

ANX-PR/CL/001-01
GUÍA DE APRENDIZAJE

ASIGNATURA

Arquitectura del data warehouse

CURSO ACADÉMICO - SEMESTRE

2016-17 - Segundo semestre

Datos Descriptivos

Nombre de la Asignatura	Arquitectura del data warehouse
Titulación	10II - Grado en Ingeniería Informática
Centro responsable de la titulación	Escuela Técnica Superior de Ingenieros Informáticos
Semestre/s de impartición	Octavo semestre
Materias	Optatividad
Carácter	Optativa
Código UPM	105000167
Nombre en inglés	Data warehouse architecture

Datos Generales

Créditos	3	Curso	4
Curso Académico	2016-17	Período de impartición	Febrero-Junio
Idioma de impartición	Castellano	Otros idiomas de impartición	

Requisitos Previos Obligatorios

Asignaturas Previas Requeridas

El plan de estudios Grado en Ingeniería Informática no tiene definidas asignaturas previas superadas para esta asignatura.

Otros Requisitos

El plan de estudios Grado en Ingeniería Informática no tiene definidos otros requisitos para esta asignatura.

Conocimientos Previos

Asignaturas Previas Recomendadas

Bases de datos

Otros Conocimientos Previos Recomendados

Conocimientos de bases de datos

Conocimientos de SQL

Competencias

CG-2/CE45 - Capacidad para el aprendizaje autónomo y la actualización de conocimientos, y reconocimiento de su necesidad en el área de la informática.

Ce 12/16 - Conocer los campos de aplicación de la informática, y tener una apreciación de la necesidad de poseer unos conocimientos técnicos profundos en ciertas áreas de aplicación; apreciación del grado de esta necesidad en, por lo menos, una situación.

Ce 13/18 - Comprender lo que pueden y no pueden conseguir las tecnologías actuales, y las limitaciones de la informática, que implica distinguir entre lo que, inherentemente, la informática no es capaz de hacer y lo que puede lograrse a través de la ciencia y la tecnología futuras.

Ce 44 - Conocimiento de tecnologías punteras relevantes y su aplicación.

Resultados de Aprendizaje

RA373 - Aplicar las técnicas y métodos relativos a una línea de especialización concreta del área tecnológica, comprendiendo sus límites tanto teóricos como prácticos, para la resolución de un problema o necesidad planteado por un consumidor o cliente real

RA276 - Dado un campo de aplicación de la informática, evaluar y diseñar el sistema informático más apropiado para resolver alguno de sus problemas, exponiendo las dificultades técnicas y los límites de la aplicación.

Profesorado

Profesorado

Nombre	Despacho	e-mail	Tutorías
Eibe García, Santiago (Coordinador/a)	4302	santiago.eibe@upm.es	L - 12:00 - 13:30 L - 15:30 - 17:00 X - 12:00 - 13:30 J - 15:30 - 17:00 Se recomienda contactar previamente mediante correo electrónico

Nota.- Las horas de tutoría son orientativas y pueden sufrir modificaciones. Se deberá confirmar los horarios de tutorías con el profesorado.

Descripción de la Asignatura

El principal objetivo de la asignatura es presentar la problemática asociada con la implantación de un Data Warehouse tanto desde el punto de vista de la arquitectura propiamente dicha como de las tecnologías sobre las que se sustenta y con las que interacciona. Por ende, se persigue exponer la problemática, más general, de los sistemas de gestión de datos en la actualidad tomando para ello el DW como hilo conductor

