



Guía docente de la asignatura

Asignatura	SISTEMAS AVANZADOS DE INTEGRACIÓN DE INFORMACIÓN		
Materia	COMPUTACIÓN		
Módulo	TECNOLOGÍAS ESPECÍFICAS		
Titulación	GRADO EN INGENIERÍA INFORMÁTICA		
Plan	545	Código	46950
Periodo de impartición	2º CUATRIMESTRE	Tipo/Carácter	OPTATIVA (Mención CO)
Nivel/Ciclo	GRADO	Curso	3º
Créditos ECTS	6 ECTS		
Lengua en que se imparte	CASTELLANO		
Profesor/es responsable/s	Mercedes Martínez González		
Datos de contacto (E-mail, teléfono...)	TELÉFONO: 983 423000 ext. 5607 E-MAIL: mercedes@infor.uva.es		
Horario de tutorías	Véase www.uva.es → Centros → Campus de Valladolid → Escuela Técnica Superior de Ingeniería Informática → Tutorías		
Departamento	Informática (ATC, CCIA y LSI)		



1. Situación / Sentido de la Asignatura

1.1 Contextualización

La asignatura *Sistemas Avanzados para la Integración de Información* forma parte del conjunto de asignaturas optativas del grado en Ingeniería Informática. Se trata de una asignatura que permite profundizar en las soluciones a un problema relacionado con la gestión de información de enorme interés y actualidad en el contexto actual de la denominada “sociedad de la información”: la integración de información procedente de fuentes heterogéneas.

1.2 Relación con otras materias

La asignatura engarza con asignaturas previas de carácter obligatorio destinadas al estudio de la gestión de información, como las introductorias a las bases de datos. Esta asignatura utiliza conocimientos y competencias adquiridos en la formación de bases de datos. Asimismo guarda relación con asignaturas de postgrado impartidas en estudios vinculados al centro, en las que se aborda la problemática de integrar información de fuentes heterogéneas dentro del marco de estudio de las tendencias actuales para compartir y explotar conocimiento y semántica en la web.

1.3 Prerrequisitos

Para abordar con garantías de éxito esta asignatura es necesario haber adquirido las competencias vinculadas a las asignaturas de bases de datos que se estudian en tercer curso del grado. El alumno que estudia esta asignatura debe tener habilidades suficientes en el manejo de bases de datos relacionales, acceso a bases de datos desde aplicaciones externas, y capacidad para entender la problemática de los datos y la importancia de los esquemas elegidos para representar información.

2. Competencias

2.1 Generales

Código	Descripción
G02	Conocimientos básicos de la profesión
G03	Capacidad de análisis y síntesis
G04	Capacidad de organizar y planificar
G05	Comunicación oral y escrita en la lengua propia
G06	Conocimiento de una segunda lengua (preferentemente inglés)
G08	Habilidades de gestión de la información
G09	Resolución de problemas
G10	Toma de decisiones
G11	Capacidad crítica y autocrítica
G12	Trabajo en equipo
G14	Responsabilidad y compromiso ético
G15	Liderazgo
G16	Capacidad de aplicar los conocimientos en la práctica
G17	Habilidades de investigación
G18	Capacidad de aprender
G19	Capacidad de adaptarse a nuevas situaciones
G20	Capacidad de generar nuevas ideas
G21	Habilidad para trabajar de forma autónoma
G22	Diseño y gestión de proyectos

2.2 Específicas

Código	Descripción
TI1	Capacidad para comprender el entorno de una organización y sus necesidades en el ámbito de las tecnologías de la información y las comunicaciones
SI1	Capacidad de integrar soluciones de Tecnologías de la Información y las Comunicaciones y procesos empresariales para satisfacer las necesidades de información de las organizaciones, permitiéndoles alcanzar sus objetivos de forma efectiva y eficiente, dándoles así ventajas competitivas
SI2	Capacidad para determinar los requisitos de los sistemas de información y comunicación de una organización atendiendo a aspectos de seguridad y cumplimiento de la normativa y la legislación vigente



SI3	Capacidad para participar activamente en la especificación, diseño, implementación y mantenimiento de los sistemas de información y comunicación
-----	--

