PROPUESTA DE ACTIVIDADES DE FORMACIÓN DOCTORAL ESPECÍFICA 2019

|  |
| --- |
|  |
| Título de la actividad |
| **Taller de Análisis de Big Data: Análisis de Redes Sociales y Análisis Semántico en Twitter** |
| Programa(s) de doctorado que proponen la actividad |
| 1. Comunicación Social
 |
| Persona de contacto |
| Nombre y apellidos **Camen Peñafiel Saiz** |  |
| Teléfono: **946012362** |  |
| Email: carmen.penafiel@ehu.eus |  |
| Breve descripción de la actividad |
| Calendario |
| **Sesión 1 (4 horas con 30 minutos de descanso)****Bloque 1: Introducción*** Almacenamiento y tratamiento de datos masivos: de reto tecnológico a reto interpretativo.
* El Análisis de Redes Sociales (ARS), breve historia, fundamentos y ejemplos de aplicación.
* ARS y Twitter: casos de estudio para las Ciencias de la Comunicación.

**Bloque 2: La obtención del dato*** Estrategias de obtención de datos tradicionales vs. computacionales.
* Captura de datos con NodeXL (Actividad Práctica)

**Bloque 3: El tratamiento del dato para el análisis relacional*** Indicadores básicos del ARS en Datos Masivos.
* Algoritmos de visualización de datos
* Análisis de datos con GEPHI (Actividad Práctica)

**Sesión 2 (4 horas con 30 minutos de descanso)****Bloque 4: La interpretación del dato*** Algunos indicadores avanzados para del ARS en Datos Masivos
* Estrategias de depuración tradicionales, semánticas y algorítmicas.
* Análisis de datos con GEPHI (Actividad Práctica).

**Bloque 5: Minería de datos para la generación de variables*** Estrategias de cruce de datos en Excel y Orange Datamining
* Introducción al Análisis de Sentimiento y al Topic Modeling con Orange Datamining (Python)

**Bloque 6. Conclusiones y cierre** **Necesidades para el taller:**El curso está planteado como un taller en el que los alumnos van a tener que descargar y analizar algunos conjuntos de datos. Por ello, son necesarios los siguientes elementos:* Excel instalado en los ordenadores + la versión que adjunto del programa NodeXL. Es importante que la versión sea exactamente la de abajo, y no la más nueva.
	+ <https://mega.nz/#!GEMy3SAR>!bGRmrZSLkR8jlIoHfed9W2xCSrDrHETxgnYBmteUbug
* Gephi instalado en los ordenadores. Requiere de Java. Que los informáticos se cercioren de que todo está instalado en versión 64 bits.
	+ <https://gephi.org/>
	+ <https://www.java.com/ES/download/>
* Orange Datamining instalado en los ordenadores. Requiere Anaconda Python que se instala con el mismo ejecutable.
	+ <https://orange.biolab.si/download/>

Esperamos una asistencia de 50 personas, por lo que dividiríamos al grupo en dos A y B, cada grupo asistiría 2 días a dos sesiones (por ejemplo: GA 16 y 17 de octubre de 2019 ; y GB 23 y 24 de octubre de 2019). |
| Lugar de impartición |
| Sala Multimedia Catilab de la Facultad de la Facultad de Ciencias Sociales y de la Comunicación |