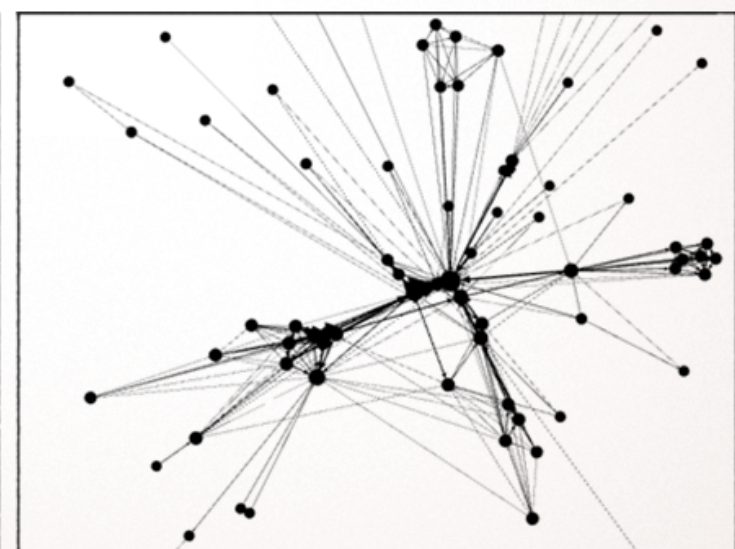
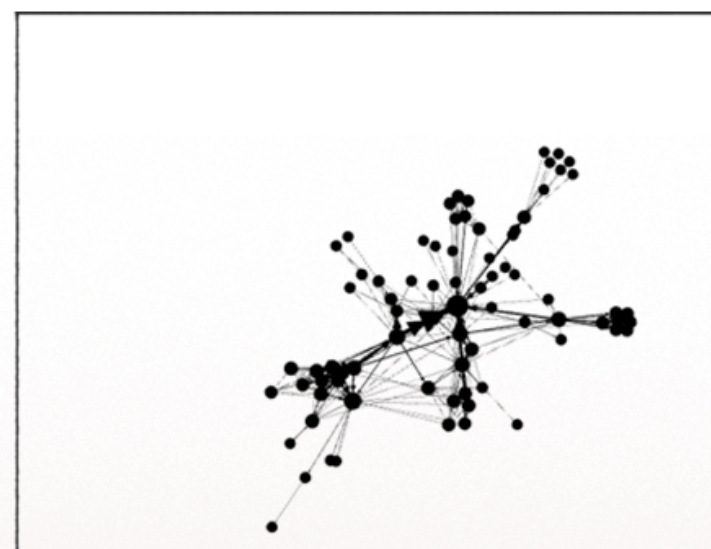
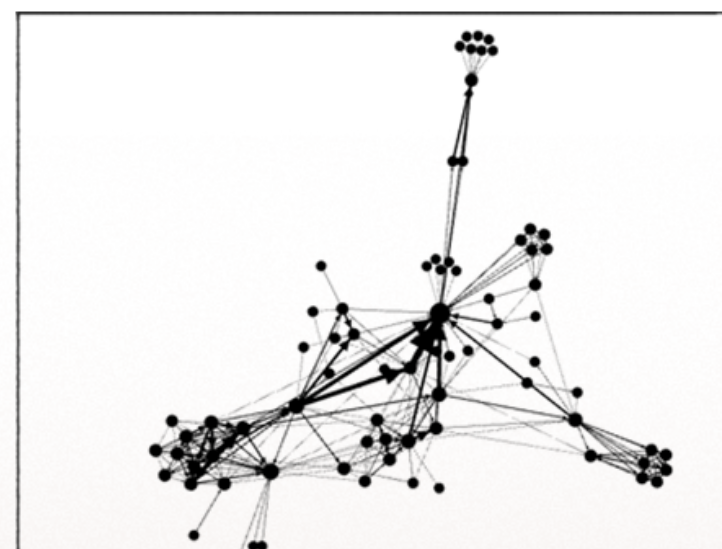
Layout settings (ForceAtlas2)

ForceAtlas2 paper
<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0098679>

Scaling



Influence of edges' weight

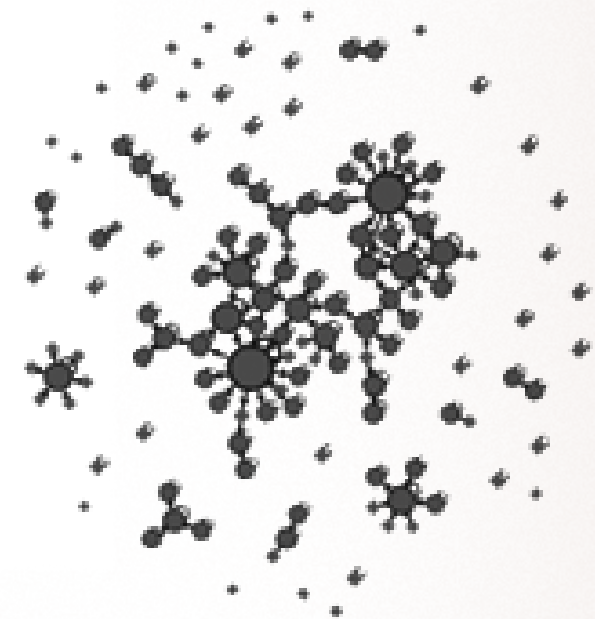
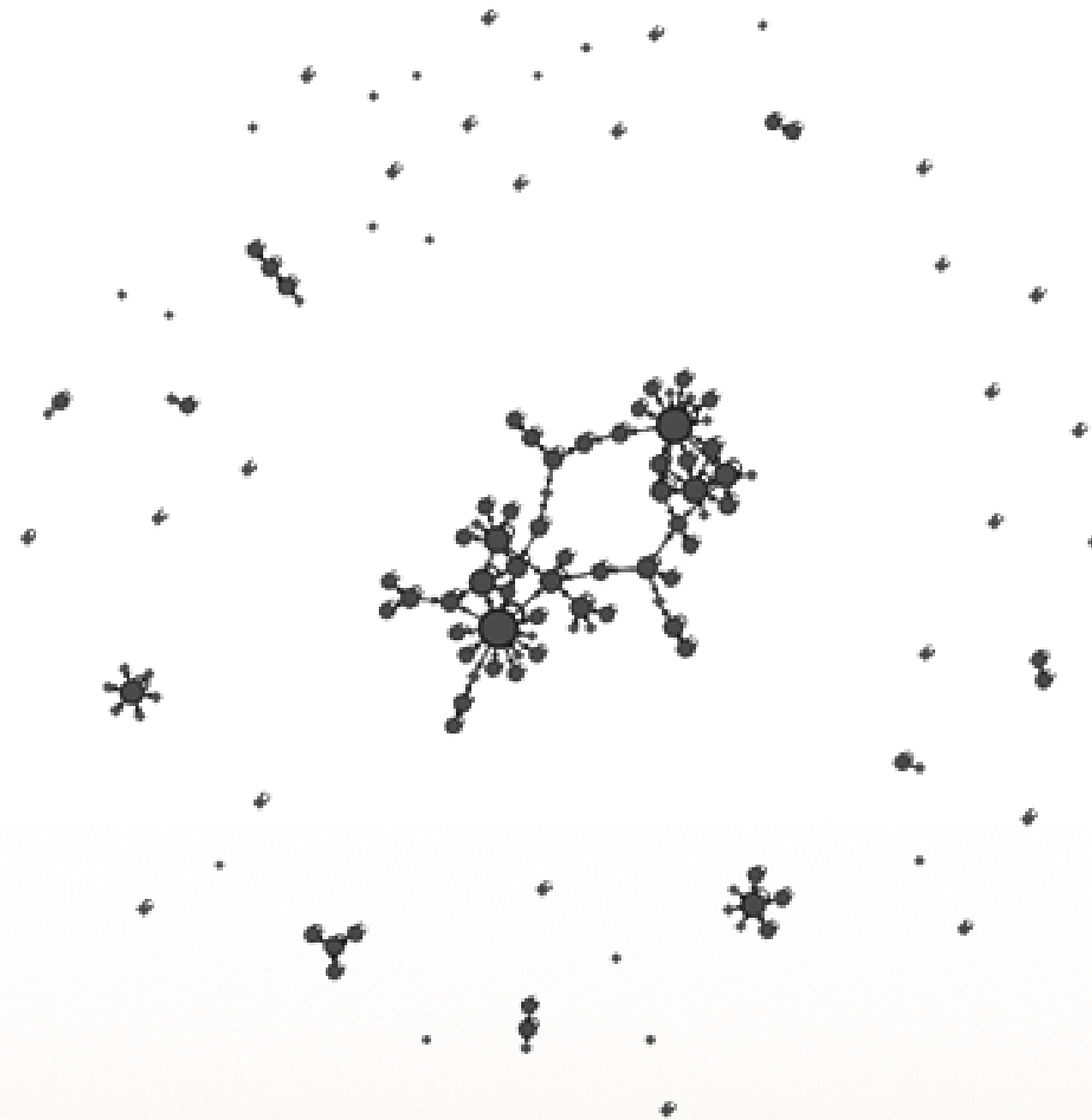


Layout settings (ForceAtlas2)

ForceAtlas2 paper

<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0098679>

Gravity

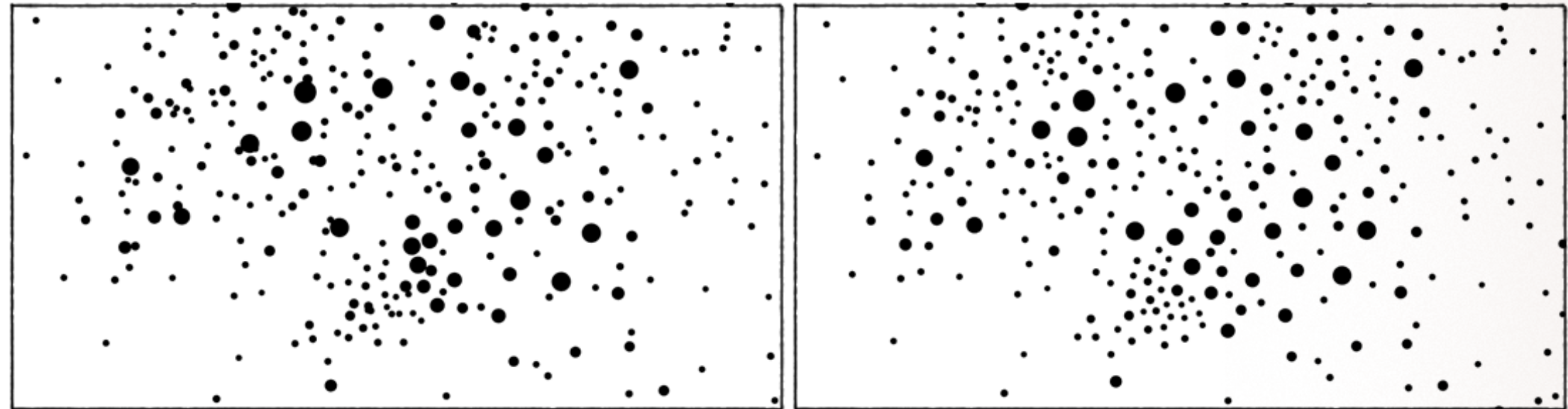


Layout settings (ForceAtlas2)

