

Introducción al análisis de redes



Alejandro González
@outliers_es
www.outliers.es



Contenidos

Grafos y métricas

Redes sociales y perspectivas humanas sobre el SNA

Introducción al análisis de redes con Gephi. Ejercicios.

Grafos y métricas

¿Qué es un grafo?

Un conjunto de nodos y un conjunto de aristas

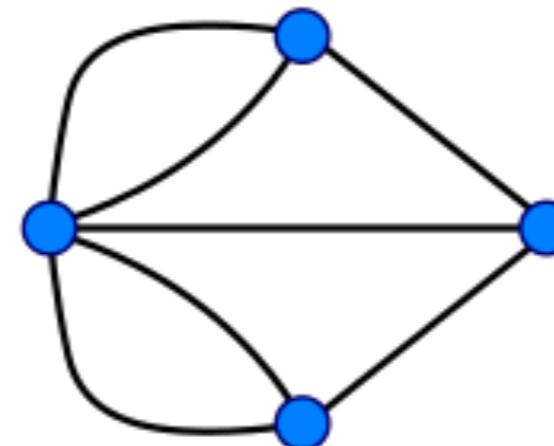
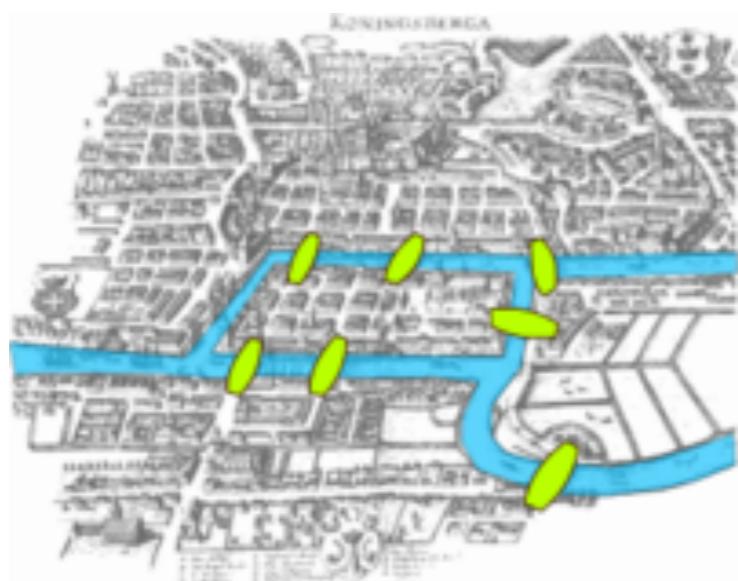
Es un modelo que podemos aplicar a multitud de problemas

Los nodos y aristas pueden representar lo que queramos

Grafos y métricas

Inicios

El primer uso conocido es el caso de los puentes de Königsberg

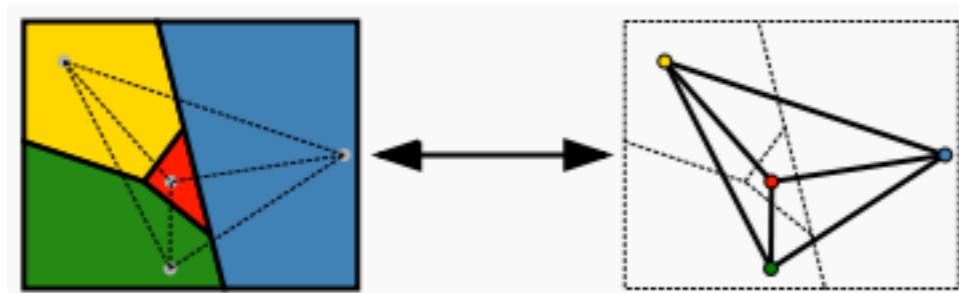


Grafos y métricas

Inicios

La primera aplicación en ingeniería la realizó
Gustav Kirchhoff

El problema de los 4 colores es otro ej
conocido



Grafos y métricas

Tipos de grafos

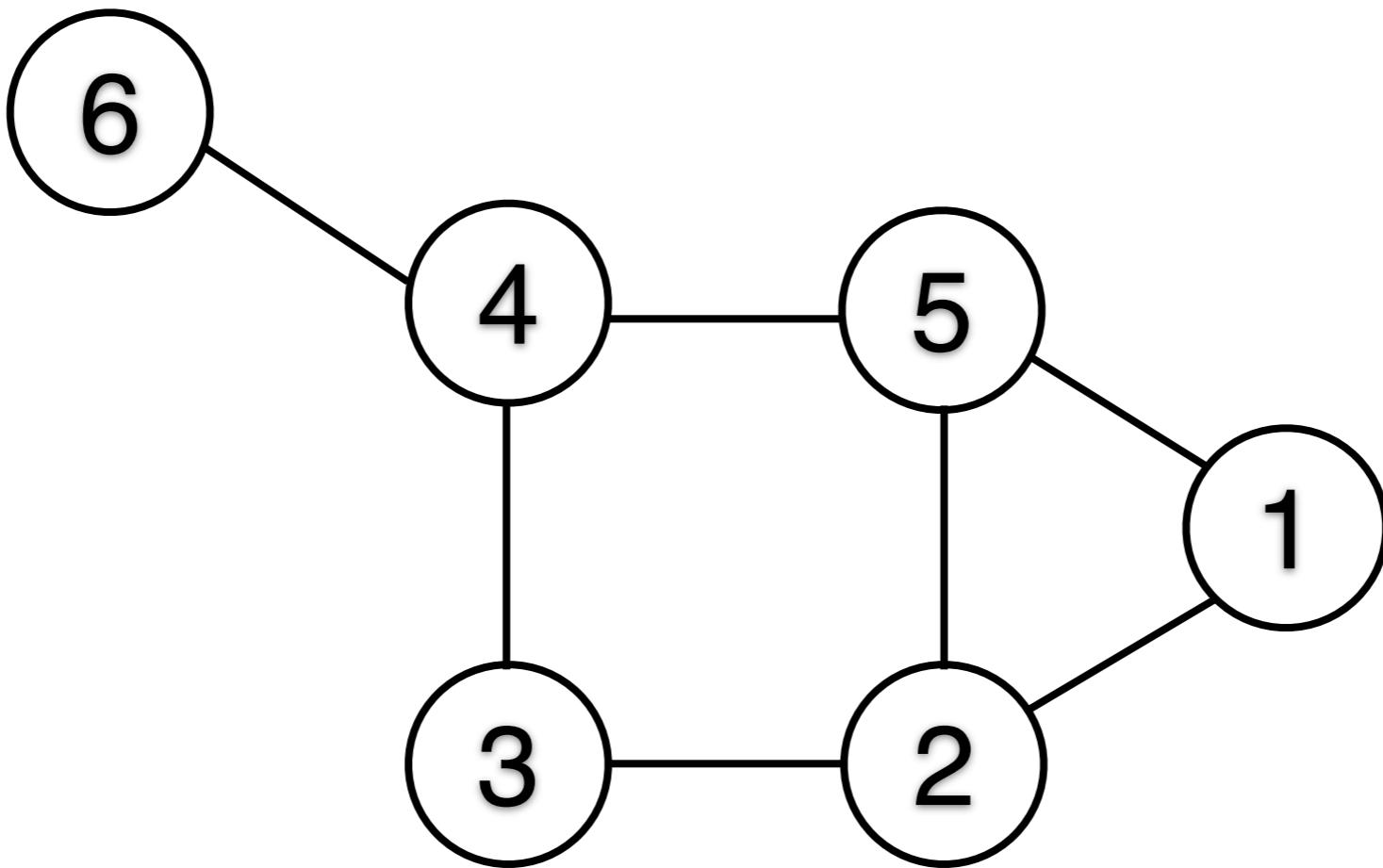
Simple o multigrafo

Dirigido o NoDirigido

Grafo denso, disperso o libre de escala

.....

Grafos y métricas



$$G = (V, E)$$

$$V = \{1, 2, 3, 4, 5, 6\}$$

$$E = \{(1, 2), (1, 5), (2, 5), (2, 3), (3, 4), (5, 4), (4, 6)\}$$

Grafos y métricas

Distancia entre dos nodos: menor número de aristas de un recorrido entre dos nodos v_1 y v_2

Diámetro de la red: máxima distancia entre dos nodos cualquiera de la red

$$\text{Densidad del grafo} = \frac{|E|}{|V|(|V| - 1)}$$

Grafos y métricas

Camino medio (Average Path Length): Media de pasos en caminos cortos para todos los pares de nodos

Coeficiente de Clustering (Clustering Coefficient): Mide el grado de conexión de los nodos

Modularidad (Modularity): Mide el tipo de conexión de los nodos en grupos o módulos

Grafos y métricas

Centralidad de grado (Degree Centrality): número de enlaces que posee un nodo con los demás

Centralidad de cercanía (Closeness Centrality): asume que la propagación de información siempre se da en la red a través del camino más corto

Centralidad de intermediación (Betweenness Centrality): cuantifica la frecuencia o el número de veces que un nodo actúa como un puente a lo largo del camino más corto entre otros dos nodos

Centralidad de vector propio (Eigenvector Centrality): mide la influencia de un nodo en una red

Aplicaciones

RED DE LIKES (FACEBOOK)

