

ANX-PR/CL/001-02
GUÍA DE APRENDIZAJE

ASIGNATURA

Recuperacion e integracion de informacion de recursos cientificos

CURSO ACADÉMICO - SEMESTRE

2015-16 - Primer semestre

Datos Descriptivos

Nombre de la Asignatura	Recuperacion e integracion de informacion de recursos cientificos
Titulación	10AN - Master Universitario en Ingenieria Informatica
Centro responsable de la titulación	E.T.S. de Ingenieros Informaticos
Semestre/s de impartición	Tercer semestre
Carácter	Optativa
Código UPM	103000653
Nombre en inglés	Information Retrieval And Integration For Scientific Resources

Datos Generales

Créditos	4.5	Curso	2
Curso Académico	2015-16	Período de impartición	Septiembre-Enero
Idioma de impartición	Castellano	Otros idiomas de impartición	

Requisitos Previos Obligatorios

Asignaturas Superadas

El plan de estudios Master Universitario en Ingenieria Informatica no tiene definidas asignaturas previas superadas para esta asignatura.

Otros Requisitos

El plan de estudios Master Universitario en Ingenieria Informatica no tiene definidos otros requisitos para esta asignatura.

Conocimientos Previos

Asignaturas Previas Recomendadas

El coordinador de la asignatura no ha definido asignaturas previas recomendadas.

Otros Conocimientos Previos Recomendados

El coordinador de la asignatura no ha definido otros conocimientos previos recomendados.

Competencias

CB10 - Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.

CB7 - Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio

CE1 - Capacidad para la integración de tecnologías, aplicaciones, servicios y sistemas propios de la Ingeniería Informática, con carácter generalista, y en contextos más amplios y multidisciplinares.

CE8 - 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.

CG4 - Capacidad para la aplicación de los conocimientos adquiridos y de resolver problemas en entornos nuevos o poco conocidos dentro de contextos más amplios y multidisciplinares, siendo capaces de integrar estos conocimientos

CG5 - Aplicación de los métodos de resolución de problemas más recientes o innovadores y que puedan implicar el uso de otras disciplinas

CG7 - Integración del conocimiento a partir de disciplinas diferentes, así como el manejo de la complejidad

CG9 - Apreciación de los límites del conocimiento actual y de la aplicación práctica de la tecnología más reciente

Resultados de Aprendizaje

RA159 - Conocer los requisitos específicos de sistemas de información en biomedicina

RA160 - Comprender y aplicar los fundamentos de los sistemas de recuperación de información

RA58 - Conocer y diseñar sistemas de extracción de Información

RA92 - Ser capaz de diseñar la arquitectura de un sistema de información

Profesorado

Profesorado

Nombre	Despacho	e-mail	Tutorías
Perez Del Rey, David (Coordinador/a)	2104	david.perez.rey@upm.es	L - 14:00 - 16:00 X - 14:00 - 16:00 V - 11:00 - 13:00
Maojo Garcia, Victor Manuel	2102	victormanuel.maojo@upm.es	M - 12:30 - 15:30 X - 12:30 - 15:30
Garcia Remesal, Miguel	2206	miguel.garcia.remesal@upm.es	M - 11:00 - 14:00 J - 11:00 - 14:00
Suarez De Figueroa Baonza, M. Carmen	3205	mdelcarmen.suarezdefigueroa@upm.es	M - 14:00 - 16:00 X - 11:00 - 13:00 J - 14:00 - 16:00

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

Personal Investigador en Formación o Similar

Nombre	e-mail	Profesor Responsable
Leitner , Florian	florian.leitner@upm.es	Perez Del Rey, David

Profesorado Externo

Nombre	e-mail	Centro de procedencia
Alonso Calvo, Raul	ralonso@fi.upm.es	ETSInf - DLSIS

Descripción de la Asignatura

La cantidad de datos disponibles en cualquier área ha crecido dramáticamente durante los últimos años. Sin embargo, este incremento no ha tenido un impacto proporcional en el conocimiento disponible para la toma de decisiones. Son necesarios métodos automáticos para la gestión de estos datos, teniendo en cuenta que la mayoría jamás serán utilizados por un humano. La asignatura de Recuperación e Información, se centrará en las tareas necesarias para extraer información, facilitar modelos de recuperación eficiente para su posterior integración. Tareas críticas a la hora de facilitar información para la toma de decisiones, y cuya complejidad aumenta cuanto mayor cantidad de datos hay disponibles. Como área de aplicación, se estudia la biomedicina, debido a su complejidad y a que cuenta con unos requisitos específicos de seguridad, localización, etc.