3. Objetivos

Código	Descripción
TI1.1	Entender las nuevas necesidades de las organizaciones y conocer las modificaciones propuestas desde el propio modelo relacional
SI3.1	Conocer los últimos avances relacionados con bases de datos: BD documentales, BD distribuidas, BD multimedia, BD espacio-temporales, etc
SI3.2	Comprender y desarrollar sistemas de procesamiento analítico en línea (OLAP), Data Warehouse y Data Mining
SI1.1	Conocer la tipología y función de los sistemas de información con respecto a los procesos de negocio (CIM, ERP, CRM, DSS,?) y comprender cómo ayudan a las organizaciones
SI2.1	Conocer la relación de la empresa y los sistemas de información para realizar una planificación estratégica (outsourcing, seguridad, privacidad)

**4. Tabla de dedicación del estudiante a la asignatura**

ACTIVIDADES PRESENCIALES	HORAS	ACTIVIDADES NO PRESENCIALES	HORAS
Clases teórico-prácticas (T/M)	30	Estudio y trabajo autónomo individual	50
Clases prácticas de aula (A)		Estudio y trabajo autónomo grupal	40
Laboratorios (L)	22		
Prácticas externas, clínicas o de campo			
Seminarios (S)	6		
Tutorías grupales (TG)			
Evaluación (fuera del periodo oficial de exámenes)	2		
Total presencial	60	Total no presencial	90

5. Bloques temáticos

Bloque 1: Técnicas para la integración de datos

Carga de trabajo en créditos ECTS:

a. Contextualización y justificación

Este bloque se dedica a presentar las técnicas en las que se apoya la integración de datos. Se trata del bloque que aporta los conocimientos fundamentales para abordar la construcción de sistemas de integración de información, estén éstos orientados a bases de datos relacionales como a otros tipos de bases de datos o fuentes de información.

b. Objetivos de aprendizaje

Código	Descripción
TI1.1	Entender las nuevas necesidades de las organizaciones y conocer las modificaciones propuestas desde el propio modelo relacional
SI3.2	Comprender y desarrollar sistemas de procesamiento analítico en línea (OLAP), Data Warehouse y Data Mining
SI1.1	Conocer la tipología y función de los sistemas de información con respecto a los procesos de negocio (CIM, ERP, CRM, DSS,?) y comprender cómo ayudan a las organizaciones
SI2.1	Conocer la relación de la empresa y los sistemas de información para realizar una planificación estratégica (outsourcing, seguridad, privacidad)

c. Contenidos

TEMA 1: El problema de la integración de información

- 1.1 Definición y características
- 1.2 Problemas específicos de la integración de información
- 1.3 La integración de información en los nuevos entornos de datos

TEMA 2: Técnicas básicas para la integración de información

- 2.1 Manipulación de consultas
- 2.2 Descripción de fuentes de datos
- 2.3 Matching de cadenas (string matching)
- 2.4 Matching (y mapping) de esquemas
- 2.5 Manipulación de esquemas
- 2.6 Matching de datos

**TEMA 3: Técnicas adicionales para la integración de información**

- 3.1 Procesamiento de consultas orientado a la integración de información
- 3.2 Wrappers para la integración
- 3.3 Data warehousing

d. Métodos docentes

Ver Anexo: Métodos docentes

e. Plan de trabajo

Para este bloque se estiman 40 horas presenciales. El tiempo de dedicación no presencial de alumno se estima en unas 55 horas. Para el desglose detallado ver el apartado 9 de esta guía.

f. Evaluación

Ver apartado 7 de esta guía.

g. Bibliografía básica

AnHai Doan, Alon Halevy, Zachary Ives. *Principles of Data Integration*. Morgan Kaufmann, 2012. ISBN (print) 978-0-12-416044-6, (web) 978-0-12-391479-8. UVA: 004 H3 DOA pri

h. Bibliografía complementaria

[Luna Dong] Luna Dong, X. y Srivastava D. *Big data integration*. Morgan & Claypool Publishers.

[Connolly] Connolly, T y Begg. C. *Sistemas de Bases de Datos. Un enfoque práctico para diseño, implementación y gestión*. Pearson Addison Wesley, 2005. 4ª edición.