TWITTER: INFLUENCIADORES

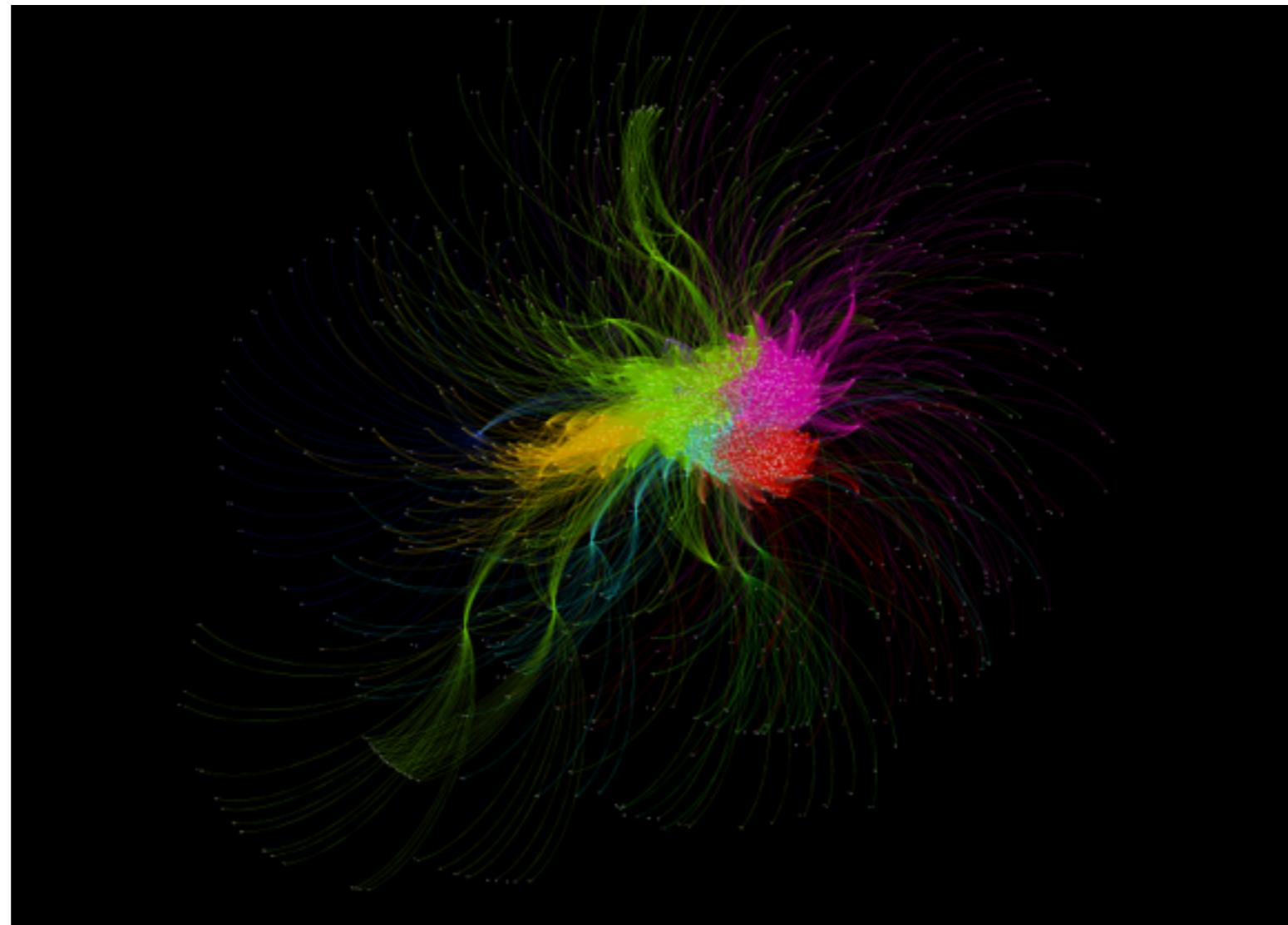
TWITTER: TIEMPO REAL

MARCAS: ASOCIACIONES

MARCAS: PERFILES

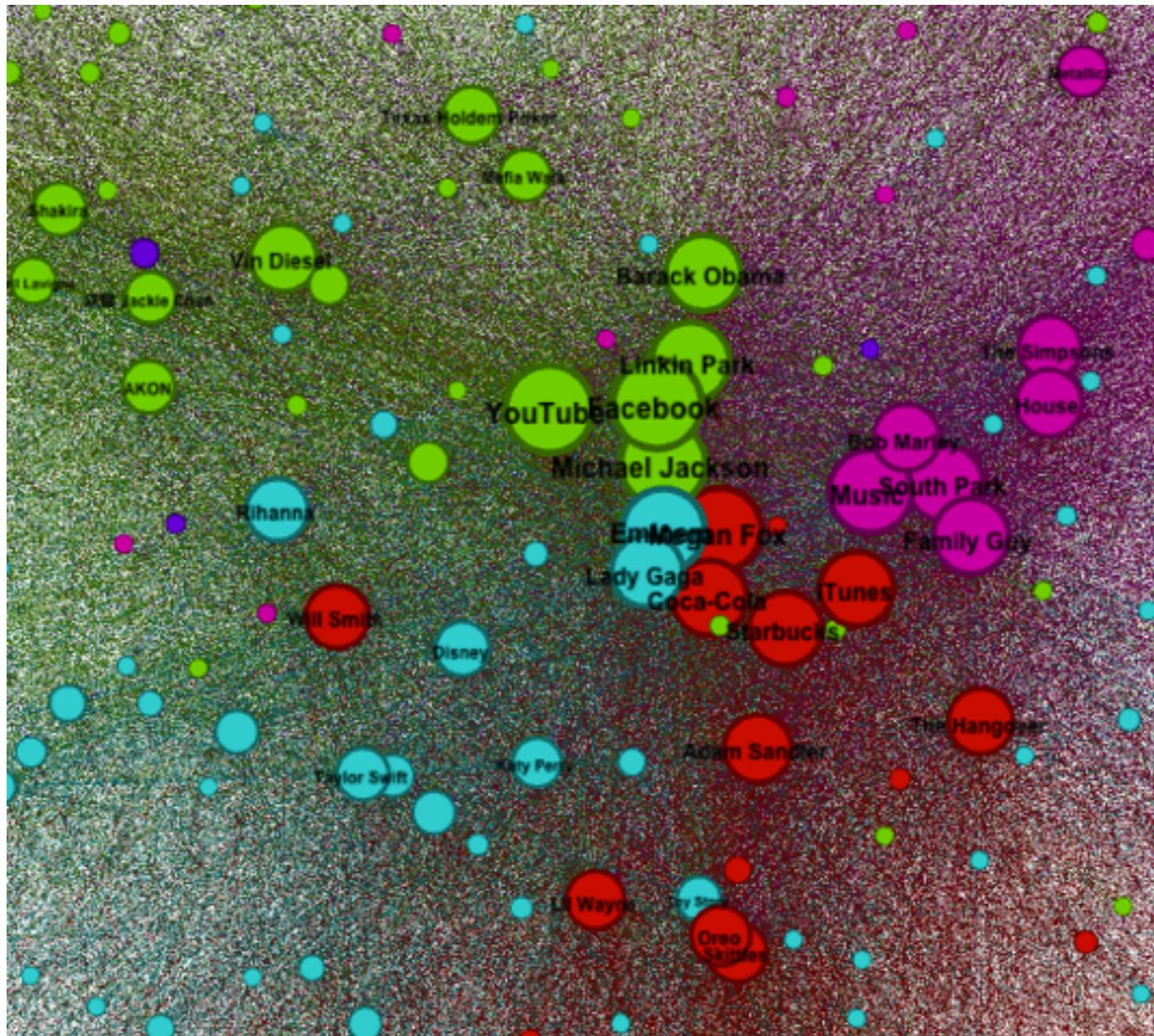
Aplicaciones

Redes de likes



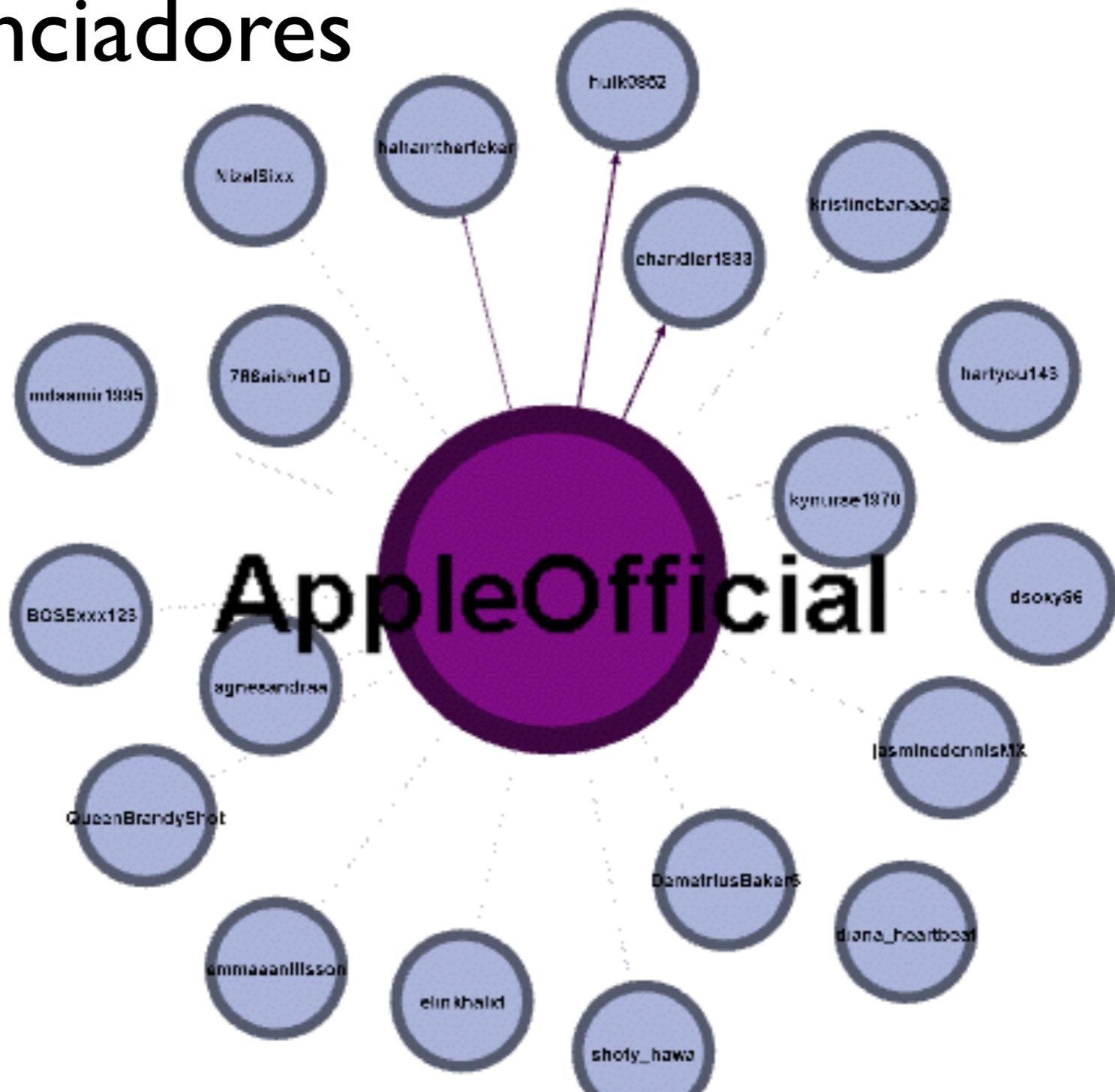
Aplicaciones

Redes de likes



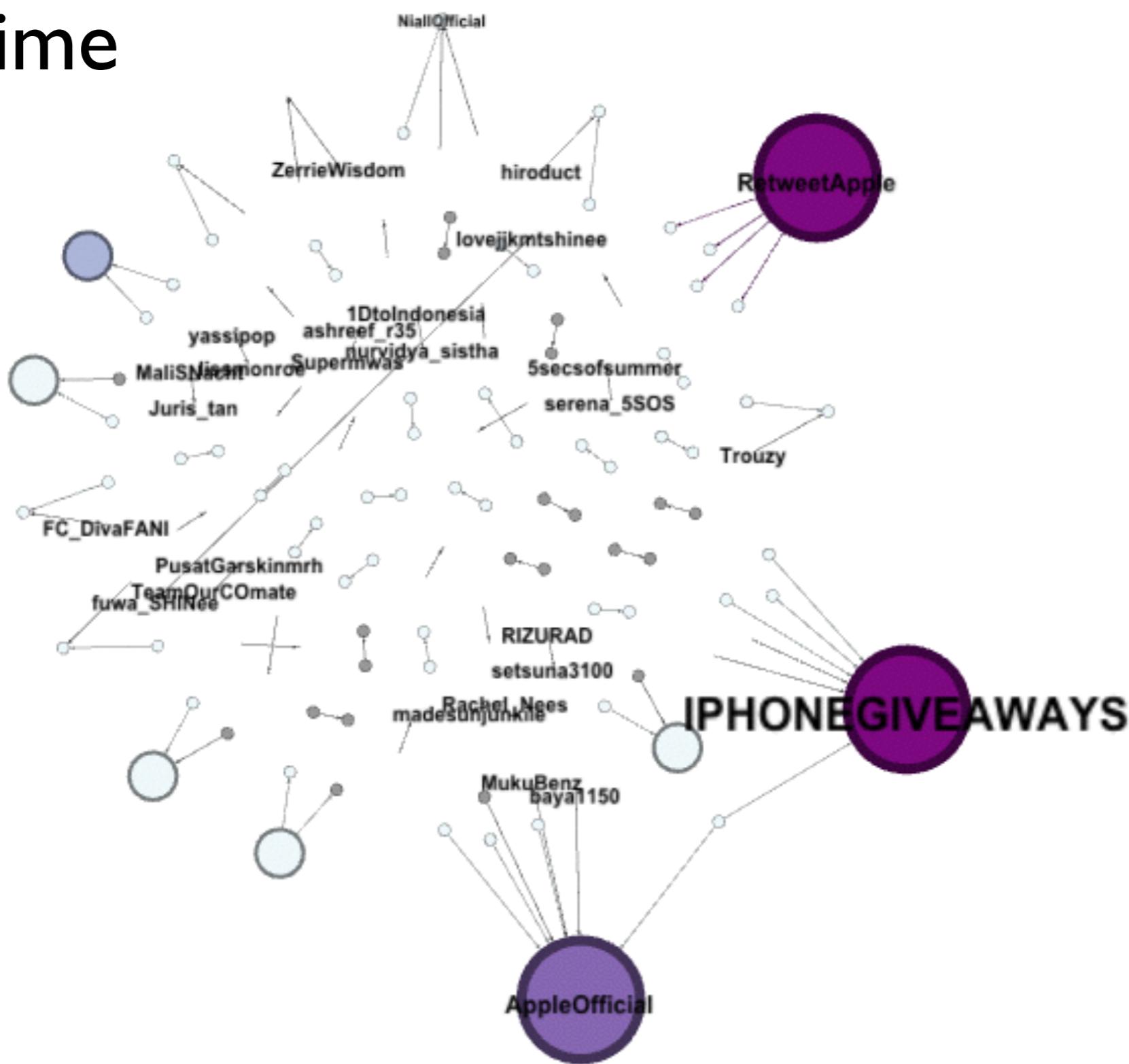
Aplicaciones

Influenciadores



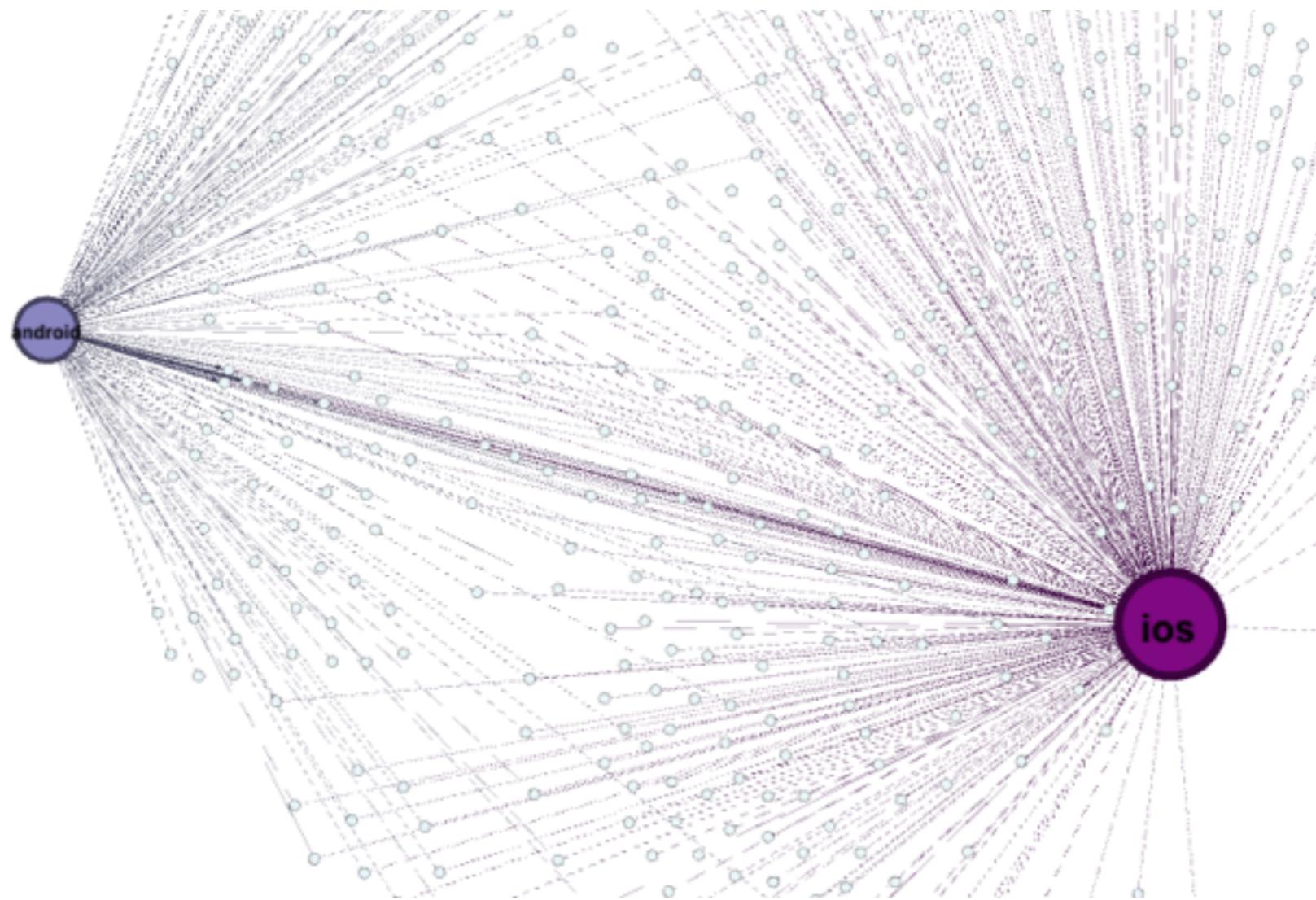
Aplicaciones

RealTime



Aplicaciones

Brand Map



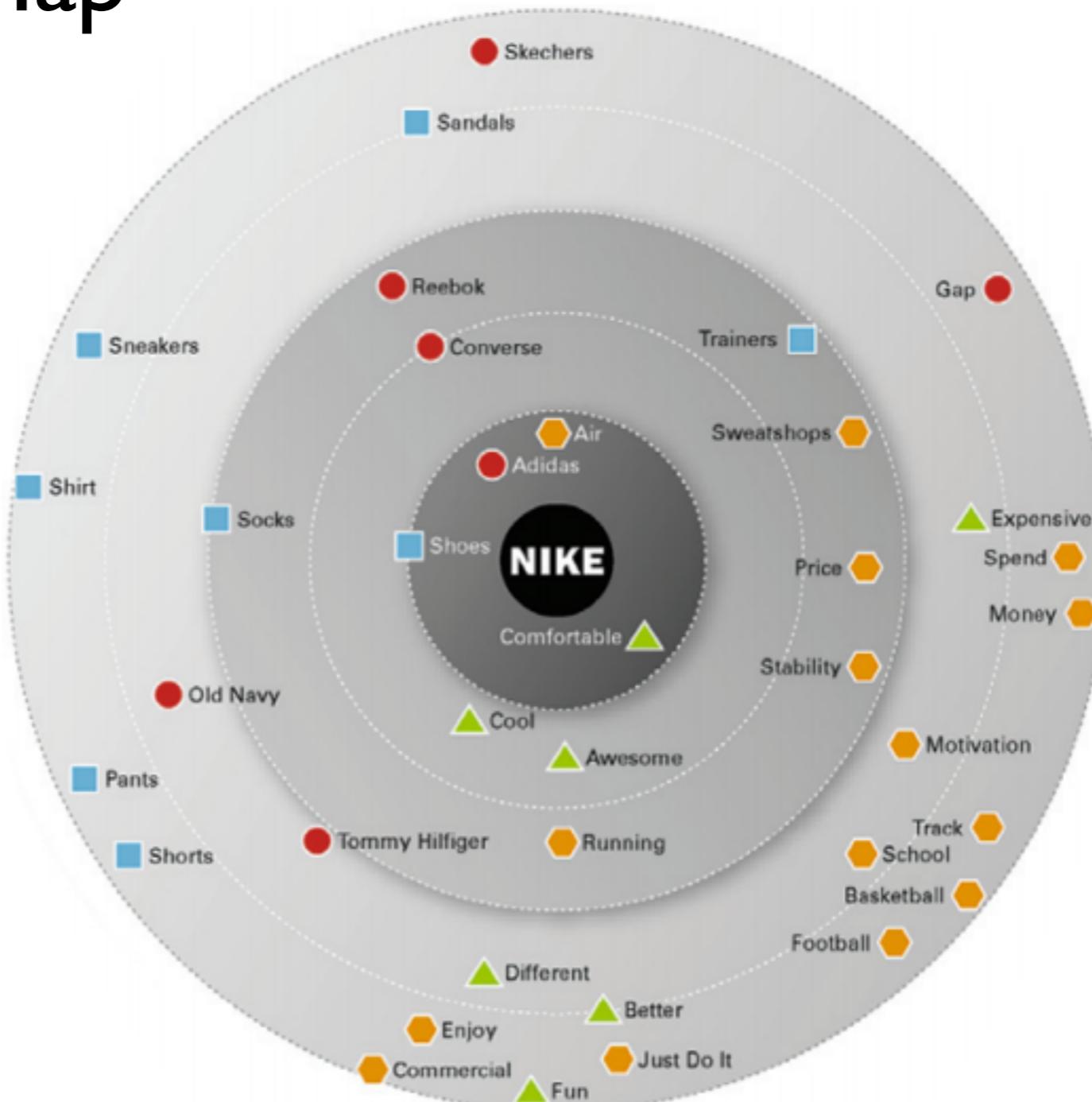
Aplicaciones

Brand Map



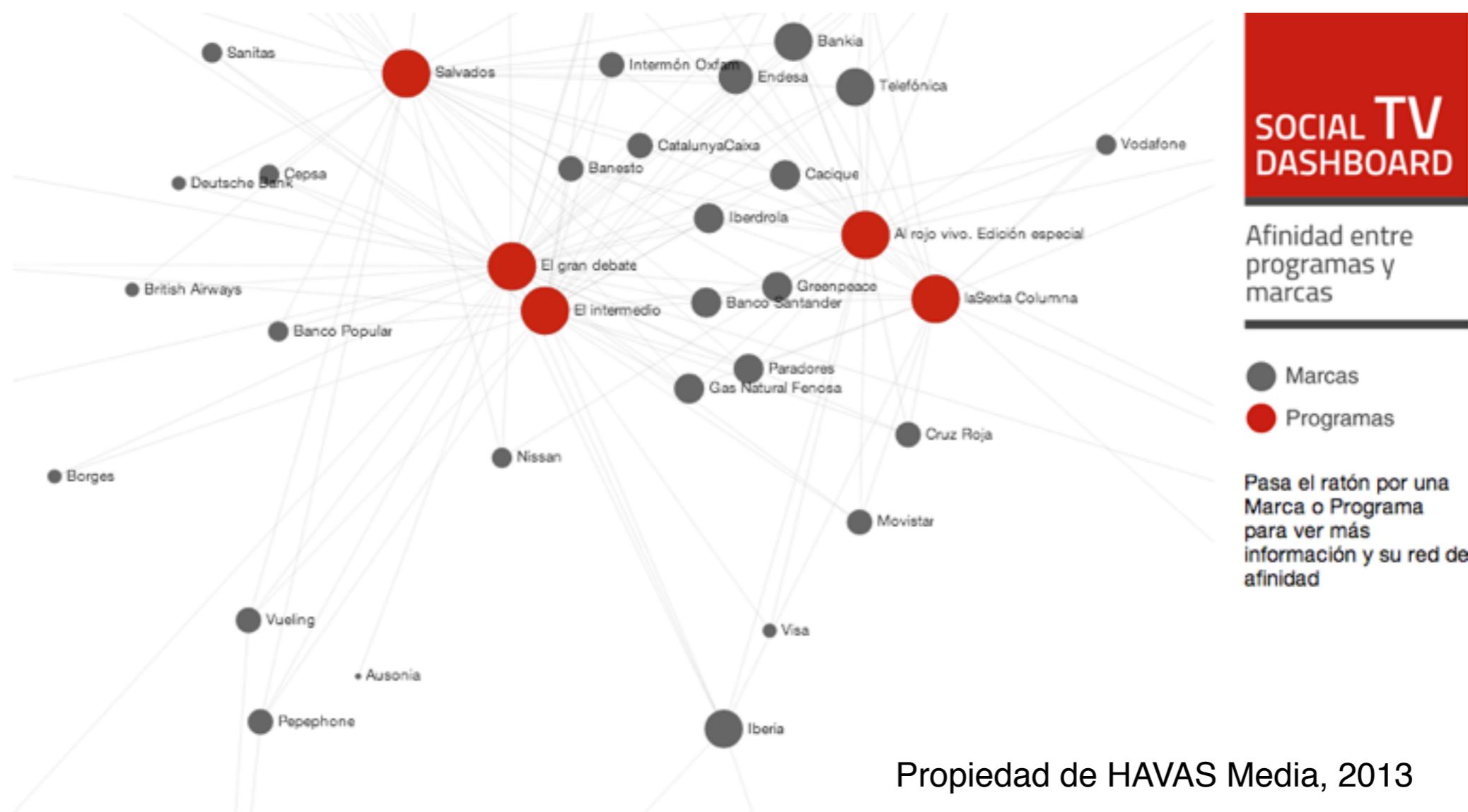
Aplicaciones

Brand Map



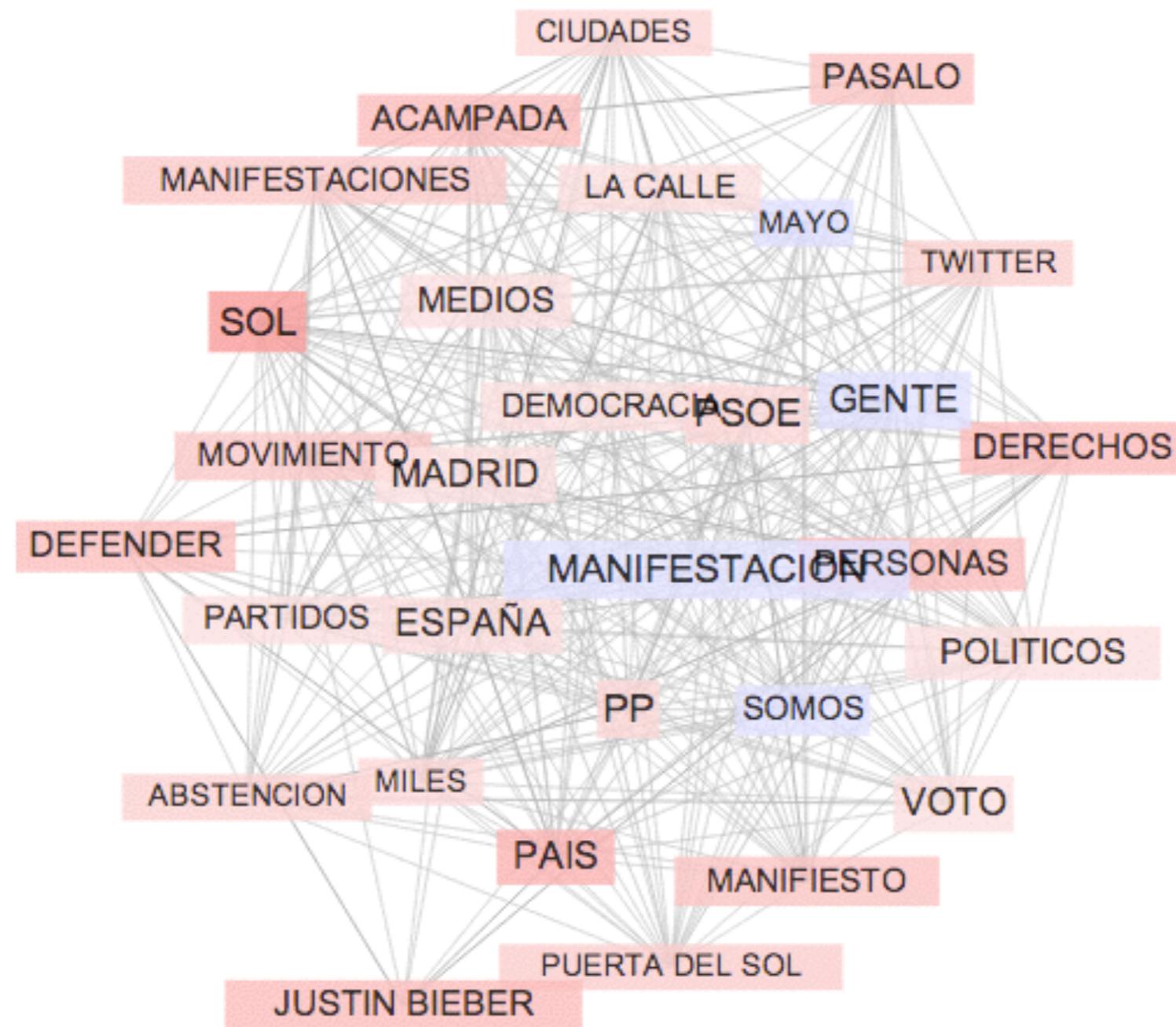
Aplicaciones

Afinidad basada en Perfiles



Aplicaciones

Cohesión de vocabulario



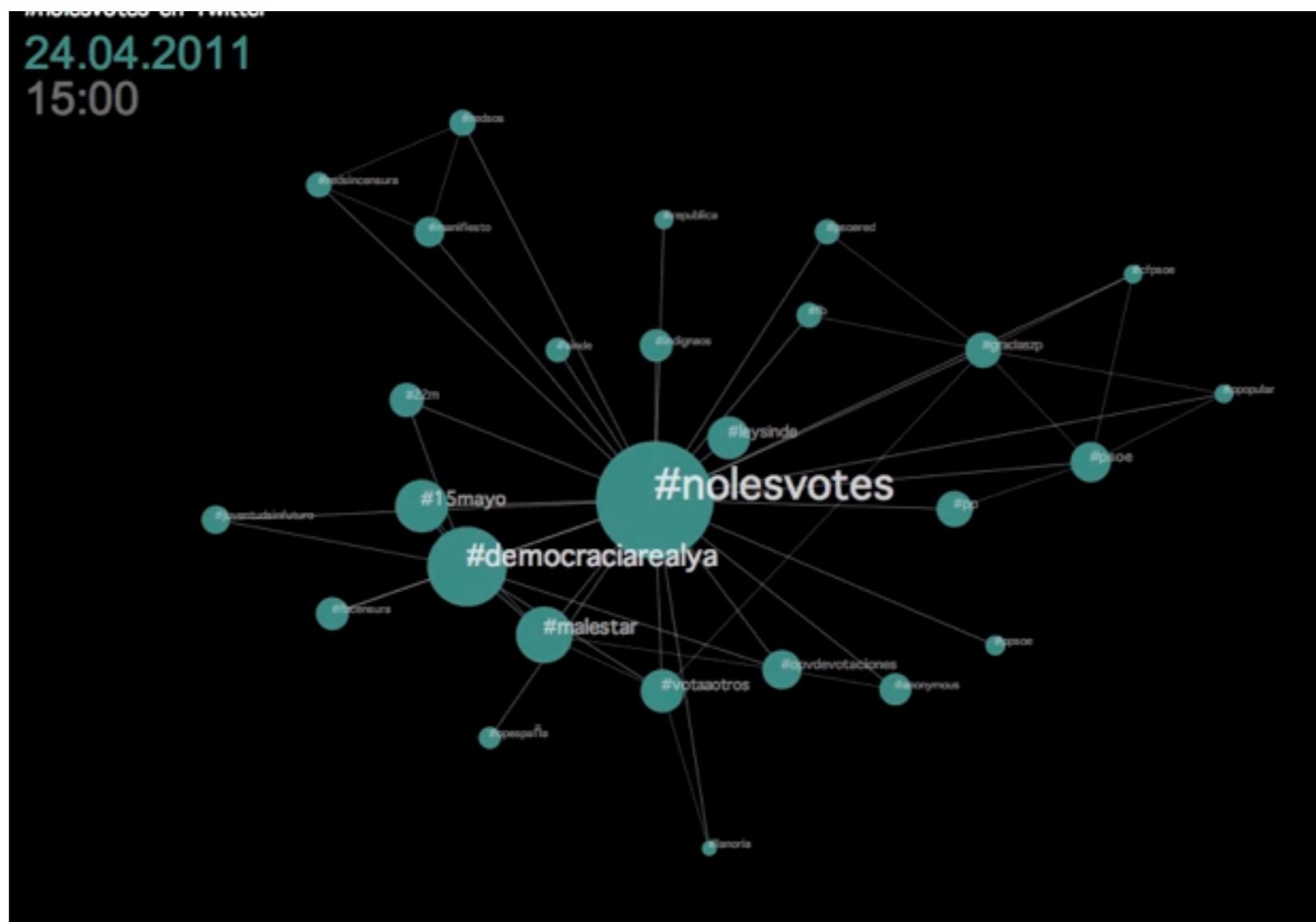
Aplicaciones

Redes de Hashtags



Aplicaciones

Redes de Hashtags



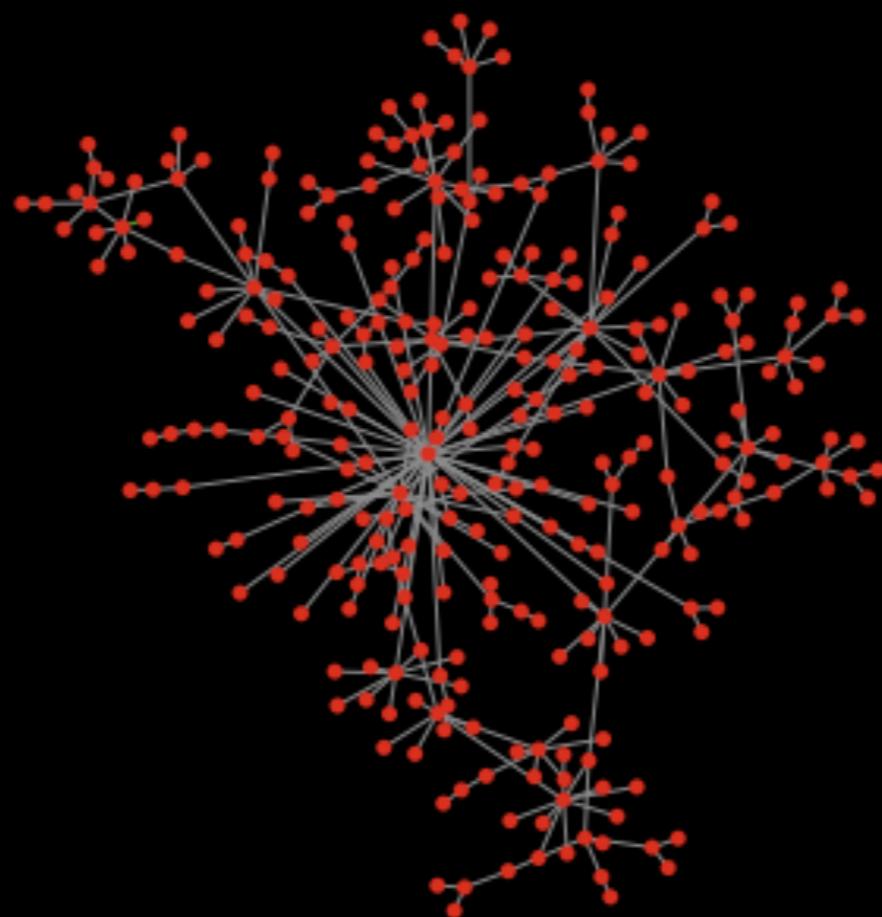
Aplicaciones

The Internet Map



Redes Sociales

Leyes



"Preferential Attachment"