ForceAtlas2 paper

<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0098679>

Prevent overlap

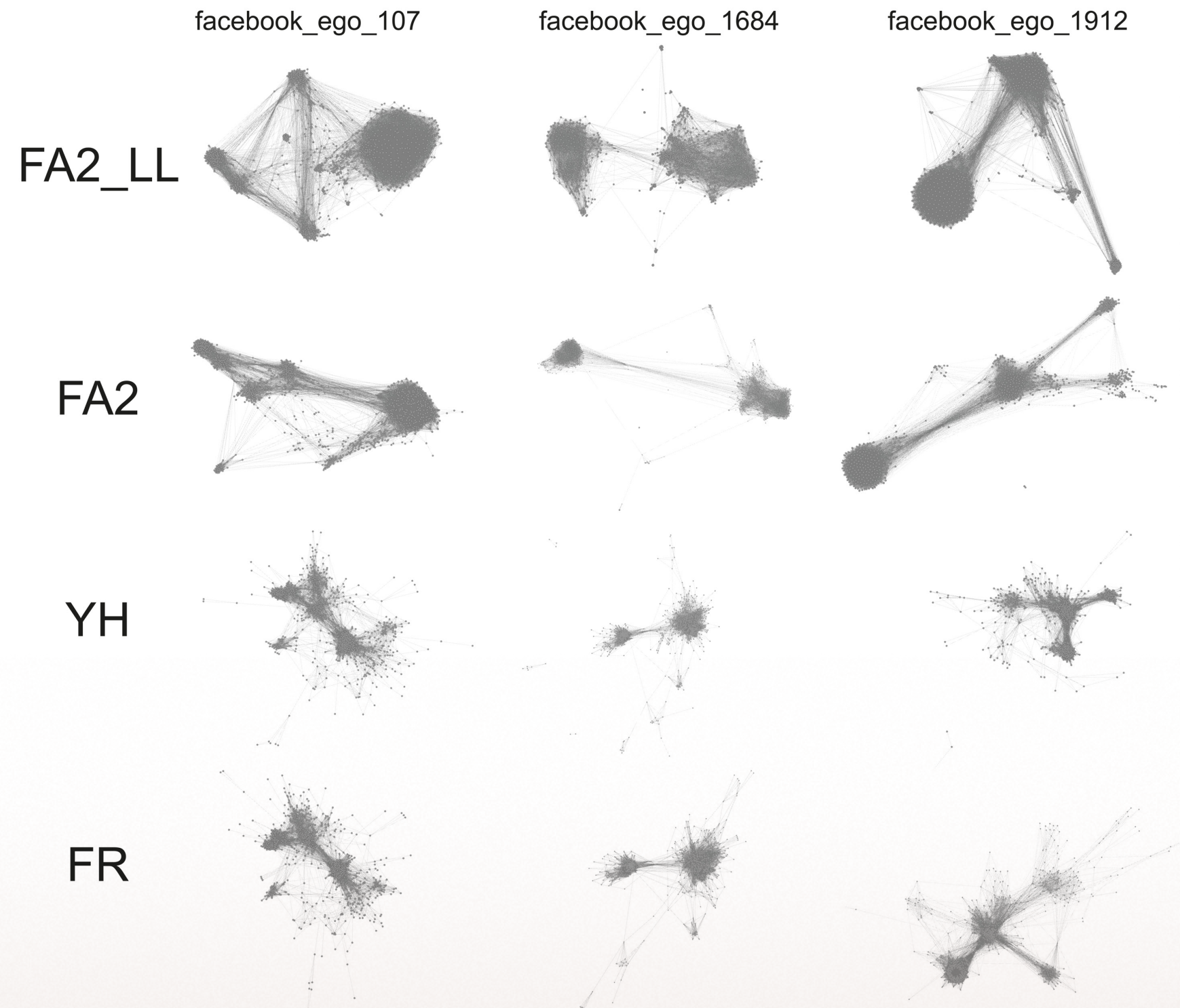


CAUTION: use only as a post-processing.

Layout settings (ForceAtlas2)

ForceAtlas2 paper
<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0098679>

Comparing different algorithms (including **LinLog**)



Framing network exploration

Exploratory data analysis

BD & Manga

Litt Generale & Etrangere

Litterature Jeunesse

How to
read such an
image?

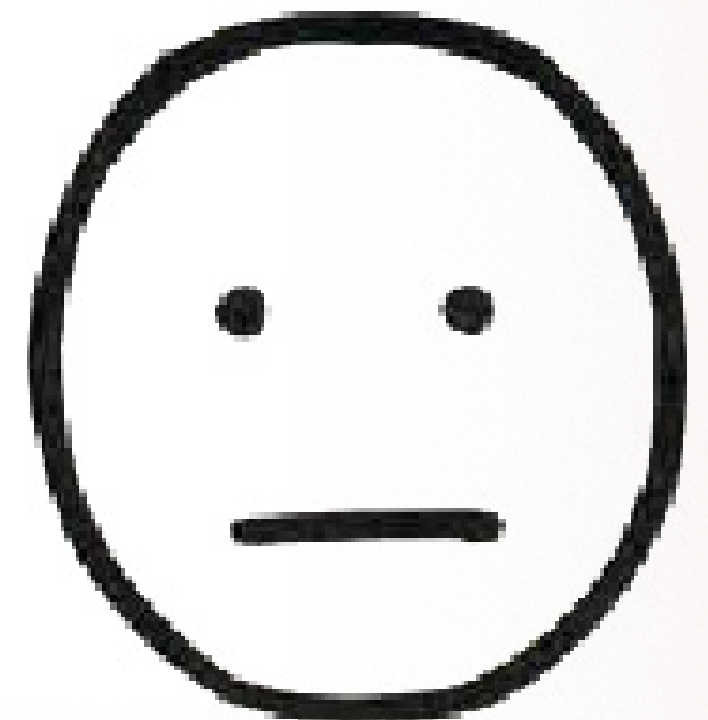
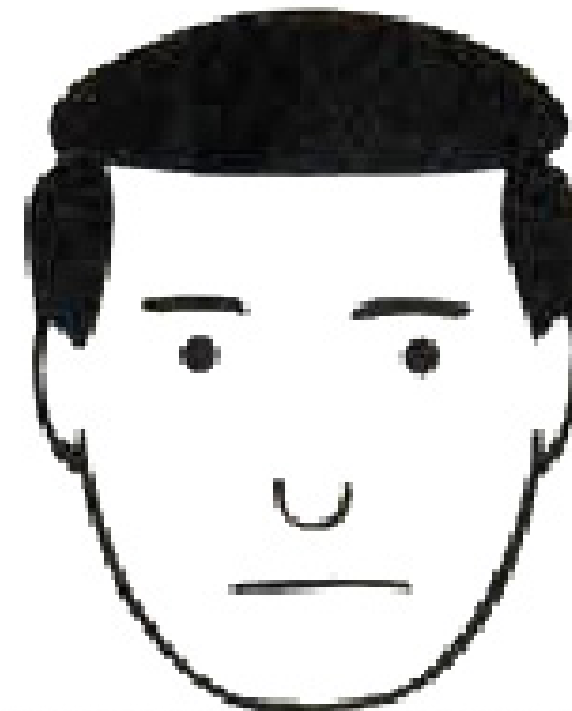
What is visible
but irrelevant?

What can you
trust?

What is hidden
but important?

