

Facultad de Ciencias

GUÍA DOCENTE DE LA ASIGNATURA

M1707 - Tecnologías de Almacenamiento de Datos No Relacionados

Máster Universitario en Ingeniería Informática
Optativa. Curso 2

Curso Académico 2021-2022

1. DATOS IDENTIFICATIVOS

Título/s	Máster Universitario en Ingeniería Informática	Tipología v Curso	Optativa. Curso 2
Centro	Facultad de Ciencias		
Módulo / materia	ASIGNATURAS OPTATIVAS		
Código y denominación	M1707 - Tecnologías de Almacenamiento de Datos No Relacionados		
Créditos ECTS	3	Cuatrimestre	Cuatrimestral (1)
Web			
Idioma de impartición	Español	English friendly	Sí
		Forma de impartición	Presencial

Departamento	DPTO. INGENIERÍA INFORMÁTICA Y ELECTRÓNICA
Profesor responsable	DIEGO GARCIA SAIZ
E-mail	diego.garcia@unican.es
Número despacho	Facultad de Ciencias. Planta: + 1. DESPACHO INVESTIGADOR (1068)
Otros profesores	ALFONSO DE LA VEGA RUIZ

2. CONOCIMIENTOS PREVIOS

Organización y gestión de datos: Ficheros. Estructuras de datos. Bases de datos relacionales. Lenguaje SQL. Modelado conceptual (UML o ER)

3. COMPETENCIAS GENÉRICAS Y ESPECÍFICAS DEL PLAN DE ESTUDIOS TRABAJADAS

Competencias Genéricas
Capacidad para proyectar, calcular y diseñar productos, procesos e instalaciones en todos los ámbitos de la Ingeniería Informática
Competencias Específicas
Capacidad para analizar las necesidades de información que se plantean en un entorno y llevar a cabo en todas sus etapas el proceso de construcción de un sistema de información
Competencias Básicas
Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio
Que los estudiantes sepan comunicar sus conclusiones -y los conocimientos y razones últimas que las sustentan- a públicos especializados y no especializados de un modo claro y sin ambigüedades
Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo
Competencias Transversales
Capacidad de análisis, síntesis y evaluación
Capacidad de resolución de problemas aplicando técnicas de ingeniería
Aprendizaje autónomo
Creatividad

3.1 RESULTADOS DE APRENDIZAJE

- Conocer y saber aplicar los principios, tecnologías y herramientas para el diseño e implementación de bases de datos orientadas a la gestión de grandes volúmenes de datos (big data).

4. OBJETIVOS

Conocer los distintos paradigmas de gestión de datos bajo el término NoSQL, sus ventajas y diferencias respecto al modelo relacional

Aprender los principios de diseño de bases de datos en estos nuevos modelos de datos

Analizar y evaluar la tecnología que mejor se adapta a las necesidades de un problema de gestión de volúmenes de datos masivos (big data)

Implementar e interrogar bases de datos bajo estas tecnologías

Conocer plataformas de la cloud que ofrecen servicios para alojar y gestionar grandes volúmenes de datos

5. MODALIDADES ORGANIZATIVAS Y MÉTODOS DOCENTES

ACTIVIDADES	HORAS DE LA ASIGNATURA
ACTIVIDADES PRESENCIALES	
HORAS DE CLASE (A)	
- Teoría (TE)	10
- Prácticas en Aula (PA)	5
- Prácticas de Laboratorio Experimental(PLE)	
- Prácticas de Laboratorio en Ordenador (PLO)	15
- Prácticas Clínicas (CL)	
Subtotal horas de clase	30
ACTIVIDADES DE SEGUIMIENTO (B)	
- Tutorías (TU)	3
- Evaluación (EV)	2
Subtotal actividades de seguimiento	5
Total actividades presenciales (A+B)	35
ACTIVIDADES NO PRESENCIALES	
Trabajo en grupo (TG)	
Trabajo autónomo (TA)	40
Tutorías No Presenciales (TU-NP)	
Evaluación No Presencial (EV-NP)	
Total actividades no presenciales	40
HORAS TOTALES	75

6. ORGANIZACIÓN DOCENTE													
CONTENIDOS		TE	PA	PLE	PLO	CL	TU	EV	TG	TA	TU-NP	EV-NP	Semana
1	Introducción: Gestión de datos: revisión histórica. NoSQL: antecedentes, características y diferencias con respecto a la gestión relacional. Taxonomía de soluciones. NoSQL vs NewSQL.	2,00	0,00	0,00	0,00	0,00	0,50	0,50	0,00	10,00	0,00	0,00	1
2	Paradigmas NoSQL: Gestores Clave-Valor, Familia de columnas, Documentales y basados en Grafos. Arquitectura y Modelo de datos. Lenguajes de interrogación. Criterios de diseño. Casos de aplicación.	6,00	5,00	0,00	13,00	0,00	2,00	1,00	0,00	25,00	0,00	0,00	1-6
3	Servicios de bases de datos en la Cloud Computing. Del modelo on-premise a la cloud. Ventajas e inconvenientes. Proveedores y tecnologías disponibles. Arquitecturas y tecnologías para la ingesta, consulta, análisis y visualización de datos masivos (big data).	2,00	0,00	0,00	2,00	0,00	0,50	0,50	0,00	5,00	0,00	0,00	7
TOTAL DE HORAS		10,00	5,00	0,00	15,00	0,00	3,00	2,00	0,00	40,00	0,00	0,00	

Esta organización tiene carácter orientativo.