Redes Sociales

Conexión Preferencial

- Los nodos que tienen mayor número de conexiones tienen candidatura más alta a recibir conexiones
- Se generan redes de ‘mundo pequeño’
- Ejemplos:
 - Pagerank
 - Followers
 - Riqueza

Redes Sociales

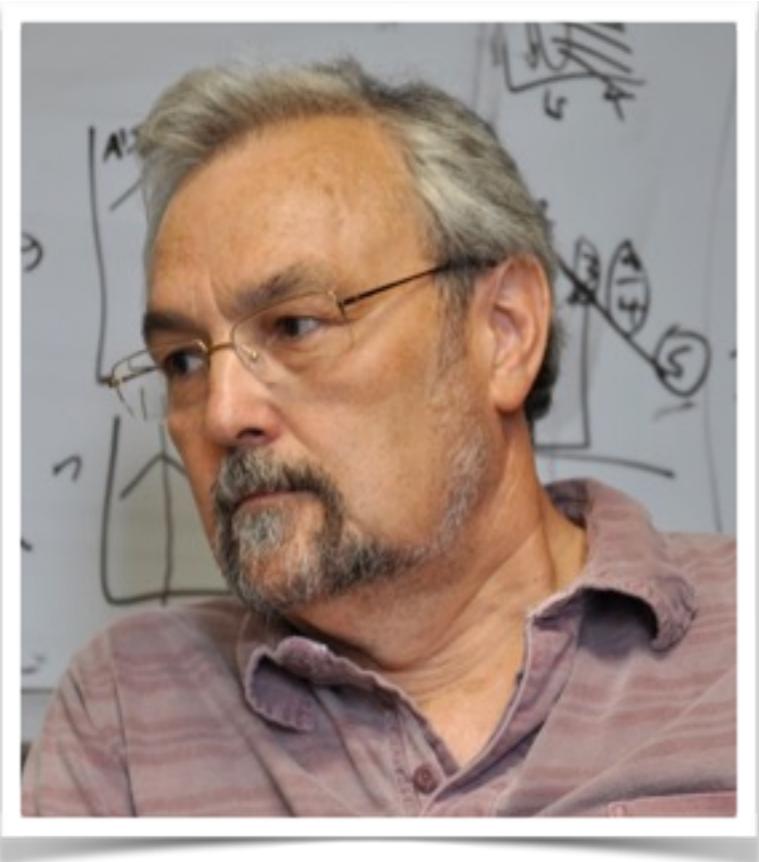
Ley de Metcalf

- El valor de una red aumenta proporcionalmente al cuadrado del número de usuarios



Redes Sociales

El número de Dunbar



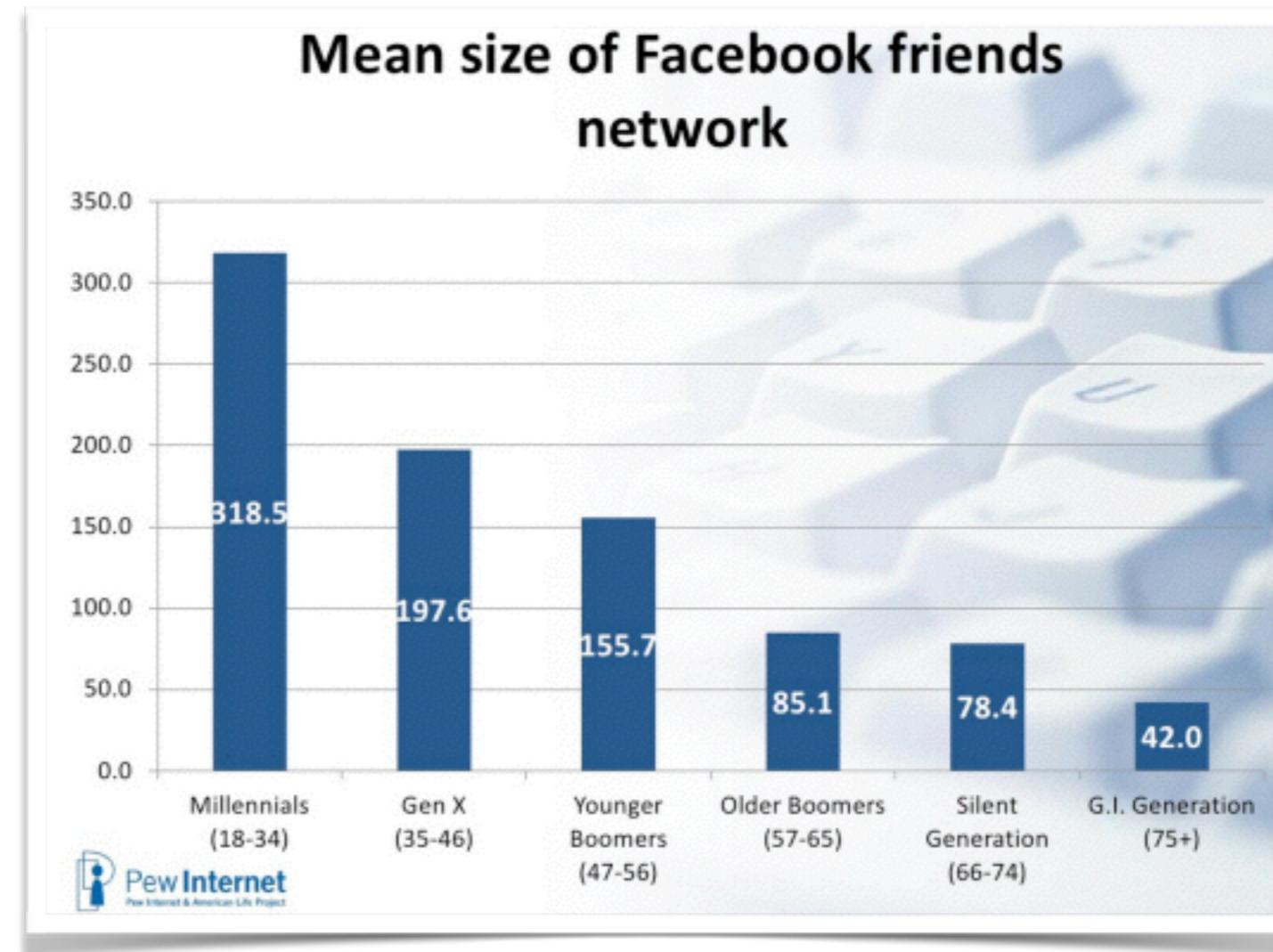
El límite cognitivo del número de personas con el que un individuo puede mantener lazos sociables estables está en torno a 150

http://www.cabdyn.ox.ac.uk/people_pages/complexity_people_dunbar.asp

http://es.wikipedia.org/wiki/N%C3%BAmero_de_Dunbar

Redes Sociales

El número de Dunbar

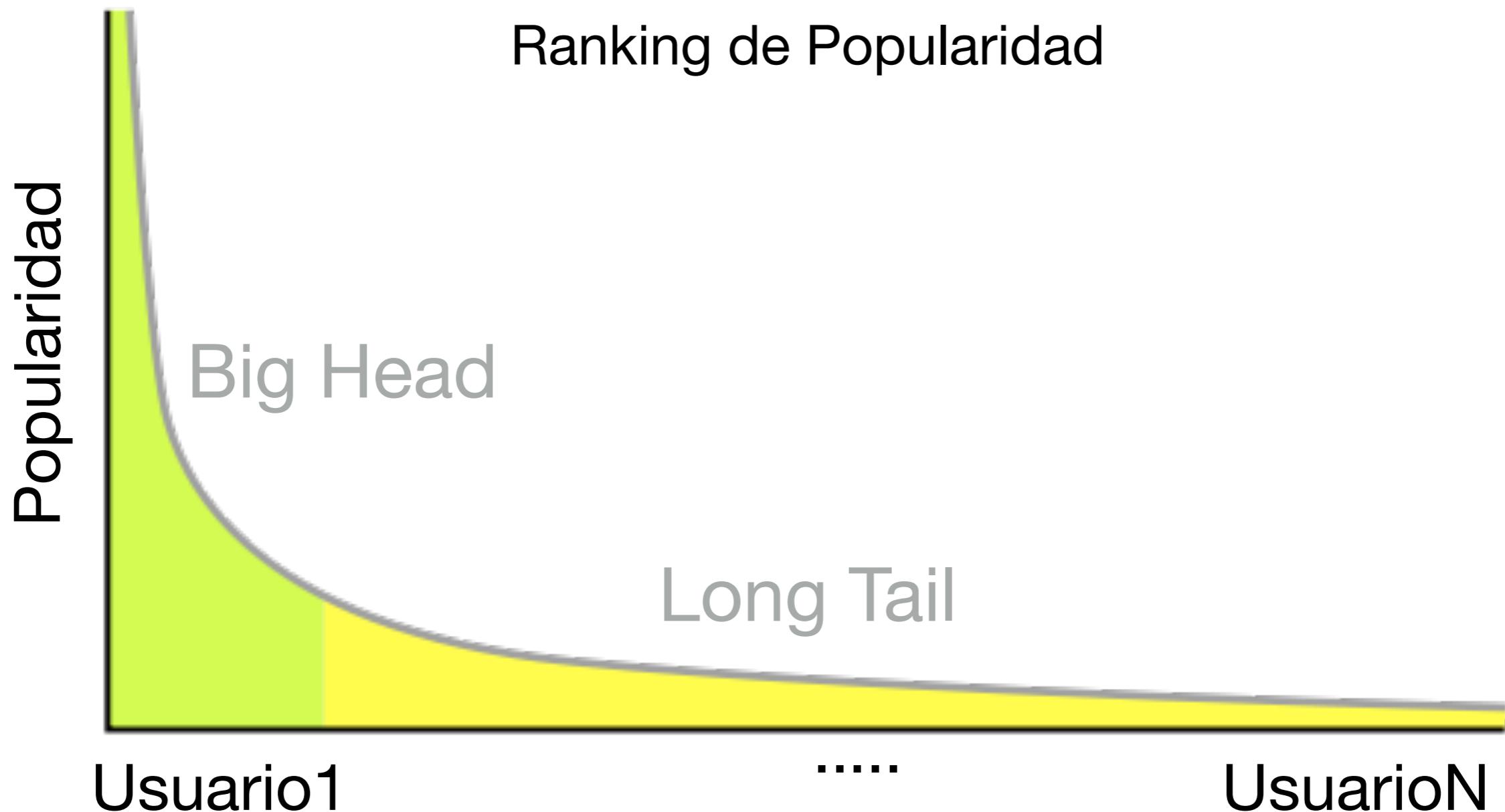


<http://pandawhale.com/convo/1634/whats-the-average-number-of-friends-a-facebook-user-has-barking-up-the-wrong-tree>

<http://www.wired.com/underwire/2012/03/dunbars-number-facebook/>

Redes Sociales

“Power Law”



Redes Sociales

“Long Tail”



Antes:

Pareto [80/20]:

Centrarse en el 20% de productos para 80% ventas

Ahora:

Dos mercados:

- Mercado de masas (hits)
- Mercados de nichos

Redes Sociales

“Long Tail”

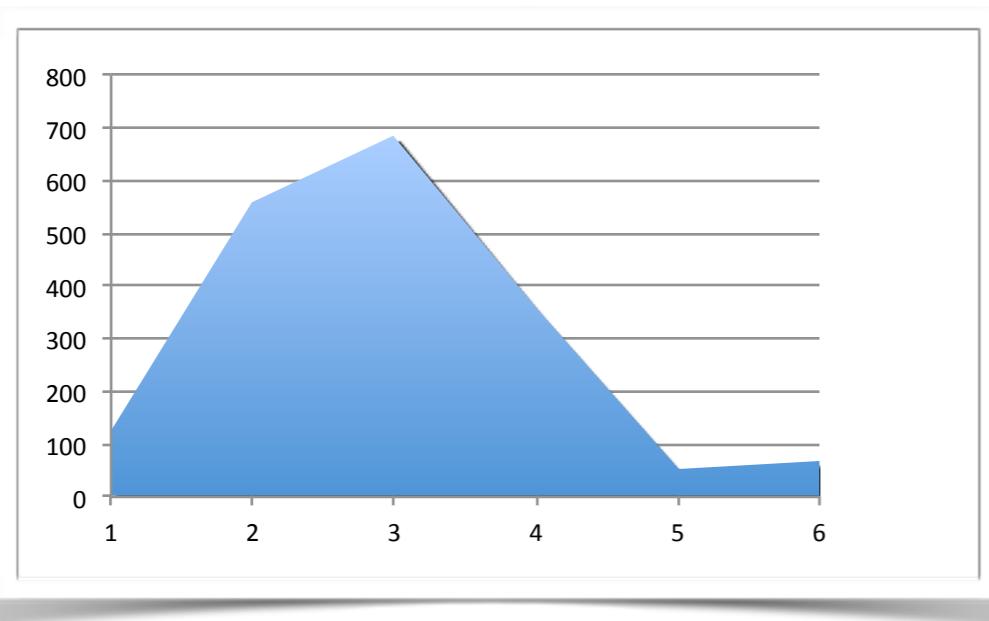
“Más del 50% de la venta de Amazon ocurre más allá de sus 130.000 “top titles”.”

Las tres fuerzas del Long Tail:

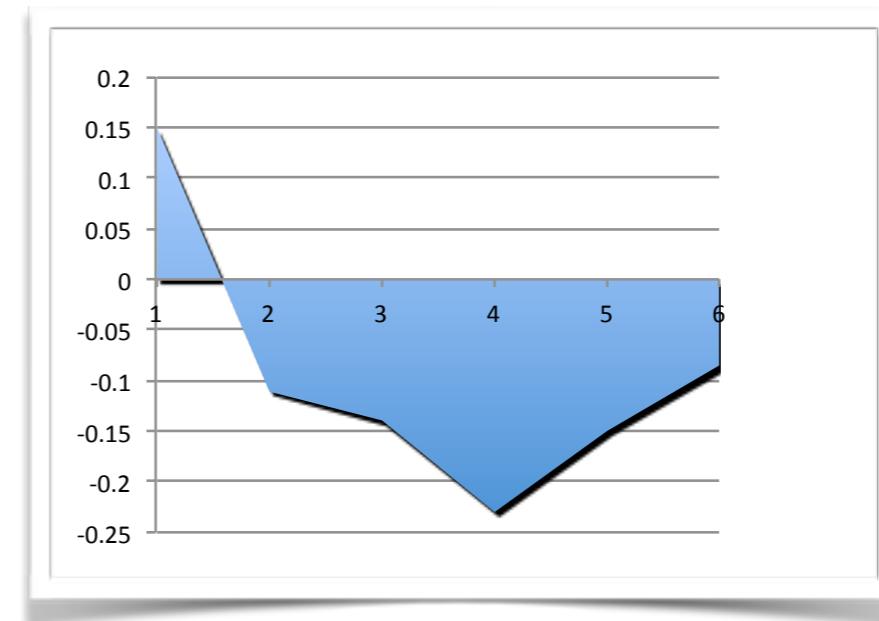
- La democratización de la producción.
- La democratización de la distribución. Ebay, You Tube, Amazon. “Todos somos retailers”.
- La conexión entre oferta y demanda.

Redes Sociales

“Long Tail”



Histograma de reputación-menciones



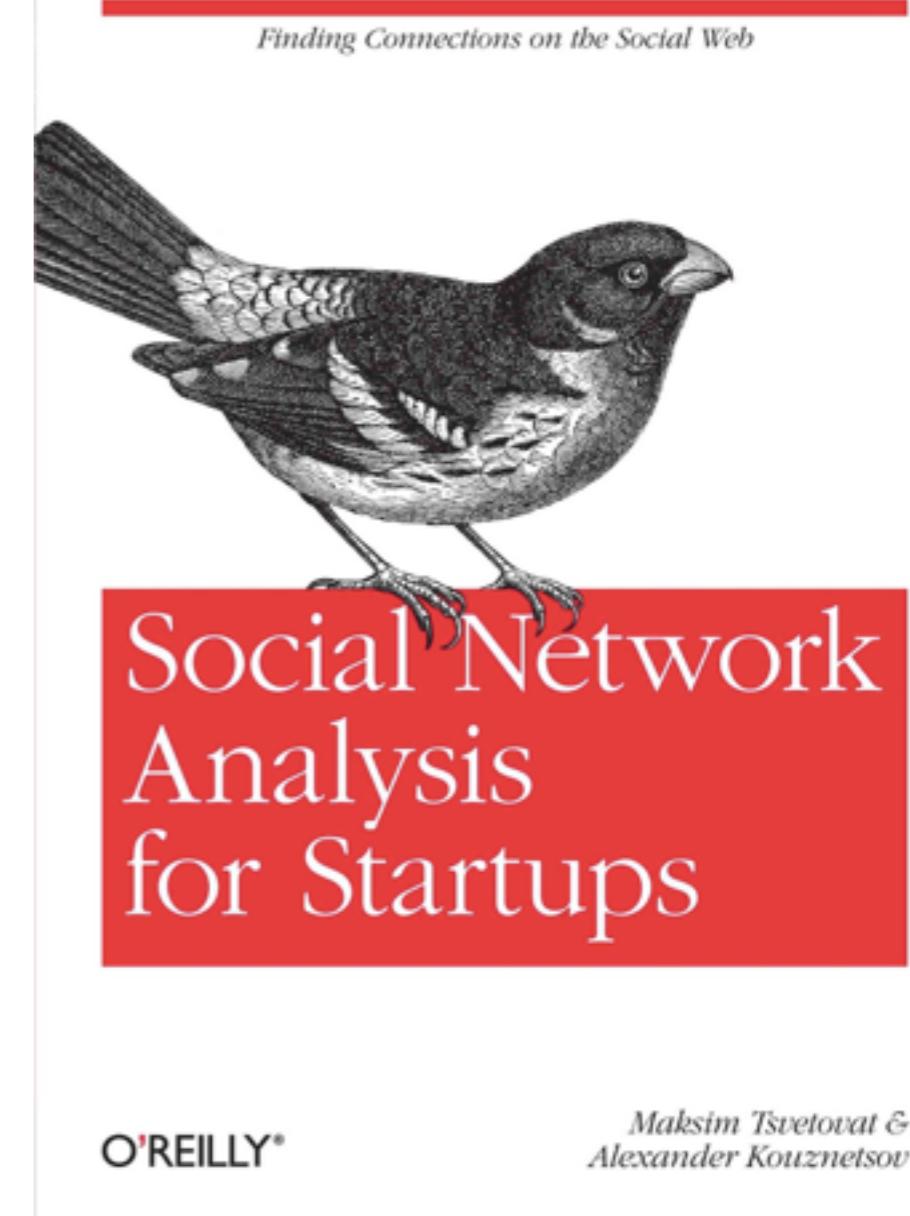
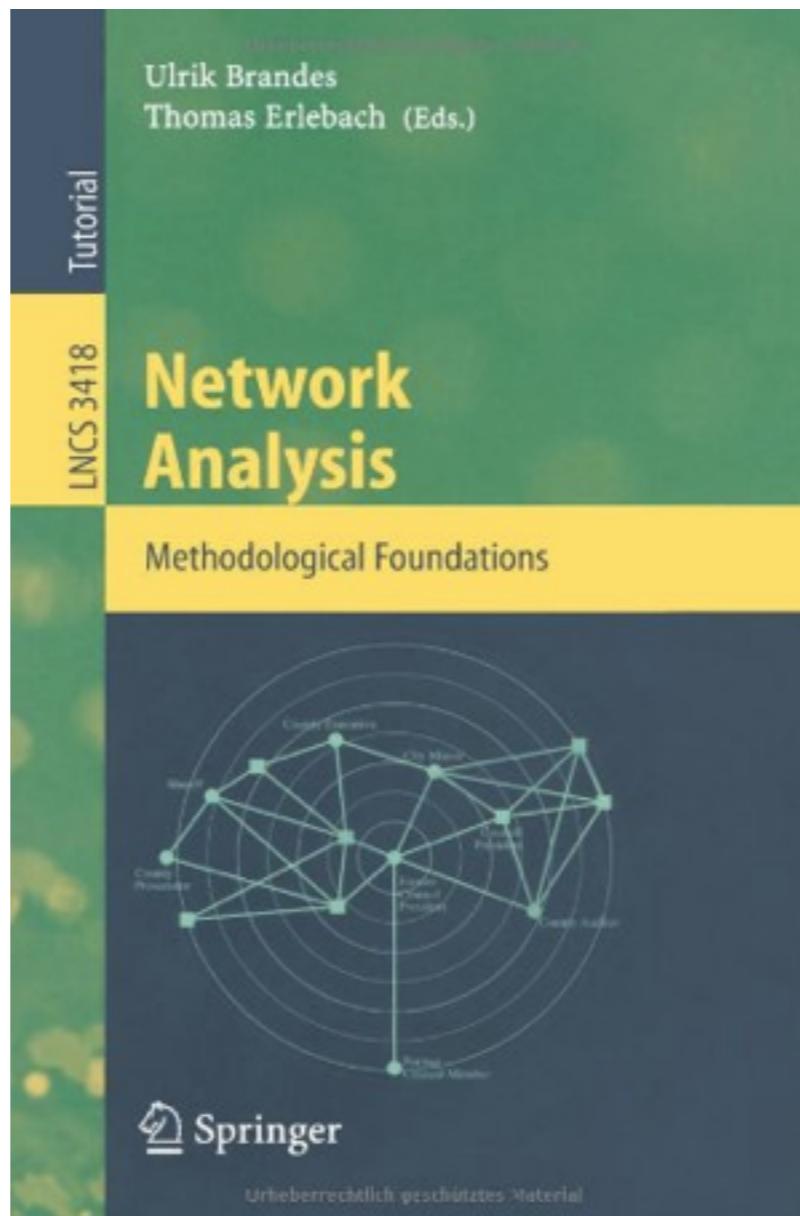
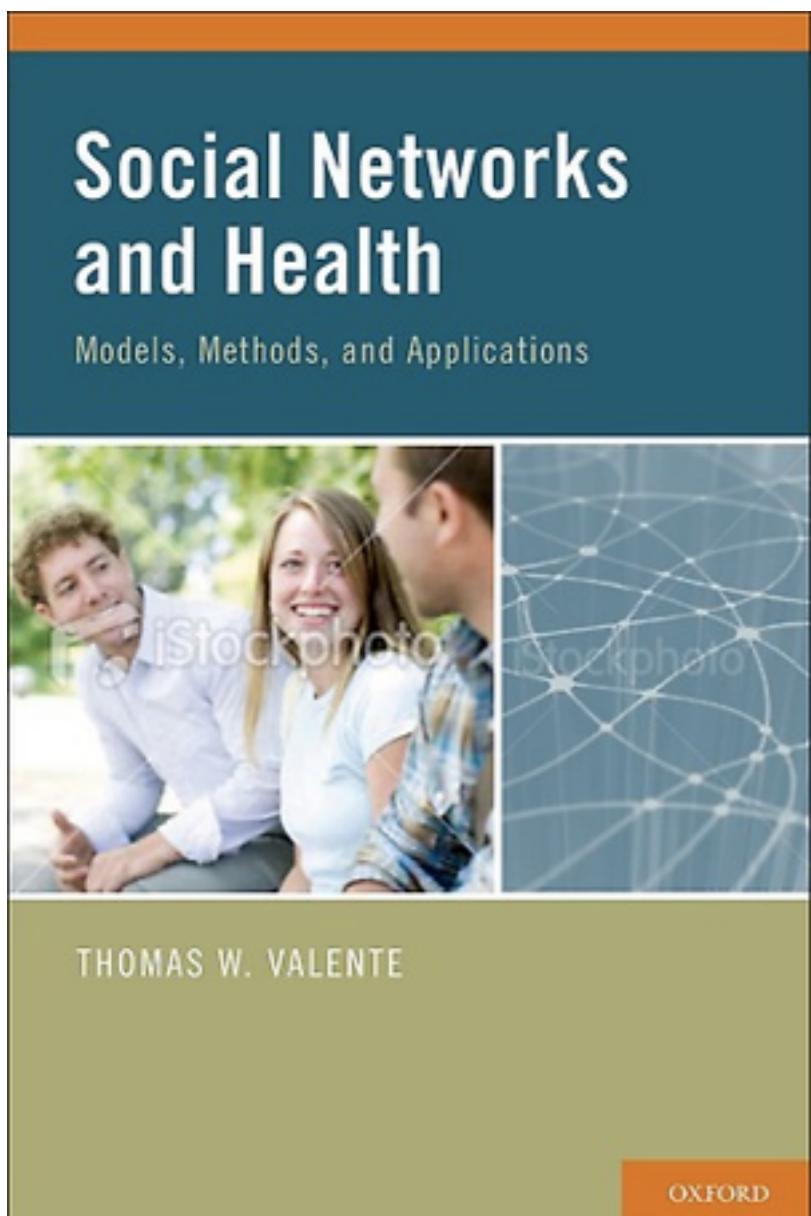
Histograma de reputación-valoración