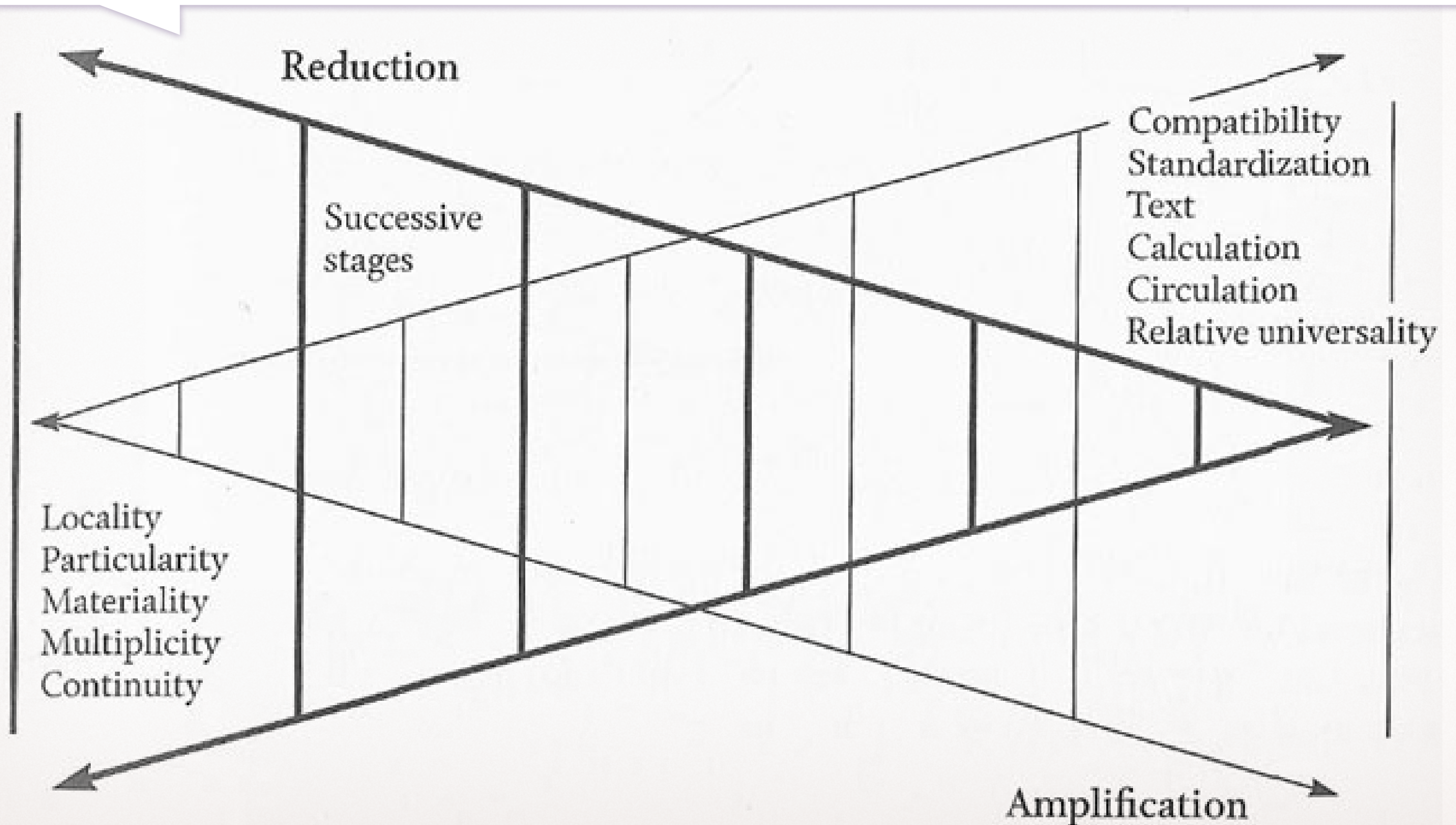
A frame for network analysis

Details



Legibility

A frame for network analysis



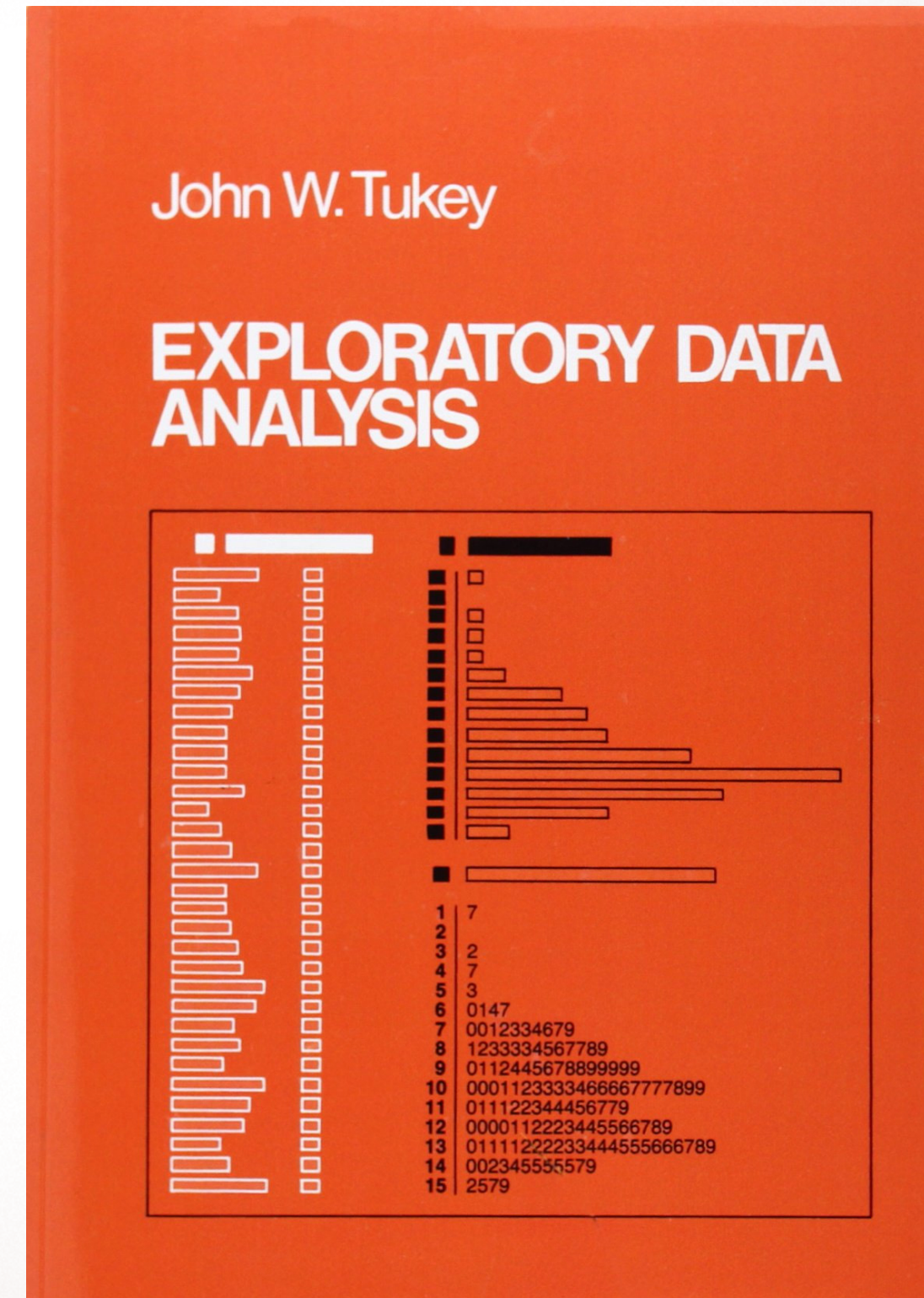
Exploratory Data Analysis

« The greatest value of a picture is when it forces us to notice what we **never expected to see.** »

— John W. Tukey

« Far better an approximate answer to **the right question**, which is often vague, than an exact answer to the wrong question, which can always be made precise. »

— John W. Tukey



Exploratory Data Analysis

Contrary to usual statistics,
EDA is **not confirmatory**

Contrary to visual communication,
EDA is **not explanatory**

EDA aims at

Discovering regularities

Detecting anomalies

Framing hypotheses

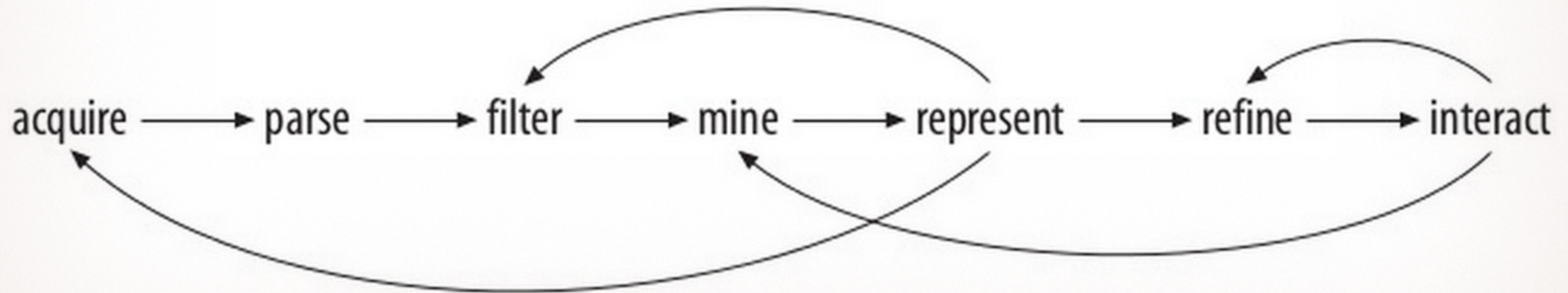
Checking preconceptions

A frame for network analysis

Example of a methodological chain

1. Identify clusters
2. Code clusters and their relations
3. Compute main metrics
4. Compare and visualize these metrics
5. Interpret considering the context

...



How to read a network

Visual networks analysis

Steps

Visualizing nodes' positions

1. Reading the variations of density
2. Interpreting the size and density of clusters
3. Detecting centers and bridges

Setting nodes' sizes

Reading the hierarchy of connectivity

Settings nodes' colors

Reading the distribution of categories

Density variations

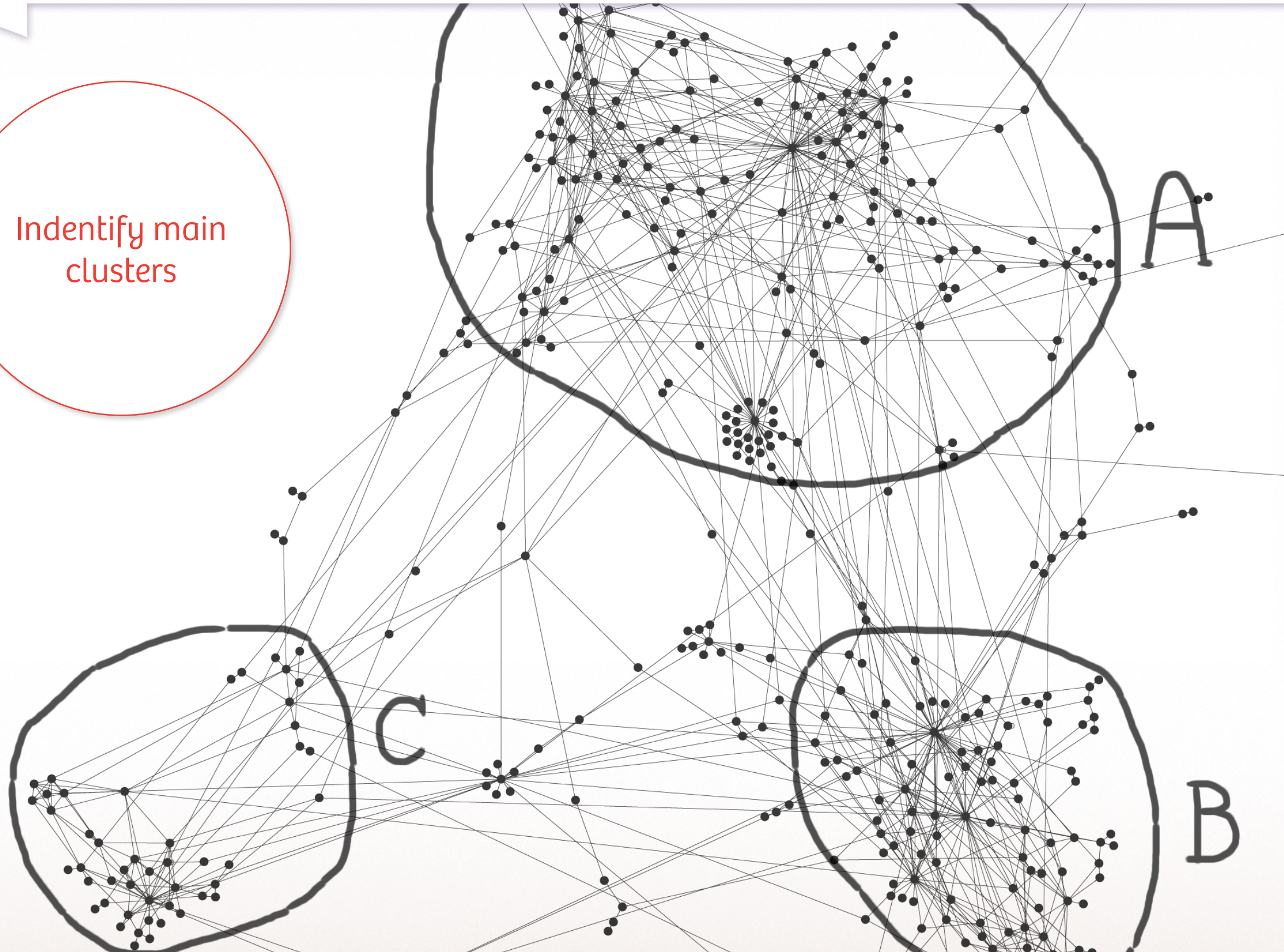
First we apply a layout



Where do nodes gather? Where are structural holes?

