

**Proyecto/Guía docente de la asignatura**

<b>Asignatura</b>	SISTEMAS AVANZADOS DE INTEGRACIÓN DE INFORMACIÓN		
<b>Materia</b>	COMPUTACIÓN		
<b>Módulo</b>	TECNOLOGÍAS ESPECÍFICAS		
<b>Titulación</b>	PROGRAMA DE ESTUDIOS CONJUNTO DE GRADO EN ESTADÍSTICA Y DE GRADO EN INGENIERÍA INFORMÁTICA (INdat)		
<b>Plan</b>	551	<b>Código</b>	46950
<b>Periodo de impartición</b>	2º CUATRIMESTRE	<b>Tipo/Carácter</b>	OBLIGATORIA (INdat)
<b>Nivel/Ciclo</b>	GRADO	<b>Curso</b>	4º
<b>Créditos ECTS</b>	6 ECTS		
<b>Lengua en que se imparte</b>	CASTELLANO		
<b>Profesor/es responsable/s</b>	Mercedes Martínez González		
<b>Datos de contacto (E-mail, teléfono...)</b>	TELÉFONO: 983 423000 ext. 5607 E-MAIL: <a href="mailto:mercedes@infor.uva.es">mercedes@infor.uva.es</a>		
<b>Horario de tutorías</b>	Véase <a href="http://www.uva.es">www.uva.es</a> → Centros → Campus de Valladolid → Escuela Técnica Superior de Ingeniería Informática → Tutorías		
<b>Departamento</b>	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.

### 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
T11	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 gestión de información y conocer las propuestas desde el modelo relacional
SI3.1	Conocer los últimos avances relacionados con bases de datos: alternativas al modelo relacional, BD distribuidas, en la web, etc.
SI3.2	Comprender y desarrollar sistemas de Data Warehouse y basados en wrappers
SI1.1	Conocer la tipología y función de los sistemas de integración de información 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 en gobernanza de la información



#### 4. Contenidos y/o 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 gestión de información y conocer las propuestas desde el modelo relacional
SI3.2	Comprender y desarrollar sistemas de Data Warehouse y basados en wrappers
SI1.1	Conocer la tipología y función de los sistemas de integración de información 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 en gobernanza de la información

#### 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 esquemas
- 2.4 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 apartado 5 de esta guía, sobre 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 sobre Evaluación de la asignatura.

#### g Material docente

---

##### g.1 Bibliografía básica

---

AnHai Doan, Alon Haley, 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 .

Serge Abiteboul, Ioana Manolescu, Philippe Rigaux, Marie-Christine Rousset, and Pierre Senellart. 2011. *Web Data Management*. Cambridge University Press, New York, NY, USA.

##### g.2 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.

##### g.3 Otros recursos telemáticos (píldoras de conocimiento, blogs, videos, revistas digitales, cursos masivos (MOOC), ...)

---

Presentaciones, páginas de referencias, chats de discusión y otros recursos disponible en el aula virtual de la asignatura.

#### h. Recursos necesarios

---

Ordenador para conectarse a los servidores de la escuela de Informática.

#### i. Temporalización

---

CARGA ECTS	PERIODO PREVISTO DE DESARROLLO
3 ECTS	Semanas 1 a 8



## 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 gestión de información y conocer las propuestas desde el modelo relacional
SI3.1	Conocer los últimos avances relacionados con bases de datos: alternativas al modelo relacional, BD distribuidas, en la web, etc.

### c. Contenidos

#### TEMA 4: Integración de información en la web y otros entornos

- 4.1 Integración de información en la Web de Datos
- 4.2 Aplicaciones prácticas: Wikipedia, Wikidata, Datos Abiertos Enlazados

### d. Métodos docentes

Ver apartado 5 de esta guía, sobre 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. Material docente

#### g.1 Bibliografía básica



AnHai Doan, Alon Haley, 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 .

Serge Abiteboul, Ioana Manolescu, Philippe Rigaux, Marie-Christine Rousset, and Pierre Senellart. 2011. *Web Data Management*. Cambridge University Press, New York, NY, USA.

## g.2 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.

## g.3 Otros recursos telemáticos (píldoras de conocimiento, blogs, videos, revistas digitales, cursos masivos (MOOC), ...)

Presentaciones, páginas de referencias, chats de discusión y otros recursos disponible en el aula virtual de la asignatura.

## h. Recursos necesarios

Ordenador para conectarse a los servidores de la escuela de Informática.

## i. Temporalización

CARGA ECTS	PERIODO PREVISTO DE DESARROLLO
3 ECTS	Semanas 9 a final

## 5. Métodos docentes y principios metodológicos

Actividad	Metodología
<b>Clase de teoría</b>	<ul style="list-style-type: none"> <li>Clase magistral participativa</li> <li>Estudio de casos en aula</li> <li>Resolución de problemas</li> </ul>
<b>Clase práctica</b>	<ul style="list-style-type: none"> <li>Clase magistral participativa</li> <li>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.</li> </ul>
<b>Seminarios</b>	<ul style="list-style-type: none"> <li>Talleres de aprendizaje</li> </ul>
<b>Tutoría</b>	<ul style="list-style-type: none"> <li>Evaluación de los contenidos teóricos y de los proyectos</li> </ul>

## 6. Tabla de dedicación del estudiante a la asignatura

ACTIVIDADES PRESENCIALES o PRESENCIALES A DISTANCIA <sup>(1)</sup>	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
TOTAL presencial + no presencial			

(1) Actividad presencial a distancia es cuando un grupo sigue una videoconferencia de forma síncrona a la clase impartida por el profesor.

## 7. Sistema y características de la evaluación

INSTRUMENTO/PROCEDIMIENTO	PESO EN LA NOTA FINAL	OBSERVACIONES
Valoración individual y supervisión de la participación del alumno en las diferentes actividades de la asignatura	10%	Se valorará la participación e implicación del alumno en las actividades de la asignatura
Entrega práctica	40%	La práctica consistirá en un proyecto, que se irá revisando en sucesivos hitos hasta llegar a la entrega final
Examen final escrito	50%	Período de exámenes

### CRITERIOS DE CALIFICACIÓN

- **Convocatoria ordinaria:**
  - Nota final = (Nota del examen escrito)\*0,50 + (Nota de la práctica)\*0,40 + (Participación)\*0,10
  - Es necesario obtener una calificación mínima de 4 en el examen final escrito para superar la asignatura.
- **Convocatoria extraordinaria:**
  - Mismo sistema que en la convocatoria ordinaria.
  - Es posible renunciar a la nota de la parte práctica y realizar únicamente un examen escrito, en cuyo caso su calificación aportaría el 100% de la nota final.

## 8. Consideraciones finales