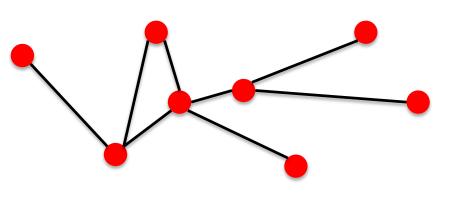


Oscar Cordón García

Dpto. Ciencias de la Computación e Inteligencia Artificial ocordon@decsai.ugr.es

FORMATOS DE FICHERO PARA ALMACENAR REDES SOCIALES



componentes: nodos

• interacciones: enlaces L

■ sistema: red, grafo (N,L)

Como sabemos, una red social representa el comportamiento de un sistema complejo

Para almacenar la estructura de la red en un fichero, que podamos recuperar, editar, procesar, etc., tenemos que guardar información sobre los nodos y los enlaces

Además, puede existir información sobre los atributos de los nodos y los enlaces. Ej:

- · Nodos: nombre, edad, género, etc.
- Enlaces: tipo de relación, frecuencia/importancia, etc.

Hay muchos formatos de fichero para almacenar redes. Son ficheros de texto con metadatos. La información de la estructura de la red puede ser interpretada y recuperada por muchas aplicaciones. La de visualización no, es específica

EL FORMATO NET DE PAJEK (1)

Pajek es otra herramienta estándar de análisis de redes sociales. Trabaja sobre Windows y es gratuita. Su nombre viene de la palabra eslovena "araña"

Su formato de almacenamiento de redes, el NET, está muy extendido. Es un formato de texto con metadatos con dos partes:

1. La primera almacena la información de los nodos, empezando con una línea con el metadato *vertices, seguida por el número de nodos de la red

Después incluye la información de cada nodo, empezando con su número que actúa como identificador único y seguido por su nombre entre comillas y el resto de atributos, con una línea para cada nodo:

```
*Vertices 82670
1 "entity"
2 "thing"
3 "anything"
4 "something"
5 "nothing"
6 "whole"
```

El formato puede incluir información de las posiciones, tamaños y colores de los nodos en la visualización pero esa información solo será interpretada por el propio Pajek, no por otras herramientas de análisis de redes como *Gephi*

EL FORMATO NET DE PAJEK (2)

- 2. La segunda almacena la información de los enlaces, con tres formatos posibles:
- Lista de enlaces: Cada enlace se guarda en una línea mediante un par de identificadores de nodos (origen y destino). Comienza con el metadato *edges para redes no dirigidas y *arcs para redes dirigidas:

Si la red es ponderada, después del nodo origen y destino del enlace incluye el peso del enlace:

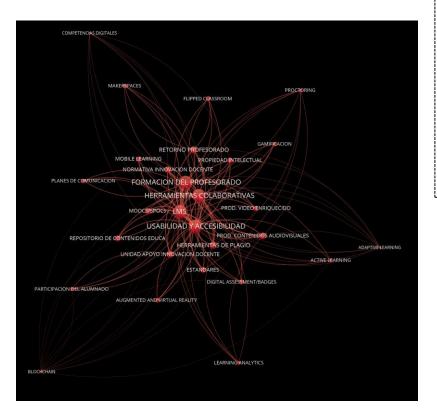
 Lista de adyacencia: El metadato de la primera línea es *edgeslist o *arcslist. Luego, cada línea tiene el identificador del nodo origen de los enlaces seguido por todos los nodos destino: *edgeslist 1 3553 3586 3587 3637 2 3583 3 4930 4 88 5 13 120

 Matriz de adyacencia: Es el menos usado al malgastar espacio guardando los enlaces vacíos. El metadato es *matrix. Luego, cada línea del fichero corresponde con una línea de la matriz de adyacencia

```
*matrix
0.000
         0.000
                  2.000
                           0.000
                                    3.000
3.000
         0.000
                  0.000
                           0.000
                                    0.000
2.000
         3.000
                                    0.000
                  0.000
                           0.000
0.000
         0.000
                  0.000
                           0.000
                                    3.000
0.000
         0.000
                  0.000
                           3.000
                                    0.000
```

EJEMPLO DE FICHERO DE PAJEK

Red de Tecnologías Educativas utilizadas en el Sistema Universitario Español



```
*vertices 28
 1 "NORMATIVA INNOVACION DOCENTE"
   "UNIDAD DE APOYO A LA INNOVACION DOCENTE"
                              0.4573 0.5124 0.5000
   "COMPETENCIAS DIGITALES"
   "MOOCS/SPOCs"
                   0.6643 0.1276 0.5000
*edaes
1 4 0.810000
   0.810000
1 12 0.810000
1 13 0.060000
20 27 0.060000
20 28 0.770000
21 26 0.020000
25 26 0.020000
26 28 0.020000
```

- Red no dirigida (*edges) y ponderada
- Los valores numéricos de detrás de los nodos son de visualización, solo valen en Pajek
- La visualización de la izquierda está hecha en Gephi, no se guarda en el fichero net que solo guarda la información propia de la red

LOS FORMATOS GraphML Y GML (1)

http://graphml.graphdrawing.org/ ; http://graphml.graphdrawing.org/primer/graphml-primer.html

El formato de fichero GraphML es un estándar especificado en XML Schema (estándar de ficheros de metadatos) para almacenar redes y grafos

Es una extensión de un formato GML anterior que no estaba basado en XML

Es parecido al formato SVG de imágenes y podría interpretarse como "el PDF de los grafos/redes"

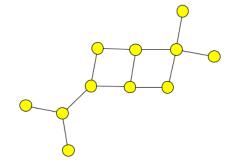
El objetivo final de diseño se resume en la siguiente frase: "El formato de intercambio de grafos debe poder representar grafos arbitrarios con información adicional arbitraria, incluyendo información de layout y visualización. La información adicional debe poder guardarse en un formato apropiado para la aplicación especifica, pero no debe complicar o interferir con la representación de datos para otras aplicaciones..."