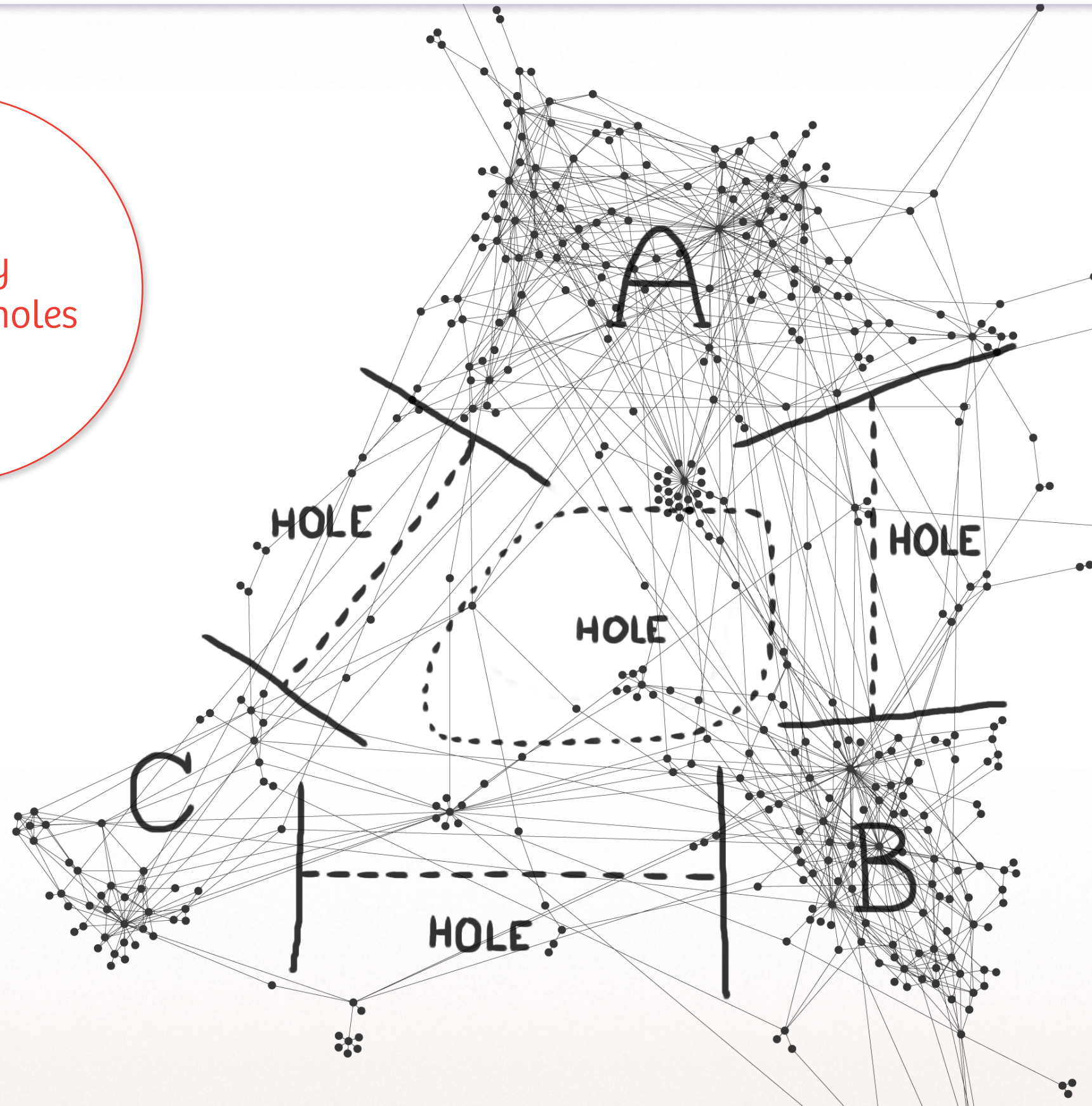
Density variations

Identify main clusters



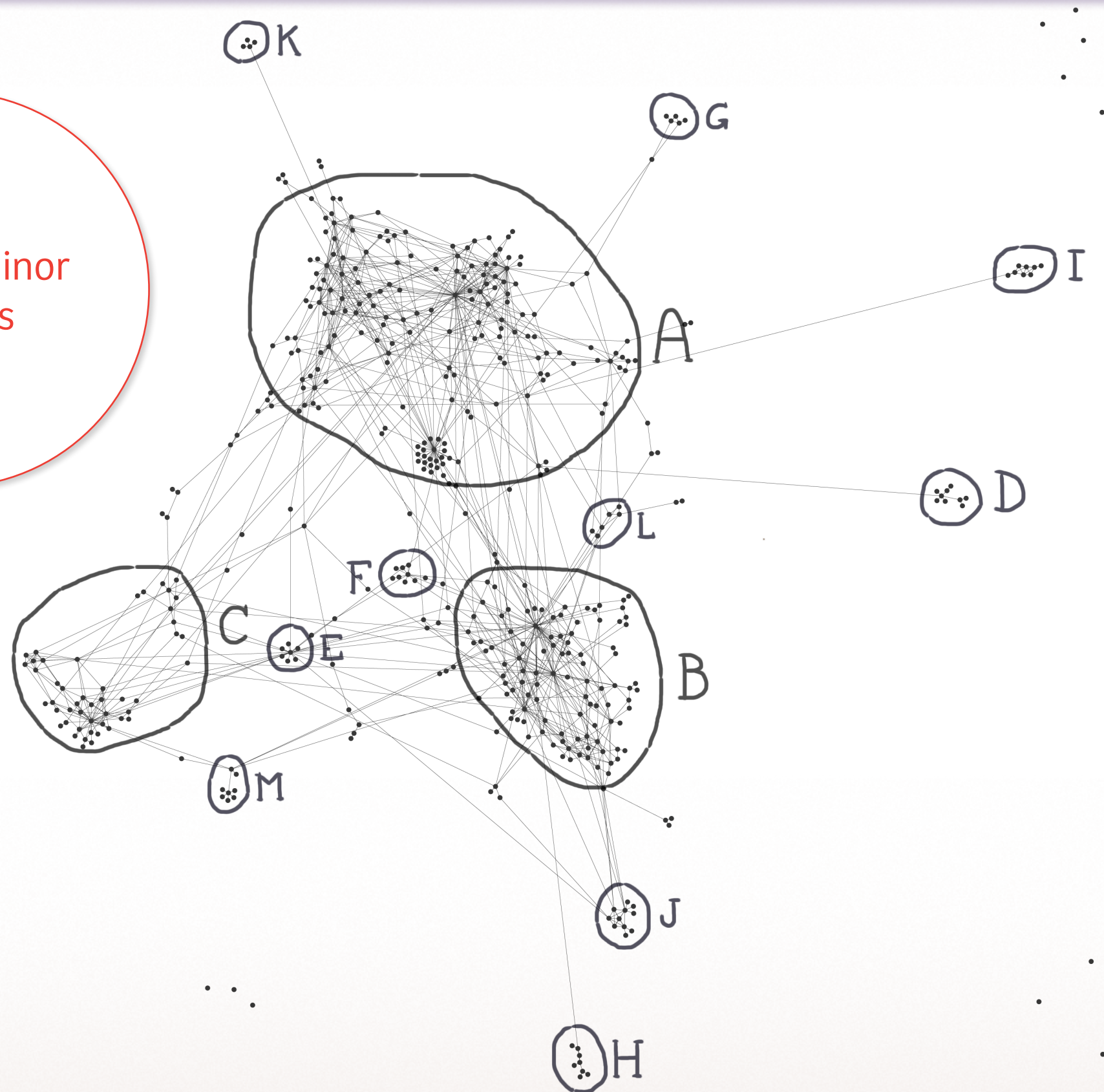
Density variations

Identify
structural holes

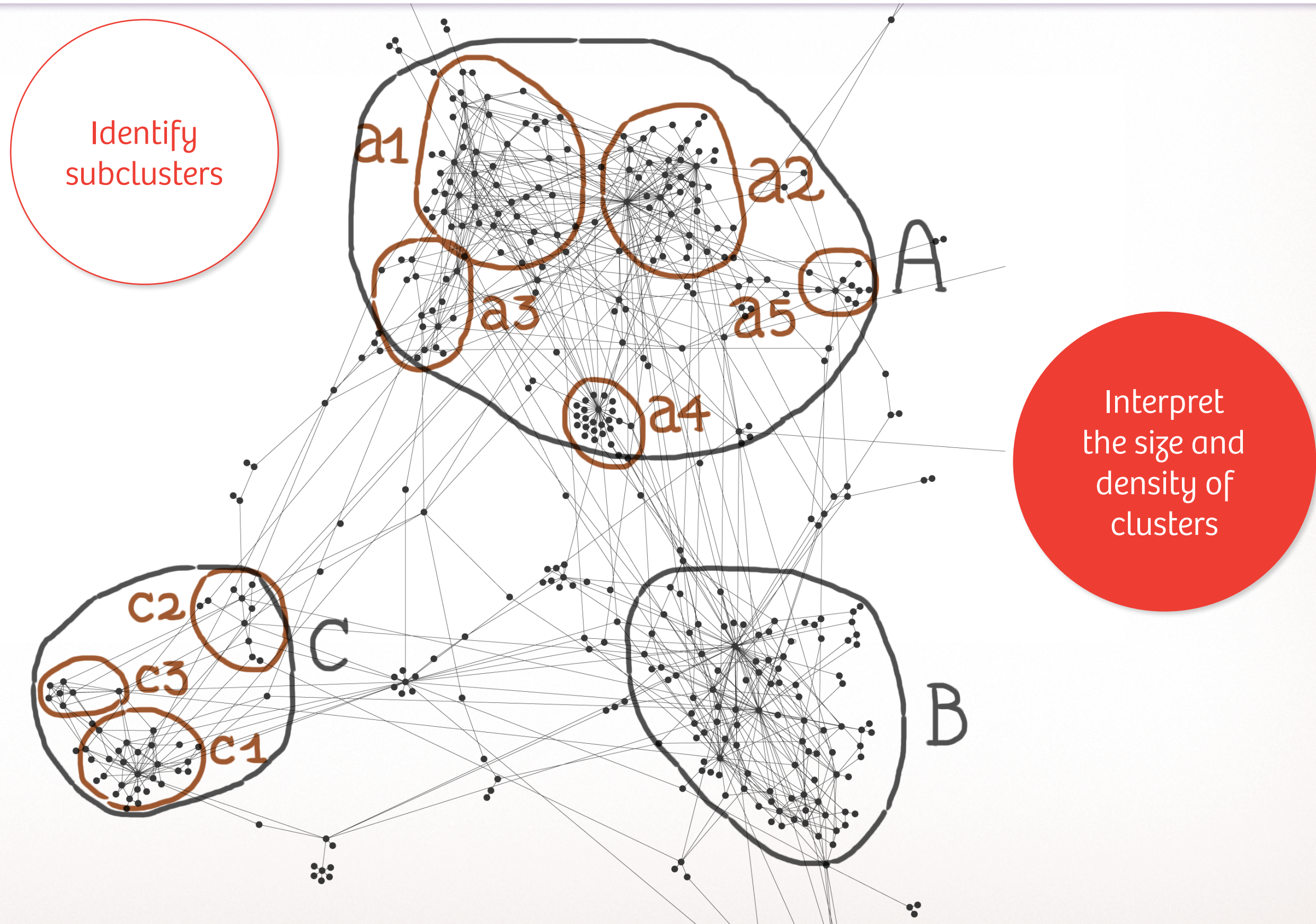


Density variations

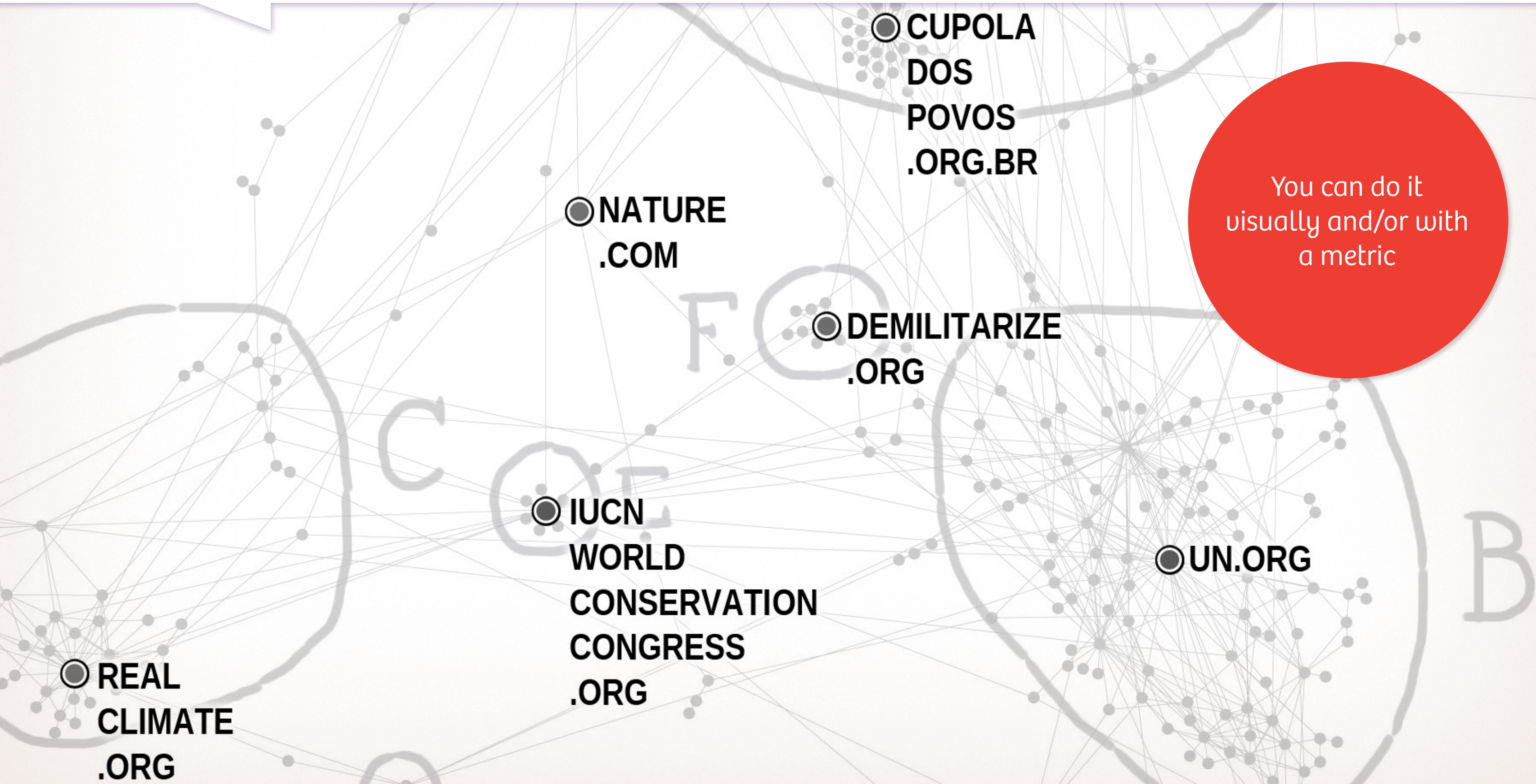
Identify minor clusters



Density variations