Redes Sociales

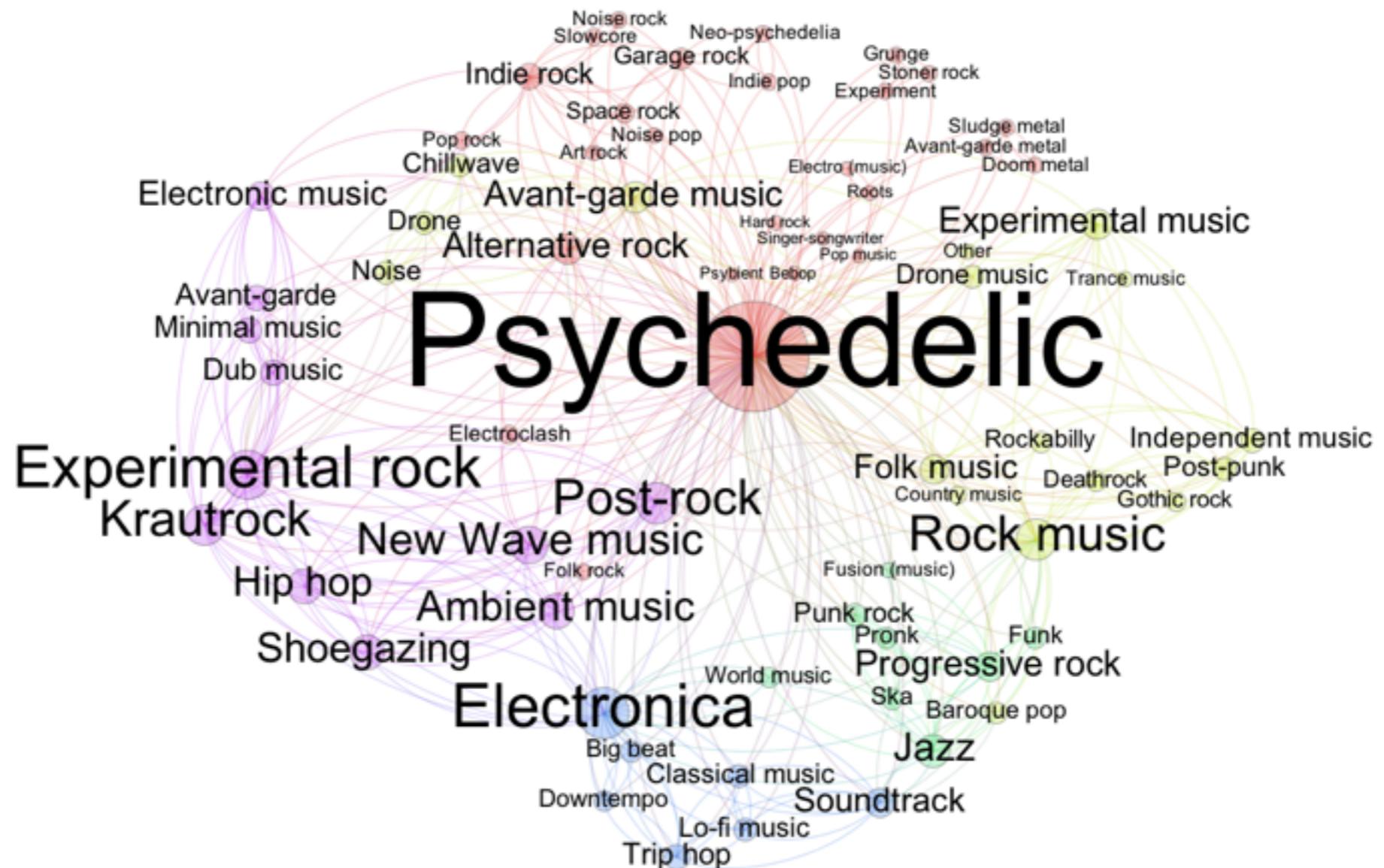
Ley de Zipf

- La distribución de palabras en un texto es del tipo “Power Law”
- Las palabras que aparecen en un 80% de un texto genérico son irrelevantes desde el punto de vista semántico

Referencias



GEPHI



<http://blog.ouseful.info/2012/07/03/visualising-related-entries-in-wikipedia-using gephi/>

GEPHI

Contenidos

INSTALACIÓN

FORMATOS Y CARGA DE UNA RED

DATA LABORATORY

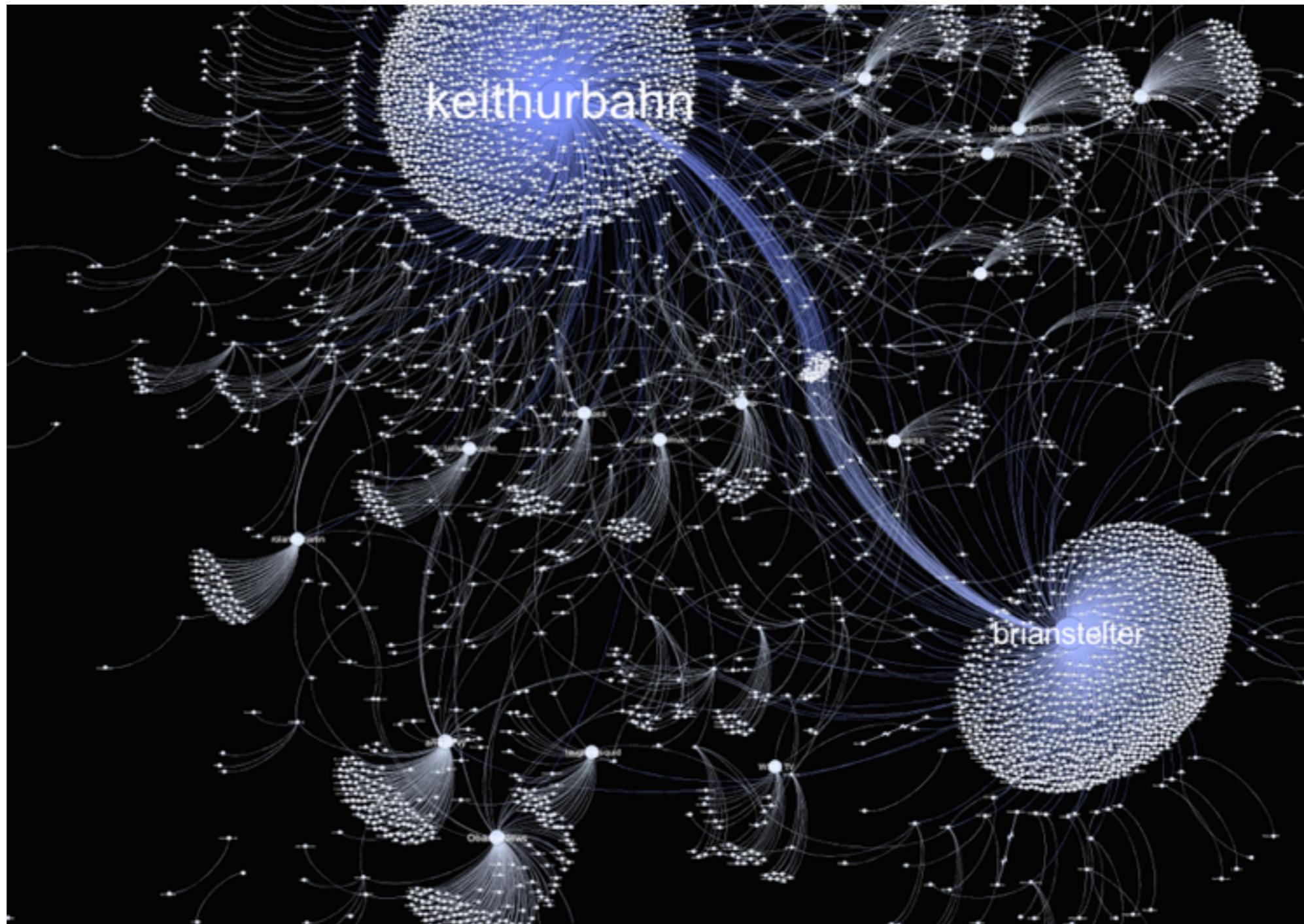
LAYOUT

MÉTRICAS Y RANKING

PRESENTACIÓN

EJERCICIOS

GEPHI



<http://mashable.com/2011/05/06/bin-laden-visualization/>

GEPHI

Herramienta para la comprensión y exploración de grafos/redes

Importación de diferentes formatos

Métricas

Filtrado

Layout

Ranking

Exportación gráfica

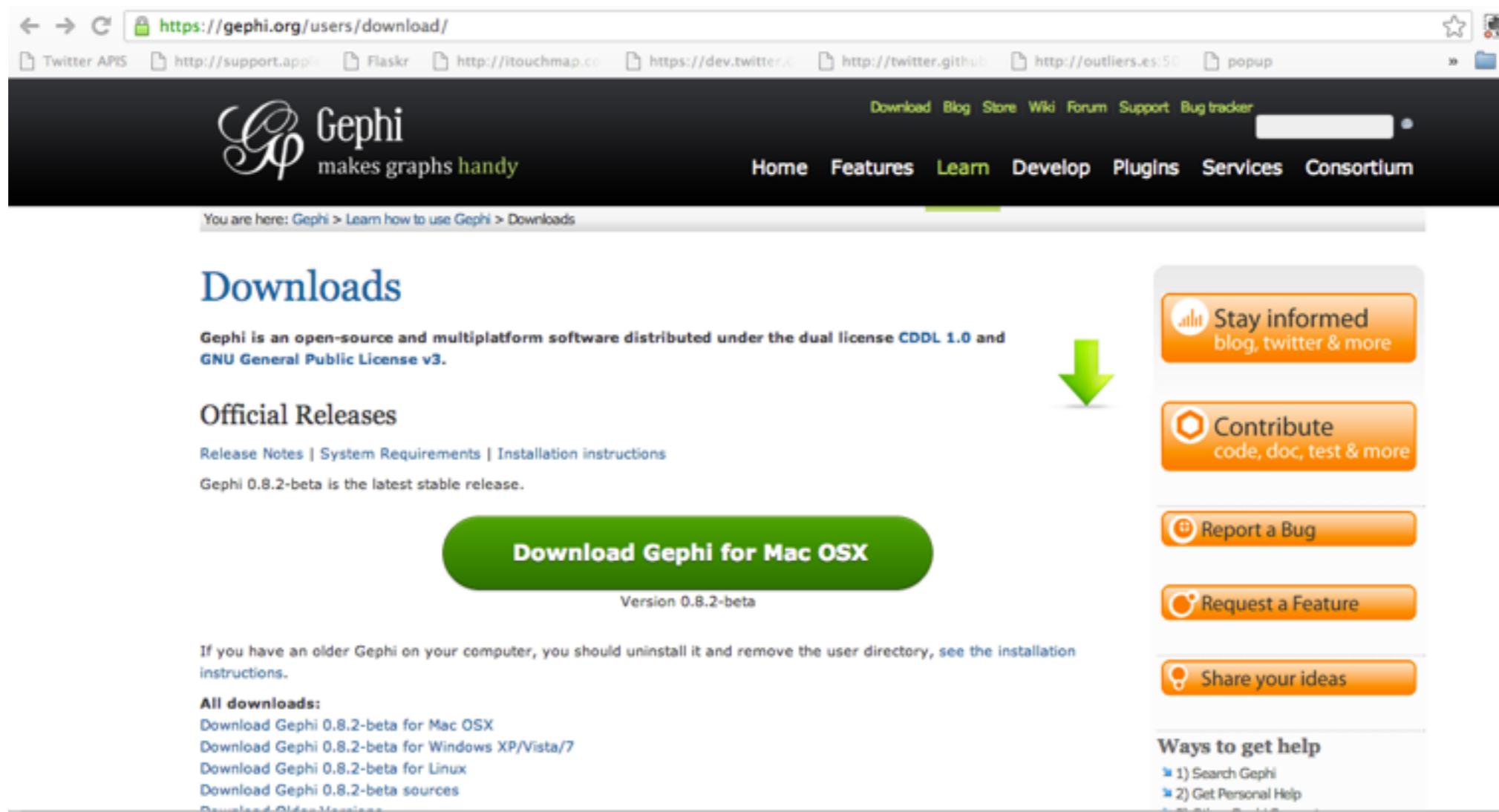
Plugins

GEPHI

Instalación

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

<https://launchpad.net/gephi/+download>



The screenshot shows a web browser displaying the Gephi download page at <https://gephi.org/users/download/>. The page has a dark header with the Gephi logo and the tagline "makes graphs handy". The navigation menu includes Home, Features, Learn (which is highlighted), Develop, Plugins, Services, and Consortium. Below the menu, a breadcrumb trail shows "You are here: Gephi > Learn how to use Gephi > Downloads".

Downloads

Gephi is an open-source and multiplatform software distributed under the dual license [CDDL 1.0](#) and [GNU General Public License v3](#).

Official Releases

[Release Notes](#) | [System Requirements](#) | [Installation instructions](#)
Gephi 0.8.2-beta is the latest stable release.

[Download Gephi for Mac OSX](#)

Version 0.8.2-beta

If you have an older Gephi on your computer, you should uninstall it and remove the user directory, see the installation instructions.

All downloads:

- [Download Gephi 0.8.2-beta for Mac OSX](#)
- [Download Gephi 0.8.2-beta for Windows XP/Vista/7](#)
- [Download Gephi 0.8.2-beta for Linux](#)
- [Download Gephi 0.8.2-beta sources](#)

Ways to get help

- 1) Search Gephi
- 2) Get Personal Help

Stay informed
blog, twitter & more

Contribute
code, doc, test & more

Report a Bug

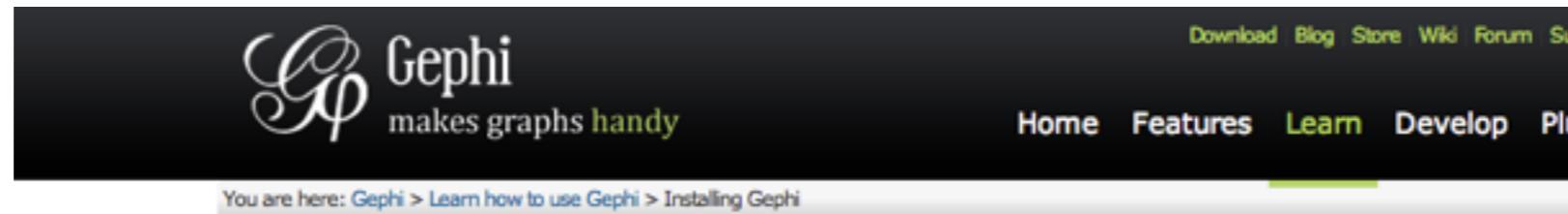
Request a Feature

Share your ideas

GEPHI

Instalación

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

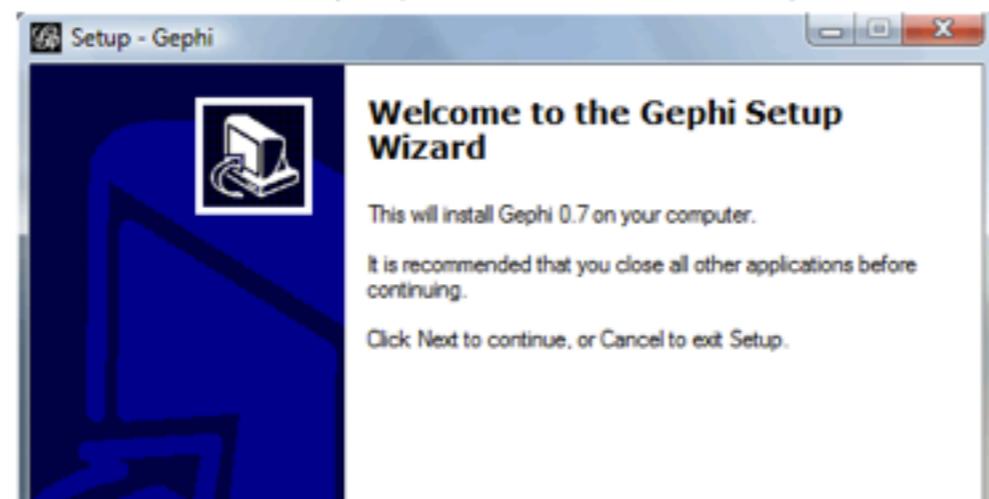


Installing Gephi

Installing the software

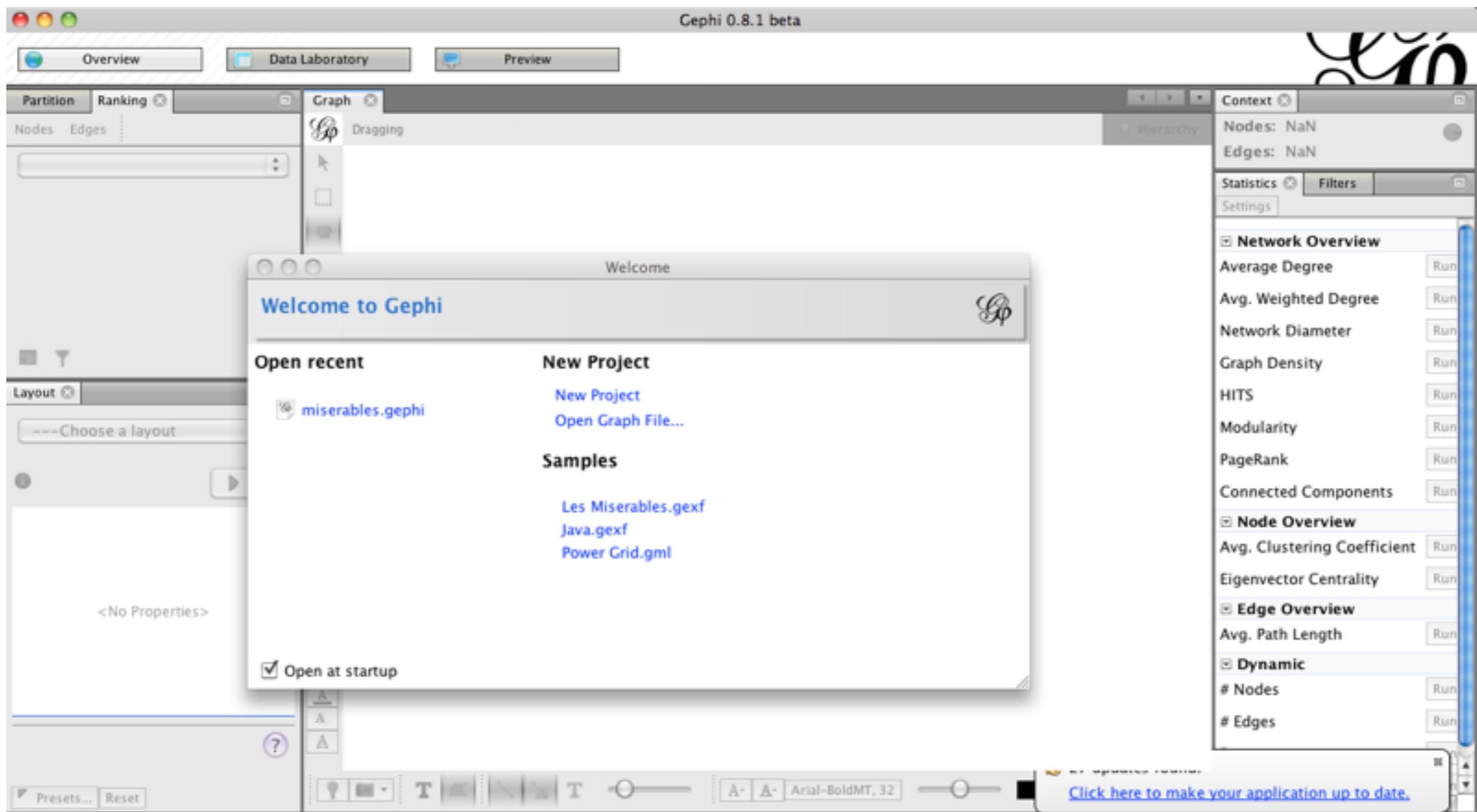
Windows

1. Be sure you have a recent Java JRE installed on your system. Download Free Java [here](#).
2. After the download completes, run the installer and follow the steps.



