

PROGRAMA DE LA ASIGNATURA. PRESENCIAL

Curso académico: 2015-2016

Identificación y características de la asignatura				
Código	502285			Créditos ECTS
				6
Denominación (español)	RECUPERACIÓN DE LA INFORMACIÓN			
Denominación (inglés)	INFORMATION RETRIEVAL			
Titulaciones	GRADO EN INFORMACIÓN Y DOCUMENTACIÓN			
Centro	FACULTAD DE BIBLIOTECONOMÍA Y DOCUMENTACIÓN			
Semestre	1º	Carácter	OBLIGATORIO	
Módulo	TECNOLOGÍAS Y APLICACIÓN PRÁCTICA DE LA INFORMACIÓN Y DOCUMENTACIÓN			
Materia	REPRESENTACIÓN Y RECUPERACIÓN DE LA INFORMACIÓN			
Profesor/es				
Nombre	Despacho	Correo-e	Página Web	
Cristina López Pujalte	309	clopez@alcazaba.unex.es		
Área de conocimiento	Biblioteconomía y Documentación			
Departamento	Información y Comunicación			
Profesor coordinador (si hay más de uno)				
Competencias				
<i>Competencias básicas</i>				
<p>CB1 - Que los estudiantes hayan demostrado poseer y comprender conocimientos en un área de estudio que parte de la base de la educación secundaria general, y se suele encontrar a un nivel que, si bien se apoya en libros de texto avanzados, incluye también algunos aspectos que implican conocimientos procedentes de la vanguardia de su campo de estudio.</p> <p>CB2 - Que los estudiantes sepan aplicar sus conocimientos a su trabajo o vocación de una forma profesional y posean las competencias que suelen demostrarse por medio de la elaboración y defensa de argumentos y la resolución de problemas dentro de su área de estudio.</p> <p>CB3 - Que los estudiantes tengan la capacidad de reunir e interpretar datos relevantes (normalmente dentro de su área de estudio) para emitir juicios que incluyan una reflexión sobre temas relevantes de índole social, científica o ética.</p> <p>CB4 - Que los estudiantes puedan transmitir información, ideas, problemas y soluciones a un público tanto especializado como no especializado.</p> <p>CB5 - Que los estudiantes hayan desarrollado aquellas habilidades de aprendizaje necesarias para emprender estudios posteriores con un alto grado de autonomía.</p>				
<i>Competencias generales</i>				
<p>CG2 - Conocimiento de los principios teóricos y metodológicos para la reunión, selección, organización, representación, preservación, recuperación, acceso, difusión e intercambio de la información.</p> <p>CG3 - Conocimiento de las tecnologías de la información que se emplean en las unidades y servicios de información.</p>				

CG4 - Habilidades en el manejo de las tecnologías como medio indispensable en los procesos de tratamiento y transferencia de la información.

Competencias específicas

CE1 - Conocimiento de los principios teóricos y metodológicos para la planificación, organización y evaluación de sistemas, unidades y servicios de información.

CE2 - Conocimiento de los principios teóricos y metodológicos para el estudio, el análisis, la evaluación y la mejora de los procesos de producción, transferencia y uso de la información y de la actividad científica.

CE4 - Comprender y aplicar los principios y las técnicas para la planificación, organización y evaluación de sistemas, unidades y de servicios de información.

CE5 - Comprender y aplicar los principios y las técnicas para la reunión, selección, organización, representación, preservación, recuperación, acceso, difusión e intercambio de la información.

CE9 - Capacidad de aplicar y valorar las técnicas de planificación, organización y evaluación de sistemas, unidades y servicios de información.

CE10 - Capacidad de usar y aplicar las técnicas, las normativas y otros instrumentos utilizados en la reunión, selección, organización, representación, preservación, recuperación, acceso, difusión e intercambio de la información.

Competencias transversales

CT1 - Capacidad de análisis y de síntesis aplicadas a la gestión y organización de la información.