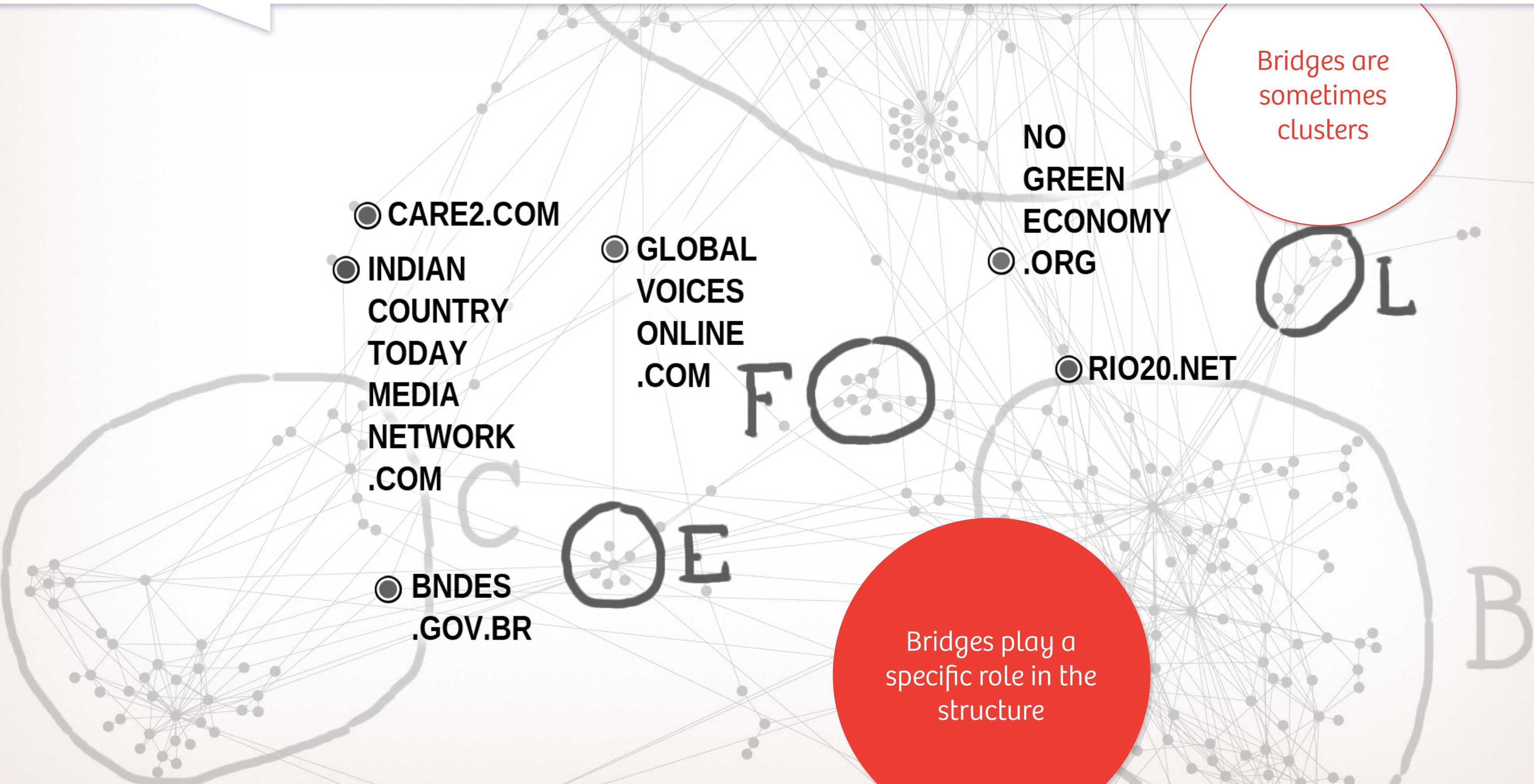


Detecting centers



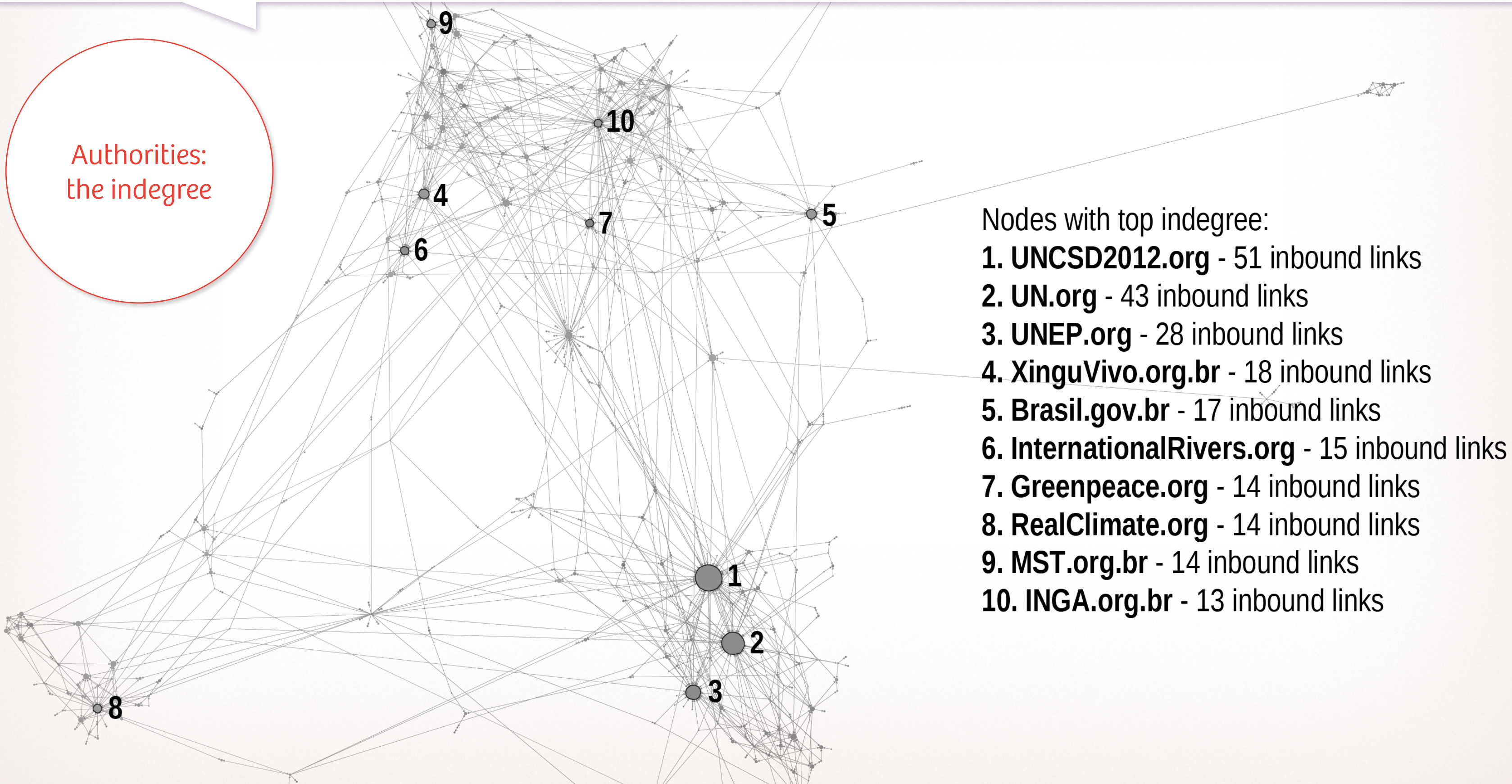
You can do it visually and/or with a metric

Detecting bridges



Hierarchy of connectivity

Authorities:
the indegree

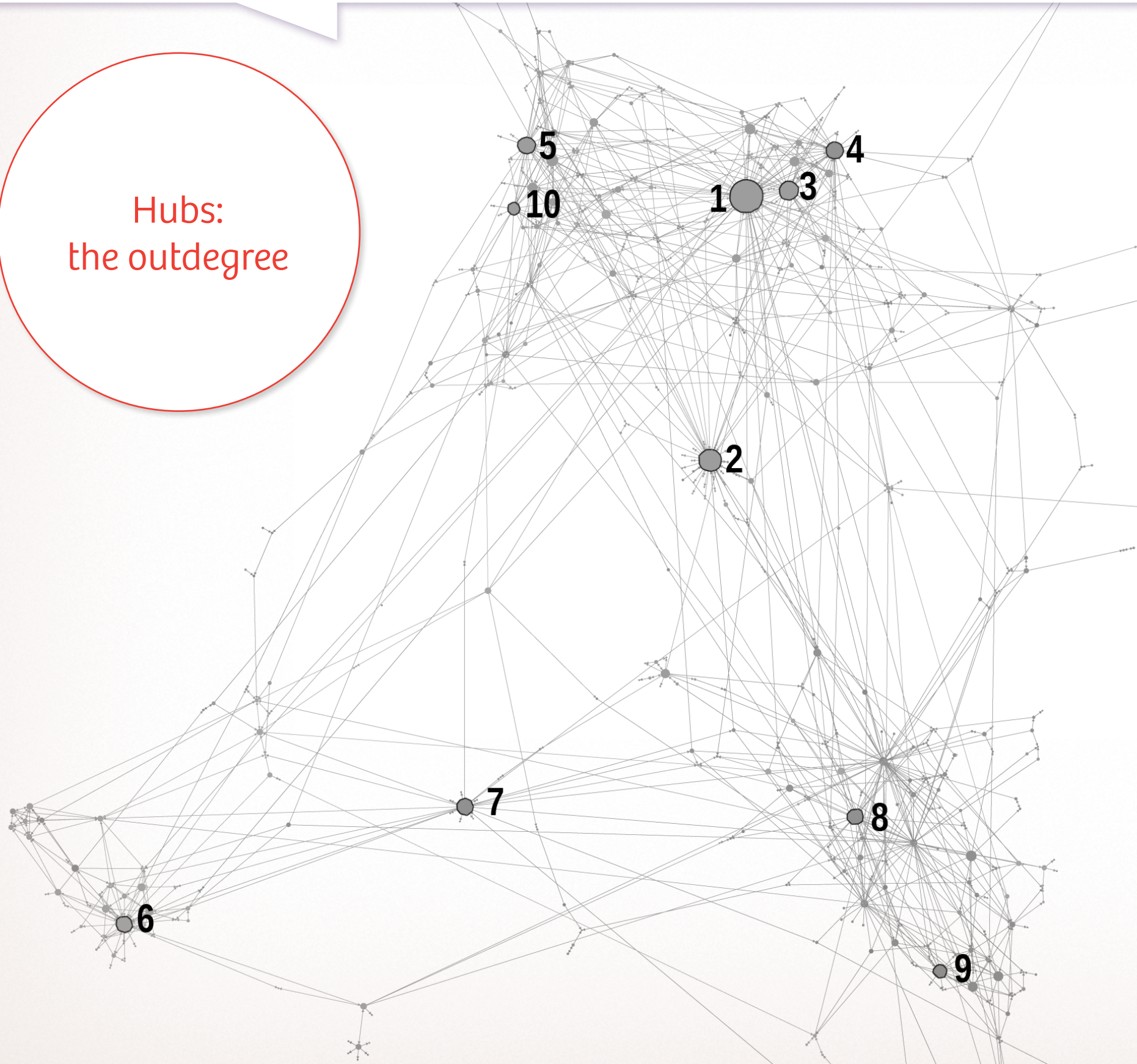


Nodes with top indegree:

1. **UNCSD2012.org** - 51 inbound links
2. **UN.org** - 43 inbound links
3. **UNEP.org** - 28 inbound links
4. **XinguVivo.org.br** - 18 inbound links
5. **Brasil.gov.br** - 17 inbound links
6. **InternationalRivers.org** - 15 inbound links
7. **Greenpeace.org** - 14 inbound links
8. **RealClimate.org** - 14 inbound links
9. **MST.org.br** - 14 inbound links
10. **INGA.org.br** - 13 inbound links

