



## Guía docente de la asignatura

<b>Asignatura</b>	TECNOLOGÍAS AVANZADAS DE GESTIÓN DE INFORMACIÓN		
<b>Materia</b>	PARADIGMAS Y TECNOLOGÍAS PARA LOS SISTEMAS Y SERVICIOS INFORMÁTICOS		
<b>Módulo</b>	ESPECIALIZACIÓN: TECNOLOGÍAS AVANZADAS DE GESTIÓN DE INFORMACIÓN (ISSI)		
<b>Titulación</b>	MÁSTER UNIVERSITARIO DE INVESTIGACIÓN EN TECNOLOGÍAS DE LA INFORMACIÓN Y LAS COMUNICACIONES		
<b>Plan</b>	371	<b>Código</b>	51324
<b>Periodo de impartición</b>	Cuatrimestre 1	<b>Tipo/Carácter</b>	OPTATIVA
<b>Nivel/Ciclo</b>	Máster	<b>Curso</b>	
<b>Créditos ECTS</b>	5 ECTS		
<b>Lengua en que se imparte</b>	CASTELLANO PARA LAS PRESENTACIONES ORALES, DOCUMENTACIÓN EN CASTELLANO / INGLÉS		
<b>Profesor/es responsable/s</b>	MANUEL BARRIO SOLÓRZANO PABLO DE LA FUENTE REDONDO MIGUEL ÁNGEL MARTÍNEZ PRIETO		
<b>Datos de contacto (E-mail, teléfono...)</b>	TELÉFONO: 983 185611 E-MAIL: <a href="mailto:mbarrio@infor.uva.es">mbarrio@infor.uva.es</a> , <a href="mailto:pfuente@infor.uva.es">pfuente@infor.uva.es</a> , <a href="mailto:migumar2@infor.uva.es">migumar2@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 Ingenieros de Telecomunicación → Tutorías		
<b>Departamento</b>	DEPARTAMENTO DE INFORMÁTICA (ATC, CCIA Y LSI)		



## 1. Situación / Sentido de la Asignatura

### 1.1 Contextualización

Es una asignatura de las tres en las que se divide la materia Paradigmas y Tecnologías para los Sistemas y Servicios informáticos (PTSSI). La materia TAG se ubica dentro del Bloque de especialización (ME) de Ingeniería de Sistemas y Servicios Informáticos (ISSI) y en particular pertenece al itinerario Ingeniería de Sistemas y Servicios Informáticos (ISSI). Esta materia optativa se apoya en las competencias generales fomentadas en el Bloque Básico (MB) del máster "Fundamentos de I+D+i en TIC" y forma la base para las materias de profundización dentro del mismo itinerario.

La considerable difusión de las aplicaciones informáticas y de la web en los últimos años plantea nuevos retos respecto a la gestión de grandes cantidades de información y la necesidad de localizar de forma rápida y eficiente. En esta asignatura se abordan los problemas de gestión de información, tanto estructurada (propia de los Sistemas de Gestión de Base de Datos) como no estructurada (XML y texto). También se revisan los mecanismos de búsqueda de dicha información, así como los nuevos retos que plantea la existencia de todo tipo de contenidos en la Web, como extracción de información, detección de opiniones, etc.

### 1.2 Relación con otras materias

Esta asignatura sienta las bases para afrontar la asignatura "Aplicaciones basadas en Servicios de Información" de la especialización ISSI.

### 1.3 Prerrequisitos

Esta asignatura optativa se apoya en las competencias generales fomentadas en el Bloque Básico (MB) del máster "Fundamentos de I+D+i en TIC".

Se supone que el alumno tiene conocimiento previo de estructuras de datos y de sistemas de gestión de bases de datos.



## 2. Competencias

Indicar las competencias que se desarrollan, de las descritas en el punto 3.2. de la memoria de verificación de la titulación y seleccionadas en el módulo, materia o asignatura correspondiente. Es conveniente identificarlas mediante letra y número, tal y como aparecen en la lista mencionada anteriormente.