Ante la situación incierta de que las medidas de distanciamiento social establecidas por las autoridades sanitarias no permitan desarrollar alguna actividad docente de forma presencial en el aula para todos los estudiantes matriculados, se adoptará una modalidad mixta de docencia que combine esta docencia presencial en el aula con docencia a distancia. De la misma manera, la tutorización podrá ser sustituida por tutorización a distancia utilizando medios telemáticos.

TE	Horas de teoría
PA	Horas de prácticas en aula
PLE	Horas de prácticas de laboratorio experimental
PLO	Horas de prácticas de laboratorio en ordenador
CL	Horas de prácticas clínicas
TU	Horas de tutoría
EV	Horas de evaluación
TG	Horas de trabajo en grupo
TA	Horas de trabajo autónomo
TU-NP	Tutorías No Presenciales
EV-NP	Evaluación No Presencial

7. MÉTODOS DE LA EVALUACIÓN

Descripción	Tipología	Eval. Final	Recuper.	%
Cuestionarios, ejercicios y actividades de discusión	Actividad de evaluación con soporte virtual	No	Sí	50,00

Calif. mínima	0,00
Duración	
Fecha realización	Durante el bimestre
Condiciones recuperación	
Observaciones	Se propondrán cuestionarios, ejercicios y actividades de discusión en cada tema que los estudiantes realizarán durante el bimestre en la plataforma Moodle. La nota se computará como la media ponderada de todas las actividades propuestas.

Trabajo individual	Trabajo	No	Sí	50,00
--------------------	---------	----	----	-------

Calif. mínima	4,00
Duración	
Fecha realización	Se realizará durante el bimestre y se entregará en la última semana de clase
Condiciones recuperación	
Observaciones	El alumno podrá elegir entre los siguientes tipos de trabajo: a) un diseño de base de datos bajo el paradigma NoSQL, b) un proceso de gestión y análisis de datos (pipeline) utilizando tecnologías big data; c) un despliegue en la nube de un servicio de datos NoSQL. En todos los casos se presentará una memoria escrita en el que se justifiquen las decisiones de diseño adoptadas, los pasos dados para la implementación de la solución, dificultades encontradas, y cualquier otro aspecto que se quiera señalar y que ponga en valor su trabajo. Se entregará asimismo los artefactos software desarrollados. El trabajo se defenderá oralmente.

TOTAL 100,00

Observaciones

Hay una única convocatoria anual. Si la asignatura no se supera en las actividades de evaluación continua realizadas durante el bimestre se podrá acceder a una evaluación de recuperación. En ésta, sólo es necesario realizar las actividades de evaluación no superadas.

Para aprobar la asignatura es necesario superar la nota mínima del trabajo individual. En caso de no superarlo, la nota final será el mínimo de 4.9 y la media obtenida como resultado de computar todas las actividades de evaluación.

Si el cupo de matrículas de honor de la asignatura se completa en la evaluación continua, los alumnos que se presenten a la recuperación no podrán optar a la calificación de matrícula de honor.

Criterios de evaluación para estudiantes a tiempo parcial

Los alumnos que por motivo justificado (estudiantes a tiempo parcial) no hayan seguido la evaluación continua se les evaluará de la siguiente manera:

- Examen escrito (cuestiones y ejercicios): 50 %
- Trabajo individual: 50 %

8. BIBLIOGRAFÍA Y MATERIALES DIDÁCTICOS

BÁSICA

Eric Redmond. Seven Databases in Seven Weeks: A Guide to Modern Databases and the NoSQL Movement. 2012 (free book)

Lee Chao. Cloud Database Development and Management. 2013. Auerbach Publications

Complementaria
Nathan Marz, James Warren. Big Data: Principles and best practices of scalable realtime data systems 1st Edition. 2015. Manning publisher
Holden Karau, Andy Konwinski, Patrick Wendell, Matei Zaharia. Learning Spark: Lightning-Fast Big Data Analysis 1st Edition. 2015. O'Reilly Media, Inc.
Eben Hewitt, Jeff Carpenter. Cassandra: The Definitive Guide, 2nd Edition. 2016.O'Reilly Media, Inc.
Kristina Chodorow. MongoDB: The Definitive Guide. 2013. O'Reilly Media, Inc. (free book)
Ian Robinson, Jim Webber, and Emil Eifrem. Graph Databases 2nd Edition. 2015. O'Reilly Media (free book)
Clinton Gormley, Zachary Tong. Elasticsearch: The Definitive Guide. A Distributed Real-Time Search and Analytics Engine. O'Reilly Media. 2015
Guy Harrison. Next Generation Databases: NoSQL, NewSQL, and Big Data. Apress, 2016

9. SOFTWARE

PROGRAMA / APLICACIÓN	CENTRO	PLANTA	SALA	HORARIO
Apache Cassandra	Facultad de Ciencias			
MongoDB	Facultad de Ciencias			
Neo4j	Facultad de Ciencias			
Oracle Cloud (DBaaS)	Facultad de Ciencias			

10. COMPETENCIAS LINGÜÍSTICAS

- Comprensión escrita Comprensión oral
 Expresión escrita Expresión oral
 Asignatura íntegramente desarrollada en inglés

Observaciones