La información de visualización de la red se incluye como una extensión, separada de los datos de la red para permitir que éstos puedan ser leídos por cualquier aplicación

EJEMPLO DE FICHERO GraphML (1)

En ambos casos, como en el formato NET, la información del grafo se guarda en dos partes, una para los nodos y otra para los enlaces, con sus correspondientes metadatos

En GraphML el fichero lleva también una cabecera:



Para almacenar una red dirigida basta con cambiar el tipo de red (el resto es igual):

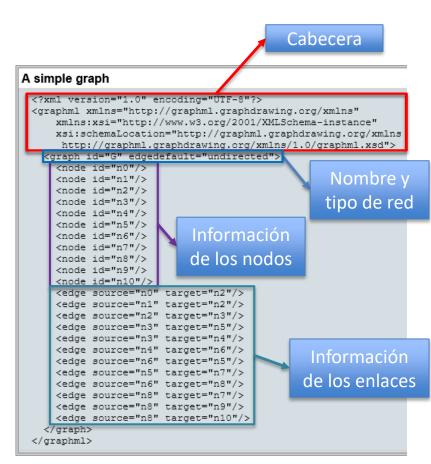
```
The definition of the graph

<graph id="G" edgedefault="directed">
```

Si es mixta, se especifica para cada enlace:

```
An edge with all XML-Attributes defined

...
<edge id="e1" directed="true" source="n0" target="n2"/>
...
```



EJEMPLO DE FICHERO GraphML (2)

Para almacenar una red ponderada hay que declarar un atributo asociado a los enlaces:

```
Declaration of a GraphML Attribute

...

<key id="d1" for="edge" attr.name="weight" attr.type="double"/>
...
```

De la misma forma, se pueden definir otros atributos como el color de los nodos e incluso darles un valor por defecto:

2.0

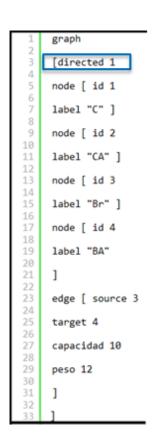
```
Example of a GraphML Document with GraphML-Attributes
 <?xml version="1.0" encoding="UTF-8"?>
 <graphml xmlns="http://graphml.graphdrawing.org/xmlns"</pre>
       xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
       xsi:schemaLocation="http://graphml.graphdrawing.org/xmlns
          http://graphml.graphdrawing.org/xmlns/1.0/graphml.xsd">
    <key id="d0" for="node" attr.name="color" attr.type="string":</pre>
      <default>yellow</default>
   <key id="d1" for="edge" attr.name="weight" attr.type="double"/:</pre>
    <graph id="G" edgedefault="undirected"</pre>
      <node id="n0">
       <data key="d0">green</data>
     </node>
      <node id="n1"/>
      <node id="n2">
                                             Asignando color
       <data kev="d0">blue</data>
      </node>
                                              azul al nodo n2
      <node id="n3">
       <data key="d0">red</data>
     </node>
      <node id="n4"/>
      <node id="n5">
       <data key="d0">turquoise</data>
      <edge id="e0" source="n0" target="n2">
       <data key="d1">1.0</data>
     <edge id="e1" source="n0" target="n1"
       <data key="d1">1.0</data>
      <edge id="e2" source="n1" target="n3">
       <data kev="d1">2.0</data>
                                                  Asignando
      <edge id="e3" source="n3" target="n2"/>
                                                  el peso 1.0
      <edge id="e4" source="n2" target="n4"/>
      <edge id="e5" source="n3" target="n5"/>
                                                 al enlace d1
      <edge id="e6" source="n5" target="n4">
       <data kev="d1">1.1</data>
     </edge>
    </graph>
 </graphml>
```

LOS FORMATOS GraphML Y GML (2)

El formato GML es similar al GraphML, con una estructura más sencilla y cambiando metadatos. Hay una cabecera simple opcional, después el tipo de red (directed 0 indica no dirigida y directed 1 indica dirigida) y la información de nodos y enlaces

Los pesos se definen también como atributos:

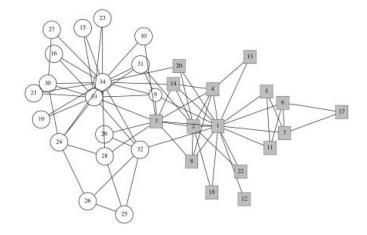
```
Creator "Diana on Wed Sep 17 15:04:20
2015"
graph
  directed 0
  node
       id 0
   label "M de Varela" ]
  node
       id 1
   label "Universidad Central"]
  node
       id 2
   label "Seminario Mayor" ]
edge
    source 0
    target 1
  edge
    source 1
    target 2
```



EJEMPLO DE FICHERO GML (1)

Red social del club de karate de Zachary

Zachary. An information flow model for conflict and fission in small groups. Journal of Anthropology Research 33: 452-473 (1977)

