

Plan 510 MÁSTER EN INGENIERÍA INFORMÁTICA

Asignatura 53181 WEB SEMÁNTICA Y EXTRACCIÓN DE INFORMACIÓN

Grupo 1

**Tipo de asignatura (básica, obligatoria u optativa)**

Optativa

**Créditos ECTS**

3

**Competencias que contribuye a desarrollar**

P { margin-bottom: 0.21cm; direction: ltr; color: rgb(0, 0, 0); widows: 2; orphans: 2; }P.western { font-family: "Times New Roman",serif; font-size: 12pt; }P.cjk { font-family: "Times New Roman",serif; font-size: 12pt; }P.cjl { font-family: "Times New Roman",serif; font-size: 12pt; }P.cjk { font-family: "Times New Roman",serif; font-size: 12pt; }P.cjl { font-family: "Times New Roman",serif; font-size: 12pt; }A:visited { color: rgb(128, 0, 128); }A.western:visited { }A.cjk:visited { }A.cjl:visited { }A:link { color: rgb(0, 0, 255); }A.cjl:link { font-family: "Times New Roman",serif; }

**CG1**

Capacidad para proyectar, calcular y diseñar productos, procesos e instalaciones en todos los ámbitos de la ingeniería informática.

**CET5**

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

**CET9**

Capacidad para aplicar métodos matemáticos, estadísticos y de inteligencia artificial para modelar, diseñar y desarrollar aplicaciones, servicios, sistemas inteligentes y sistemas basados en el conocimiento

P { margin-bottom: 0.21cm; direction: ltr; color: rgb(0, 0, 0); widows: 2; orphans: 2; }P.western { font-family: "Times New Roman",serif; font-size: 12pt; }P.cjk { font-family: "Times New Roman",serif; font-size: 12pt; }P.cjl { font-family: "Times New Roman",serif; font-size: 12pt; }A:visited { color: rgb(128, 0, 128); }A.western:visited { }A.cjk:visited { }A.cjl:visited { }A:link { color: rgb(0, 0, 255); }A.cjl:link { font-family: "Times New Roman",serif; }

**CA1**

Capacidad para comprender y responder a las necesidades de representación de información, conocimiento y semántica en entornos de compartición masiva de información, tales como la web

**CA2**

Capacidad para analizar las necesidades de extracción de información, y aplicar las técnicas adecuadas para extraer información de fuentes de datos no estructuradas

**Objetivos/Resultados de aprendizaje**

P { margin-bottom: 0.21cm; direction: ltr; color: rgb(0, 0, 0); widows: 2; orphans: 2; }P.western { font-family: "Times New Roman",serif; font-size: 12pt; }P.cjk { font-family: "Times New Roman",serif; font-size: 12pt; }P.cjl { font-family: "Times New Roman",serif; font-size: 12pt; }A:visited { color: rgb(128, 0, 128); }A.western:visited { }A.cjk:visited { }A.cjl:visited { }A:link { color: rgb(0, 0, 255); }A.cjl:link { font-family: "Times New Roman",serif; }

**R1**

Entender los retos de interoperabilidad que supone la compartición de información en la web

**R2**

Ser capaz de conocer y utilizar los modelos de datos asociados a la web para compartir datos y semántica

**R3**

Conocer y saber aplicar las tecnologías (estándares, lenguajes de consulta, otros) asociados a estos modelos de datos

**R4**

Reconocer las necesidades de extracción de información en un problema real y proponer soluciones tecnológicas capaces de resolverla

P { margin-bottom: 0.21cm; direction: ltr; color: rgb(0, 0, 0); widows: 2; orphans: 2; }P.western { font-family: "Times New Roman",serif; font-size: 12pt; }P.cjk { font-family: "Times New Roman",serif; font-size: 12pt; }P.cjl { font-family: "Times New Roman",serif; font-size: 12pt; }P.cjk { font-family: "Times New Roman",serif; font-size: 12pt; }P.cjl { font-family: "Times New Roman",serif; font-size: 12pt; }A:visited { color: rgb(128, 0, 128); }A.western:visited { }A.cjk:visited { }A.cjl:visited { }A:link { color: rgb(0, 0, 255); }A.cjl:link { font-family: "Times New Roman",serif; }