Temario

1. Sistemas gestores de bases de datos
 - 1.1. Introducción a SGBD
 - 1.2. Funcionalidades básicas de un SGBD
2. Almacenamiento de datos
 - 2.1. Jerarquía de almacenamiento
 - 2.2. Técnicas de indexación. Rendimiento
3. Procesamiento de interrogaciones
 - 3.1. Preprocesamiento y análisis sintáctico. Árboles de interrogación
 - 3.2. Análisis de costes. Estimación
 - 3.3. Plan de ejecución. Operadores
4. Caso práctico de MySQL
 - 4.1. Arquitectura y administración básica
 - 4.2. Gestión de almacenamiento
 - 4.3. Control de acceso
5. Introducción a data warehouse
 - 5.1. Introducción a data warehouse
 - 5.2. Arquitectura de un data warehouse
 - 5.3. Módulos de un data warehouse
6. ETL
 - 6.1. Funcionalidad ETL
 - 6.2. Herramientas ETL
7. Base de datos del data warehouse
 - 7.1. Diseño multidimensional básico
 - 7.2. Diseño multidimensional extendido
 - 7.3. Sistemas gestores de bases de datos para el data warehouse

- 8. Metodologías de diseño del data warehouse
 - 8.1. Metodologías de diseño del data warehouse
 - 8.2. Bus común del data warehouse
 - 8.3. DW 2.0
- 9. Replicación y particionado
 - 9.1. Replicación y particionado
- 10. Consulta de información del data warehouse
 - 10.1. OLAP
 - 10.2. Herramientas OLAP

Cronograma

Horas totales: 32 horas

Horas presenciales: 32 horas (41%)

Peso total de actividades de evaluación continua:
100%

Peso total de actividades de evaluación sólo prueba final:
100%

Semana	Actividad Presencial en Aula	Actividad Presencial en Laboratorio	Otra Actividad Presencial	Actividades Evaluación
Semana 1	1 sistemas gestores de bases de datos Duración: 02:00 LM: Actividad del tipo Lección Magistral			
Semana 2	2 Almacenamiento de datos Duración: 02:00 LM: Actividad del tipo Lección Magistral			
Semana 3	3 Procesamiento de interrogaciones Duración: 02:00 LM: Actividad del tipo Lección Magistral			Tecnologías en el DW Duración: 00:00 TG: Técnica del tipo Trabajo en Grupo Evaluación continua Actividad no presencial
Semana 4		4 Caso práctico: MySQL Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio		
Semana 5		4 Caso práctico: MySQL Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio		
Semana 6	5.1 Introducción a data warehouse Duración: 02:00 LM: Actividad del tipo Lección Magistral			Integración de procesos en el DW: ETLs Duración: 00:00 TG: Técnica del tipo Trabajo en Grupo Evaluación continua Actividad no presencial
Semana 7	5.2 Arquitectura del data warehouse Duración: 02:00 LM: Actividad del tipo Lección Magistral			
Semana 8	5.3 Módulos del data warehouse Duración: 02:00 LM: Actividad del tipo Lección Magistral			
Semana 9	6.1 Funcionalidad ETL Duración: 02:00 LM: Actividad del tipo Lección Magistral			
Semana 10	7.1 Diseño multidimensional básico Duración: 02:00 LM: Actividad del tipo Lección Magistral			