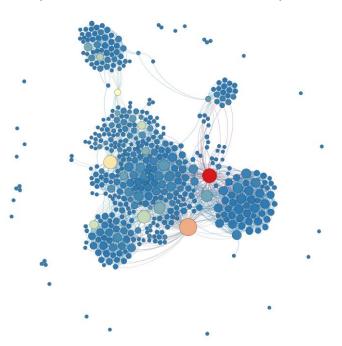


- Red no dirigida y no ponderada
- La visualización mostrada está hecha en Gephi, no se almacena en el fichero GML que solo guarda información estructural de la propia red

```
Creator "Mark Newman on Fri Jul 21 12:39:27 2006"
graph
directed
 node
    id 1
  node
    id 2
 node
    id 34
 edge
    source 2
    target 1
  edge
    source 3
    target 1
 edge
    source 34
    target 33
```

EJEMPLO DE FICHERO GML (2)

Red social de amistad en Facebook de Oscar Cordón (la veremos en el Tema 3)



- Red no dirigida y no ponderada (no hace falta poner tipo)
- La visualización mostrada está hecha en Gephi, no se guarda en el fichero GML

```
directed 0
              id 0
              label "Xavier Llorá "
      node [
              id 1
              label "Elena Rodriguez"
node [
              id 514
              label "Diego Santiago Camposo"
      node [
              id 515
             label "Rubén Díaz"
       edge
              source 0
              target 10
      edge [
              source 1
              target 22
       edge [
              source 512
              target 513
```

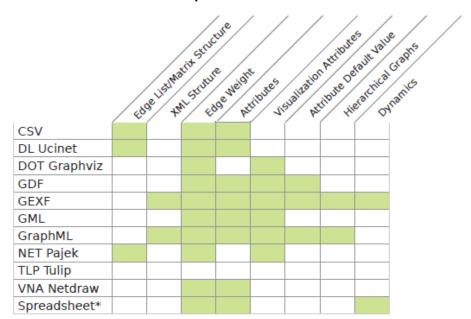
LOS FORMATOS de GEPHI (1)

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



Gephi soporta muchos tipos de archivo diferentes, lo que lo hace compatible con muchas herramientas de análisis de redes sociales. Cada formato permite almacenar cierta información:

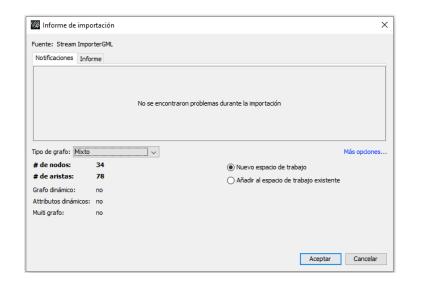
- <u>GEXF</u>
- GDF
- GML
- GraphML
- Pajek NET
- GraphViz DOT
- CSV
- UCINET DL
- Tulip TPL
- Netdraw VNA
- Spreadsheet

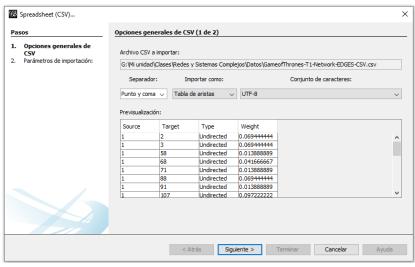


Algunos ya los hemos visto. Del resto, vamos a estudiar el formato GEFX, que es el formato XML propio de *Gephi*, y el CSV, por su facilidad de manejo y potencialidad

LOS FORMATOS de GEPHI (2)

Para abrir cualquier fichero que no sea un GEFX se escoge la opción *Abrir* del menú *Archivo*. *Gephi* abrirá una ventana de importación (*wizard*) que nos guiará en el proceso:





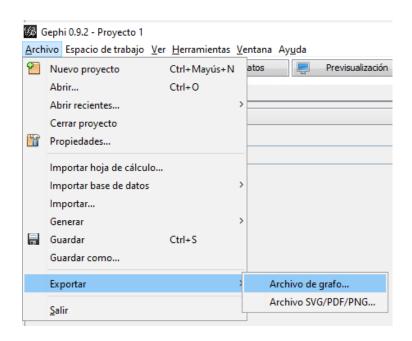
En general, este wizard importa directamente toda la información del fichero de red

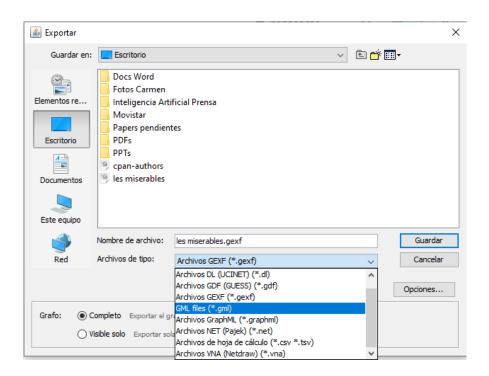
Como mucho, tendremos que decirle el tipo de red, si queremos importar toda la información o solo una parte, y si queremos añadir la red al espacio de trabajo (ventana de *Gephi*) actual o a uno nuevo

LOS FORMATOS de GEPHI (3)

También podemos exportar nuestra red a cualquiera de los formatos manejados usando la opción *Exportar* del menú *Archivo*

Basta con elegir la opción *Archivo de grafo...* y luego escoger el formato de fichero deseado en la lista *Archivos de tipo* de la ventana *Exportar*.