Temario

1. Conceptos básicos
 - 1.1. Introducción
 - 1.2. Datos, información y conocimiento
 - 1.3. Tipos de datos
2. Extracción y recuperación de información
 - 2.1. Extracción de información
 - 2.2. Modelos de recuperación de información
 - 2.3. Procesamiento de lenguaje natural
 - 2.4. Buscadores web
 - 2.5. Tratamiento de datos no textuales
3. Integración de datos
 - 3.1. Arquitecturas de integración
 - 3.2. Traducción de consultas
 - 3.3. Interoperabilidad semántica
 - 3.4. Procedencia de datos
4. Aplicaciones en biomedicina
 - 4.1. Sistemas de información biomédica
 - 4.2. Estándares de interoperabilidad clínica
 - 4.3. Vocabularios biomédicos
 - 4.4. Sistemas de recuperación de literatura científica

Cronograma

Horas totales: 119 horas

Horas presenciales: 50 horas (42.7%)

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	<p>Presentación de la asignatura Duración: 01:00 LM: Actividad del tipo Lección Magistral</p> <p>Desarrollo del tema 1.1 Duración: 02:00 LM: Actividad del tipo Lección Magistral</p>			<p>Estudio autónomo Duración: 04:00 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial</p>
Semana 2	<p>Desarrollo del Tema 1.2 Duración: 02:00 LM: Actividad del tipo Lección Magistral</p> <p>Desarrollo del Tema 1.3 Duración: 01:00 LM: Actividad del tipo Lección Magistral</p>			<p>Estudio autónomo Duración: 04:00 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial</p>
Semana 3	<p>Desarrollo del Tema 1.3 Duración: 02:00 LM: Actividad del tipo Lección Magistral</p> <p>Desarrollo del Tema 2.1 Duración: 01:00 LM: Actividad del tipo Lección Magistral</p>			<p>Estudio autónomo Duración: 04:00 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial</p>
Semana 4	<p>Desarrollo del Tema 2.1 Duración: 02:00 LM: Actividad del tipo Lección Magistral</p> <p>Desarrollo del Tema 2.2 Duración: 01:00 LM: Actividad del tipo Lección Magistral</p>			<p>Estudio autónomo Duración: 05:00 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial</p>
Semana 5	<p>Desarrollo del Tema 2.2 Duración: 02:00 LM: Actividad del tipo Lección Magistral</p> <p>Desarrollo del Tema 2.3 Duración: 01:00 LM: Actividad del tipo Lección Magistral</p>			<p>Entrega trabajo individual Duración: 06:00 TI: Técnica del tipo Trabajo Individual Evaluación continua y sólo prueba final Actividad no presencial</p>

Semana 6	<p>Desarrollo del Tema 2.3 Duración: 01:00 LM: Actividad del tipo Lección Magistral</p>			<p>Exposición del trabajo individual Duración: 02:00 TI: Técnica del tipo Trabajo Individual Evaluación continua y sólo prueba final Actividad presencial Estudio autónomo Duración: 03:00 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial</p>
Semana 7	<p>Desarrollo del Tema 2.3 Duración: 01:00 LM: Actividad del tipo Lección Magistral Desarrollo del Tema 2.4 Duración: 02:00 LM: Actividad del tipo Lección Magistral</p>			<p>Estudio autónomo Duración: 04:00 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial</p>
Semana 8	<p>Desarrollo del Tema 2.5 Duración: 02:00 LM: Actividad del tipo Lección Magistral Trabajo en grupo supervisado Duración: 01:00 AC: Actividad del tipo Acciones Cooperativas</p>			<p>Estudio autónomo Duración: 04:00 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial</p>
Semana 9	<p>Desarrollo del Tema 2.5 Duración: 02:00 LM: Actividad del tipo Lección Magistral Trabajo en grupo supervisado Duración: 01:00 AC: Actividad del tipo Acciones Cooperativas</p>			<p>Entrega trabajo en grupo Duración: 06:00 TG: Técnica del tipo Trabajo en Grupo Evaluación continua Actividad no presencial</p>
Semana 10	<p>Desarrollo del Tema 3.1 Duración: 01:00 LM: Actividad del tipo Lección Magistral</p>			<p>Exposición de trabajos en grupo Duración: 02:00 TG: Técnica del tipo Trabajo en Grupo Evaluación continua Actividad presencial Estudio autónomo Duración: 03:00 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial</p>
Semana 11	<p>Desarrollo del Tema 3.1 Duración: 03:00 LM: Actividad del tipo Lección Magistral</p>			<p>Estudio autónomo Duración: 04:00 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial</p>

