

Datos básicos de la materia

Carácter	<input type="checkbox"/> Obligatoria	<input checked="" type="checkbox"/> Optativa	<input type="checkbox"/> Prácticas Externas	<input type="checkbox"/> Trabajo Fin de Máster	<input type="checkbox"/> Mixta
Créditos ECTS	12 créditos ECTS				
Unidad Temporal	<input checked="" type="checkbox"/> Cuatrimestral				
ECTS Cuatrimestre 1	ECTS Cuatrimestre 2	12	ECTS Cuatrimestre 3		
Lenguas en que se imparte	<input checked="" type="checkbox"/> Castellano	<input checked="" type="checkbox"/> Inglés		<input type="checkbox"/> Otra:	

Contenidos

M06

1. Portales de acceso a datos en abierto
2. Registro de datos y metadatos, estándares y herramientas.
3. Descarga de datos, acceso mediante web services y APIs
4. Ejemplos relevantes de portales globales de datos en abierto(GEOSS)
5. Infraestructuras de procesado de datos en abierto.
6. Preservación de datos: preservación física (bit)
7. Preservación de datos: uso de identificadores (PID y DOI)
8. Preservación del software: gestión de código y uso de VM
9. Preservación del análisis y cadena de provenance
10. Estándares de ingestión de datos.
11. Repositorios digitales: ejemplos institucionales.
12. Implementación de un repositorio digital: herramienta Invenio.

13. Herramientas analíticas de explotación de repositorios.

Observaciones

Objetivos:

Esta materia tiene como objetivo proporcionar al estudiante un contacto directo con las tecnologías utilizadas para la implementación de repositorios de datos y su posterior explotación. Se presentan los estándares más relevantes y las iniciativas más visibles en diferentes áreas de conocimiento dentro del contexto Open Science.

Metodología:

En cada asignatura se comenzará por una exposición de los conceptos básicos, incluyendo ejemplos sencillos pero relevantes, que serán analizados individualmente y discutidos en común.

Se revisarán los diferentes componentes de una solución, y los actores que participan en el desarrollo de la misma.

Los estudiantes, organizados en grupos, realizarán un análisis detallado de un caso de estudio empleando las diferentes tecnologías discutidas, por ejemplo el portal GEOSS, el repositorio institucional del CERN basado en Invenio, o los portales de preservación de DPHEP.

Resultados del aprendizaje:

- Identificar la forma de acceder a diferentes repositorios de datos en abierto para abordar un problema de Data Science
- Implementar ejemplos sencillos pero útiles de acceso a portales de datos relevantes directamente o mediante APIs en un contexto de web services.
- Operar un repositorio sencillo con datos históricos. Implementar una versión ya disponible y proponer mejoras en el diseño del mismo.
- Ilustrar cómo el software existente puede usarse para mejorar la calidad de los datos, y en particular su accesibilidad.

Competencias específicas de la especialización:

Hay que indicar que en esta materia se adquirirán la competencia específica “DSDM04 - Desarrollar y mantener un repositorio histórico de datos y de resultados de su análisis, garantizando su procedencia (“provenance”)” que será única para esta especialidad:

Información sobre las asignaturas

Denominación	M06-01 PORTALES Y SERVICIOS DE ACCESO A DATOS EN ABIERTO/ Open Data portals and access services		
Carácter OPTATIVA	ECTS 4	Lengua impartición	castellano
<input checked="" type="checkbox"/> Cuatrimestral			
Denominación	M06-02 PRESERVACIÓN DE DATOS/ Data preservation		
Carácter OPTATIVA	ECTS 4	Lengua impartición	castellano
<input checked="" type="checkbox"/> Cuatrimestral			
Denominación	M06-03 REPOSITORIOS EN ABIERTO/ Open data repositories		
Carácter OPTATIVA	ECTS 4	Lengua impartición	castellano
<input checked="" type="checkbox"/> Cuatrimestral			