[Elmasri] Elmasri, R. y Navathe, S. *Fundamentos de Sistemas de Bases de Datos*. Editorial Addison-Wesley. 2002. 3ª edición.

[Silberschatz] Silberschatz, A., Korth, H.F. y Suddarshan, S. *Fundamentos de Bases de Datos*. McGraw-Hill, 2006. 5ª edición.

i. Recursos necesarios

Libros de texto, presentaciones audiovisuales, material disponible en el aula virtual de la asignatura.

Bloque 2: La integración de información con bases de datos no relacionales y en la web

Carga de trabajo en créditos ECTS:

a. Contextualización y justificación

La integración de información, si bien es un problema de larga tradición en las bases de datos relacionales, ha adquirido aún mayor peso con la aparición de nuevos modelos de datos destinados a la publicación y almacenamiento de información no estructurada y de información en la web. En este bloque se introduce al alumno en el conocimiento de esta problemática, incidiendo en las diferencias y similitudes con lo ya visto en el

bloque anterior para bases de datos relacionales. Este bloque es una introducción a uno de los problemas de mayor actualidad a los que se enfrentan los grandes consumidores de información (portales web, buscadores varios, etc.).

b. Objetivos de aprendizaje

Código	Descripción
TI1.1	Entender las nuevas necesidades de las organizaciones y conocer las modificaciones propuestas desde el propio modelo relacional
SI3.1	Conocer los últimos avances relacionados con bases de datos: BD documentales, BD distribuidas, BD multimedia, BD espacio-temporales, etc

c. Contenidos

TEMA 4: Integración de información usando XML

- 4.1 Características de XML utilizadas en la integración de información
- 4.2 Otras herramientas para la integración de información con XML

TEMA 5: Integración de información en la web

- 5.1 Características de los datos en la web
- 5.2 La web profunda (deep web)
- 5.3 Sistemas para la integración de información en la web

TEMA 6: Integración de información en la web de datos y big data

- 4.1 Integración de información en la web de datos con linked data
- 4.2 Integración de información en big data

d. Métodos docentes

Ver Anexo: Métodos docentes

e. Plan de trabajo

Para este bloque se estiman 20 horas presenciales. El tiempo de dedicación no presencial de alumno se estima en unas 35 horas. Para el desglose detallado ver el apartado 9 de esta guía.

f. Evaluación

Ver apartado 7 de esta guía.

g. Bibliografía básica

AnHai Doan, Alon Halevy, Zachary Ives. *Principles of Data Integration*. Morgan Kaufmann, 2012. ISBN (print) 978-0-12-416044-6, (web) 978-0-12-391479-8. UVa: 004 H3 DOA pri

h. Bibliografía complementaria

[Luna Dong] Luna Dong X. y Srivastava D. *Big Data Integration*. Morgan & Claypool Publishers. 2015.

[Connolly] Connolly, T y Begg. C. *Sistemas de Bases de Datos. Un enfoque práctico para diseño, implementación y gestión*. Pearson Addison Wesley, 2005. 4ª edición.

[Elmasri] Elmasri, R. y Navathe, S. *Fundamentos de Sistemas de Bases de Datos*. Editorial Addison-Wesley. 2002. 3ª edición.

[Silberschatz] Silberschatz, A., Korth, H.F. y Suddarshan, S. *Fundamentos de Bases de Datos*. McGraw-Hill, 2006. 5ª edición.

i. Recursos necesarios

Libros de texto, presentaciones audiovisuales, material disponible en el aula virtual de la asignatura.

6. Temporalización (por bloques temáticos)

BLOQUE TEMÁTICO	CARGA ECTS	PERIODO PREVISTO DE DESARROLLO
Bloque 1: Técnicas para la integración de datos	4 ECTS	Semanas 1 a 10
Bloque 2: La integración de información con bases de datos no relacionales y en la web	2 ECTS	Semanas 11 a 15

7. Sistema de calificaciones – Tabla resumen

INSTRUMENTO/PROCEDIMIENTO	PESO EN LA NOTA FINAL	OBSERVACIONES
Entrega práctica	70%	La práctica consistirá en un proyecto, que se irá revisando en sucesivos hitos hasta llegar a la entrega final.
Examen final escrito	30%	Período de exámenes

CRITERIOS DE CALIFICACIÓN

- **Convocatoria ordinaria:**
 - o Nota final = (Nota del examen escrito)*0,30 + (Nota de la práctica)*0,70
- **Convocatoria extraordinaria:**
 - o Mismo sistema que en la convocatoria ordinaria.