Semana 12	<p>Desarrollo del Tema 3.1 Duración: 01:00 LM: Actividad del tipo Lección Magistral</p> <p>Desarrollo del Tema 3.2 Duración: 02:00 LM: Actividad del tipo Lección Magistral</p>			<p>Estudio autónomo Duración: 04:00 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial</p>
Semana 13	<p>Desarrollo del Tema 3.2 Duración: 02:00 LM: Actividad del tipo Lección Magistral</p> <p>Desarrollo del Tema 3.3 Duración: 01:00 LM: Actividad del tipo Lección Magistral</p>			<p>Estudio autónomo Duración: 04:00 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial</p>
Semana 14	<p>Desarrollo del Tema 3.3 Duración: 02:00 LM: Actividad del tipo Lección Magistral</p> <p>Desarrollo del Tema 3.4 Duración: 01:00 LM: Actividad del tipo Lección Magistral</p>			<p>Estudio autónomo Duración: 04:00 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial</p>
Semana 15	<p>Desarrollo del Tema 3.4 Duración: 03:00 LM: Actividad del tipo Lección Magistral</p>			<p>Entrega de trabajo en grupo Duración: 06:00 TG: Técnica del tipo Trabajo en Grupo Evaluación continua Actividad no presencial</p>
Semana 16				<p>Exposición de trabajos en grupo Duración: 03:00 TG: Técnica del tipo Trabajo en Grupo Evaluación continua Actividad presencial</p> <p>Estudio autónomo Duración: 04:00 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial</p>
Semana 17				<p>Examen individual Duración: 02:00 EX: Técnica del tipo Examen Escrito Evaluación continua Actividad presencial</p> <p>Examen final Duración: 02:00 EX: Técnica del tipo Examen Escrito Evaluación sólo prueba final 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
1	Estudio autónomo	04:00	Evaluación continua	OT: Otras técnicas evaluativas	No			
2	Estudio autónomo	04:00	Evaluación continua	OT: Otras técnicas evaluativas	No			
3	Estudio autónomo	04:00	Evaluación continua	OT: Otras técnicas evaluativas	No			
4	Estudio autónomo	05:00	Evaluación continua	OT: Otras técnicas evaluativas	No			
5	Entrega trabajo individual	06:00	Evaluación continua y sólo prueba final	TI: Técnica del tipo Trabajo Individual	No	15%	3 / 10	CE1, CG9, CE8, CG5, CB10
6	Exposición del trabajo individual	02:00	Evaluación continua y sólo prueba final	TI: Técnica del tipo Trabajo Individual	Sí	10%	3 / 10	CG5, CB10
6	Estudio autónomo	03:00	Evaluación continua	OT: Otras técnicas evaluativas	No			
7	Estudio autónomo	04:00	Evaluación continua	OT: Otras técnicas evaluativas	No			
8	Estudio autónomo	04:00	Evaluación continua	OT: Otras técnicas evaluativas	No			
9	Entrega trabajo en grupo	06:00	Evaluación continua	TG: Técnica del tipo Trabajo en Grupo	No	15%	3 / 10	CE1, CB7, CG7
10	Exposición de trabajos en grupo	02:00	Evaluación continua	TG: Técnica del tipo Trabajo en Grupo	Sí	10%	3 / 10	CG7, CE1, CB7
10	Estudio autónomo	03:00	Evaluación continua	OT: Otras técnicas evaluativas	No			
11	Estudio autónomo	04:00	Evaluación continua	OT: Otras técnicas evaluativas	No			
12	Estudio autónomo	04:00	Evaluación continua	OT: Otras técnicas evaluativas	No			
13	Estudio autónomo	04:00	Evaluación continua	OT: Otras técnicas evaluativas	No			
14	Estudio autónomo	04:00	Evaluación continua	OT: Otras técnicas evaluativas	No			
15	Entrega de trabajo en grupo	06:00	Evaluación continua	TG: Técnica del tipo Trabajo en Grupo	No	15%	3 / 10	CG4, CG9, CE8
16	Exposición de trabajos en grupo	03:00	Evaluación continua	TG: Técnica del tipo Trabajo en Grupo	Sí	10%	3 / 10	CG4, CG9, CE8
16	Estudio autónomo	04:00	Evaluación continua	OT: Otras técnicas evaluativas	No			
17	Examen individual	02:00	Evaluación continua	EX: Técnica del tipo Examen Escrito	Sí	25%	3 / 10	CG7, CB10, CG4, CE8
17	Examen final	02:00	Evaluación sólo prueba final	EX: Técnica del tipo Examen Escrito	No	75%	5 / 10	CB7, CG7, CG4, CB10

