

Línea de Especialización: Ingeniería de Datos, Información y Conocimiento

Responsable: Claudio Gutiérrez

20 de Julio de 2012

1. Profesores Integrantes

Pablo Barceló, Benjamín Bustos, Claudio Gutiérrez, Gonzalo Navarro, Jorge Pérez y Bárbara Poblete.

2. Objetivos

Las tecnologías de generación y captura de datos, así como las de producción de información, han dado un salto exponencial en la última década, en un fenómeno conocido como *Diluvio de Datos*.

Se hace necesario tener especialistas (usualmente conocidos en inglés como *Data Scientists*) que sean capaces de trabajar con ellos. Esto involucra todo el ciclo de manejo de datos: publicación, presentación, intercambio, procesamiento, consulta, análisis, almacenamiento y archivo.

El objetivo de esta línea de especialización es la formación de un especialista con expertise y dominio de herramientas conceptuales y prácticas para asumir esos desafíos.

En particular, el alumno debería comprender y conocer las siguientes áreas, y dependiendo de los cursos obligatorios y electivos que elija, tener competencias en una o más de ellas:

- Datos en la Web
 - Estándares de Publicación de datos masivos;
 - Presentación y Visualización de datos masivos;
 - Almacenamiento y Archivo de datos masivos;
 - Organización, Procesamiento y Consulta de datos masivos.

- Minería de Datos
 - Intercambio y Transmisión de datos masivos;
 - Extracción de Conocimiento/Información de datos en bruto;
 - Limpieza y Preparación de datos masivos;
 - Organización, Procesamiento y Consulta de datos masivos;
 - Procesamiento y Análisis de datos masivos.

- Teoría
 - Modelamiento conceptuales sobre datos;
 - Organización, Procesamiento y Consulta de datos;
 - Integración e Intercambio de datos.

Todo lo anterior para diferentes formatos y tipos de datos (texto, multimedia, datos estructurados, no estructurados, etc.).

3. Campo Laboral

Los egresados de esta línea de especialización podrán ejercer en empresas, centros de investigación y desarrollo, organizaciones y proyectos que requieran el dominio de procesos y tecnologías para el manejo y procesamiento de grandes volúmenes de datos.

4. Cursos

A continuación la lista de cursos del área que se han dictado en los últimos dos años. El alumno debe completar 50 UD con los cursos indicados más abajo, de los cuales 20 UD deben ser de los indicados como “obligatorios”:

4.1 Cursos Obligatorios (≥ 20 UD)

- CC5201 Bases de Datos Multimedia
- CC5203 Recuperación de la Información
- CC5206 Minería de Datos
- CC5207 Teoría de Bases de Datos
- CC6202 La Web de Datos (Linked Data)

4.2 Cursos Optativos (30 UD)

El alumno debe completar 30 UD con cursos optativos. (Si el alumno completó $n > 20$ UD con cursos obligatorios, entonces necesita sólo completar las 50 UD con cursos optativos).

- CC4501 Astroinformática
- CC5108 Estructuras de Datos Comprimidas
- CC5204 Búsqueda por Contenido de Imágenes y Videos
- CC5312 Seguridad de Datos
- CC6102 Seminario de Estructuras de Datos Compactas

Y otros Talleres de Bases de Datos no convencionales, visualización de datos, técnicas recientes en el área.