



GESTIÓN DE DATOS EXPERIMENTALES

PRESENTACIÓN

En esta asignatura aprendemos a organizar los datos experimentales de forma que se facilite su análisis, tratamiento, archivo y recuperación. Se presta especial atención a la estructuración lógica de los datos en campos y registros, a la optimización de la información, y a la estandarización y descripción de los metadatos para facilitar su integración, localización y reutilización en repositorios y estructuras más amplias de información. La asignatura es eminentemente práctica y basada en casos.

- **Titulación: MÁSTER UNIVERSITARIO EN CIENCIA DE DATOS PARA CIENCIAS EXPERIMENTALES**
- **Módulo y materia:** Módulo III Optativo. Materia 3.1. Optativas
- **Carácter:** Optativa
- **ECTS:** 3
- **Curso y semestre:** Curso 1º y semestre 2º
- **Idioma:** Español, con documentación principalmente en inglés
- **Profesor responsable de la asignatura:** Arturo H. Ariño
- **Profesores:** Arturo H. Ariño, David Galicia
- **Horario y aula:** consultar calendario del máster

RESULTADOS DE APRENDIZAJE

RA2 Diseñar un experimento científico para que sea rico en información, recogiendo gran cantidad de datos de manera estructurada que faciliten su procesamiento posterior

RA07 Integrar la información geoespacial con variables ambientales para abordar cuestiones relacionadas con la crisis de la biodiversidad en un escenario de cambio climático y desarrollo sostenible.

PROGRAMA

1. El ciclo de generación y utilización de datos
 - a. Naturaleza y utilidad de los datos en Ciencias Experimentales.
 - b. El flujo de información en Ciencia.
 - c. Adquisición de datos, digitalización y minería.
 - d. Análisis y explotación.
 - e. Archivo y documentación.
2. Plan de Gestión de Datos (DMP)
 - a. Análisis de necesidades y recursos.
 - b. Contenidos del DMP.
 - c. Herramientas de diseño para DMPs.
 - d. Registro y normativa.
3. Formatos de datos y criterios de almacenamiento
 - a. Codificación binaria y sus derivados. Bits y bytes. Niveles lógico y léxico.
 - b. Almacenamiento de datos numéricos: enteros, racionales, signo.
 - c. Codificaciones léxicas comunes para alfanuméricos: ASCII, UTF.
 - d. Formatos de imagen, audio y vídeo.



- e. Compresión de datos.
- f. Medios físicos, en línea, removibles y en nube.
- 4. Organización lógica y estructuras de datos
 - a. Tablas, archivos y bases de datos.
 - b. Organización eficiente en bloques funcionales.
 - c. Plantillas de adquisición de datos.
 - d. Relaciones y vínculos.
 - e. Metadatos y documentación.
 - f. Entornos colaborativos.
- 5. Sistemas de gestión de bases de datos
 - a. Características.
 - b. Modelos relacionales. Diagramas de relación. Claves e índices.
 - c. Reglas de integridad.
 - d. Transacciones.
- 6. Estándares de intercambio de información
 - a. Concepto, utilidad del estándar. Interoperabilidad y accesibilidad.
 - b. Organismos reguladores
 - c. Ejemplos temáticos: casos en distintas ciencias
 - d. El acceso abierto.
- 7. Seguridad, blindaje y preservación de la información
 - a. Sistemas de copia y salvaguardia. Paradigmas lógicos.
 - b. Automatización y documentación de copias de seguridad.
 - c. Datos sensibles
 - d. Ofuscación, ocultación y anonimización.
 - e. Fundamentos de la criptografía. Sistemas de claves simétricas y asimétricas.
 - f. Gestores de identidad y contraseñas. Política de seguridad.
 - g. Control de acceso a la información
- 8. Publicación y compartición, repositorios.
 - a. El paradigma Findable, Accesible, Interoperable, Reusable (principios FAIR).
 - b. Principales repositorios globales: figshare, zenodo, OSF...
 - c. Repositorios temáticos: GBIF, WDCC, DAAC, GenBank, NCBI, omics...
 - d. Repositorios regionales: EUDAT, NEON...
 - e. Artículos de datos.

Al término de la asignatura se requerirá la creación de un modelo de gestión de datos adecuado al TFM elegido por el alumno

ACTIVIDADES FORMATIVAS

Las clases son teórico-prácticas y es necesario contar con un ordenador personal en el aula. A medida que se imparte el contenido teórico se irán planteando actividades de resolución de problemas y valoración de distintos aspectos relacionados con la gestión de información de conjuntos de datos experimentales. Uno de los objetivos de la asignatura es ayudar al alumno en el desarrollo de su trabajo fin de máster por lo que algunas de las tareas estarán orientadas al caso particular del proyecto que esté desarrollando cada alumno.

Actividades prácticas programadas y evaluables:



El ciclo de generación y utilización de datos: digitalización de información a partir de las hojas primarias de registro de datos y extracción de información de fuentes públicas (minería).

Diseño y gestión de tablas básicas de datos. Tablas planas, pivots y tuplas.

Organización lógica y estructuras de datos: diseño de una base de datos relacional y extracción de información mediante consultas.

Estándares de intercambio de información: exportación de la información de una base de datos dedicada a formatos estandarizados de intercambio de información.

Seguridad, blindaje y preservación de la información: ofuscación, ocultación y anonimización.

Publicación y compartición, repositorios: creación de un artículo de datos (datapaper)

Creación de un modelo de gestión de datos (SDM) adecuado al TFM elegido por el alumno.

Creación de un dashboard con importación de datos de fuentes externas.

ACTIVIDADES FORMATIVAS				
ACTIVIDADES FORMATIVAS	Horas totales	Horas lectivas*	% Presencialidad**	% Interacción virtual***
AF1 - Clase expositiva	10	10	100%	0%
AF2 - Clases prácticas, seminarios y/o talleres	20	20	100%	0%
AF3- Trabajos dirigidos, individuales o en equipo	20	0	0%	0%
AF4 - Estudio personal	20	0	0%	0%
AF6 - Evaluación	3	3	100%	0%
AF7 - Tutorías	2	2	100%	0%

SISTEMAS DE EVALUACIÓN

SISTEMAS DE EVALUACIÓN		
SISTEMA DE EVALUACIÓN	PONDERACIÓN MIN	PONDERACIÓN MÁX
SE2 - Resolución de problemas, casos prácticos y trabajos	50	80
SE3 - Evaluaciones parciales y/o evaluación final	20	50



Universidad
de Navarra

Facultad de Ciencias

OBSERVACIONES

Todos los ejercicios realizados en las sesiones prácticas, trabajos dirigidos y casos son evaluables por separado, con ponderaciones individuales variables en función de su carga de trabajo. Algunos ejercicios son individuales y otros a realizar en grupos de dos a cuatro personas. Los ejercicios realizados en grupo reciben una evaluación única igual para todos los miembros del equipo. La evaluación final puede incluir un cuestionario y la resolución de uno o dos casos.