EL FORMATO GEFX

http://gexf.net/

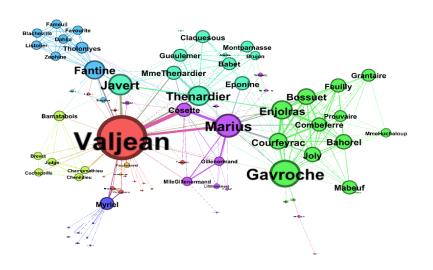
El formato de fichero GEFX es un estándar de XML desarrollado por la propia comunidad de *Gephi*, que lo presenta como un formato simple y específico para redes y grafos

Es bastante similar a los ya estudiados con otros metadatos. A la parte de datos específicos de la red, se puede añadir información de visualización que *Gephi* interpreta directamente:

```
<?xml version="1.0" encoding="UTF-8"?>
<qexf xmlns="http://www.qexf.net/1.2draft" version="1.2">
    <meta lastmodifieddate="2009-03-20">
        <creator>Gexf.net</creator>
        <description>A hello world! file</description>
    </meta>
   <graph mode="static" defaultedgetype="directed">
        <nodes>
            <node id="0" label="Hello" />
            <node id="1" label="Word" />
        </nodes>
        <edges>
            <edge id="0" source="0" target="1" />
        </edges>
   </graph>
</gexf>
```

EJEMPLO DE FICHERO GEFX

Red social de personajes de la novela Los Miserables de Victor Hugo



- Red no dirigida y ponderada (frecuencia de co-ocurrencias de personajes)
- En este caso, la información de visualización y de detección de comunidades está incluida en el propio fichero para Gephi

```
<?xml version="1.0" encoding="UTF-8"?>
<gexf xmlns="http://www.gexf.net/1.3" version="1.3"</pre>
xmlns:viz="http://www.gexf.net/1.3/viz"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.gexf.net/1.3
http://www.gexf.net/1.3/gexf.xsd">
  <meta lastmodifieddate="2021-09-07">
    <creator>Gephi 0.9</creator>
    <description></description>
  </meta>
 <graph defaultedgetype="undirected" mode="static">
    <attributes class="node" mode="static">
      <attribute id="modularity class" title="Modularity Class"
type="integer"></attribute>
    </attributes>
    <nodes>
      <node id="11" label="Valjean">
        <attvalues>
          <attvalue for="modularity class" value="1"></attvalue>
        </attvalues>
        <viz:size value="100.0"></viz:size>
        \langle viz:position x="-207.62332" y="-8.691125"></viz:position>
        <viz:color r="245" q="91" b="91"></viz:color>
      </node>
 </nodes>
    <edges>
      <edge id="0" source="1" target="0"></edge>
      <edge id="1" source="2" target="0" weight="8.0"></edge>
      <edge id="2" source="3" target="0" weight="10.0"></edge>
      <edge id="3" source="3" target="2" weight="6.0"></edge>
      <edge id="4" source="4" target="0"></edge>
    </edges>
  </graph>
</gexf>
```

EL FORMATO CSV (1)

https://es.wikipedia.org/wiki/Valores_separados_por_comas

El formato CSV es un estándar de intercambio de datos, por ejemplo entre programas de bases de datos o de hojas de cálculo como *Microsoft Excel*

Se basa en almacenar los datos de una tabla como un fichero de texto que incluye los datos en cada línea separados por coma (,) o punto y coma (;)

La primera línea contiene las cabeceras (nombres) de los campos (columnas) de la tabla. El resto de líneas incluyen los datos de cada fila, siempre con el mismo tamaño y separados por

coma/punto y coma:

Año	Marca	Modelo	Descripción	Precio
1997	Ford	E350	ac, ABS, moon	3000.00
1999	Chevy	Venture	Extended Edition	4900.00
1999	Chevyr	Venture	Extended Edition, Very Large	5000.00
1996	Jeep	Grand Cherokee	MUST SELL! air, moon roof, loaded	4799.00

Año, Marca, Modelo, Descripción, Precio
1997, Ford, E350, "ac, abs, moon", 3000.00
1999, Chevyr, Venture, Extended Edition, 4900.00
1999, Chevy, Venture, "Extended Edition, Very Large", 5000.00
1996, Jeep, Grand Cherokee, "MUST SELL! air, moon roof, loaded", 4799.00

EL FORMATO CSV (2)

https://gephi.org/users/supported-graph-formats/csv-format/

Este formato se puede usar para almacenar redes. Es muy cómodo, dado que podemos crear una red directamente desde *Excel* o adquirir datos de medios sociales como *Twitter* (hacer *scrapping*), guardarlos en formato CSV y leerlos con *Gephi* para trabajar con la red descargada

Gephi admite distintos tipos de CSV para definir redes. En el más básico, se almacenan solo los enlaces, con tres formatos distintos:

- Listas de enlaces (pares nodo origen nodo destino):
 Ej. Dos enlaces (a,b) y (b,c)
- Listas de adyacencia (lista de enlaces) para cada nodo: Ej. Tres enlaces (a,b), (b,c) y (b,d)
- Matriz de adyacencia. Ej: Seis enlaces (a,b),
 (a,d), (b,a), (c,c), (d,b) y (d,d)

;A;B;C;D;E A;0;1;0;1; B;1;0;0;0; C;0;0;1;0; D;0;1;0;1; E;0;0;0;0;0;

b;c;d

EL FORMATO CSV (3)

También se pueden guardar atributos de nodos, pesos de enlaces, ... usando columnas adicionales. En ese caso, lo mejor es usar dos ficheros CSV, uno de nodos y otro de enlaces:

- Tabla de nodos: Una lista con los nodos de la red con dos campos: Id y Label. El campo Id es numérico y sirve como identificador único del nodo para Gephi, es obligatorio y no debe tener repetidos. Label almacena la etiqueta del nodo
 - Se pueden añadir campos adicionales con otros atributos de los nodos
- Tabla de enlaces: Una lista con los enlaces de la red con dos campos: Source y Target, nodo origen y nodo destino, respectivamente
 - Los enlaces también pueden tener atributos adicionales como el peso, Weight, el tipo, Type (directed, undirected); o la etiqueta, Label