GEPHI

Instalación



GEPHI

Formatos y carga de una red

<https://gephi.org/users/supported-graph-formats/>

	Edge List/Matrix Structure	XML Structure	Edge Weight	Attributes	Visualization Attributes	Attribute Default Value	Hierarchical Graphs	Dynamics
CSV	Light Green							
DL Ucinet	Light Green							
DOT Graphviz								
GDF			Light Green		Light Green			
GEXF		Light Green	Light Green	Light Green	Light Green	Light Green		
GML		Light Green	Light Green	Light Green	Light Green	Light Green		
GraphML	Light Green	Light Green	Light Green	Light Green	Light Green	Light Green		
NET Pajek	Light Green							
TLP Tulip								
VNA Netdraw		Light Green	Light Green					
Spreadsheet*							Light Green	

GEPHI

Formatos y carga de una red

<https://wiki.gephi.org/index.php?title=Datasets>

Datasets

Gephi sample datasets, in various format (GEXF, GDF, GML, NET, GraphML, DL, DOT). Feel free to add new datasets. Be sure you cite original authors.

Supported graph formats are described [here](#).

Note that Gephi can open these files without the need to be unzipped.

[Contents](#) [hide]

- [1 Web and Internet](#)
- [2 Social networks](#)
- [3 Biological networks](#)
- [4 Infrastructure networks](#)
- [5 Other networks](#)
- [6 Sources](#)
- [7 Other network data repositories](#)

Web and Internet

[GEXF] [EuroSiS web mapping study](#): Mapping interactions between Science in Society actors on the Web of 12 European countries. Original report and data can be found [here](#).

[GML] [Internet](#): a symmetrized snapshot of the structure of the Internet at the level of autonomous systems, reconstructed from BGP tables posted by the University of Oregon R Views Project. This snapshot was created by Mark Newman from data for July 22, 2006 and is not previously published.

Social networks

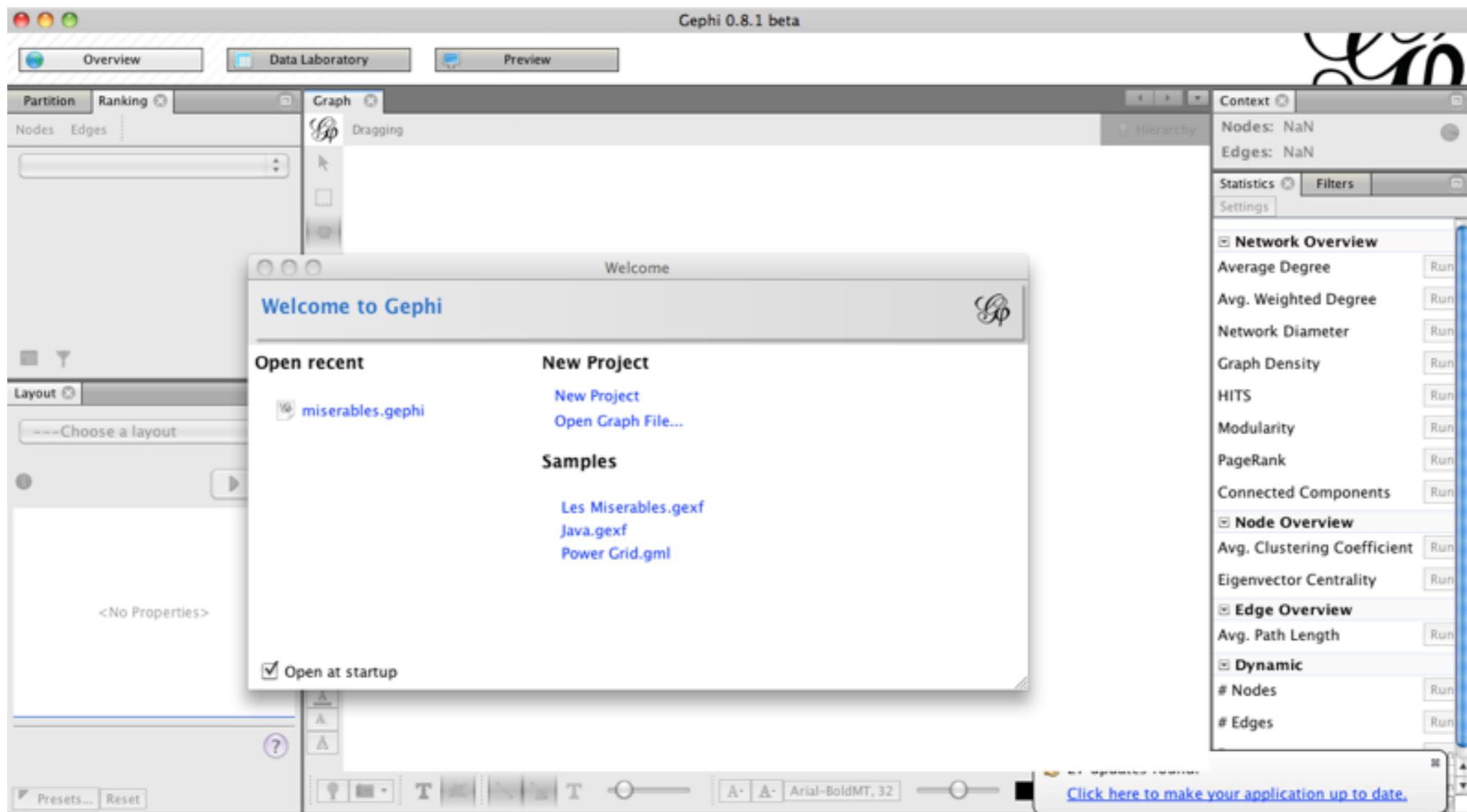
[GML] [Les Misérables](#): coappearance weighted network of characters in the novel Les Misérables. D. E. Knuth, The Stanford GraphBase: A Platform for Combinatorial Computing Addison-Wesley, Reading, MA (1993).

[GEXF] [Hypertext 2009 dynamic contact network](#): contact network during the Hypertext 2009 conference. Source: [Sociopatterns.org](#).

[GML] [Zachary's karate club](#): social network of friendships between 34 members of a karate club at a US university in the 1970s. W. W. Zachary, An information flow model for

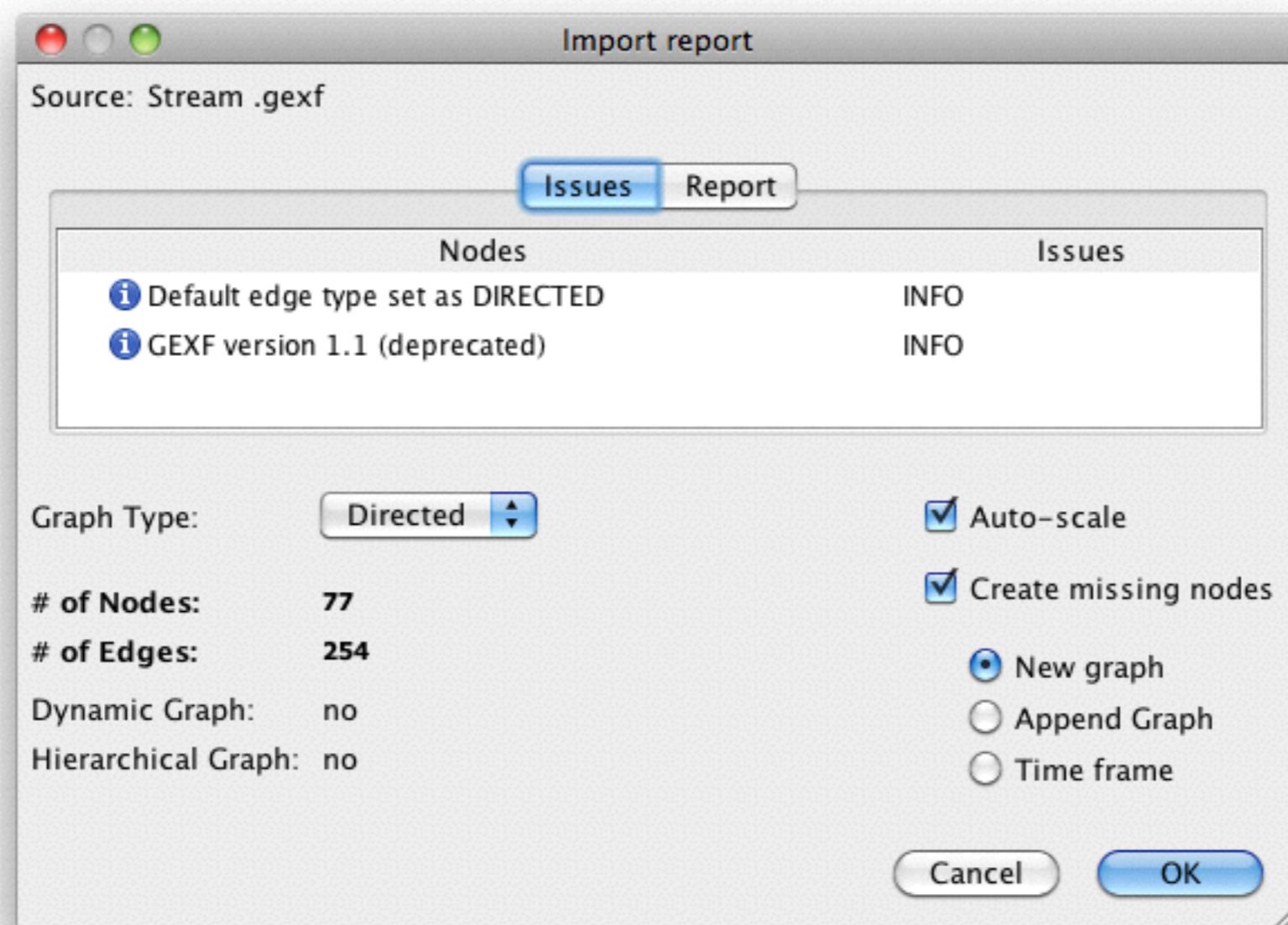
GEPHI

Formatos y carga de una red



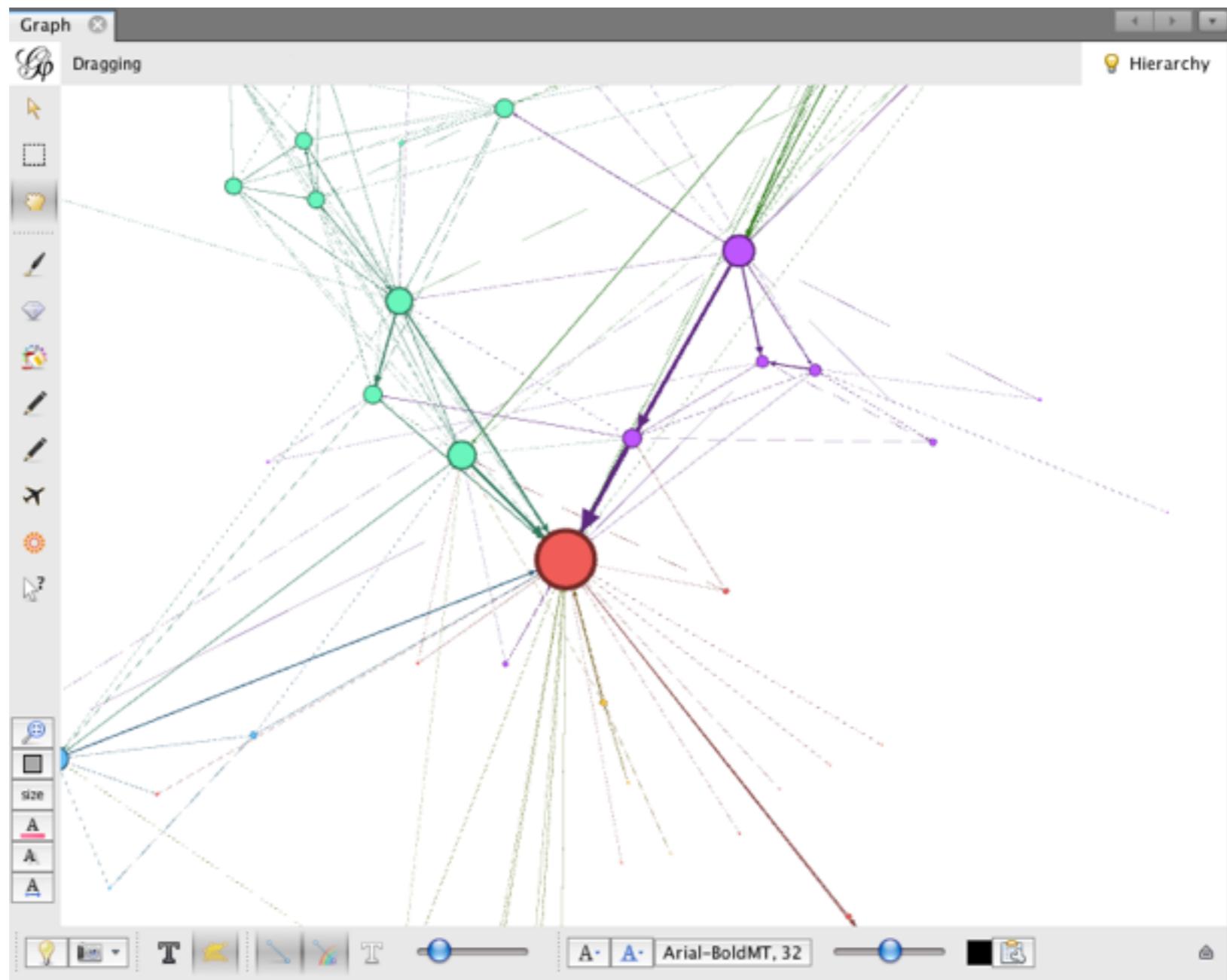
GEPHI

Formatos y carga de una red



GEPHI

Formatos y carga de una red



GEPHI

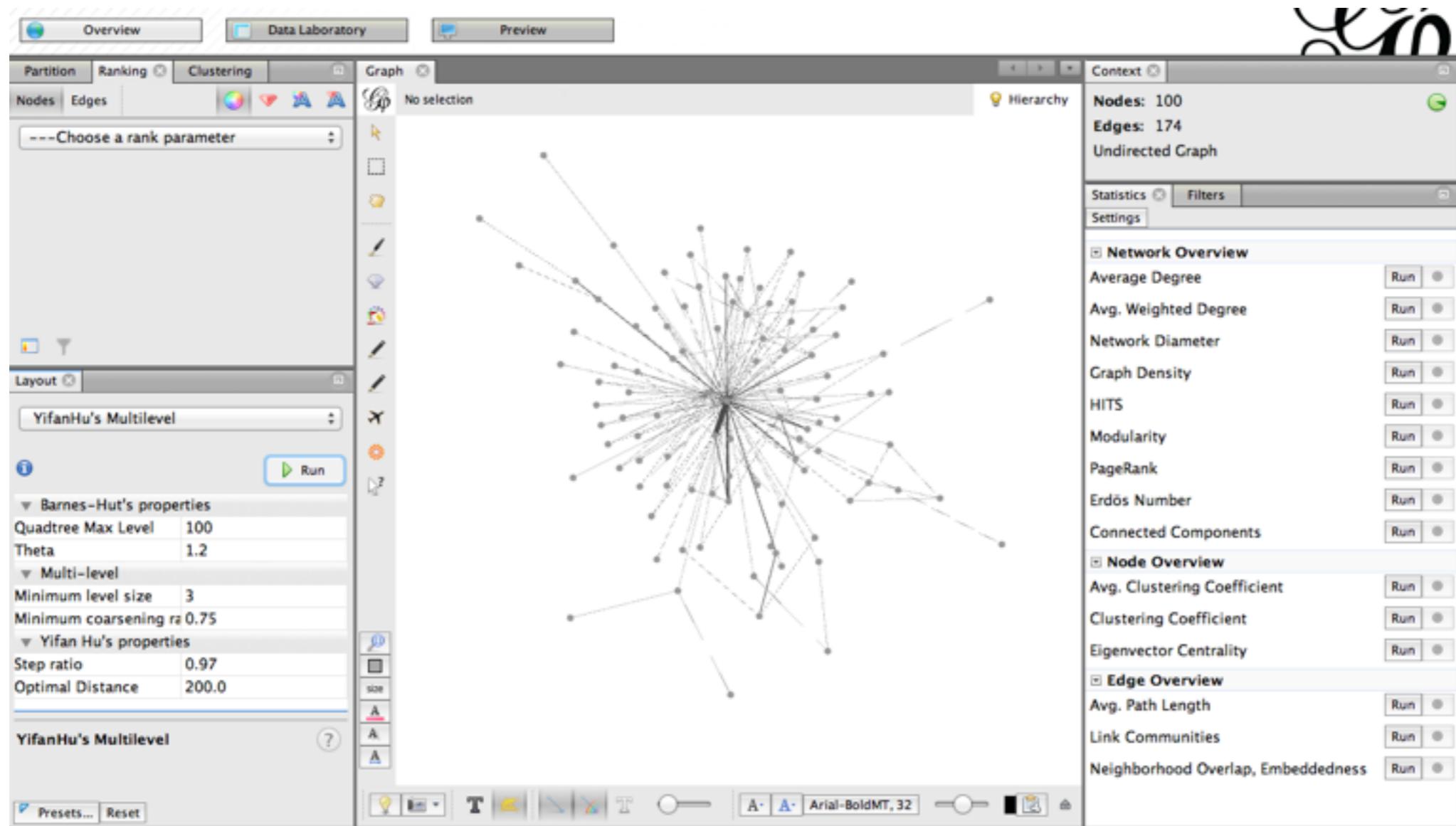
Laboratorio de datos

The screenshot shows the Gephi Data Laboratory interface. At the top, there are three tabs: Overview, Data Laboratory (which is selected), and Preview. Below the tabs is a toolbar with buttons for Overview, Data Laboratory, and Preview. The main area is a data table titled "Data Table". The table has three columns: "Nodes", "Id", and "Label". The "Nodes" column lists 20 characters from "Les Misérables": Myriel, Napoleon, MlleBaptistine, MmeMagloire, CountessDeLo, Geborand, Champtercier, Cravatte, Count, OldMan, Labarre, Valjean, Marguerite, MmeDeR, Isabeau, Gervais, Tholomyes, Listolier, Faneuil, Blacheville, and Favourite. The "Id" column contains numerical values from 0.0 to 20.0. The "Label" column contains the names of the characters. The bottom of the table has a row of buttons for data manipulation: Add column, Merge columns, Delete column, Clear column, Copy data to other column, Fill column with a value, Duplicate column, Create a boolean column from regex match, and Create column with list of regex matching groups.