CT3 - Habilidades en el uso de Internet y software genérico (ofimática).

CT5 - Capacidad de organización y planificación del trabajo propio.

CT6 - Capacidad de trabajar en equipo y de integración en equipos multidisciplinares.

CT8 - Razonamiento crítico en el análisis y la valoración de alternativas.

CT10 - Capacidad para el aprendizaje autónomo.

Temas y contenidos

Breve descripción del contenido

Descriptores del Título:

Historia, Evaluación y Leyes aplicadas a la Recuperación y Representación Documental. Estrategias Booleanas de Búsqueda.

Temario de la asignatura

PROGRAMA TEÓRICO

Denominación del tema 1: *Introducción. Conceptos básicos.*

Contenidos del tema 1:

- 1.1. Introducción. Acotación del ámbito abarcado por la asignatura
- 1.2. Concepto de Recuperación de Información (RI). Su papel dentro del Proceso Documental
- 1.3. Recuperación de información e indización
- 1.4. Sistemas de Recuperación de Información
- 1.5. Breve historia de la RI: *Astia, Cranfield, SMART...*

Denominación del tema 2: *Evaluación de la Recuperación de Información*

Contenidos del tema 2:

- 2.1. Criterios de evaluación. Eficacia. Concepto de relevancia
- 2.2. Medidas de evaluación
- 2.3. Evaluación según la técnica de recuperación empleada
- 2.4. Tests de diferencias de medias

Denominación del tema 3: *Modelos de RI y Representación documental*
 Contenidos del tema 3:

- 3.1. Sistemas basados en índices invertidos
- 3.2. Leyes de Zipf
- 3.3. Valor de información
- 3.4. Modelos de RI
- 3.5. Representación documental. Modelo del Espacio Vectorial
- 3.6. Técnicas de recuperación. Medidas de similitud
- 3.7. Selección de las componentes y ponderación

Denominación del tema 4: *Mejoras de la RI*
 Contenidos del tema 4:

- 4.1. Concepto de Unitérminos. Limitaciones
- 4.2. Mejoras a los unitérminos
- 4.3. La importancia de las consultas. Estrategias de búsqueda. Técnicas de modificación automática de consultas: *Relevance Feedback*
- 4.4. Tendencias de investigación relacionadas

PROGRAMA PRÁCTICO

Todos los temas (salvo el introductorio) tienen una batería de ejercicios prácticos propuestos. Estos ejercicios se estructuran de la siguiente forma:

1. Resolución de problemas tipo sobre evaluación de SRIs
2. Resolución de problemas tipo sobre implementación del modelo del espacio vectorial
3. Resolución de problemas tipo sobre técnicas de representación documental y su posterior recuperación

Asimismo, se recomendarán lecturas complementarias.

Actividades formativas

Horas de trabajo del alumno por tema		Presencial		Actividad de seguimiento	No presencial
Tema	Total	GG	SL	TP	EP
1	14	5			9
2	40	15	5		20
3	40	15	5		20
4	34,5	12,5	2		20
Preparación del examen	21,5			1,5	20
Evaluación del conjunto	150	47,5	12	1,5	89

GG: Grupo Grande (100 estudiantes).

SL: Seminario/Laboratorio (prácticas clínicas hospitalarias = 7 estudiantes; prácticas laboratorio o campo = 15; prácticas sala ordenador o laboratorio de idiomas = 30, clases problemas o seminarios o casos prácticos = 40).

TP: Tutorías Programadas (seguimiento docente, tipo tutorías ECTS).

EP: Estudio personal, trabajos individuales o en grupo, y lectura de bibliografía.

Metodologías docentes

- Explicación en clase de los temas programados
- Utilización de material docente en diferentes tipos y formatos.
- Discusión de los contenidos.
- Aplicación práctica de los conocimientos teóricos a través de los laboratorios, talleres, etc.
- Análisis y resolución de problemas prácticos propuestos.
- Actividades de seguimiento del aprendizaje.
- Trabajo en grupo.