Criterios de Evaluación

Modalidad de evaluación continua (ordinaria)

La calificación final de la asignatura se calculará a partir de las calificaciones de un trabajo individual,

dos trabajos en grupo y un examen individual.

Primer trabajo individual. Tras la presentación del primer tema de la asignatura, los alumnos deberán realizar un trabajo bibliográfico de los contenidos, con un tema que tendrán que exponer en clase. La calificación de este trabajo será un 15% sobre el contenido y un 10% sobre la presentación.

Trabajo en grupo. Trabajos en grupo sobre la aplicación de sistemas de integración al campo de la biomedicina. Se deberá realizar una implementación, un breve informe y una exposición en clase. La calificación de este trabajo será un 15% sobre el informe y un 10% sobre la presentación.

Segundo trabajo individual. Implementación de un sistema de recuperación de información de recursos científicos utilizando soluciones libres (e.g. Lucene). La calificación de este trabajo será un 15% sobre el informe y un 10% sobre la presentación.

El examen será el 25% restante de la nota

Modalidad de evaluación extraordinaria

En la convocatoria ordinaria, la elección entre el sistema de evaluación continua o el sistema de evaluación mediante prueba final corresponde al estudiante. Quien desee seguir el sistema de evaluación mediante sólo prueba final, deberá **OBLIGATORIAMENTE** comunicarlo **DURANTE LOS 15 PRIMEROS DÍAS** a contar desde el inicio de la actividad docente de la asignatura, mediante escrito dirigido al Prof. Coordinador de la asignatura que entregará dentro del plazo establecido.

El 25% es el trabajo individual que deben exponer en clase y el examen final 75%

Recursos Didácticos

Descripción	Tipo	Observaciones
Modern Information Retrieval	Bibliografía	Baeza-Yates, Ricardo, and Berthier Ribeiro-Neto. Modern information retrieval. New York: ACM press, 1999.
The data warehouse toolkit	Bibliografía	Kimball, Ralph, and Margy Ross. The data warehouse toolkit: the complete guide to dimensional modeling. John Wiley & Sons, 2011.
Introduction to Information Retrieval	Bibliografía	Manning CD, Raghavan P, Schütze H. Introduction to Information Retrieval. Cambridge University Press. 2008
Managing Gigabytes	Bibliografía	Witten IH, Moffat A, Bell TC. Managing Gigabytes: Compressing and Indexing Documents and Images, 2nd Edition. Morgan Kaufmann. 1999.
Natural Language Processing with Python	Bibliografía	7. Bird S, Klein E, Loper E. Natural Language Processing with Python. O'Reilly 2009. In successive academic years the individual work prepared by E1 students will be also available for other students? cohorts.