8. Anexo: Métodos docentes

Actividad	Metodología
Clase de teoría	<ul style="list-style-type: none">• Clase magistral participativa• Estudio de casos en aula• Resolución de problemas
Clase práctica	<ul style="list-style-type: none">• Clase magistral participativa• Realización de un proyecto guiado por el profesor, que encargará y guiará el trabajo que se realizará en grupos (2/3 alumnos), siguiendo un enfoque colaborativo.
Seminarios	<ul style="list-style-type: none">• Talleres de aprendizaje
Tutoría	<ul style="list-style-type: none">• Evaluación de los contenidos teóricos y de los proyectos

9. Anexo: Cronograma de actividades previstas

Semana	Contenido	Actividad presencial	Actividad no presencial
Semana 1	Tema 1: El problema de la integración de información.	Sesiones en aula (4h): 4 h.	Estudio.
Semana 2	Tema 1: El problema de la integración de información.	Sesiones en aula y laboratorio: 4 h.	Estudio y solución de problemas. Estudio de casos.
Semana 3	Tema 1: El problema de la integración de información.	Sesiones en aula y laboratorio: 4 h.	Estudio y solución de problemas. Estudio de casos.
Semana 4	Tema 2: Técnicas básicas para la integración de información.	Sesiones en aula (2h) y laboratorio (2h): 4 h.	Estudio y solución de problemas. Preparación de trabajos prácticos.
Semana 5	Tema 2: Técnicas básicas para la integración de información.	Sesiones en aula (2h) y laboratorio (2h): 4h.	Estudio y solución de problemas. Preparación de trabajos prácticos.
Semana 6	Tema 2: Técnicas básicas para la integración de información.	Sesiones en aula (2h) y laboratorio (2h): 4h.	Estudio y solución de problemas. Preparación de trabajos prácticos.
Semana 7	Tema 2: Técnicas básicas para la integración de información.	Sesiones en aula: 2h. Seminario (sobre trabajo práctico): 2h	Estudio y solución de problemas. Preparación de trabajos prácticos. Revisión de la evolución del trabajo práctico.
Semana 8	Tema 2: Técnicas básicas para la integración de información.	Sesiones en aula (2h) y laboratorio (2h): 4h.	Estudio y solución de problemas. Preparación de trabajos prácticos.
Semana 9	Tema 3: Técnicas adicionales para la integración de información.	Sesiones en aula (2h) y laboratorio (2h): 4h.	Estudio y solución de problemas. Prácticas con el SGBD. Preparación de trabajos prácticos.
Semana 10	Tema 3: Técnicas adicionales para la integración de información.	Sesiones en aula (2h): 2h. Seminario (sobre el trabajo práctico): 2h	Estudio y solución de problemas. Revisión de la evolución del trabajo práctico.
Semana 11	Tema 4: Integración de información usando XML.	Sesiones en aula (2h) y laboratorio (2h): 4h.	Estudio y solución de problemas. Preparación de trabajos prácticos.
Semana 12	Tema 4: Integración de información usando XML.	Sesiones en aula y laboratorio: 3h. Seminario (sobre trabajo práctico): 1h.	Estudio y solución de problemas. Preparación de trabajos prácticos. Revisión de la evolución del trabajo práctico.
Semana 13	Tema 5: Integración de información usando XML.	Sesiones en aula y laboratorio: 4h.	Estudio y solución de problemas. Preparación de trabajos prácticos.
Semana 14	Tema 5: Integración de información en la web.	Sesiones en aula (2h) y laboratorio: 3h. Seminario (sobre trabajo práctico): 1h.	Estudio y solución de problemas. Preparación de trabajos prácticos. Revisión de la evolución del trabajo práctico.
Semana 15	Tema 6: Integración de información en la web de datos y big data.	Sesiones en aula y laboratorio: 2h. Evaluación (trabajo práctico): 2h.	Estudio y solución de problemas. Preparación de trabajos prácticos. Entrega final de trabajo práctico.