### 2.1 Generales

- Capacidad de asumir una postura crítica hacia las tecnologías relacionadas con los sistemas y servicios informáticos como medio imprescindible para la detección de nuevos retos a resolver. [CG 1]
- Capacidad de analizar y aplicar los conocimientos técnicos específicos de los sistemas y servicios informáticos en nuevos entornos y contextos, teniendo en cuenta los parámetros y variables más significativas de cada nueva situación. [CG 5]
- Capacidad de comprender las implicaciones éticas y sociales de las decisiones adoptadas, así como las implicaciones relacionadas con la igualdad de sexo, raza o religión, la cultura de la paz, durante el ejercicio de las labores profesionales y de investigación. [CG 6, CG7]
- Capacidad de comunicar los resultados de investigación mediante artículos o comunicaciones escritos [CG8]
- Capacidad de emplear las técnicas y medios más adecuados para la comunicación oral en diversos foros de la comunidad académica, científica o empresarial, así como para su divulgación en general en la sociedad, o en otros términos, preparar y realizar presentaciones orales correctas tanto para audiencias expertas como en contextos divulgativos. [CG 9]
- Capacidad de aprendizaje y trabajo en grupo y en ámbitos multi-disciplinares, involucrados en el desarrollo de soluciones informáticas. [CG11]

### 2.2 Específicas

- Capacidad de comprender el campo de la ingeniería de sistemas y servicios informáticos y sus principales elementos, situándole dentro del sistema global de I+D+i [CE-ISSI 1]
- Capacidad de entender las principales tecnologías emergentes de la Internet de futuro y plantear preguntas de investigación sobre su idoneidad y sostenibilidad [CE-ISSI 3]
- Capacidad para comprender los modelos de datos, así como de comparar y seleccionar los mecanismos de representación de información más adecuados en entornos avanzados y servicios informáticos, así como los utilizados en la Web semántica [CE-ISSI 5]
- Capacidad para analizar de forma crítica los diferentes modelos y sistemas de recuperación y gestión de información y de evaluar, en términos de eficiencia y calidad de los resultados, diferentes alternativas de los mismos. [CE-ISSI 6]
- Capacidad para entender el alcance de nuevos paradigmas y estándares de representación de información y de evaluar sus posibilidades de contribuir al desarrollo de la sociedad digital [CE-ISSI 7]
- Capacidad para utilizar las técnicas de extracción y de gestión de información que permiten automatizar los tratamientos de grandes masas de información [CE-ISSI 8]



### 3. Objetivos

1. Poder comprender el campo de ingeniería de sistemas y servicios informáticos en el sistema global de I+D+i
2. Capacidad para analizar y aplicar los conocimientos técnicos específicos del área de sistemas y servicios informáticos en nuevos entornos y contextos
3. Tener una postura crítica hacia las tecnologías relacionadas con sistemas y servicios informáticos, así como de conceptos emergentes relacionados con el área
4. Poder comunicar los resultados de investigación mediante artefactos escritos y comunicación oral en el ámbito de la ingeniería de sistemas y servicios informáticos
5. Ser capaz de analizar las implicaciones éticas y sociales de las decisiones adoptadas de esta área, así como las relacionadas con la igualdad de sexo, raza o religión, y con la cultura de paz en las soluciones informáticas desarrolladas
6. Ser capaz de trabajar en grupo y especialmente en ámbitos multi-disciplinares.
7. Ser capaz de situar los diversos paradigmas, así como las arquitecturas, dentro del ámbito de los sistemas y servicios informáticos.
8. Entender las principales tecnologías emergentes de la Internet de Futuro y plantear preguntas de investigación sobre su idoneidad y sostenibilidad.
9. Conocer y utilizar los modelos de datos y mecanismos de representación más adecuados para entornos avanzados y servicios informáticos, incluyendo la web semántica.
10. Ser capaz de analizar de forma crítica los diferentes modelos y sistemas de recuperación y manipulación de información y de evaluar, en términos de eficiencia y calidad de los resultados, diferentes alternativas de los mismos.
11. Entender el alcance de nuevos paradigmas y estándares de representación y acceso a la información y de evaluar sus posibilidades de contribuir al desarrollo de la sociedad digital
12. Ser capaz de utilizar las técnicas de extracción y de manipulación de información que permiten automatizar los tratamientos de grandes masas de información.



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

ACTIVIDADES PRESENCIALES	HORAS	ACTIVIDADES NO PRESENCIALES	HORAS
Teoría	20	Estudio y trabajo autónomo individual	45
Prácticas en aula	24	Estudio y trabajo autónomo individual	30
Seminarios y Tutorías	6		
<b>Total presencial</b>	<b>50</b>	<b>Total no presencial</b>	<b>75</b>







## 5. Bloques temáticos<sup>1</sup>

### Bloque 1: Mecanismos eficientes de almacenamiento y acceso a la información

Carga de trabajo en créditos ECTS: 2

#### a. Contextualización y justificación

Véase Contextualización de la Asignatura.