Hierarchy of connectivity

Hubs:
the outdegree

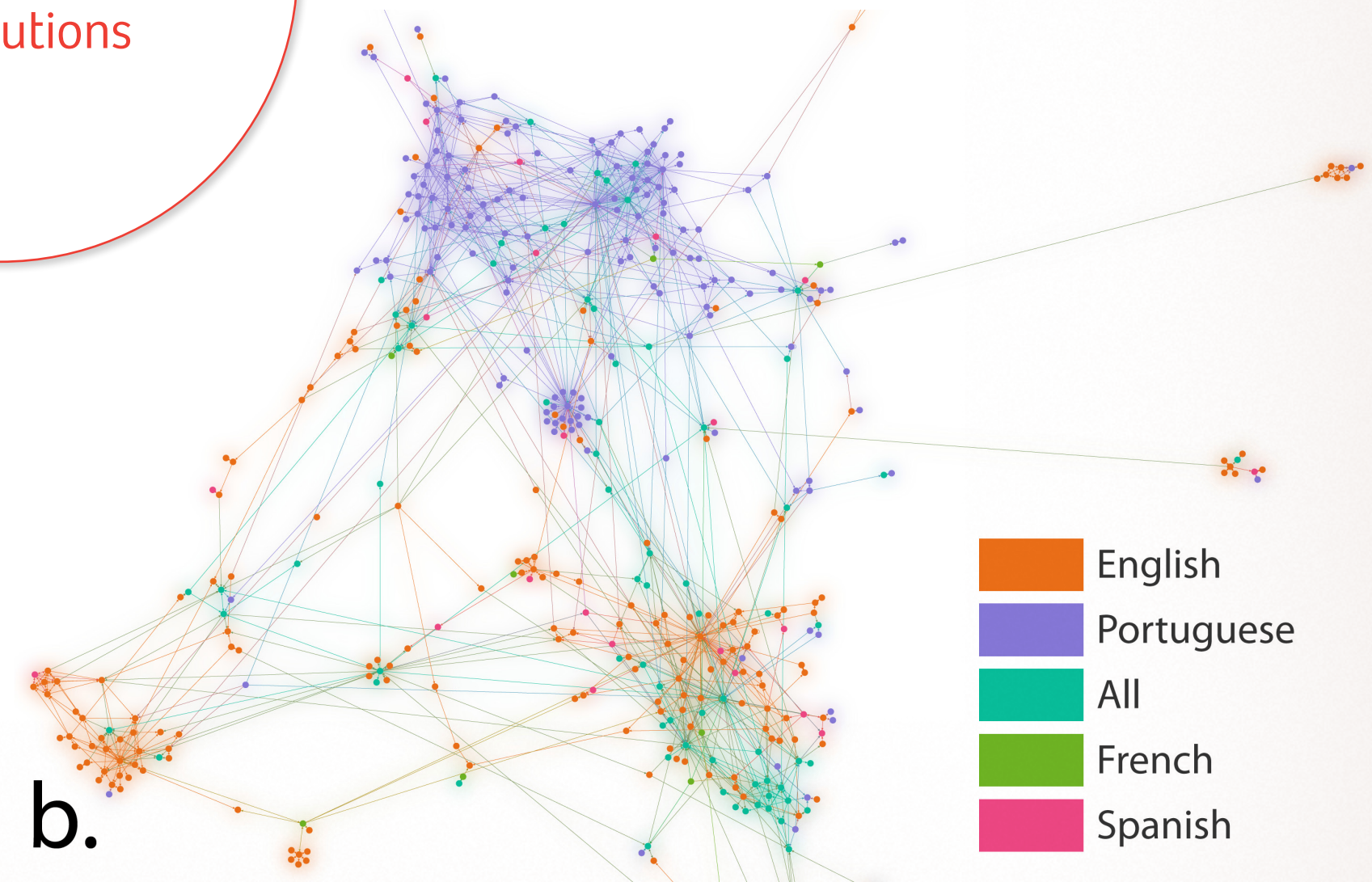
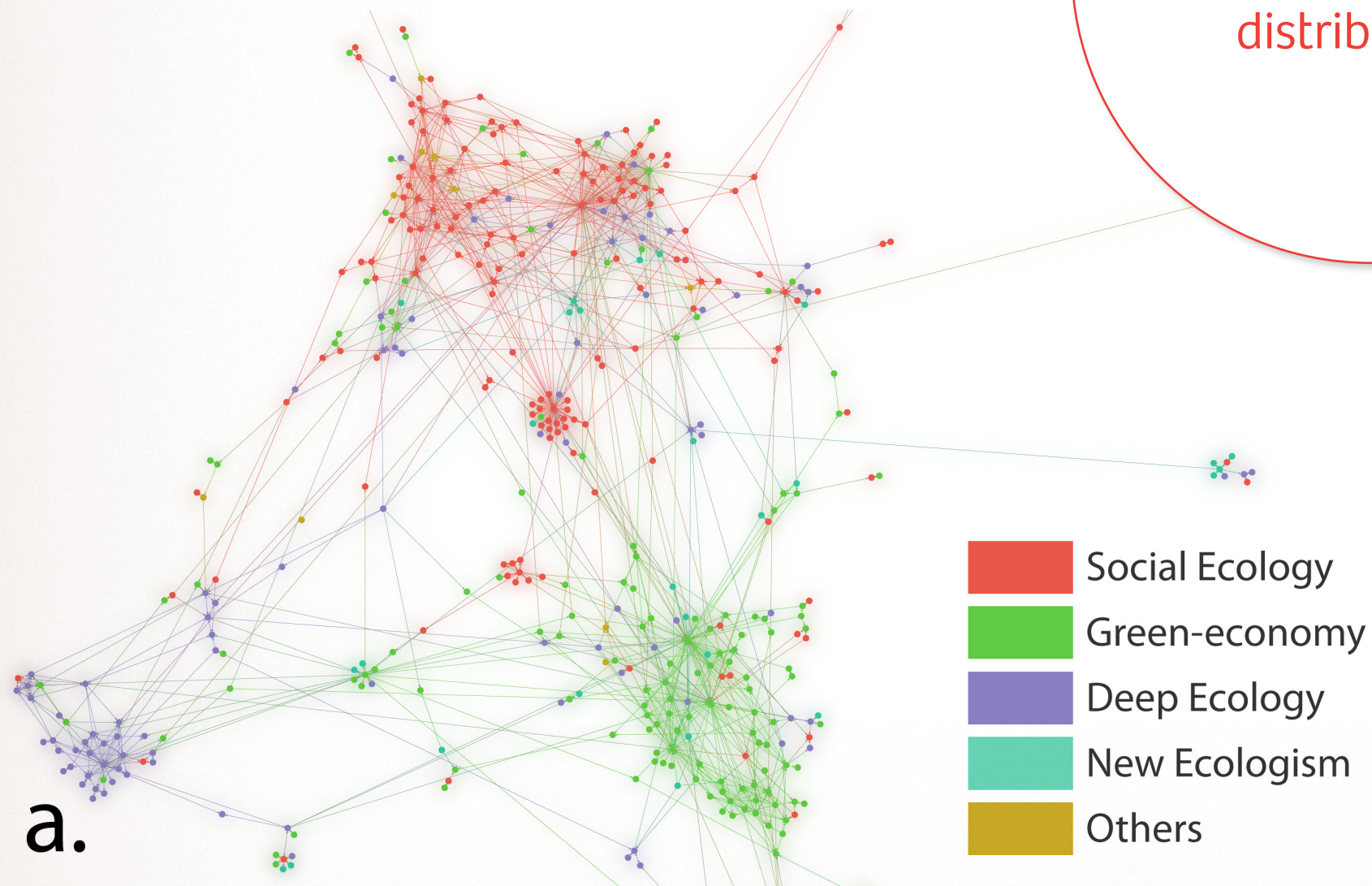


Nodes with top outdegree:

1. **INGA.org.br** - 47 outbound links
2. **CupulaDosPovos.org.br** - 31 outbound links
3. **RuralPovertyPortal.org** - 26 outbound links
4. **OECD.com.br** - 23 outbound links
5. **AdVivo.com.br** - 23 outbound links
6. **RealClimate.org** - 22 outbound links
7. **IUCNWorldConservationCongress.org** - 22 o. links
8. **GreenEconomyMyCoalition.org** - 21 outbound links
9. **UNEMG.org** - 18 outbound links
10. **forumbr163.blogspot.fr** - 16 outbound links

Distribution of categories

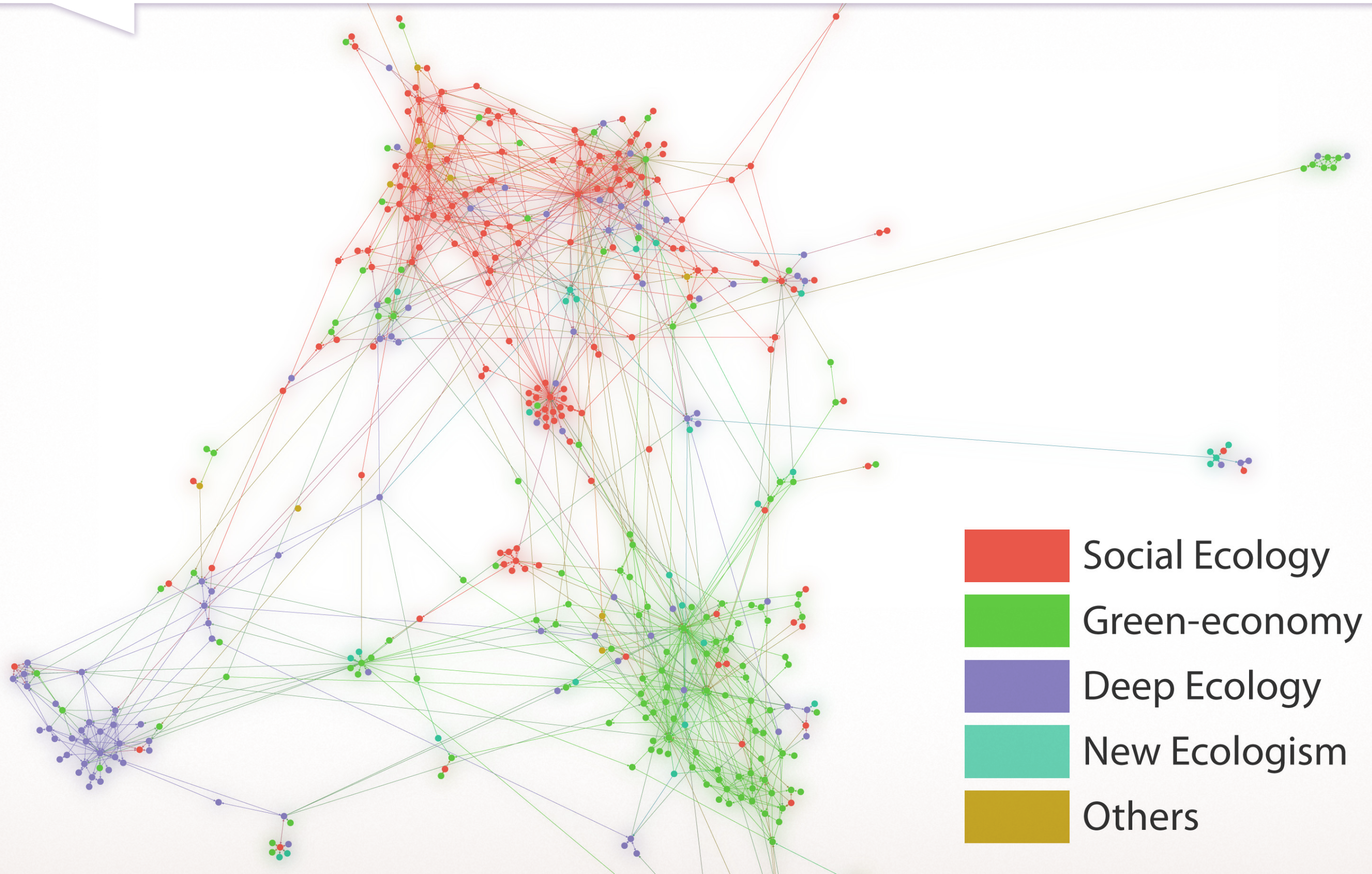
Different categories
give different
distributions