Resultados de aprendizaje

- Encuadrar el concepto de recuperación de información en del proceso documental.
- Comprender la importancia de la representación documental (indización) y su relación con la recuperación de información.
- Conocer la historia de los sistemas de recuperación de información y cómo han evolucionado a lo largo del tiempo.
- Comprender la importancia de una correcta evaluación de la recuperación de la información, y aplicar las técnicas necesarias para llevarla a cabo.
- Conocer y emplear los distintos modelos de recuperación de información.
- Utilizar las metodologías de automatización de la representación documental y determinar su repercusión en la posterior recuperación a partir de sus ventajas e inconvenientes.
- Aplicar las técnicas básicas de recuperación y de representación documental, tanto estadísticas como lingüísticas.
- Comprender la importancia de la indización de las consultas en la mejora de los resultados de la recuperación, así como de las estrategias de búsquedas.

Sistemas de evaluación

Medidas de evaluación

CRITERIOS GENERALES DE EVALUACIÓN DEL MÓDULO III ESTABLECIDOS EN EL TÍTULO DE GRADO EN DOCUMENTACIÓN E INFORMACIÓN DE LA UEX SEGÚN EL DOCUMENTO OFICIAL APROBADO POR LA ANECA:

La evaluación consistirá en la superación de uno o varios exámenes de los conocimientos teóricos adquiridos, dichos exámenes podrán consistir en una prueba oral o de desarrollo escrito, de preguntas cortas o tipo test. Para la valoración final de la prueba se determinará previamente el valor de cada pregunta.

Para el cálculo final de la nota podrá atenderse, también, a la participación y asistencia del alumnado a las clases magistrales y prácticas realizadas. En todo caso el examen deberá aprobarse para superar la asignatura.

En cuanto al sistema de calificación en las prácticas, se atenderá a la participación y el resultado de la resolución de problemas planteados a los alumnos en las actividades prácticas de la asignatura y, en su caso, podrá realizarse un examen práctico.

En cada caso, el profesor encargado de la asignatura hará públicos los criterios específicos de evaluación en los plazos que se determinen por la UEx.

La calificación es la suma de la evaluación continua y la evaluación final atendiendo a los siguientes criterios:

Evaluación continua es la media ponderada de la calificación obtenida en los trabajos y tareas estipuladas en cada asignatura siendo el total no superior al 50% de la evaluación final de las asignaturas de corte teórico y al 70% de la evaluación de las asignaturas de corte práctico.

Examen final consiste en la realización de una o varias pruebas orales o escritas cuyo valor no debe ser inferior al 50% de la calificación de las asignaturas de corte teórico y al 30% de las asignaturas de corte práctico.

Se aplicará el sistema de calificaciones vigente en cada momento (RD 1125/2003, artículo 5º). Los resultados obtenidos por el alumno en cada una de las materias del plan de estudios se calificarán en función de la siguiente escala numérica de 0 a 10, con expresión de un decimal, a la que podrá añadirse su correspondiente calificación cualitativa: 0 - 4,9: Suspenso (SS), 5,0 - 6,9: Aprobado (AP), 7,0 - 8,9: Notable (NT), 9,0 - 10: Sobresaliente (SB). La mención de Matrícula de Honor (MH) podrá ser otorgada a alumnos que hayan obtenido una calificación igual o superior a 9.0. Su número no podrá exceder del 5 % de los alumnos matriculados, salvo que el número de alumnos matriculados sea inferior a 20, en cuyo caso se podrá conceder una sola Matrícula de Honor.

Criterios de evaluación

ACTIVIDADES/INSTRUMENTOS DE EVALUACIÓN

Dentro de las competencias y objetivos de la asignatura están tanto la adquisición de unos conocimientos como el desarrollo de una serie de habilidades, por lo tanto se evaluarán para la misma:

- Conocimientos teóricos y conocimientos prácticos
- Prácticas realizadas (asistencia, participación y realización)
- Trabajo en equipo obligatorio

Esta evaluación se realizará a partir de la evaluación continua y un examen final.