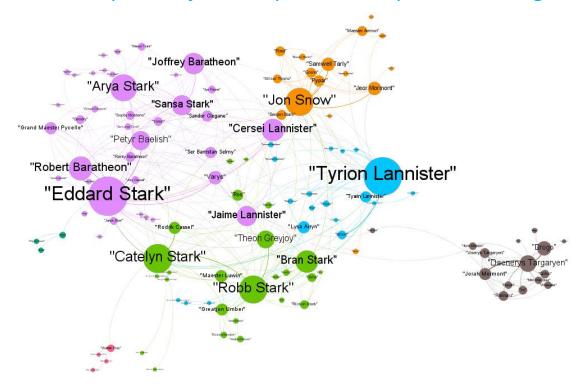
Mientras que en los ficheros de enlaces se suele usar el separador punto y coma (;), en este caso se pueden usar tanto el separador punto y coma (;) como el coma (,)

Hay disponible un tutorial muy completo en:

https://seinecle.github.io/gephi-tutorials/generated-html/importing-csv-data-in-gephi-en.html

EJEMPLO DE FICHERO CSV (1)

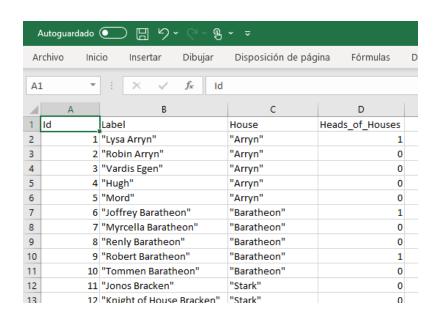
Red social de personajes de la primera temporada de Juego de Tronos



- Red no dirigida y ponderada (frecuencia de co-ocurrencias de personajes)
- La información de visualización (posiciones, colores y tamaños de nodos) y de detección de comunidades no está incluida en los ficheros CSV, se ha hecho en Gephi

EJEMPLO DE FICHERO CSV (2)

Fichero CSV de nodos

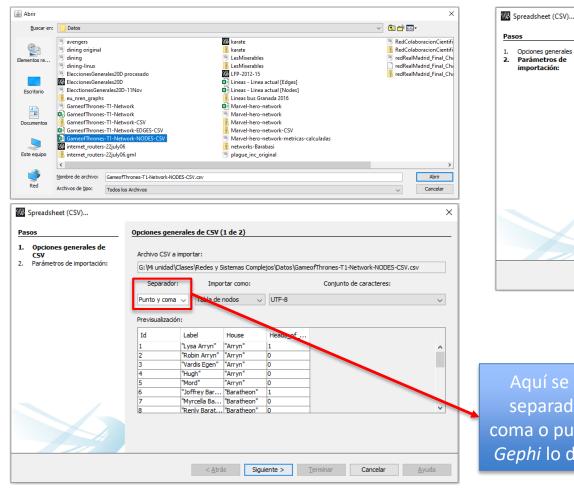


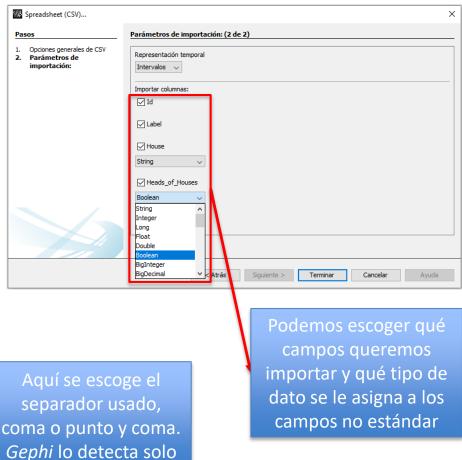
```
GameofThrones-T1-Network-NODES-CSV: Blo...
Archivo Edición Formato Ver Ayuda
Id;Label;House;Heads of Houses
1;"""Lysa Arryn"";"""Arryn""";1
2; """Robin Arryn""; """Arryn""; 0
3;"""Vardis Egen""";"""Arryn""";0
4;"""Hugh""";"""Arryn""";0
5;"""Mord""";"""Arryn""";0
6;"""Joffrey Baratheon""";"""Baratheon""";1
7;"""Myrcella Baratheon""";"""Baratheon""";0
8;"""Renly Baratheon""";"""Baratheon""";0
9;"""Robert Baratheon""";""Baratheon""";1
10;"""Tommen Baratheon""";"""Baratheon""";0
11; """Jonos Bracken""; """Stark"""; 0
12;"""Knight of House Bracken""";"""Stark""";0
13;"""Janos Slynt""";"""Baratheon""";0
14;"""Gregor Clegane""";"""Clegane""";1
15;"""Sandor Clegane""";"""Clegane""";0
16;"""Joss Stilwood""";"""Clegane""";0
                  100%
Línea 1, columna 1
                        Windows (CRLF)
                                        UTF-8
```

Además de los campos Id y Label, se incluyen otros dos atributos con información de los nodos:
 House, que almacena la Casa Nobiliaria a la que pertenece el personaje, y Head of Houses, valor
 binario que vale 1 si el personaje es el Jefe de la casa y 0 en caso contrario