Nodes	Id	Label
Myriel	0.0	Myriel
Napoleon	1.0	Napoleon
MlleBaptistine	2.0	MlleBaptistine
MmeMagloire	3.0	MmeMagloire
CountessDeLo	4.0	CountessDeLo
Geborand	5.0	Geborand
Champtercier	6.0	Champtercier
Cravatte	7.0	Cravatte
Count	8.0	Count
OldMan	9.0	OldMan
Labarre	10.0	Labarre
Valjean	11.0	Valjean
Marguerite	12.0	Marguerite
MmeDeR	13.0	MmeDeR
Isabeau	14.0	Isabeau
Gervais	15.0	Gervais
Tholomyes	16.0	Tholomyes
Listolier	17.0	Listolier
Faneuil	18.0	Faneuil
Blacheville	19.0	Blacheville
Favourite	20.0	Favourite

GEPHI

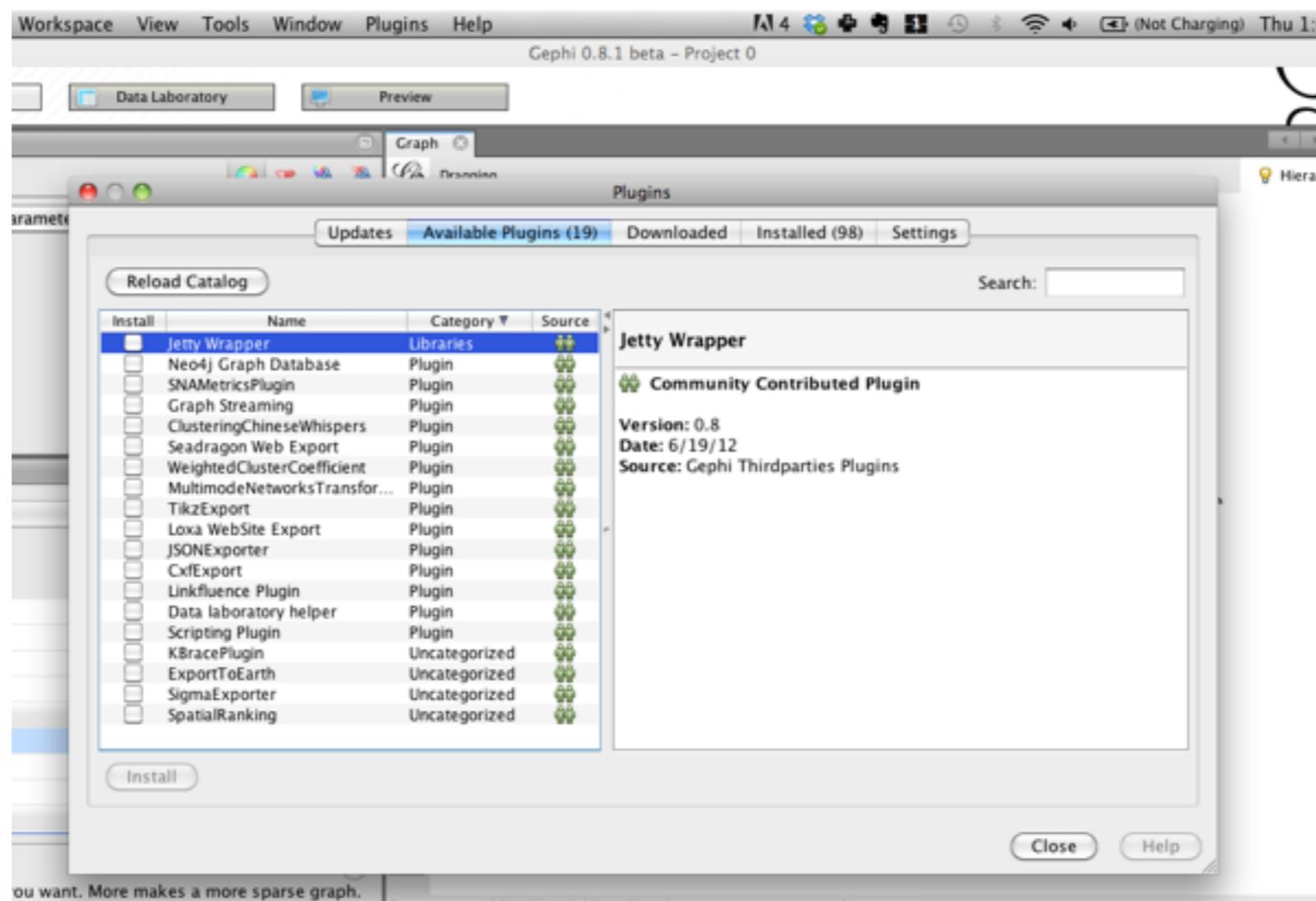
Layout



GEPHI

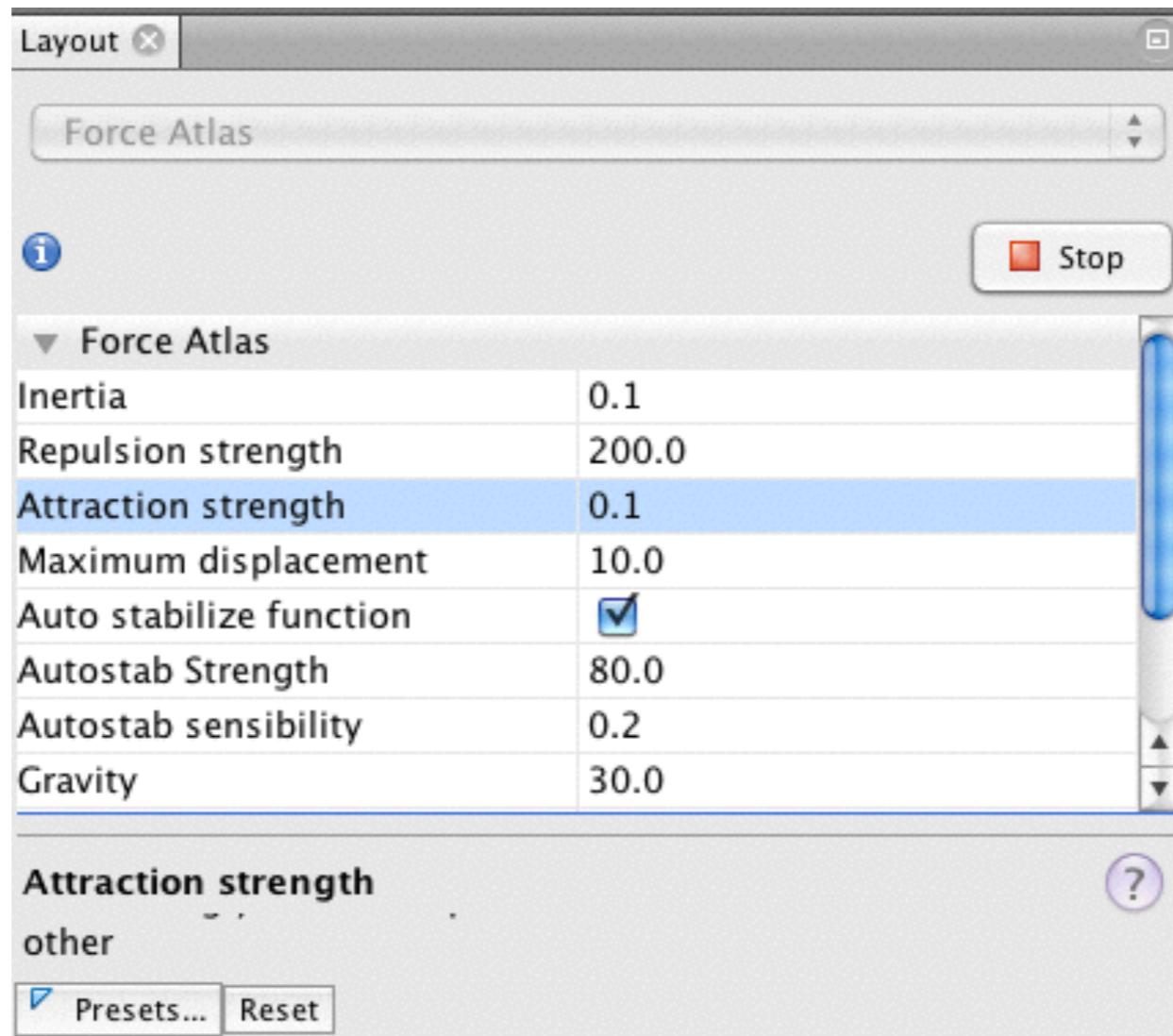
Layout

En Tools/Plugins: Instalar los de Layout



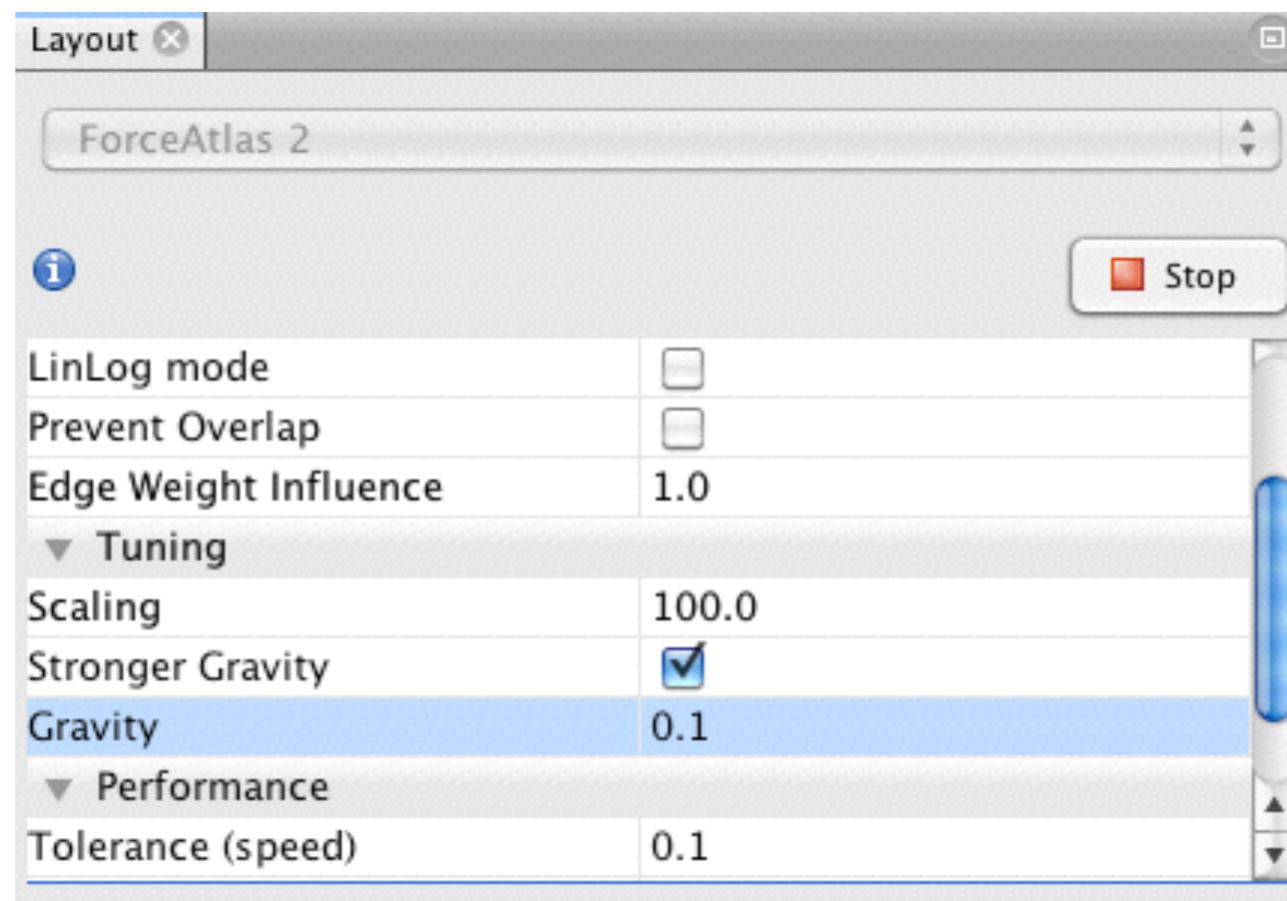
GEPHI

Layout



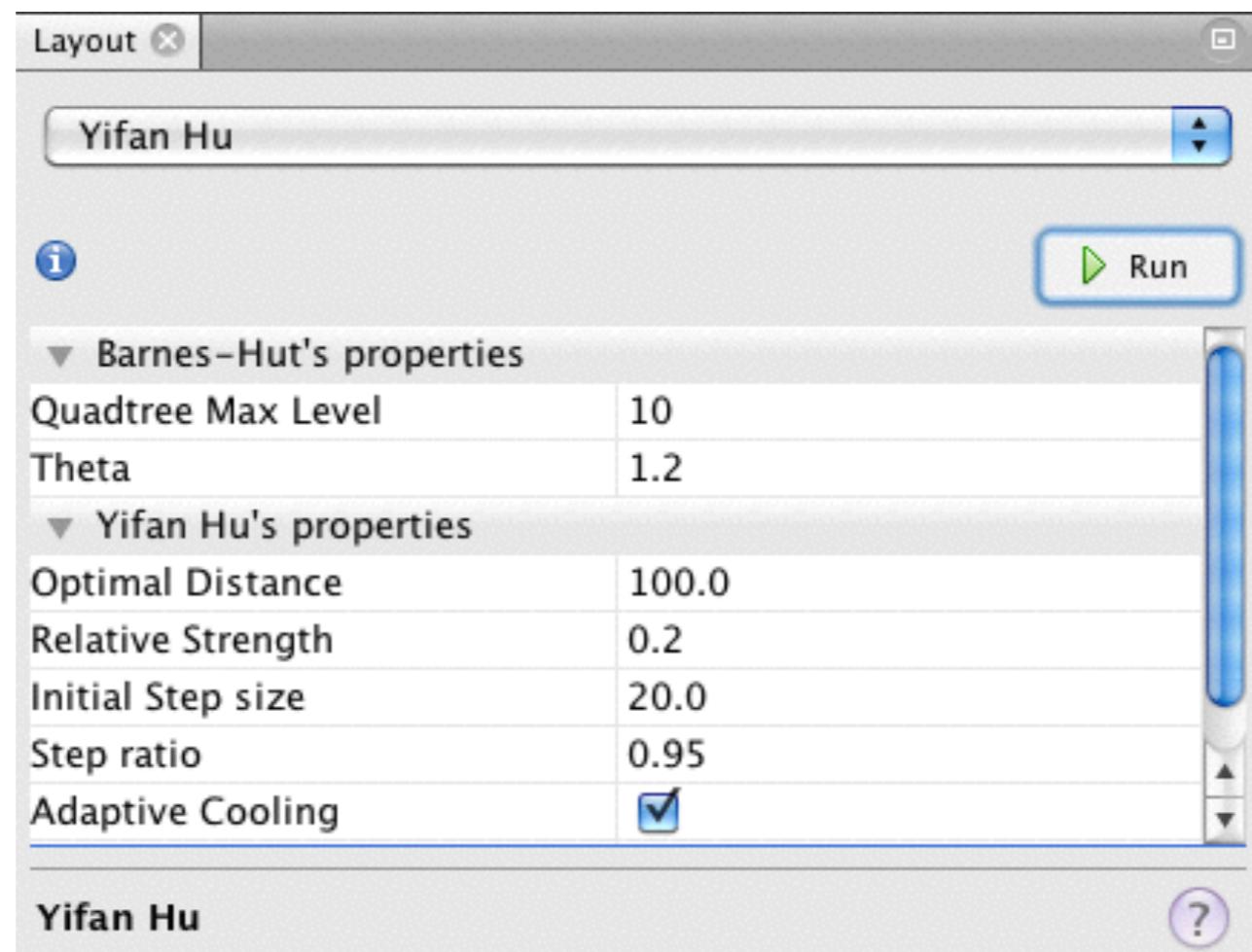
GEPHI

Layout



GEPHI

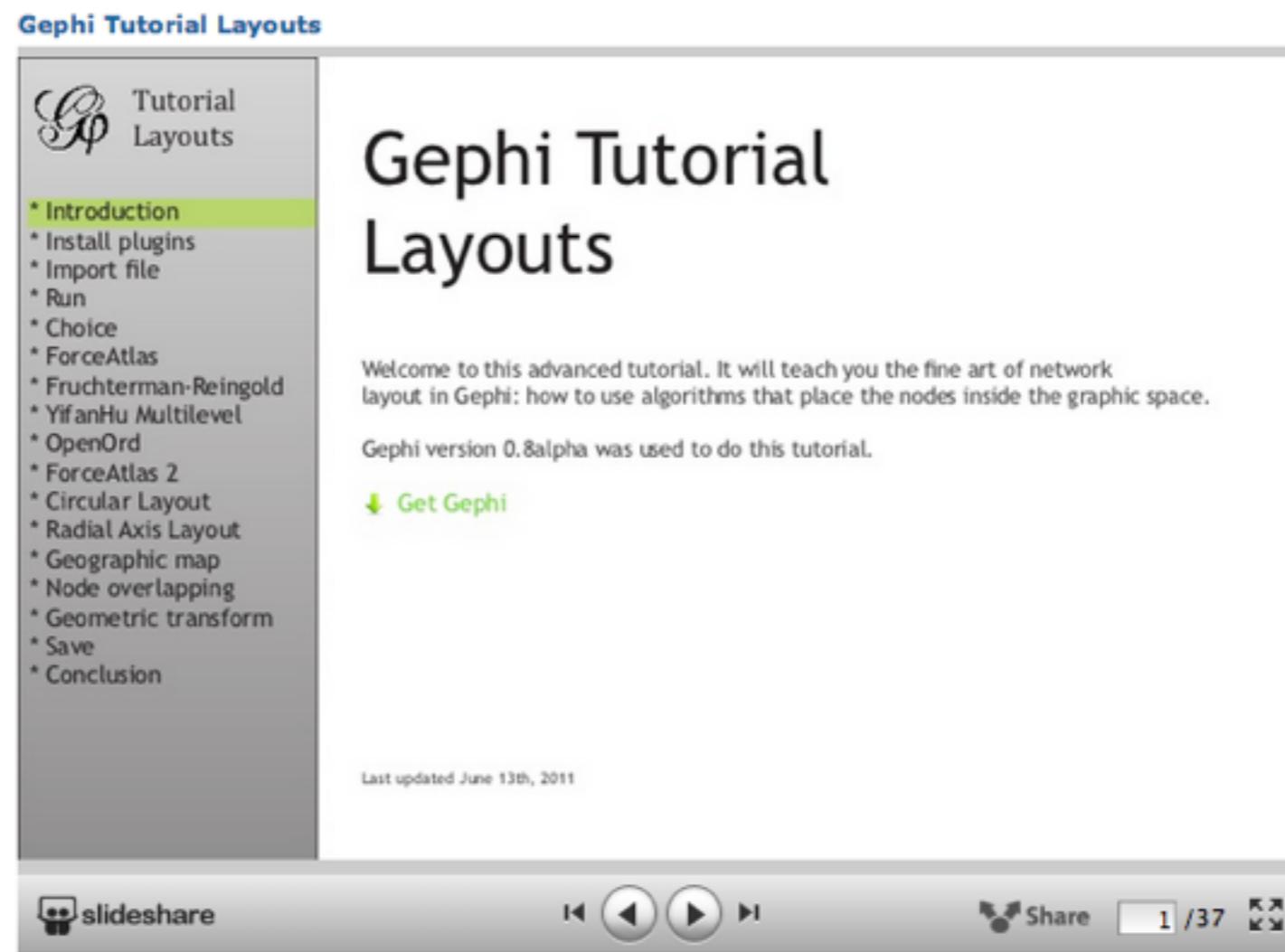
Layout



GEPHI

Layout

<https://gephi.org/2011/new-tutorial-layouts-in-gephi/>



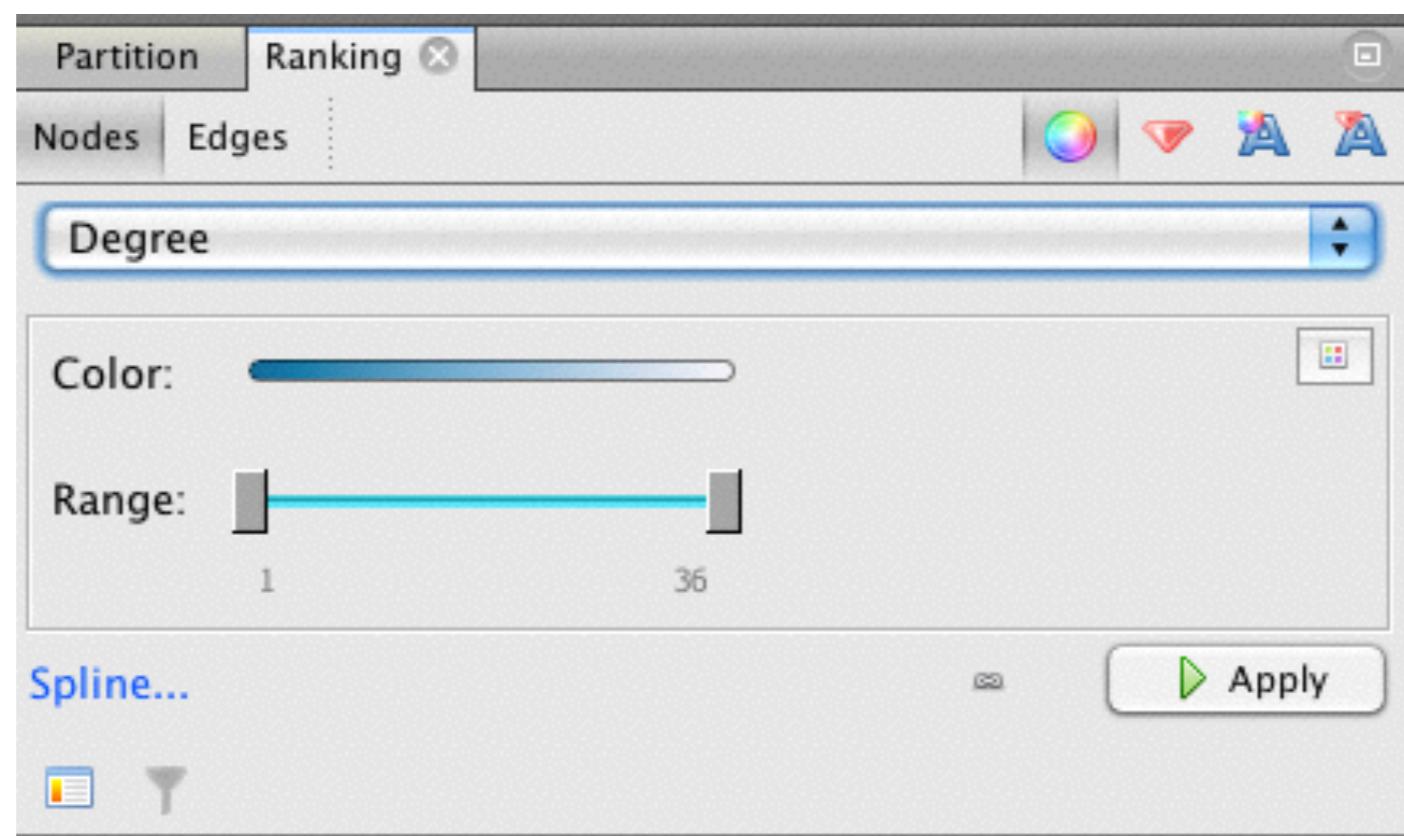
The screenshot shows a web-based tutorial for Gephi layouts. The left sidebar has a dark grey background and contains a navigation menu with the following items:

- * Introduction (highlighted in green)
- * Install plugins
- * Import file
- * Run
- * Choice
- * ForceAtlas
- * Fruchterman-Reingold
- * YifanHu Multilevel
- * OpenOrd
- * ForceAtlas 2
- * Circular Layout
- * Radial Axis Layout
- * Geographic map
- * Node overlapping
- * Geometric transform
- * Save
- * Conclusion

The main content area has a white background and features the title "Gephi Tutorial Layouts" in large, bold, black font. Below the title is a welcome message: "Welcome to this advanced tutorial. It will teach you the fine art of network layout in Gephi: how to use algorithms that place the nodes inside the graphic space." Underneath that, it says "Gephi version 0.8alpha was used to do this tutorial." At the bottom of the main content, there is a link "Get Gephi" with a small arrow icon. At the very bottom of the slide, there is a footer bar with icons for Slideshare, navigation arrows, a share button, and the text "1 / 37".