#### b. Objetivos de aprendizaje

Al finalizar este bloque temático, el alumno deberá ser capaz de:

- Comprender los modelos de datos, así como de comparar y seleccionar los mecanismos de representación de información más adecuados en entornos avanzados y servicios informáticos.
- Analizar de forma crítica los diferentes modelos y sistemas de recuperación y gestión de información y de evaluar, en términos de eficiencia y calidad de los resultados, diferentes alternativas de los mismos.
- Entender el alcance de nuevos paradigmas y estándares de representación de información y de evaluar sus posibilidades de contribuir al desarrollo de la sociedad digital

#### c. Contenidos

- Bases de datos relacionales y sus limitaciones. Otros sistemas de bases de datos.
- Información estructurada, semi-estructurada y no-estructurada.
- Sistemas de información basados en lenguajes etiquetados. XML y modelo semi-estructurado.
- Lenguajes de consulta sobre XML.
- Estructuras de datos para sistemas de información estructurada, semi-estructurada y textual.

#### d. Métodos docentes

1. Actividades presenciales (0,8 ECTS, 40% del tiempo total = 20 horas)
  - a. Presentación en el aula mediante método de clase magistral participativa y no participativa
  - b. Seminarios de discusión entre los integrantes de la clase y con agentes externos, tutorías individuales y grupales, y actividades de evaluación: empleo principal del método de estudio de casos y de técnicas de aprendizaje colaborativo
  - c. Exposiciones orales de trabajos de alumnos
2. Actividades no presenciales (1,2 ECTS, 60% del tiempo total = 30 horas)
  - a. Trabajo individual: estudio de material proporcionado, selección de material relacionado, redacción de contribuciones y especialmente de informes sobre los casos de estudio
  - b. Trabajo en grupo hacia la elaboración de posturas comunes y la preparación de los seminarios presenciales: aprendizaje colaborativo, estudio de casos, método de proyectos

#### e. Plan de trabajo

Véase el Anexo, publicado en la web del MUITIC.

#### f. Evaluación

La evaluación de la adquisición de competencias y sistema de calificaciones se basará en los siguientes elementos:

- Trabajos e informes realizados por el alumno de forma individual y en grupo de trabajo (40%)
- Presentaciones orales de los trabajos (20%)
- Prueba escrita sobre conceptos fundamentales (30%)
- Valoración de la actitud y participación del alumno en las actividades formativas (10%)

#### g. Bibliografía básica

- G. Vossen. *Data Models, Database Languages and Database Management Systems*. Addison-Wesley 1990
- Serge Abiteboul, Peter Buneman, Dan Suciu. *Data on the web: from relations to semistructured data and XML*.



Morgan Kaufmann, 2000

#### h. Bibliografía complementaria

- W3C XML homepage, <http://www.w3.org/XML/>

### Bloque 2: Recuperación de Información

Carga de trabajo en créditos ECTS:

#### a. Contextualización y justificación

Véase Contextualización de la Asignatura.

#### b. Objetivos de aprendizaje

Al finalizar este bloque temático, el alumno deberá ser capaz de:

- Analizar de forma crítica los diferentes modelos y sistemas de recuperación y gestión de información y de evaluar, en términos de eficiencia y calidad de los resultados, diferentes alternativas de los mismos. [CE-ISSI 6]
- Entender el alcance de nuevos paradigmas y estándares de representación de información y de evaluar sus posibilidades de contribuir al desarrollo de la sociedad digital [CE-ISSI 7]

#### c. Contenidos

##### Recuperación de información.

- Modelos
- Operaciones sobre el texto
- Evaluación de modelos y sistemas de búsqueda

#### d. Métodos docentes

1. Actividades presenciales (0,4 ECTS, 40% del tiempo total = 10 horas)
  - a. Presentación en el aula mediante método de clase magistral participativa y no participativa
  - b. Seminarios de discusión entre los integrantes de la clase y con agentes externos, tutorías individuales y grupales, y actividades de evaluación: empleo principal del método de estudio de casos y de técnicas de aprendizaje colaborativo
  - c. Exposiciones orales de trabajos de alumnos
2. Actividades no presenciales (0,6 ECTS, 60% del tiempo total = 15 horas)
  - a. Trabajo individual: estudio de material proporcionado, selección de material relacionado, redacción de contribuciones y especialmente de informes sobre los casos de estudio
  - b. Trabajo en grupo hacia la elaboración de posturas comunes y la preparación de los seminarios presenciales: aprendizaje colaborativo, estudio de casos, método de proyectos

#### e. Plan de trabajo

Véase el Anexo, publicado en la web del MUITIC.