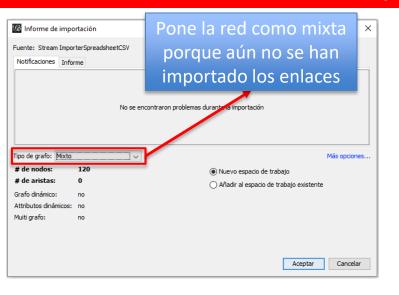
EJEMPLO DE FICHERO CSV (3)

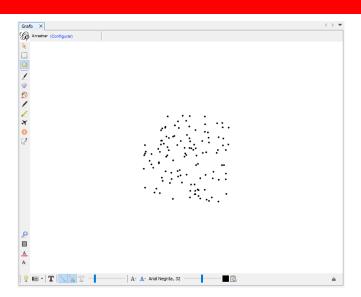
Primero abrimos el fichero CSV con la información de los nodos:

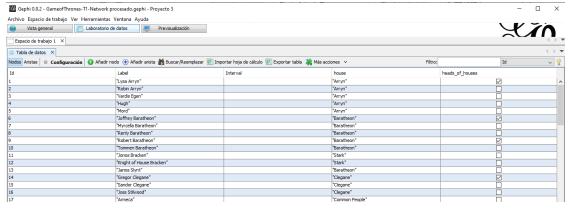


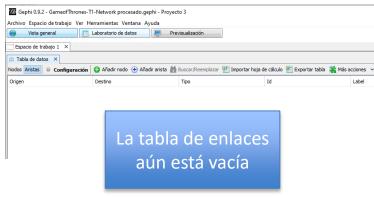


EJEMPLO DE FICHERO CSV (4)





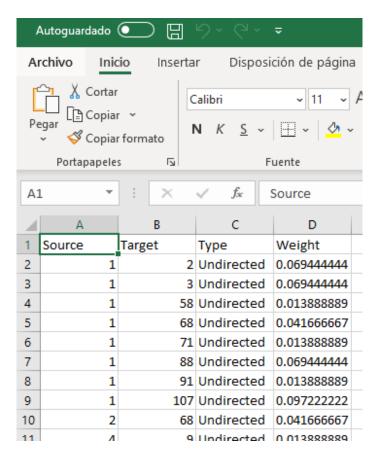


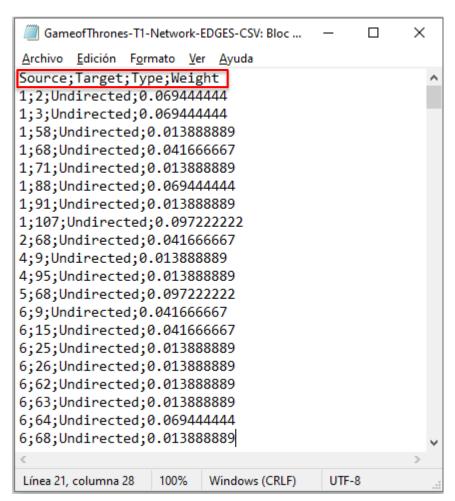


EJEMPLO DE FICHERO CSV (5)

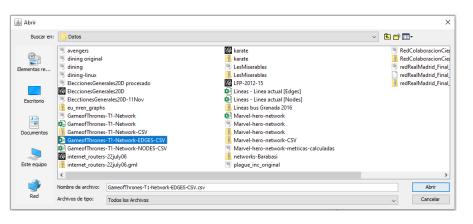
Ahora abrimos el fichero CSV con la información de los enlaces:

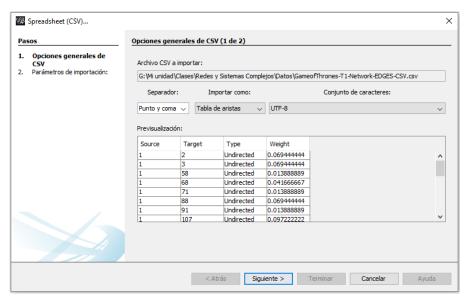
Fichero CSV de enlaces

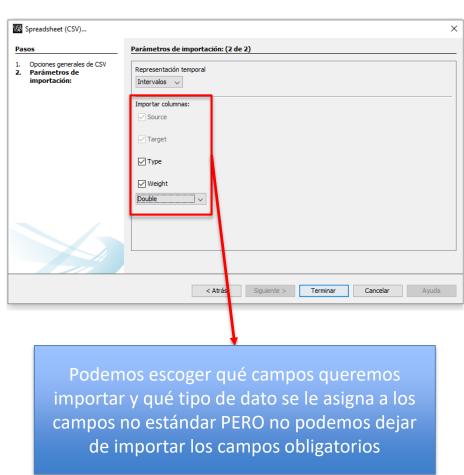




EJEMPLO DE FICHERO CSV (6)

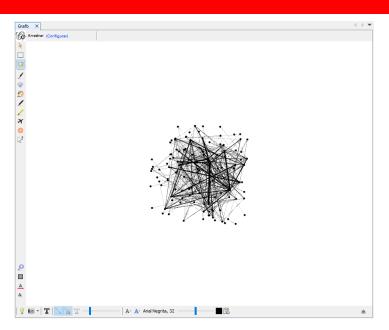


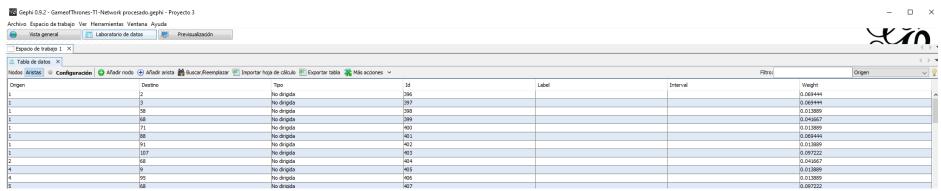




EJEMPLO DE FICHERO CSV (7)