- R5  
Seleccionar la técnica de extracción de información más adecuada para cada problema
- R6  
Combinar diferentes técnicas básicas para obtener sistemas más eficientes de extracción de información

## Contenidos

P { margin-bottom: 0.21cm; direction: ltr; color: rgb(0, 0, 0); widows: 2; orphans: 2; }P.western { font-family: "Times New Roman",serif; font-size: 12pt; }P.cjk { font-family: "Times New Roman",serif; font-size: 12pt; }P.cjl { font-family: "Times New Roman",serif; font-size: 12pt; }A:visited { color: rgb(128, 0, 128); }A.western:visited { }A.cjk:visited { }A.cjl:visited { }A:link { color: rgb(0, 0, 255); }A.cjl:link { font-family: "Times New Roman",serif; }

TEMA 1: Introducción al problema de la información en la web y sus soluciones

- 1.1 Deficiencias actuales del acceso a la información en la web.
- 1.2 Web de datos: semántica en la web.
- 1.2 Estudio de aplicaciones reales de web semántica y su valor añadido.
- 1.3 Perspectivas de desarrollos futuros.

TEMA 2: Tecnologías para la representación, manipulación y compartición de semántica en la Web

P { margin-bottom: 0.21cm; direction: ltr; color: rgb(0, 0, 0); widows: 2; orphans: 2; }P.western { font-family: "Times New Roman",serif; font-size: 12pt; }P.cjk { font-family: "Times New Roman",serif; font-size: 12pt; }P.cjl { font-family: "Times New Roman",serif; font-size: 12pt; }A:visited { color: rgb(128, 0, 128); }A.western:visited { }A.cjk:visited { }A.cjl:visited { }A:link { color: rgb(0, 0, 255); }A.cjl:link { font-family: "Times New Roman",serif; }

- 2.1 Modelos de datos usados para la representación de semántica en la web. Sus posibilidades y uso en casos reales.
- 2.2 Acceso a la información semántica: lenguajes de consulta, estrategias para la integración.
- 2.3 Casos de estudio: datos enlazados, mashup semánticos.

P { margin-bottom: 0.21cm; direction: ltr; color: rgb(0, 0, 0); widows: 2; orphans: 2; }P.western { font-family: "Times New Roman",serif; font-size: 12pt; }P.cjk { font-family: "Times New Roman",serif; font-size: 12pt; }P.cjl { font-family: "Times New Roman",serif; font-size: 12pt; }A:visited { color: rgb(128, 0, 128); }A.western:visited { }A.cjk:visited { }A.cjl:visited { }A:link { color: rgb(0, 0, 255); }A.cjl:link { font-family: "Times New Roman",serif; }

TEMA 3: Introducción al tratamiento de cadenas:

- 3.1 Definiciones: alfabeto, cadenas, lenguajes, términos, distancia entre cadenas.
- 3.2 Representación de lenguajes: autómatas y gramáticas.
- 3.3 Representación de distribuciones de cadenas mediante autómatas y gramáticas.
  - 3.3.1 Autómatas probabilísticos
  - 3.3.2 Gramáticas independientes del contexto probabilísticas
  - 3.3.3 Distancia entre gramáticas
  - 3.3.4 Distancia entre distribuciones de cadenas

TEMA 4: Técnicas de aprendizaje automático de lenguajes y su aplicación en la extracción de información

P { margin-bottom: 0.21cm; direction: ltr; color: rgb(0, 0, 0); widows: 2; orphans: 2; }P.western { font-family: "Times New Roman",serif; font-size: 12pt; }P.cjk { font-family: "Times New Roman",serif; font-size: 12pt; }P.cjl { font-family: "Times New Roman",serif; font-size: 12pt; }A:visited { color: rgb(128, 0, 128); }A.western:visited { }A.cjk:visited { }A.cjl:visited { }A:link { color: rgb(0, 0, 255); }A.cjl:link { font-family: "Times New Roman",serif; }