#### f. Evaluación

La evaluación de la adquisición de competencias y sistema de calificaciones se basará en los siguientes elementos:

- Trabajos e informes realizados por el alumno de forma individual y en grupo de trabajo (40%)
- Presentaciones orales de los trabajos (20%)
- Prueba escrita sobre conceptos fundamentales (30%)
- Valoración de la actitud y participación del alumno en las actividades formativas (10%)

#### g. Bibliografía básica

- W. Bruce Croft, Donal Metzler, Trevor Strohman. *Search Engines. Information Retrieval in practice*. Pearson 2010
- Ricardo Baeza-Yates, Berthier Ribeiro-Neto. *Modern Information Retrieval: The Concepts and Technology behind Search* (2nd Edition). 2011
- Manning, Raghavan, and Schütze. *Introduction to Information Retrieval*, Cambridge University Press, 2007



#### **h. Bibliografía complementaria**

<http://nlp.stanford.edu/IR-book/information-retrieval-book.html>

<http://people.ischool.berkeley.edu/~hearst/irbook/>

[http://www.ir.iit.edu/~dagr/cs529/ir\\_book.html](http://www.ir.iit.edu/~dagr/cs529/ir_book.html)

Se proporcionarán otros enlaces así como diversos artículos científicos para su lectura y análisis durante el curso.

#### **i. Recursos necesarios**

Se utilizará software gratuito para las distintas partes de los trabajos prácticos. Las aplicaciones serán proporcionadas por los profesores de la asignatura.

5.

### **Bloque 3: Búsqueda en la Web**

Carga de trabajo en créditos ECTS:

#### **a. Contextualización y justificación**

Véase contextualización de la asignatura.

#### **b. Objetivos de aprendizaje**

Al finalizar este bloque temático, el alumno deberá ser capaz de:

- Analizar de forma crítica las diferentes máquinas de búsqueda en la web y de evaluar, en términos de eficiencia y calidad de los resultados, diferentes alternativas de los mismos. [CE-ISSI 6]
- Entender el alcance de nuevos paradigmas y estándares de representación de información y de evaluar sus posibilidades de contribuir al desarrollo de la sociedad digital [CE-ISSI 7]
- Utilizar las técnicas de extracción y de gestión de información que permiten automatizar los tratamientos de grandes masas de información [CE-ISSI 8]

#### **c. Contenidos**

Representación y extracción de información semántica y su utilización en la web:

- Búsqueda en la Web
- Mecanismos de extracción de información y aplicaciones: detección de entidades con nombre, extracción de opiniones, etc.

#### **d. Métodos docentes**

1. Actividades presenciales (0,4 ECTS, 40% del tiempo total = 10 horas)
  - a. Presentación en el aula mediante método de clase magistral participativa y no participativa
  - b. Seminarios de discusión entre los integrantes de la clase y con agentes externos, tutorías individuales y grupales, y actividades de evaluación: empleo principal del método de estudio de casos y de técnicas de aprendizaje colaborativo
  - c. Exposiciones orales de trabajos de alumnos
2. Actividades no presenciales (0,6 ECTS, 60% del tiempo total = 15 horas)
  - a. Trabajo individual: estudio de material proporcionado, selección de material relacionado, redacción de contribuciones y especialmente de informes sobre los casos de estudio
  - b. Trabajo en grupo hacia la elaboración de posturas comunes y la preparación de los seminarios presenciales: aprendizaje colaborativo, estudio de casos, método de proyectos

#### **e. Plan de trabajo**

Véase el Anexo, publicado en la web del MUITIC.

#### **f. Evaluación**

La evaluación de la adquisición de competencias y sistema de calificaciones se basará en los siguientes elementos:

- Trabajos e informes realizados por el alumno de forma individual y en grupo de trabajo (40%)





- Presentaciones orales de los trabajos (20%)
- Prueba escrita sobre conceptos fundamentales (30%)
- Valoración de la actitud y participación del alumno en las actividades formativas (10%)

#### g. Bibliografía básica

- Manning, Raghavan, and Schütze. *Introduction to Information Retrieval*, Cambridge University Press, 2007
- W. Bruce Croft, Donal Metzler, Trevor Strohman. *Search Engines. Information Retrieval in practice*. Pearson 2010
- Ricardo Baeza-Yates, Berthier Ribeiro-Neto. *Modern Information Retrieval: The Concepts and Technology behind Search* (2nd Edition). 2011

#### h. Bibliografía complementaria

<http://nlp.stanford.edu/IR-book/information-retrieval-book.html>

<http://www.pearsonhighered.com/croft1epreview/>

Se proporcionarán otros enlaces así como diversos artículos científicos para su lectura y análisis durante el curso.

#### i. Recursos necesarios

Se utilizará software gratuito para las distintas partes de los trabajos prácticos. Las aplicaciones serán proporcionadas por los profesores de la asignatura.