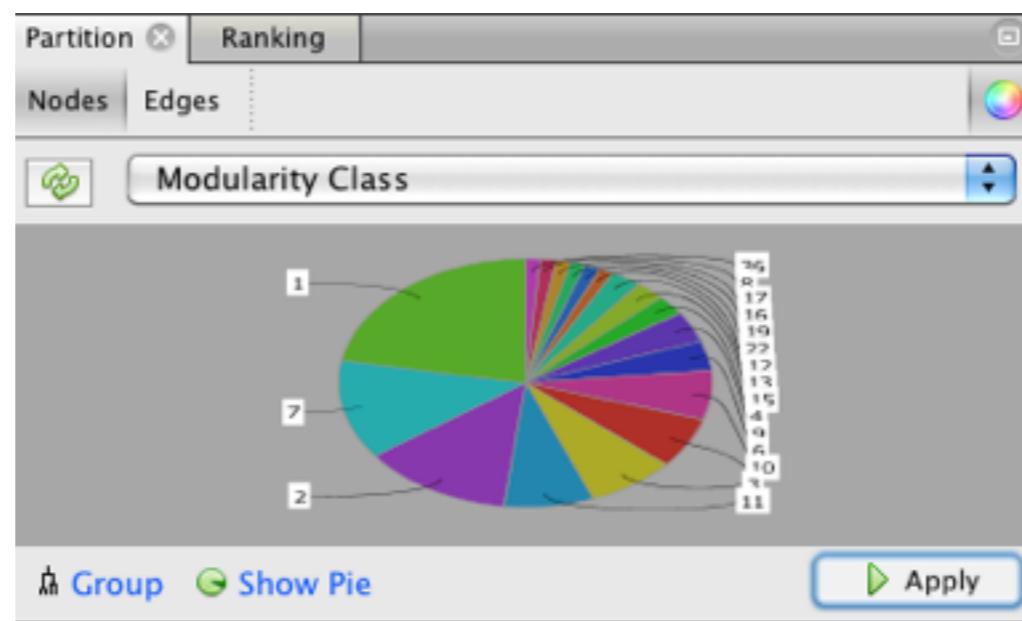
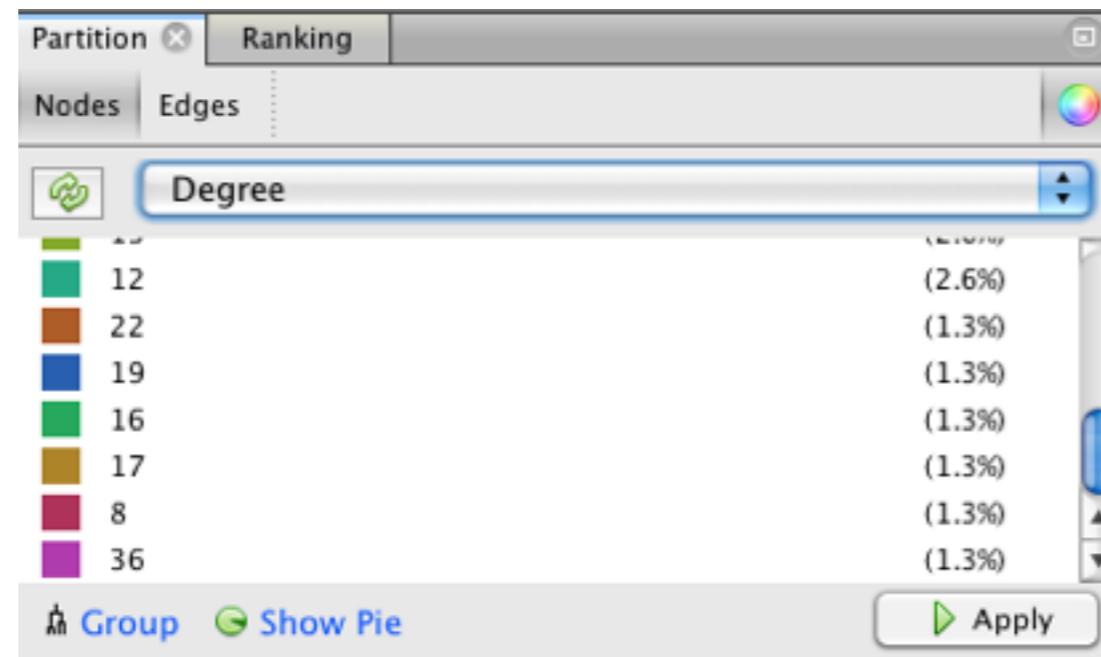
GEPHI

Métricas y ranking



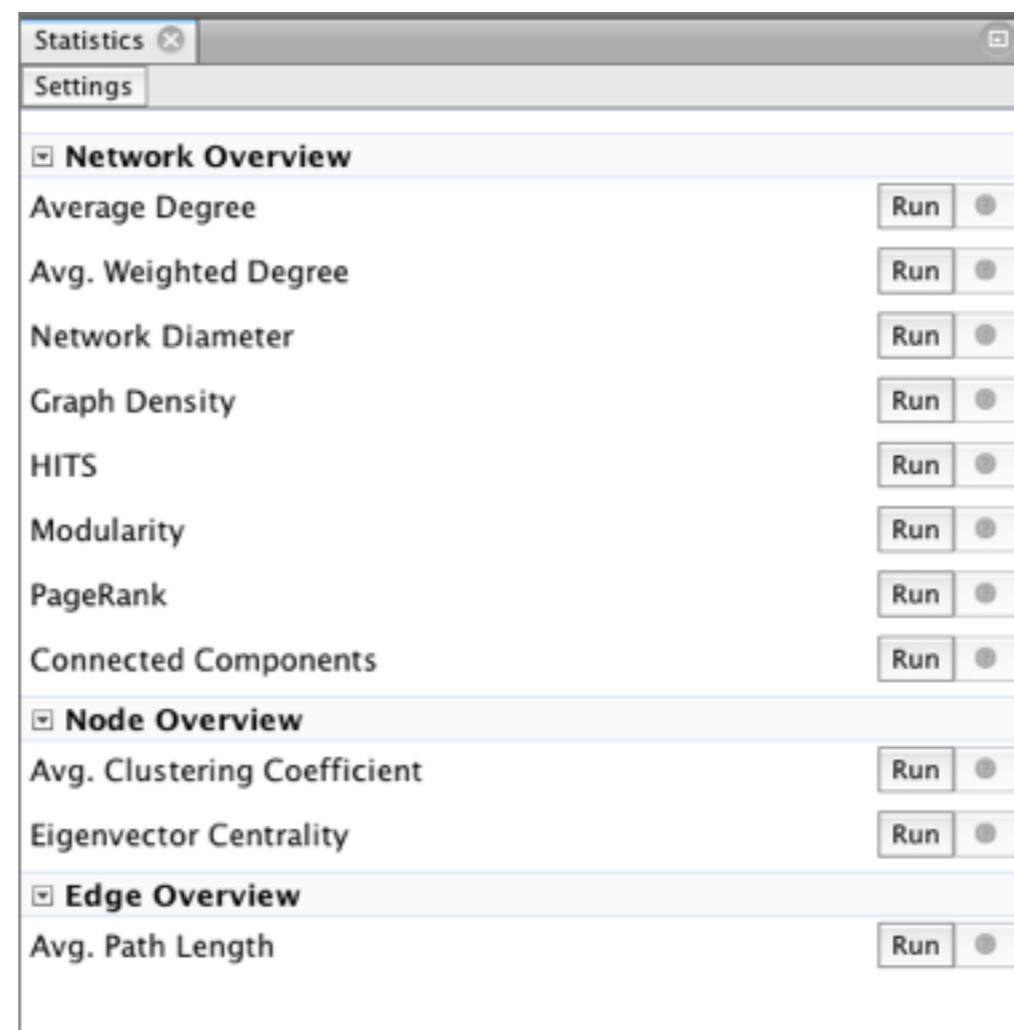
GEPHI

Métricas y ranking



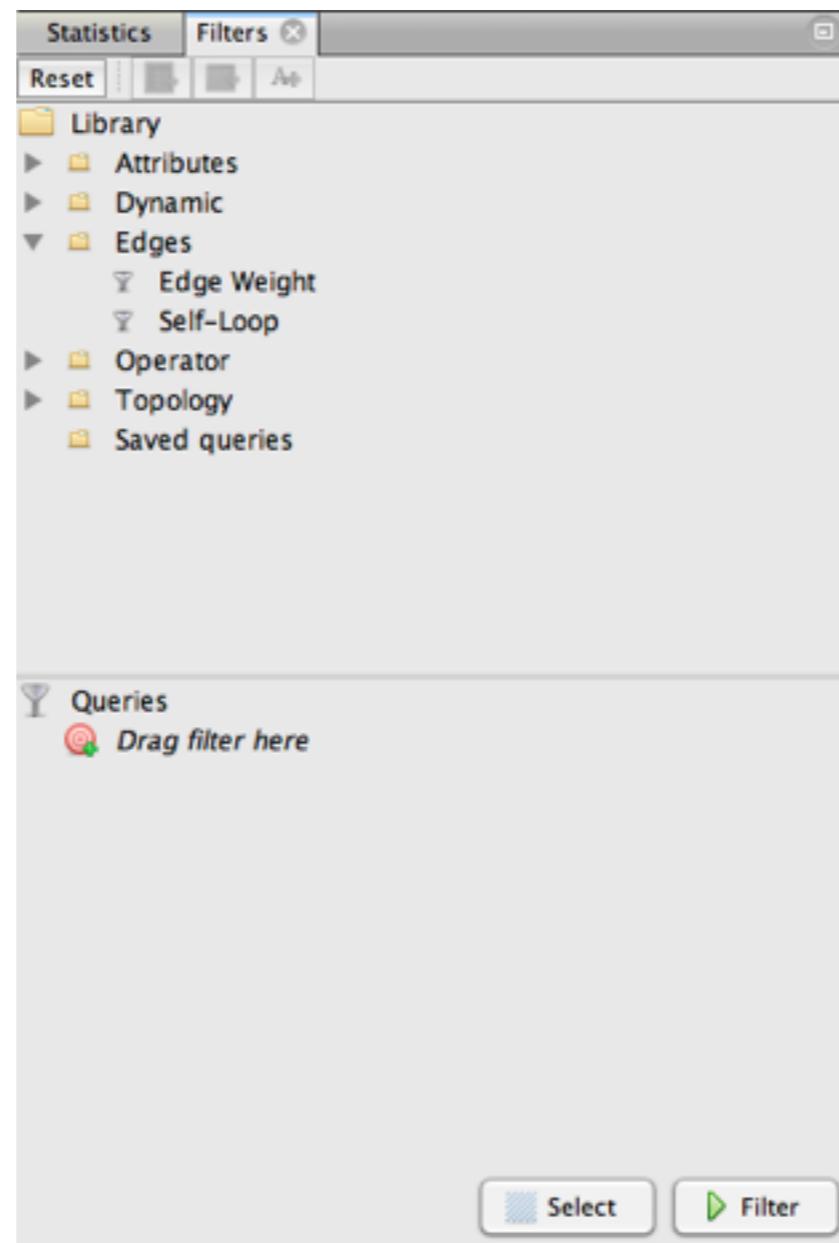
GEPHI

Métricas y ranking



GEPHI

Métricas y ranking



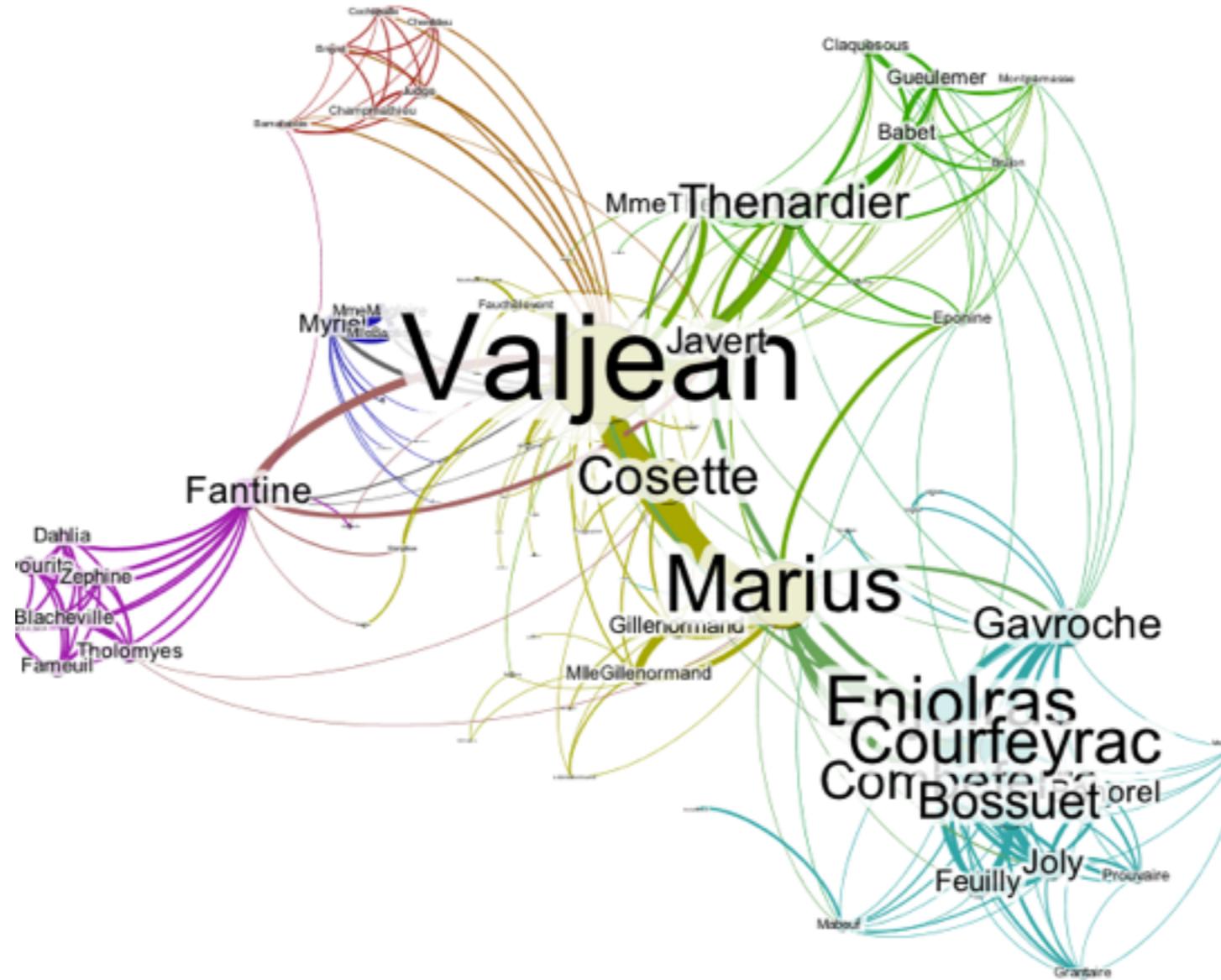
GEPHI

Presentación

Default	
Border Width	1.0
Border Color	custom [0,0,0] ...
Opacity	100.0
▼ Node Labels	
Show Labels	<input checked="" type="checkbox"/>
Font	Arial 12 Plain ...
Proportional size	<input checked="" type="checkbox"/>
Color	custom [0,0,0] ...
Shorten label	<input type="checkbox"/>
Max characters	30
Outline size	10.0
Outline color	custom [255,255,255] ...
Outline opacity	80.0
Box	<input type="checkbox"/>
Box color	parent ...
Box opacity	100.0
▼ Edges	
Show Edges	<input checked="" type="checkbox"/>
Thickness	1.0
Rescale weight	<input type="checkbox"/>
Color	mixed ...
Opacity	100.0
Curved	<input checked="" type="checkbox"/>