EL FORMATO JSON

https://en.wikipedia.org/wiki/JSON



Existe otro formato de fichero muy común, el JavaScript Object Notation (JSON)

Es interesante conocerlo porque es el formato habitual de los datos extraídos de Twitter

Es similar a muchos de los estudiados, un fichero de texto con metadatos en el que la información de los nodos está en el bloque nodes y la de los enlaces en el bloque edges con metadatos como source, target y weight

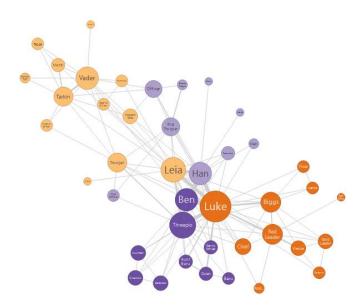
Puede incorporar información específica de visualización y datos adicionales que solo pueden manejar aplicaciones específicas

Gephi no maneja este formato pero puede exportar redes a JSON usando un plugin adicional

Para cargar una red en formato JSON en Gephi tendremos que transformarla a un formato que *Gephi* reconozca

EJEMPLO DE FICHERO JSON

Red social de personajes de la película Star Wars: Episode IV - A New Hope https://moviegalaxies.com/



- Red no dirigida y ponderada (frecuencia de co-ocurrencias de personajes)
- La información de visualización, detección de comunidades y otras medidas está incluida en el propio fichero JSON

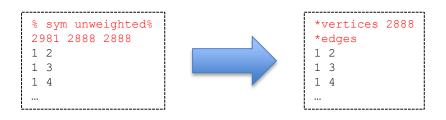
```
"network": {
  "edges": [
      "source": "BEN",
      "target": "GANTRY OFFICER",
      "weight": 1
  "nodes": [
      "id": "BEN",
      "name": "BEN",
      "color": "#5E3C99",
      "group": 0,
      "degree": 13,
      "pagerank": 0.07092876242260801,
      "object id": "4752562",
      "triangles": 35,
      "eccentricity": 3.0,
      "closeness centrality": 1.7073170731707317,
      "betweenness centrality": 61.05190540453698,
      "eigenvector centrality": 0
  "positions": [
      "x": -9.6373625,
      "v": 68.279976,
      "name": "BEN"
```

CONVERSIÓN DE FORMATOS: EDICIÓN MANUAL DEL FICHERO

Los ficheros de redes son de texto. Por tanto, podemos modificarlos directamente con un editor de texto para transformarlos al formato deseado y así leerlos con *Gephi*

P.ej., podemos pasar de cualquier fichero de texto que incluya una lista de enlaces a un formato NET editándolo para incluir el número de nodos con un *vertices y luego incorporar un *edges para redes no dirigidas (*arcs para redes dirigidas)

Ejemplo: Red social de amistad en *Facebook* (http://konect.cc/networks/ego-facebook/), fichero de texto *out.ego-facebook.txt* transformado a formato NET:



Si el fichero original tiene una matriz de adyacencia (poco habitual) incluiremos un *matrix

Si además tuviéramos información sobre los nodos como su etiqueta, la incorporaríamos detrás del *vertices:

```
*Vertices 82670
1 "entity"
2 "thing"
3 "anything"
4 "something"
5 "nothing"
6 "whole"
```

CONVERSIÓN DE FORMATOS (2): JSON A CSV

Para manejar ficheros de red almacenados en formato JSON podemos transformarlos a CSV y luego importarlos en *Gephi*

Para ello usamos un conversor de formato JSON a CSV. Hay muchos gratuitos disponibles online, p.ej:

https://www.aconvert.com/es/document/json-to-csv/

https://data.page/json/csv

https://www.convertcsv.com/json-to-csv.htm

Usando el primer y el tercer conversor, el fichero JSON de la red de *Star Wars* quedaría así

Se pierde la información de los nodos pero *Gephi* puede cargarlo directamente y recupera los nombres

Con el segundo conversor se puede recuperar el resto de información pero Gephi no los carga directamente, así que hay que editar el CSV a mano para darle el formato esperado

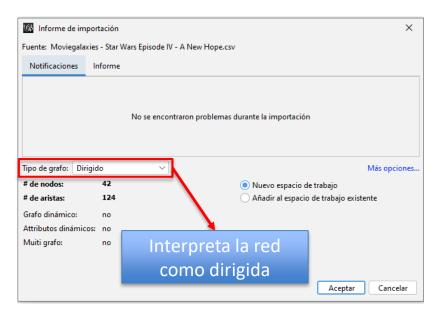
```
source, target, weight
BEN, GANTRY OFFICER, 1
BEN, BIGGS, 1
BEN, HUMAN, 1
BEN, RED LEADER, 1
BEN, HAN, 8
BEN, BARTENDER, 1
BEN, VADER, 1
BEN, TROOPER, 1
BEN, CREATURE, 1
BEN, CHIEF, 1
BEN, THREEPIO, 5
BEN, LEIA, 2
BEN, LUKE, 15
GANTRY OFFICER, HAN, 1
GANTRY OFFICER, LUKE, 1
GANTRY OFFICER, THREEPIO, 1
OWEN, LUKE, 3
OWEN, THREEPIO, 1
OWEN, BERU, 1
OWEN, AUNT BERU, 2
BIGGS, CAMIE, 1
BIGGS, WEDGE, 1
BIGGS, FIXER, 1
BIGGS, CHIEF, 1
BIGGS, RED LEADER, 2
BIGGS, THREEPIO, 1
BIGGS, LEIA, 1
BIGGS, LUKE, 4
BIGGS, GOLD LEADER, 1
HUMAN, LUKE, 1
HUMAN, THREEPIO, 1
HUMAN, CREATURE, 1
HUMAN, BARTENDER, 1
```