Semana 11	<p>7.2 Diseño multidimensional extendido</p> <p>Duración: 02:00</p> <p>LM: Actividad del tipo Lección Magistral</p>			
Semana 12	<p>8.1 Metodologías de diseño del data warehouse</p> <p>Duración: 01:00</p> <p>LM: Actividad del tipo Lección Magistral</p> <p>8.2 Bus común del data warehouse</p> <p>Duración: 01:00</p> <p>LM: Actividad del tipo Lección Magistral</p>			<p>Procesos de Implantación del DW: Diseño y Carga del Data Warehouse</p> <p>Duración: 00:00</p> <p>TG: Técnica del tipo Trabajo en Grupo</p> <p>Evaluación continua</p> <p>Actividad no presencial</p>
Semana 13	<p>8.3 DW 2.0</p> <p>Duración: 01:00</p> <p>LM: Actividad del tipo Lección Magistral</p> <p>10 Consulta de información del data warehouse</p> <p>Duración: 01:00</p> <p>LM: Actividad del tipo Lección Magistral</p>			
Semana 14	<p>7.3 Sistemas gestores de bases de datos para el data warehouse</p> <p>Duración: 02:00</p> <p>LM: Actividad del tipo Lección Magistral</p>			<p>Procesos de Implantación del DW: Análisis y Explotación del Data Warehouse</p> <p>Duración: 00:00</p> <p>TG: Técnica del tipo Trabajo en Grupo</p> <p>Evaluación continua</p> <p>Actividad no presencial</p>
Semana 15	<p>9 Replicación y particionado</p> <p>Duración: 02:00</p> <p>LM: Actividad del tipo Lección Magistral</p>			
Semana 16	<p>9 Replicación y particionado</p> <p>Duración: 02:00</p> <p>LM: Actividad del tipo Lección Magistral</p>			
Semana 17				<p>Examen</p> <p>Duración: 02:00</p> <p>EX: Técnica del tipo Examen Escrito</p> <p>Evaluación sólo prueba final</p> <p>Actividad presencial</p> <p>Proyecto práctico</p> <p>Duración: 00:00</p> <p>OT: Otras técnicas evaluativas</p> <p>Evaluación sólo prueba final</p> <p>Actividad no presencial</p>

Nota.- El cronograma sigue una planificación teórica de la asignatura que puede sufrir modificaciones durante el curso.

Nota 2.- Para poder calcular correctamente la dedicación de un alumno, la duración de las actividades que se repiten en el tiempo (por ejemplo, subgrupos de prácticas") únicamente se indican la primera vez que se definen.

Actividades de Evaluación

Semana	Descripción	Duración	Tipo evaluación	Técnica evaluativa	Presencial	Peso	Nota mínima	Competencias evaluadas
3	Tecnologías en el DW	00:00	Evaluación continua	TG: Técnica del tipo Trabajo en Grupo	No	35%	3 / 10	CG-2/CE45, Ce 12/16, Ce 13/18
6	Integración de procesos en el DW: ETLs	00:00	Evaluación continua	TG: Técnica del tipo Trabajo en Grupo	No	30%	3 / 10	CG-2/CE45, Ce 12/16, Ce 13/18
12	Procesos de Implantación del DW: Diseño y Carga del Data Warehouse	00:00	Evaluación continua	TG: Técnica del tipo Trabajo en Grupo	No	20%	3 / 10	CG-2/CE45, Ce 12/16, Ce 13/18
14	Procesos de Implantación del DW: Análisis y Explotación del Data Warehouse	00:00	Evaluación continua	TG: Técnica del tipo Trabajo en Grupo	No	15%	3 / 10	CG-2/CE45, Ce 12/16, Ce 13/18
17	Examen	02:00	Evaluación sólo prueba final	EX: Técnica del tipo Examen Escrito	Sí	80%	4 / 10	Ce 12/16, Ce 13/18
17	Proyecto práctico	00:00	Evaluación sólo prueba final	OT: Otras técnicas evaluativas	No	20%	3.5 / 10	Ce 12/16, CG-2/CE45

Criterios de Evaluación

La asignatura de Arquitectura del Data Warehouse se evaluará mediante 4 proyectos:

- tres cubren específicamente el desarrollo de un proyecto de implantación de un Data Warehouse
- uno se enfoca al análisis de alguna de las diversas tecnologías partícipes en la Arquitectura de un Data Warehouse

Los proyectos se realizarán en grupos de **hasta 4 alumnos** de entre los matriculados de la asignatura al inicio del curso.