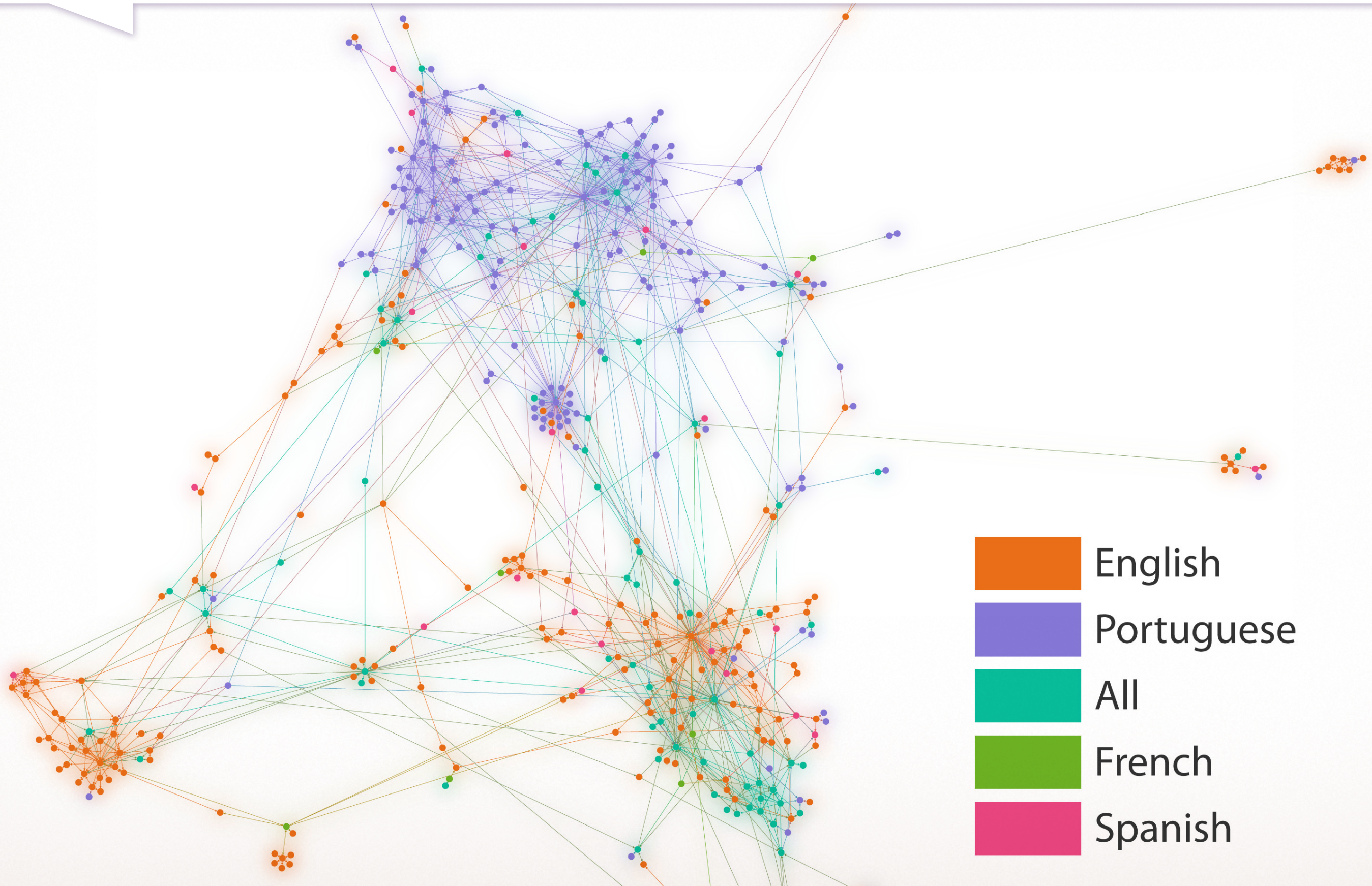
Approche de l'écologie

Langage

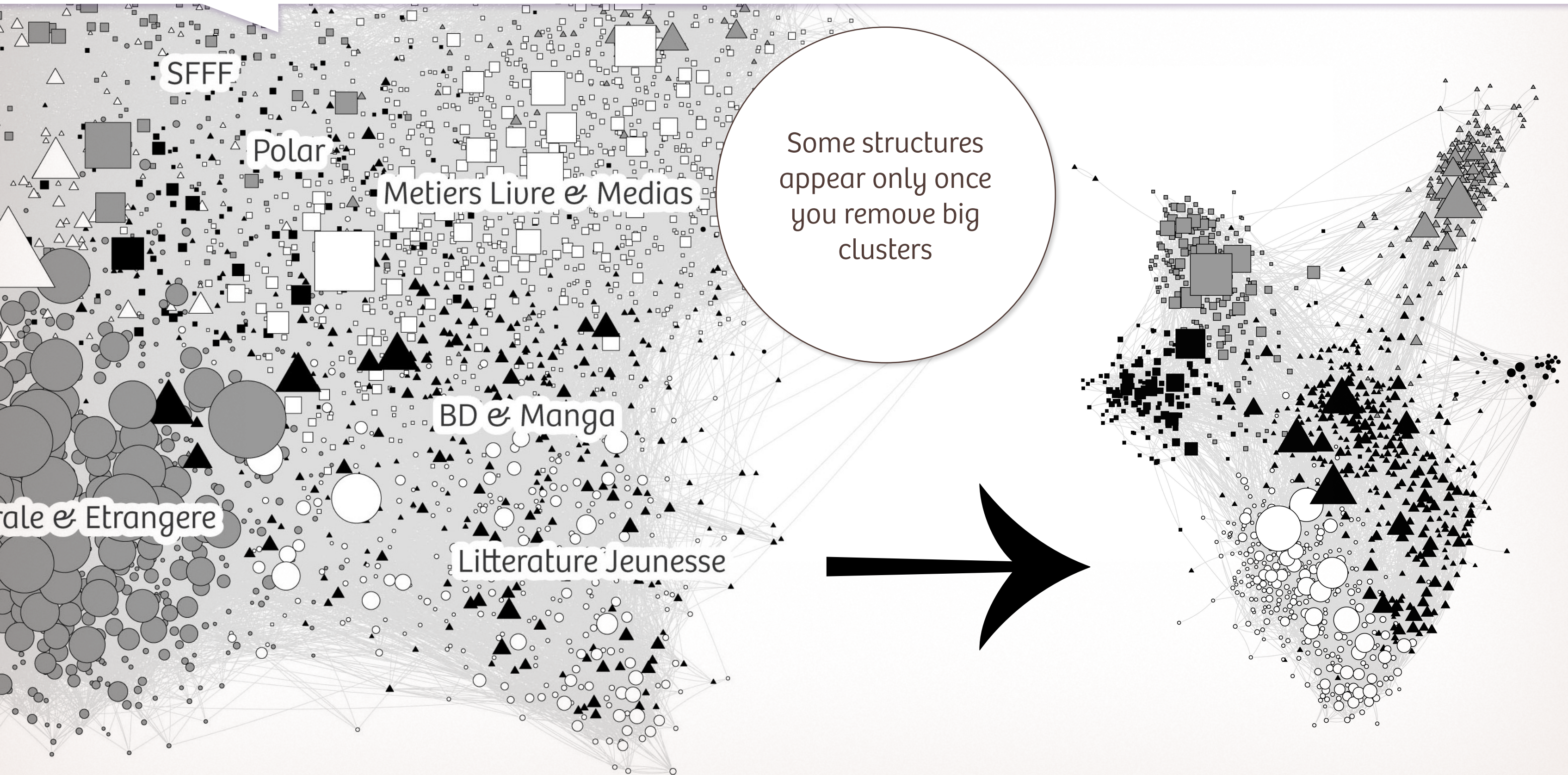
Distribution of categories



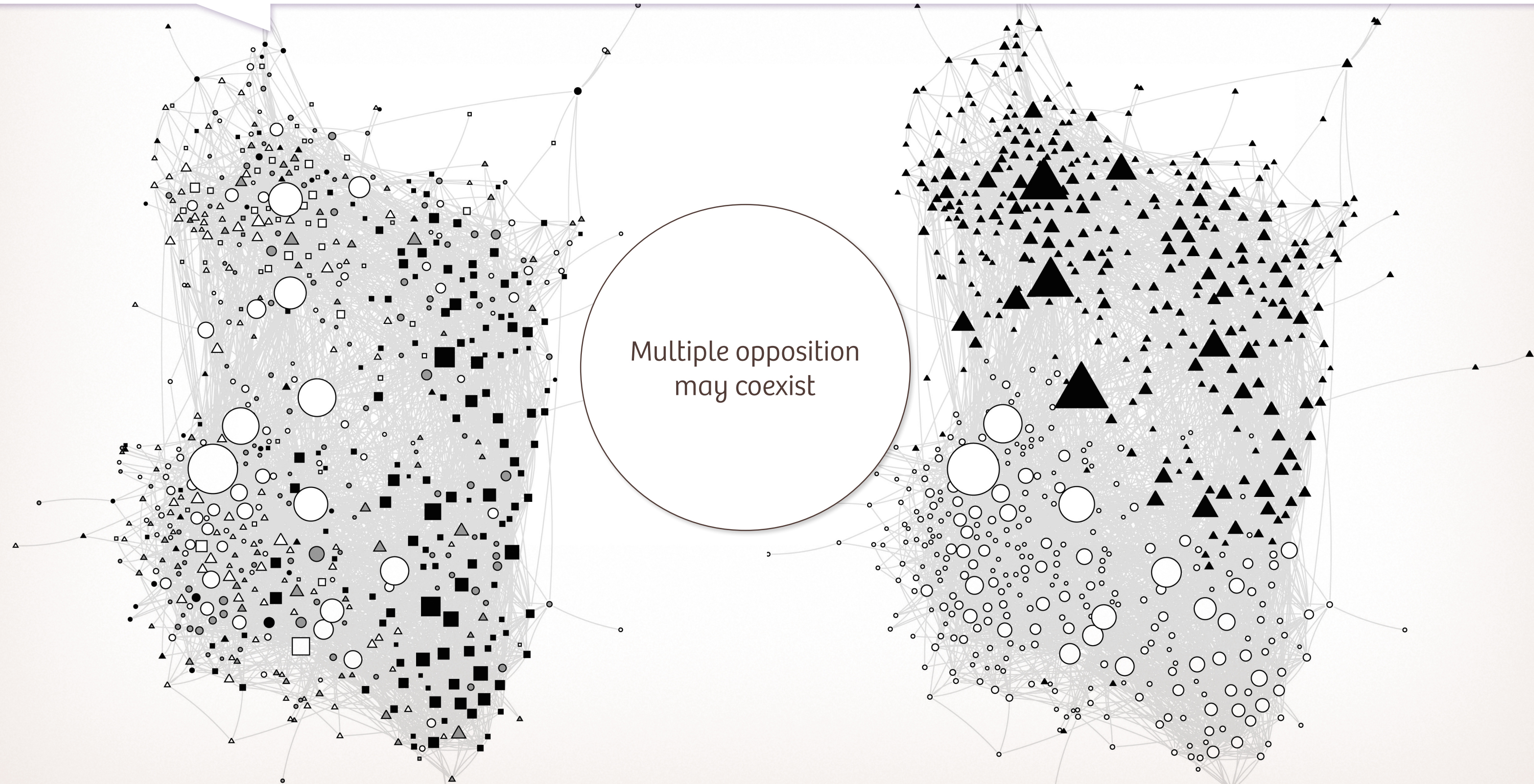
Distribution of categories



Removing clusters



Varying facets



Thanks for your attention

Mathieu.Jacomy@sciencespo.fr

SciencesPo
MÉDIALAB

<http://medialab.sciences-po.fr>