CONVERSIÓN DE FORMATOS (3): JSON A CSV

Una vez creado el fichero CSV, podemos luego importarlos en *Gephi* de forma sencilla usando el *wizard* que hemos visto antes

¡OJO! Como al hacer la conversión del fichero JSON no se indica explícitamente de qué tipo es la red, puede ocurrir que *Gephi* la interprete de forma incorrecta

Por ejemplo, esto es lo que ocurre al cargar el fichero CSV transformado de la red social de personajes de la película Star Wars: Episode IV - A New Hope anterior:



No hay que dejar que Gephi elija el tipo de red. Puede equivocarse

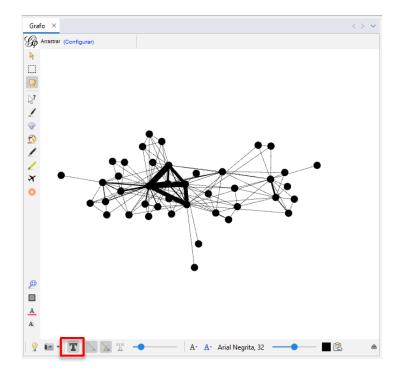
Vosotros sois los sociólogos, los que conocéis el sistema social representado en la red y sois los que tenéis que indicárselo

CONVERSIÓN DE FORMATOS (4): JSON A CSV

Otro pequeño problema que puede ocurrir es que al importar los nombres de los nodos, los incluya en el campo *Id* en vez de en el campo *Label*, que es lo habitual

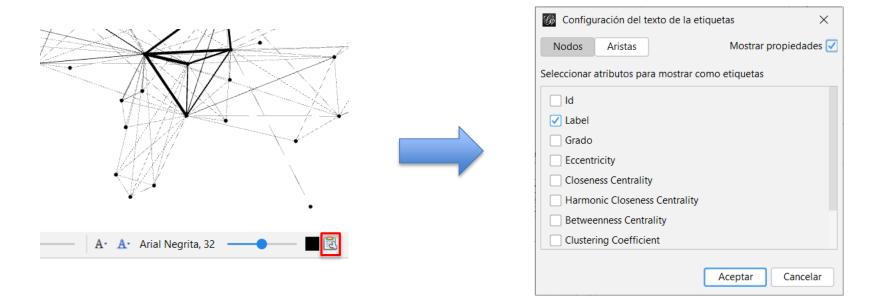


En ese caso, si intentamos visualizar las etiquetas, nos saldrán en blanco:



CONVERSIÓN DE FORMATOS (5): JSON A CSV

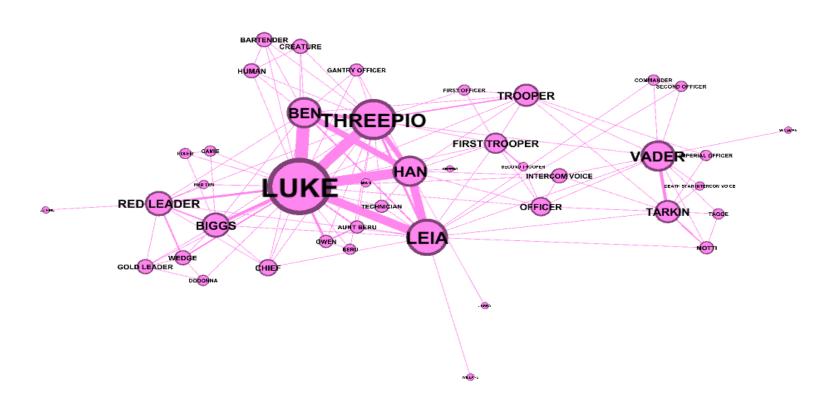
Esto se soluciona de forma sencilla indicándole a *Gephi* cuál es el campo que queremos visualizar como etiqueta de los nodos. Para ello, pinchamos en el icono con la llave inglesa que aparece en la parte inferior derecha:



Con eso se abre una nueva ventana en la que escoger el campo (o campos) de la tabla de datos que queremos visualizar como etiqueta de los nodos

CONVERSIÓN DE FORMATOS (6): JSON A CSV

Después de escalar el tamaño de los nodos de acuerdo con su grado, cambiar su color a rosa, escalar el tamaño de las etiquetas con el tamaño del nodo y ejecutar de nuevo el algoritmo de distribución, quedaría:



ULTIMO PASO: GRABAR EN FORMATO GEPHI

Una vez que importemos un fichero de red en cualquiera de los formatos anteriores en *Gephi*, ya podremos empezar a trabajar con él en la herramienta

Como sabemos, cualquier cálculo que hagamos generará resultados en el cuadro *Estadísticas* y nuevas columnas en las tablas del *Laboratorio de Datos*, y cualquier operación de visualización se representará en la ventana *Grafo* de la *Vista General*

Para seguir trabajando con esa red las siguientes veces, tendremos que guardarla en formato Gephi. Pulsaremos en *Guardar* del menú *Archivo* y le pondremos el nombre que queramos

Con esto ya se almacenará como proyecto Gephi y mantendrá todo lo que hayamos hecho

Cuando la carguemos no hará falta importar información de nuevo pero el fichero solo será válido para *Gephi*