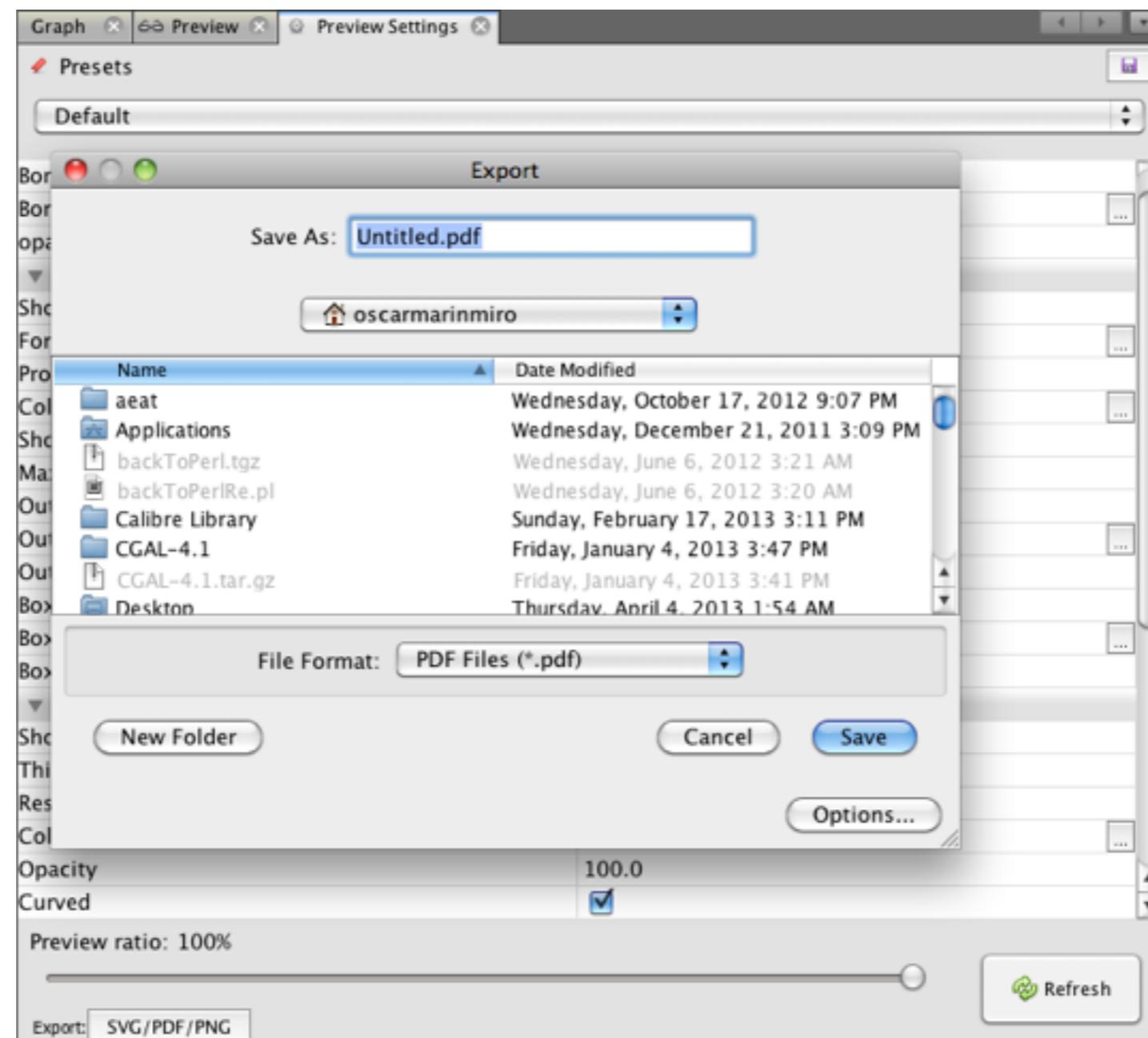
GEPHI

Presentación



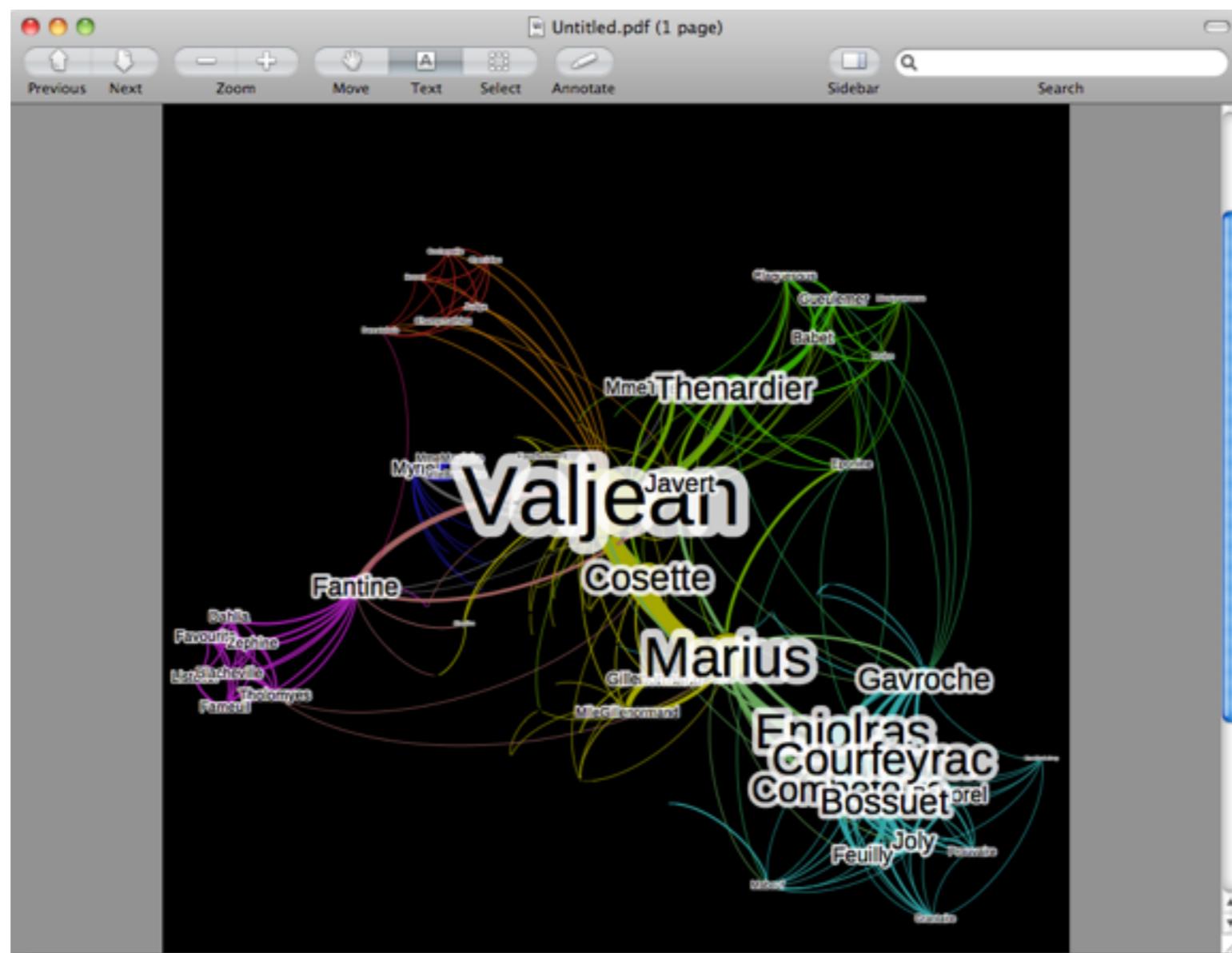
GEPHI

Presentación



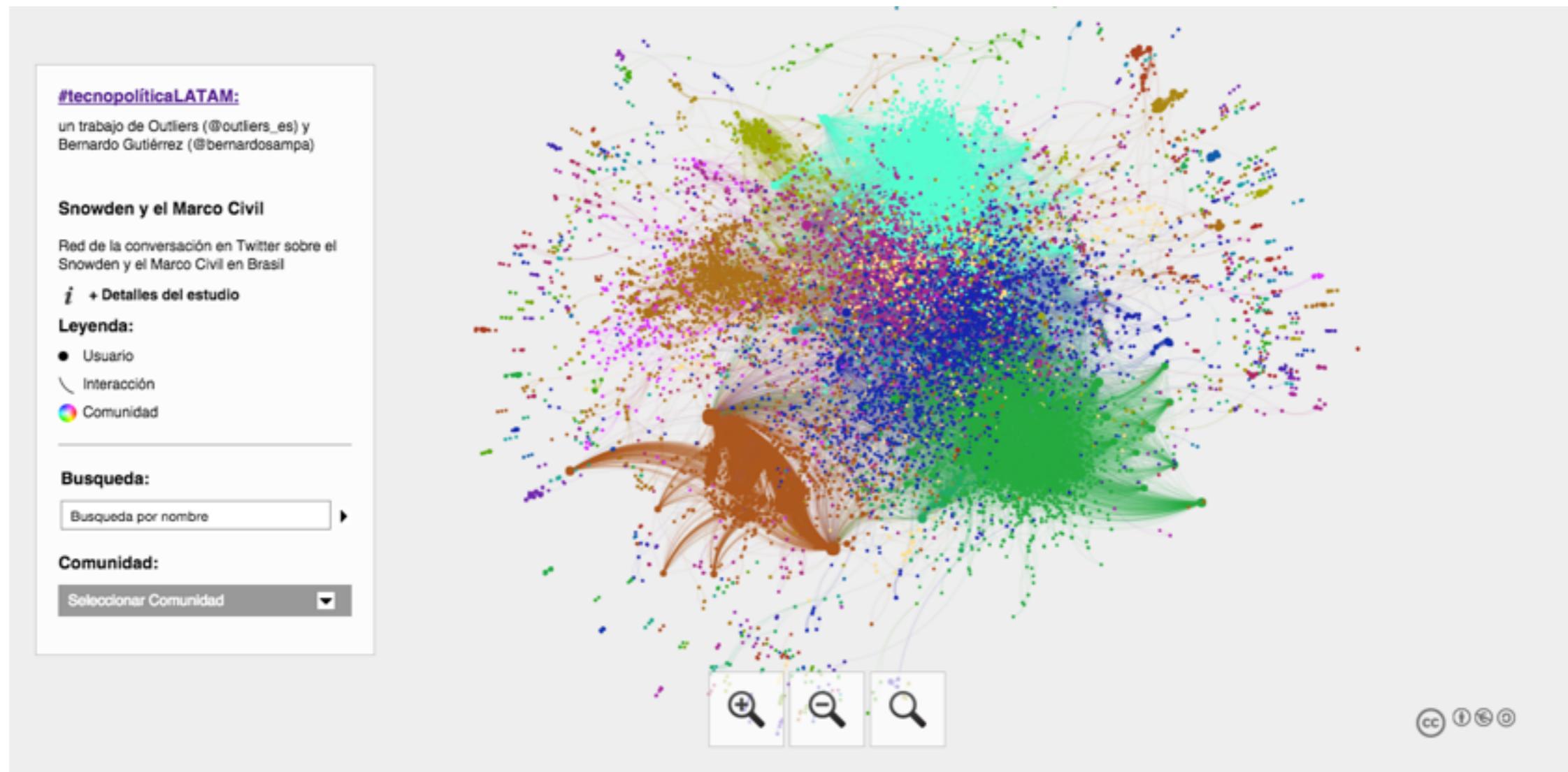
GEphi

Presentación



GEPHI

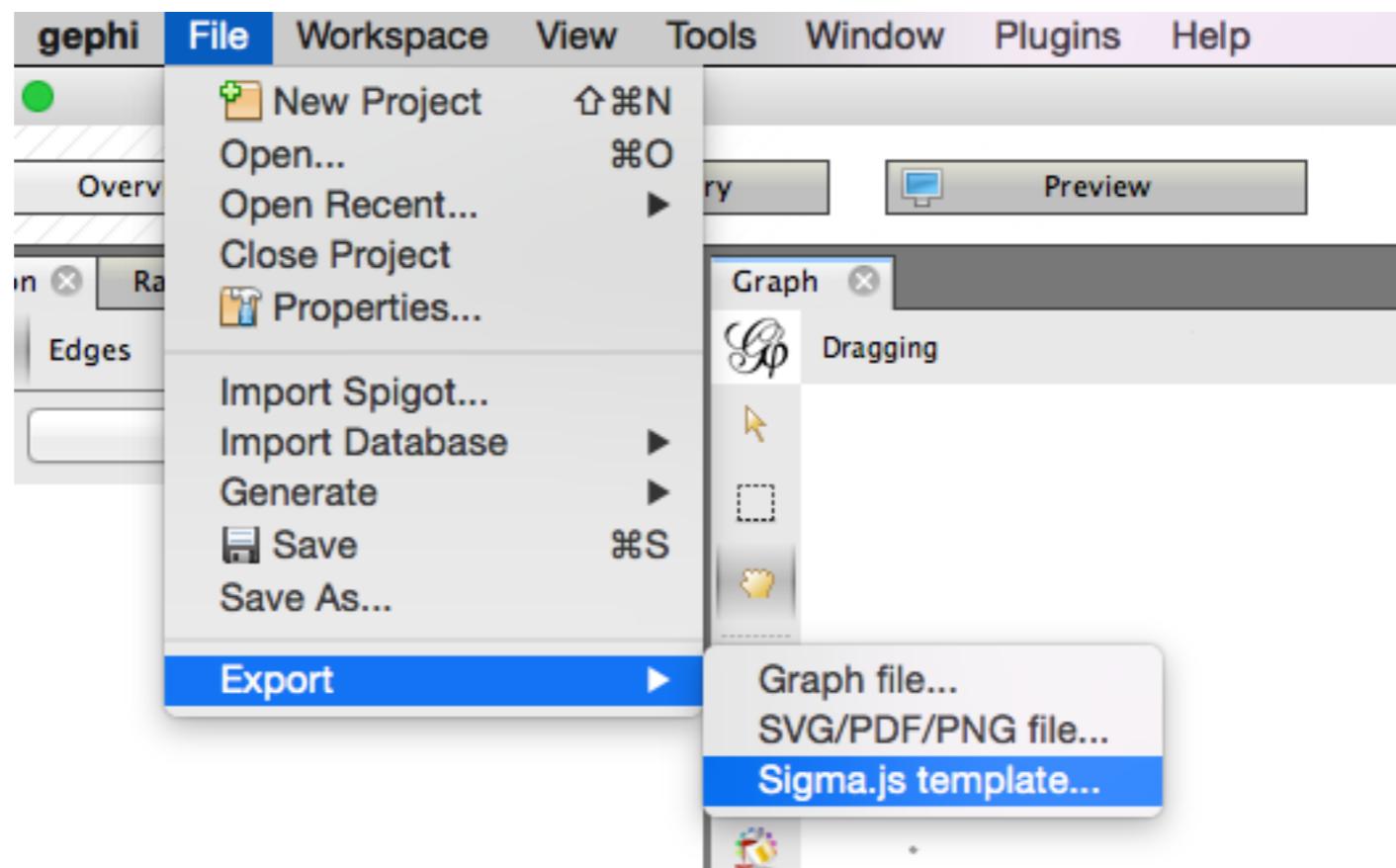
Presentación



<http://tecnopolitica.outliers.es/>

GEPHI

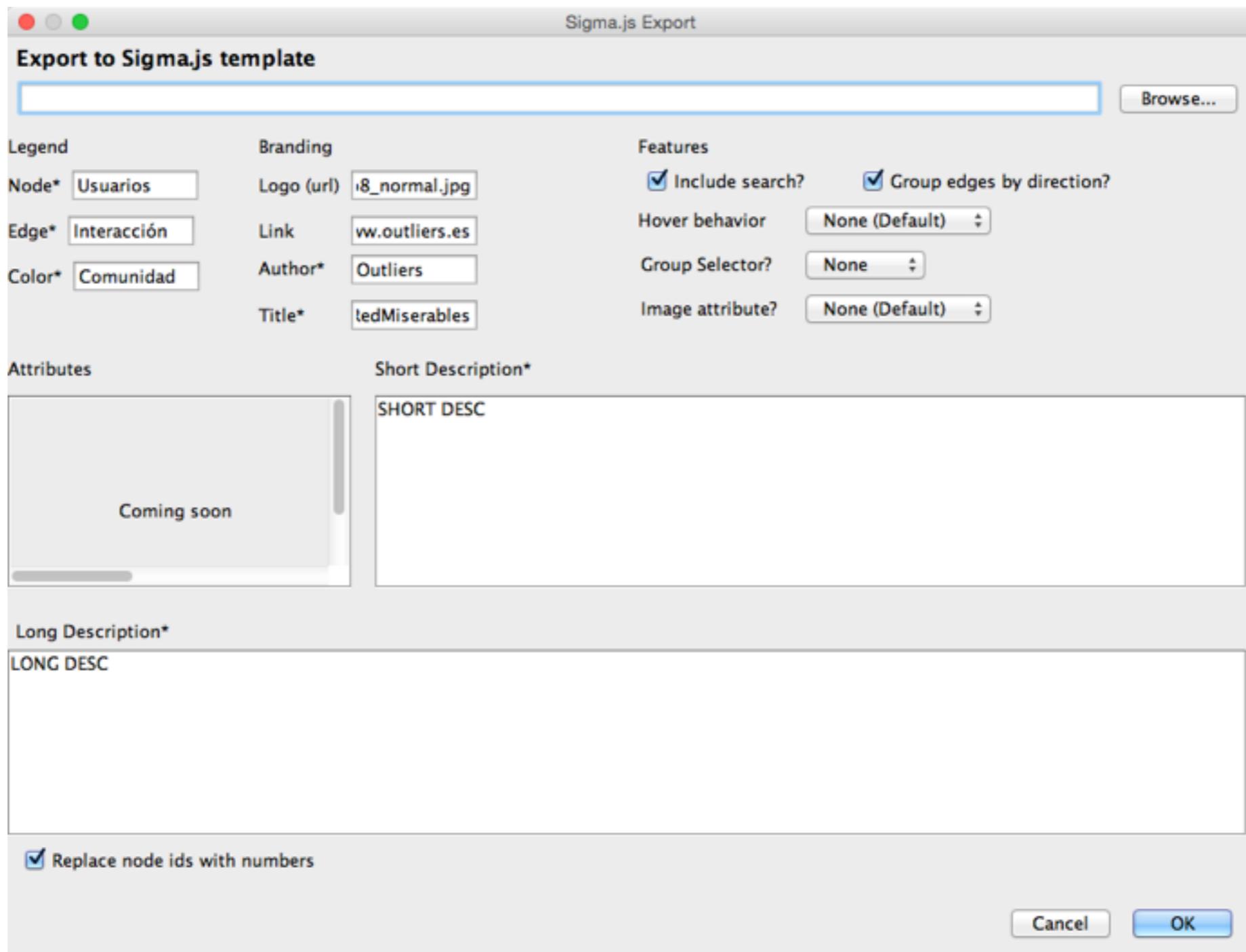
Presentación



<http://sigmajs.org/>

GEPHI

Presentación



GEPHI

Ejercicio I

Cargar red LesMiserables (Menú de Gephi)

Aplicar diferentes layouts y jugar con los parámetros

GEPHI

Ejercicio I

Cargar red LesMiserables (Menú de Gephi)

Cargarla en Gephi

Aplicar diferentes layouts y jugar con los parámetros

GEPHI

Ejercicio II

Obtener un CSV con entradas Autor-Retuiteado a partir del dataset de panamapapers como vimos en el primer día (u obtenerlo del github: [panama_stream_tabakalera_red.csv](#))

Cargarla en Gephi

GEPHI

Ejercicio II

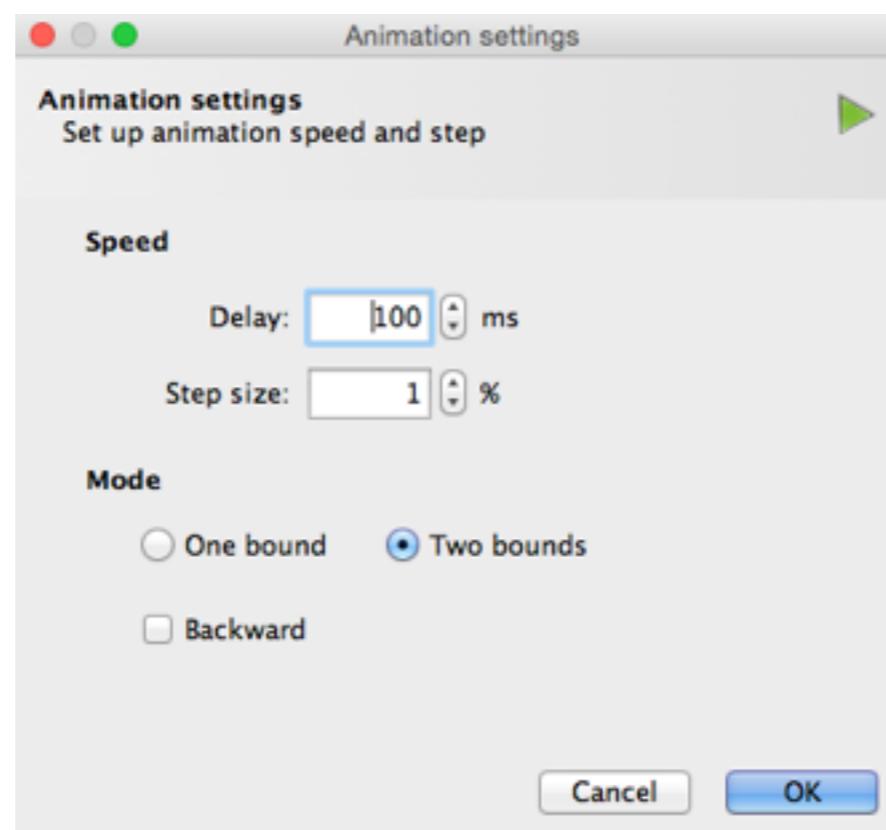
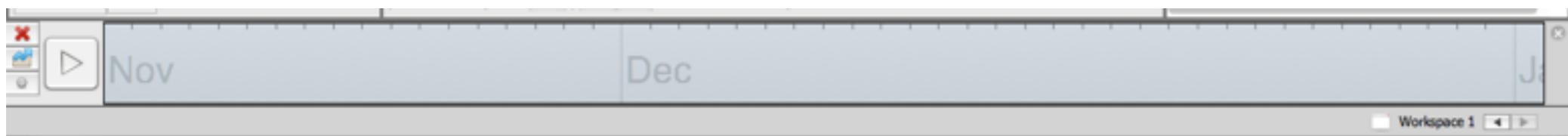
Aplicar un layout.

Aplicar rankings y calcular clases de modularidad

Explorar la red

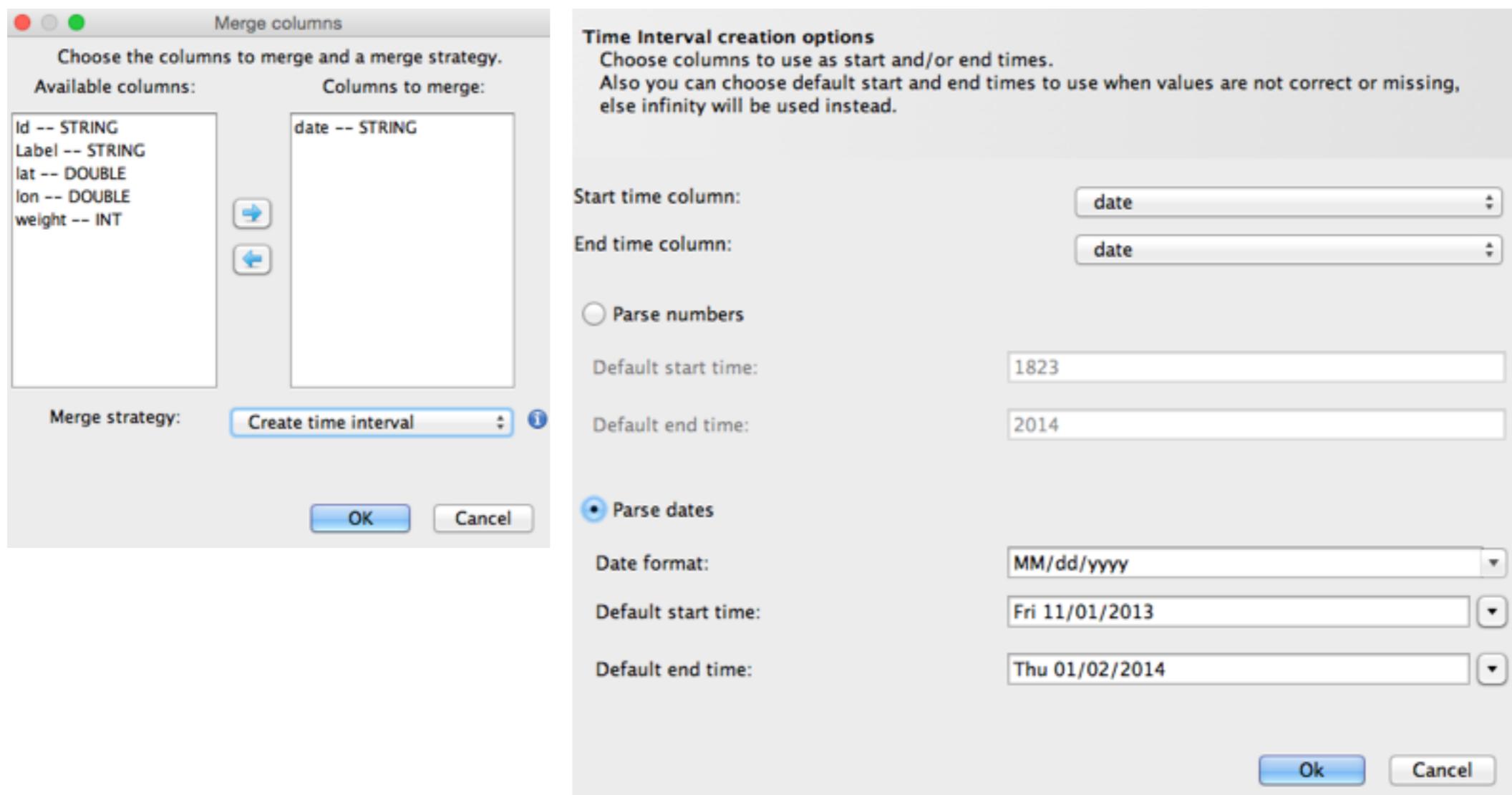
GEPHI

Gráficos Dinámicos



GEPHI

Gráficos Dinámicos



GEPHI

Ejercicio III

Abrir networkGeoDatetime.gexf

Aplicar layout geo

Generar columna de TimeInterval

Generar animación de creación de la red

GEPHI

Ejercicio IV

Descargar fichero
panama_papers_icij_red.gexf

Cargarla en Gephi

GEPHI

Ejercicio IV

Cambiar weight a 1 de los edges

Aplicar layout FORCE ATLAS 2 con parámetro scale=100.0

Calcular Modularity Class. ¿Salen muchas comunidades? ¿Pocas? ¿Porqué será?

Buscar Pilar de Borbón, con quien comparte clase de modularidad

Buscar el nodo African Resources Investment Fund SPC. Buscar con quien comparte clase de modularidad. Buscar en Google el/los nombre/s propio/s que veamos. ¿Quienes son? ¿Son cargos públicos?

GEPHI

Ejercicio IV

Buscar Micaela Domecq Solís-Beaumont (mujer de Arias Cañete). Buscar con quien comparte clase de modularidad

Buscar la clase de SIERRA LEONA S.L. Buscar con quien comparte clase de modularidad. Buscar en Google el/los nombre/s propio/s que veamos. ¿Quienes son?

Buscar algún grupo de los que se separan. Identificar su clase y buscar quienes la componen.

GEPHI

Ejercicio Personal I

Obtener una red de Twitter en:

<http://flocker.outliers.es>

Cargarla en Gephi

Explorarla y detectar insights

GEPHI

Ejercicio Personal II

Descargar las redes flocker-#felizjueves.gexf y flocker-#RiveraAR.gexf

Cargar en Gephi la primera

Explorar la red y sus nodos

¿Cómo parece la conversación? ¿Hay muchas interacciones? Es un conversación centralizada mucha gente enlazada?

GEPHI

Ejercicio Personal II

Cargar en Gephi la segunda

Explorar la red y sus nodos

¿Cómo parece la conversación? ¿Se parece a la anterior? ¿Hay nodos “centrales”?

Any questions?
Thanks for your attention!