### Bloque 4: Compresión de grandes volúmenes de datos

Carga de trabajo en créditos ECTS:

#### a. Contextualización y justificación

Véase contextualización de la asignatura.

#### b. Objetivos de aprendizaje

Al finalizar este bloque temático, el alumno deberá ser capaz de:

- Entender el alcance de nuevos paradigmas y estándares de representación de información y de evaluar sus posibilidades de contribuir al desarrollo de la sociedad digital [CE-ISSI 7]
- Capacidad para utilizar las técnicas de extracción y de gestión de información que permiten automatizar los tratamientos de grandes masas de información [CE-ISSI 8]

#### c. Contenidos

**La Revolución de los Datos (Volumen, Velocidad, Variedad)**

**Compresión de Datos:**

1. Conceptos Fundamentales de Teoría de la Información.
2. Compresión Estadística, basada en Diccionario y Otros Métodos de Compresión.

**Estructuras de Datos Compactas:**

3. Conceptos Fundamentales y Operaciones Primitivas.
4. Secuencias de Bits, Enteros y Otras Estructuras de Datos.

**Auto-índices:**

Conceptos Fundamentales y Técnicas Principales

#### d. Métodos docentes

##### 1. Actividades presenciales (0,4 ECTS, 40% del tiempo total = 10 horas)

1. Presentación en el aula mediante método de clase magistral participativa y no participativa.
2. Seminarios de discusión entre los integrantes de la clase y con agentes externos, tutorías individuales y grupales, y actividades de evaluación: empleo principal del método de estudio de casos y de técnicas de aprendizaje colaborativo

##### 2. Actividades no presenciales (0,6 ECTS, 60% del tiempo total = 15 horas)

Trabajo individual: estudio de material proporcionado, selección de material relacionado, redacción de contribuciones y especialmente de informes sobre los casos de estudio



Trabajo en grupo hacia la elaboración de posturas comunes y la preparación de los seminarios presenciales: aprendizaje colaborativo, estudio de casos, método de proyectos

#### e. Plan de trabajo

Véase el Anexo, publicado en la web del MUITIC.

#### f. Evaluación

La evaluación de la adquisición de competencias y sistema de calificaciones se basará en los siguientes elementos:

- Trabajos e informes realizados por el alumno de forma individual y en grupo de trabajo (40%)
- Presentaciones orales de los trabajos (20%)
- Prueba escrita sobre conceptos fundamentales (30%)
- Valoración de la actitud y participación del alumno en las actividades formativas (10%)

#### g. Bibliografía básica

- K. Sayood. Introduction to Data Compression 3rd Edition. Morgan Kaufmann, 2005
- G. Navarro, V. Mäkinen. Compressed Full-Text Indexes, ACM Computing Surveys 39(1), article 2, 2007.

#### h. Bibliografía complementaria

<http://www.datacompression.info/>

<http://www.dcc.uchile.cl/~gnavarro/tutorial.pdf>

<http://dataweb.infor.uva.es/aprendiendo-a-nadar-en-el-diluvio-de-datos/>

Se proporcionarán otros enlaces así como diversos artículos científicos para su lectura y análisis durante el curso.

#### i. Recursos necesarios

Se utilizará software gratuito para las distintas partes de los trabajos prácticos.

<https://code.google.com/p/libcds/>

<http://pizzachili.dcc.uchile.cl/>

### 6. Temporalización (por bloques temáticos)

BLOQUE TEMÁTICO	CARGA ECTS	PERIODO PREVISTO DE DESARROLLO
Mecanismos eficientes de almacenamiento y acceso a la información	2	
Recuperación de información	1	
Búsqueda en la web	1	
Compresión de grandes volúmenes de datos	1	

### 7. Tabla resumen de los instrumentos, procedimientos y sistemas de evaluación/calificación

INSTRUMENTO/PROCEDIMIENTO	PESO EN LA NOTA FINAL	OBSERVACIONES
Valoración de los trabajos e informes realizados por el alumno de forma individual y en grupo de trabajo	40%	Será necesario entregar todos los trabajos encargados en cada bloque.
Valoración de las presentaciones orales de los trabajos	20%	Las sesiones de presentación podrán realizarse en inglés para poder dar contenido a la competencia transversal correspondiente.



Prueba escrita sobre conceptos fundamentales	30%	Es condición necesaria (pero no suficiente) para superar la asignatura alcanzar una calificación igual o superior a 5 para superar el bloque
Valoración de la actitud y participación del alumno en las actividades formativas	10%	

## 8. Consideraciones finales

