

Técnicas Basadas en Grafos Aplicadas al Procesamiento del Lenguaje Natural

Tema 2

La primera tarea del curso consiste en el análisis de una serie de grafos con el objetivo de conocer algunas de sus características. La herramienta que vamos a utilizar para analizar los grafos se llama *Gephi*. *Gephi* es una herramienta para la visualización y consulta de grafos, así como el análisis de redes sociales. Es una aplicación de código abierto que incluye un paquete de software de visualización escrito en lenguaje Java que se puede descargar gratuitamente desde el siguiente enlace: <http://gephi.org>. También se puede encontrar documentación sobre la herramienta en la sección “Learn” de dicha página Web. Ahí podréis encontrar tutoriales, video-tutoriales, ejemplo de grafos, foros e incluso una lista de correo.

Para esta tarea disponéis de 4 grafos que podréis encontrar en el curso virtual:

- **EuroSiSGenerale.gephi:** Grafo que refleja las relaciones entre diferentes asociaciones científicas Europeas generado por EuroSIS.
- **RetGraph.gephi:** Grafo que muestra las apariciones comunes o co-ocurrencias entre palabras en una misma frase sobre un conjunto de documentos.
- **LesMiserables.gephi:** Red de co-aparición ponderada de los personajes de la novela Los Miserables.
- **BoundaryCountries.gephi:** Grafo generado por un proyecto Europeo que tenía el objetivo de asignar países a un determinado distrito.

Se pide un documento PDF en el que se incluya:

- Un informe para cada uno de los grafos, en el que se represente sus principales características. Estas características podrán variar dependiendo de si el grafo es dirigido o no, o si se trata de un grafo con pesos o sin pesos. Se deben mostrar obligatoriamente los siguientes datos:
 - Nodos y Aristas
 - Si se trata de un grafo dirigido o no.
 - Si se trata de un grafo con pesos o no.
 - Longitud Media de los Caminos entre nodos
 - Coeficiente de Clustering Medio
 - Diámetro de la Red
 - Grado Medio
 - Grado Medio con Pesos
 - Número de componentes conexas fuerte y débilmente conectadas.
- Basándose en las características anteriormente citadas y alguna otra que el alumno crea oportuna, será necesario completar el informe de cada grafo con una reflexión acerca de su distribución del grado y una posible clasificación según los modelos de red o tipos de grafo presentados en el curso.