- 4.1 Identificación de lenguajes.
- 4.2 Aprendizaje a partir de texto.
- 4.3. Aprendizaje activo aplicado a la Web Semántica
- 4.4 Aprendizaje de distribuciones de cadenas.
- 4.5 Aprendizaje de texto corregido mediante aproximaciones probabilísticas.

## Principios Metodológicos/Métodos Docentes

P { margin-bottom: 0.21cm; direction: ltr; color: rgb(0, 0, 0); widows: 2; orphans: 2; }P.western { font-family: "Times New Roman",serif; font-size: 12pt; }P.cjk { font-family: "Times New Roman",serif; font-size: 12pt; }P.cjl { font-family: "Times New Roman",serif; font-size: 12pt; }A:visited { color: rgb(128, 0, 128); }A.western:visited { }A.cjk:visited { }A.cjl:visited { }A:link { color: rgb(0, 0, 255); }A.cjl:link { font-family: "Times New Roman",serif; }

- Sesiones de aula:
  - Clases magistrales participativas y expositivas
  - Aprendizaje basado en problemas
  - Estudios de caso.
- Laboratorio y Prácticas supervisadas:

- Resolución de problemas y casos prácticos.
- Aprendizaje basado en problemas.
- Aprendizaje cooperativo.
- Estudios del caso.
- Método de proyectos.
- Seminarios:
  - Aprendizaje basado en problemas
  - Estudios del caso
  - Aprendizaje cooperativo

## Crterios y sistemas de evaluaci3n

P { margin-bottom: 0.21cm; direction: ltr; color: rgb(0, 0, 0); widows: 2; orphans: 2; }P.western { font-family: "Times New Roman",serif; font-size: 12pt; }P.cjk { font-family: "Times New Roman",serif; font-size: 12pt; }P.cnt { font-family: "Times New Roman",serif; font-size: 12pt; }A:visited { color: rgb(128, 0, 128); }A.western:visited { }A.cjk:visited { }A.cnt:visited { }A:link { color: rgb(0, 0, 255); }A.cnt:link { font-family: "Times New Roman",serif; }

### INSTRUMENTO/PROCEDIMIENTO

#### PESO EN LA NOTA FINAL

#### OBSERVACIONES

Entrega pr3ctica

40%

La pr3ctica consistir3 en un proyecto, que se ir3 revisando en sucesivos hitos hasta llegar a la entrega final.

Valoraci3n del trabajo realizado en el laboratorio

5%

Cada tema tiene una serie de ejercicios que los alumnos deben realizar para comprobar su nivel de comprensi3n y evoluci3n.

Participaci3n en las actividades

5%

## Recursos de aprendizaje y apoyo tutorial

P { margin-bottom: 0.21cm; direction: ltr; color: rgb(0, 0, 0); widows: 2; orphans: 2; }P.western { font-family: "Times New Roman",serif; font-size: 12pt; }P.cjk { font-family: "Times New Roman",serif; font-size: 12pt; }P.cnt { font-family: "Times New Roman",serif; font-size: 12pt; }A:visited { color: rgb(128, 0, 128); }A.western:visited { }A.cjk:visited { }A.cnt:visited { }A:link { color: rgb(0, 0, 255); }A.cnt:link { font-family: "Times New Roman",serif; }

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

## Calendario y horario

Segundo cuatrimestre. Para detalles ver p3gina web de la escuela:

<https://www.inf.uva.es/2013/04/16/master-en-ingenieria-informatica/>

## Tabla de Dedicaci3n del Estudiante a la Asignatura/Plan de Trabajo

P { margin-bottom: 0.21cm; direction: ltr; color: rgb(0, 0, 0); widows: 2; orphans: 2; }P.western { font-family: "Times New Roman",serif; font-size: 12pt; }P.cjk { font-family: "Times New Roman",serif; font-size: 12pt; }P.cnt { font-family: "Times New Roman",serif; font-size: 12pt; }A:visited { color: rgb(128, 0, 128); }A.western:visited { }A.cjk:visited { }A.cnt:visited { }A:link { color: rgb(0, 0, 255); }A.cnt:link { font-family: "Times New Roman",serif; }

TD P { margin-bottom: 0cm; direction: ltr; color: rgb(0, 0, 0); widows: 2; orphans: 2; }TD P.western { font-family: "Times New Roman",serif; font-size: 12pt; }TD P.cjk { font-family: "Times New Roman",serif; font-size: 12pt; }TD P.cnt { font-family: "Times New Roman",serif; font-size: 12pt; }P { margin-bottom: 0.21cm; direction: ltr; color: rgb(0, 0, 0); widows: 2; orphans: 2; }P.western { font-family: "Times New Roman",serif; font-size: 12pt; }P.cjk { font-family: "Times New Roman",serif; font-size: 12pt; }P.cnt { font-family: "Times New Roman",serif; font-size: 12pt; }A:visited { color: rgb(128, 0, 128); }A.western:visited { }A.cjk:visited { }A.cnt:visited { }A:link { color: rgb(0, 0, 255); }A.cnt:link { font-family: "Times New Roman",serif; }

Semana

Contenido

Actividad presencial

Actividad no presencial

Semana 1

Tema 1

Sesiones en aula y laboratorio: 4h.

---

Semana 2  
Temas 1 y 2  
Sesiones en aula y laboratorio: 4 h.  
Estudio y solución de problemas. Estudio de casos.

Semana 3  
Tema 2  
Sesiones en aula y laboratorio: 4 h.  
Estudio y solución de problemas. Estudio de casos.

Semana 4  
Tema 2  
Seminario (2h)  
Seguimiento del trabajo práctico del bloque 1.

Semana 5

Semana 6

Semana 7

Semana 8

Seminario (1h)  
Seguimiento del trabajo práctico del bloque 1.

```
P { margin-bottom: 0.21cm; direction: ltr; color: rgb(0, 0, 0); widows: 2; orphans: 2; }P.western { font-family: "Times New Roman", serif; font-size: 12pt; }P.cjk { font-family: "Times New Roman", serif; font-size: 12pt; }P.cnt { font-family: "Times New Roman", serif; font-size: 12pt; }A:visited { color: rgb(128, 0, 128); }A.western:visited { }A.cjk:visited { }A.cnt:visited { }A:link { color: rgb(0, 0, 255); }A.cnt:link { font-family: "Times New Roman", serif; }
```

Semana  
Contenido  
Actividad presencial  
Actividad no presencial  
Semana 1

Semana 2

Semana 3

Semana 4  
Tema 3  
Sesiones en aula y laboratorio: 2h.  
Estudio y solución de problemas. Estudio de casos.  
Semana 5  
Tema 3  
Sesiones en aula y laboratorio: 4h  
Estudio y solución de problemas. Estudio de casos.

---

---

Semana 6

Tema 4

Sesiones en aula y laboratorio: 4h

Estudio y solución de problemas. Estudio de casos.

Semana 7

Tema 4

Sesiones en aula y laboratorio: 4h

Estudio y solución de problemas. Estudio de casos.

Semana 8

Tema 4

Seminario (1h)

Evaluación (2h)

Seguimiento del trabajo práctico del bloque 2.

---

Responsable de la docencia (recomendable que se incluya información de contacto y breve CV en el que aparezcan sus líneas de investigación y alguna publicación relevante)

Mercedes Martínez González y Teodoro Calonge Cano.

Información de contacto y otros disponible en la web del departamento de Informática: <http://www.infor.uva.es/>

---

Idioma en que se imparte

Castellano