Para poder superar la asignatura, en la **convocatoria de junio**, se establecen los siguientes requisitos:

1. Obtener un mínimo de 50 puntos sobre los 100 disponible en el cómputo global de la evaluación sumativa
2. Es OBLIGATORIO realizar todos los proyectos
3. En los proyectos se debe obtener una nota mínima igual o superior al 30% de la valoración del mismo (ver tabla de valoración sumativa)
4. Es OBLIGATORIO superar una prueba de evaluación individual de cada proyecto

Para poder superar la asignatura en la **convocatoria extraordinaria de julio**, se establecen los siguientes requisitos:

1. No habrá proyectos, sólo se realizará un examen que cubrirá los aspectos teóricos y prácticos de la asignatura
2. Obtener un mínimo de 50 puntos sobre los 100 disponibles en el cómputo global

Para aquellos alumnos que de forma extraordinaria, no puedan realizar la evaluación continua, y previa petición por escrito durante los primeros 15 días del curso, la forma de evaluación de la asignatura será la siguiente, **siendo excluyente con la evaluación continua**.

1. Examen en junio en la fecha establecida en el calendario oficial de exámenes por jefatura de estudios.
2. Realización de un proyecto práctico, en grupos de hasta cuatro alumnos, compuesto por tres entregas, valorado en 20 puntos.

Para poder superar la asignatura en la **convocatoria de junio (mediante la forma extraordinaria)**, se establecen los siguientes requisitos:

1. La valoración de examen es de 80 puntos.
2. Para poder aprobar la asignatura en esta convocatoria es necesario obtener una nota mínima igual o superior al 40% de la valoración del examen.
3. En el proyecto práctico es necesario obtener una valoración mínima del 35% de la valoración del mismo.

4. Obtener un mínimo de 50 puntos sobre los 100 disponibles en el cómputo global (examen + proyecto práctico)

A los alumnos que opten por este sistema de evaluación extraordinario en la convocatoria de junio y no superen la asignatura **no se les guardará ninguna de la calificaciones** para la convocatoria de julio y sucesivas.

Recursos Didácticos

Descripción	Tipo	Observaciones
Building the data warehouse. W.H. Immon. 1996. Willey	Bibliografía	
Managing the data warehouse. W.H. Immune. 1997. Willey	Bibliografía	
Building the operational Data Store. W. H. Immon. 1999. Willey	Bibliografía	
Exploration Datawarehouse. W. Immon. 2000. Willey	Bibliografía	
The data warehouse lifecycle toolkit. R. Kimball. 2000. Willey	Bibliografía	
Improving Data Warehouse and Business Information Quality. Methods for reducing cost and increasing profits. L. English. 1999 Willey	Bibliografía	
Principles of Data Base Systems (Second Edition), Jeffrey D. Ullman, Ed. Computer Science Press, Rockville, Maryland, 1982	Bibliografía	
First Course in Database Systems, Jeffrey D. Ullman, Jennifer Widom, ISBN-10: 013600637X. 2007	Bibliografía	
Sistemas de Bases de Datos, R. Elmasri y S.B.Navathe, 2ª edición, Addison-Wesley Iberoamericana, 1997	Bibliografía	
Fundamentos de bases de datos, A. Silberschatz, H. Korth, S. Sudarsham, 5ª edición, Mcgraw-Hill, 2006	Bibliografía	
Database Systems-A Practical Approach to Design, Implementation and Management. 4th ed., Connolly, T., Begg, C. AND Strachan, A., 2004. Addison-Wesley	Bibliografía	
Fundamentals of Database Systems, 5th ed ., Elmasri, R. Navathe, S., 2006.. Addison-Wesley	Bibliografía	
MySQL Administrator?s Bible, Sheeri Cabral, Keith Murphy. Wiley Publishing 2009	Bibliografía	
The Complete Reference MySQL, Vikram Vaswani.McGraw-Hill/Osborne 2007	Bibliografía	
Database Systems: The Complete Book, Hector Garcia-Molina, Jeff Ullman, and Jennifer Widom. (DS-CB), 2008, 2nd edition	Bibliografía	
Mondrian 3.0 Technical Guide	Bibliografía	
Pentaho Data Integration 2ªEd María Carina Roldán	Bibliografía	
Pentaho Data Integration 4 Cookbook A.S.Pulvirenti, M.C. Roldán	Bibliografía	