

# 62225 - Manipulación y análisis de grandes volúmenes de datos

## Información del Plan Docente

**Año académico:** 2024/25

**Asignatura:** 62225 - Manipulación y análisis de grandes volúmenes de datos

**Centro académico:** 110 - Escuela de Ingeniería y Arquitectura

**Titulación:** 534 - Máster Universitario en Ingeniería Informática

**Créditos:** 6.0

**Curso:** 1

**Periodo de impartición:** Segundo semestre

**Clase de asignatura:** Obligatoria

**Materia:**

## 1. Información básica de la asignatura

Con el desarrollo de la asignatura se pretenden lograr los siguientes objetivos:

- Que el estudiante analice, dado un problema que implica grandes volúmenes de datos, los requisitos necesarios para su gestión (almacenamiento, transferencia, procesamiento, visualización e interacción).
- Que el estudiante desarrolle los elementos necesarios para integrar fuentes de datos heterogéneas, utilizando técnicas clásicas para la representación, tratamiento, análisis, visualización e interacción con repositorios de datos heterogéneos.
- Que el estudiante desarrolle una aplicación para un contexto dado, donde sea necesaria la gestión de grandes volúmenes de datos, y teniendo en cuenta criterios de escalabilidad, usabilidad y normativos.

## 2. Resultados de aprendizaje

El estudiante, para superar esta asignatura, deberá demostrar los siguientes resultados...

1. Comprender y especificar los requisitos necesarios para la interacción, almacenamiento, transferencia y procesado de grandes volúmenes de datos.
2. Conocer, comprender y aplicar las técnicas más comunes para la representación, tratamiento, análisis e interacción con repositorios de datos heterogéneos.
3. Diseñar, desarrollar y evaluar una aplicación que facilite la elaboración y gestión de grandes volúmenes de datos, conforme a criterios de escalabilidad y normativa existente.

## 3. Programa de la asignatura

- Introducción y motivación al problema de los grandes volúmenes de datos (*Big Data*).
- Almacenamiento de grandes volúmenes de datos:
  - Almacenes de datos (*data warehouses*). Diseño en estrella.
  - Bases de datos NoSQL.
- Gestión de grandes volúmenes de datos:
  - Distribución de los datos.
  - Integración de información con fuentes de datos heterogéneas.
  - Uso de técnicas de representación del conocimiento (ontologías) para la representación de fuentes de datos y su acceso e integración.
  - Técnicas de procesamiento paralelo: MapReduce (Hadoop).
  - Sistemas de gestión de flujos de datos (*data streams*).
  - Otras técnicas: agentes móviles.
- Interacción con grandes volúmenes de datos:
  - Técnicas de visualización.
  - Diseño de interfaces de usuario apropiados.
  - Usabilidad.
- Análisis de grandes volúmenes de datos:
  - Minería de datos.
  - Minería de textos, análisis del sentimiento.
- Casos de uso y aplicaciones, tales como:
  - Datos provenientes de sensores.
  - Datos no estructurados en la Web.

- Sistemas de recomendación.
- Análisis de blogs y redes sociales.
- Ciudades inteligentes (*smart cities*).
- Sistemas de Transporte Inteligentes.

#### 4. Actividades académicas

La asignatura consta de 6 créditos ECTS que suponen una dedicación estimada por parte del estudiante de 150 horas distribuidas del siguiente modo:

- Actividades dirigidas por el profesorado de la asignatura: 60 h (Clase magistral, Resolución de problemas y casos, Prácticas de laboratorio, Prácticas especiales)
- Realización de trabajos de aplicación o investigación prácticos: 65 h
- Estudio de teoría: 20 h
- Pruebas de evaluación: 5 h

#### 5. Sistema de evaluación

El estudiante deberá demostrar que ha alcanzado los resultados de aprendizaje previstos mediante las siguientes actividades de evaluación

- **Realización y presentación de trabajos.** Estudio de un tema relacionado con la asignatura, elaboración de un informe sobre el mismo, y su presentación en clase. [20%]. Resultados de aprendizaje: 1, 2 y 3.
- **Proyecto.** Un proyecto de grupo con ordenadores, en el que se podrán en práctica los conocimientos y habilidades adquiridos en la asignatura. [50%]. Resultados de aprendizaje: 1, 2 y 3.
- **Prueba final escrita incluyendo preguntas de respuesta corta y de respuesta extensa.** [30%]. Resultados de aprendizaje: 1, 2 y 3.

Para promediar, será necesario aprobar individualmente cada una de las actividades de evaluación.

El estudiante que no opte por el procedimiento de evaluación descrito anteriormente, que no supere dichas pruebas durante el periodo docente o que quisiera mejorar su calificación, tendrá derecho a realizar una prueba global que será programada dentro del periodo de exámenes correspondiente a la primera o segunda convocatoria, según corresponda. Dicha prueba global constará de las siguientes actividades de evaluación: 1) realización y presentación del trabajo indicado previamente (estudio de un tema) [20%], 2) realización y defensa del proyecto (prácticas de la asignatura) [50%], y 3) realización de una prueba final escrita incluyendo preguntas de respuesta corta y de respuesta extensa [30%]. Para promediar, será necesario aprobar individualmente cada una de las actividades de evaluación.

#### 6. Objetivos de Desarrollo Sostenible

10 - Reducción de las Desigualdades