CRITERIOS DE EVALUACIÓN ESPECÍFICOS PARA LA ASIGNATURA “RECUPERACIÓN DE INFORMACIÓN”:

La calificación final será la suma de la evaluación continua y del examen atendiendo a los siguientes criterios:

- Evaluación continua (30%): realización de prácticas tuteladas relacionadas con la materia desarrolladas durante la impartición de la asignatura que se expondrán en clase. La evaluación continua no es recuperable en ningún caso, y deberá realizarse en los plazos establecidos por el profesor durante el semestre correspondiente (1º semestre). La calificación obtenida en la evaluación continua será vigente para las convocatorias del curso académico correspondiente.
- Examen (70%): realización de una prueba escrita sobre conocimientos teóricos y prácticos de la materia. En todo caso el examen deberá aprobarse para poder superar la asignatura, y solo en este caso se sumará a la nota obtenida en el examen la nota de la evaluación continua.

La calificación final mínima para superar la asignatura será de Aprobado (5).

Bibliografía (básica y complementaria)

Bibliografía básica:

Baeza-Yates, R.; Ribeiro-Neto, B. Modern information retrieval. Addison-Wesley. Essex, UK, 1999.

Büttcher, S.; Clarke, C.L.A.; Cormack, G.V. Information Retrieval: Implementing and Evaluating Search Engines. MIT Press. Cambridge, 2010.

Cacheda Sijo, F.; Manuel Fernández Luna, J.; Huete Guadix, J.F. (coord.). Recuperación de información: un enfoque práctico y multidisciplinar. RA-MA, cop. Madrid, 2011.

Chu, H. Information Representation and Retrieval in the Digital Age. Information Today. Medford, New Jersey, 2007.

Ellis, D. Progress and problems in information retrieval. The Library Association. London, 1996.

Frakes, W. B. & Baeza-Yates, R. Information retrieval: data structures & algorithms. Upper Saddle River. New Jersey, 1992.

Harman, D.K. Relevance feedback and other query modification techniques. In: Frakes, W.B.; Baeza-Yates, R. (Eds.) Information retrieval: data structures and algorithms. Prentice Hall. NJ, 1992. 241-263.

Manning, C.D.; Raghavan, P.; Schütze, H. Introduction to Information Retrieval. Cambridge University Press, 2008.

Moya Anegón, F. de. Los Sistemas Integrados de Gestión Bibliotecarias. ANABAD. Madrid, 1995.

Salton, G. Automatic text procesing: the transformation, analysis, and retrieval of information by computer. Addison-Wesley. Reading, Massachusetts, 1989.

Salton, G. & McGill, M.J. Introduction to modern information retrieval. McGraw-Hill. New York, 1983.

Bibliografía complementaria:

Keyser, P. de. Indexing From Thesauri to the Semantic Web. Chandos Publishing. 2007.

Mayer-Schönberger, V. & Cukier, K. (2013). Big data. La revolución de los datos masivos. Madrid: Turner Publicaciones S.L.

Pestana Caldes, A.I.; López-Pujalte, C. Algoritmos Genéticos aplicados a la Recuperación de Información. En Borges, M.M. & Sanz Casado, E. *A Ciência da Informação Criadora de Conhecimento*. Vol. II, 405-408. Coimbra

University Press. Coimbra, 2009.

Strzalkowski, T. (Ed.). Natural Language Information Retrieval. Kluwer Academic Publishers. Dordrecht, Netherlands, 1999.

Otros recursos y materiales docentes complementarios

- Se utilizarán colecciones de test específicas de evaluación de recuperación de información (*Cranfield, CISI, Medline, etc.*) accesibles on-line:
<http://www.filewatcher.com/b/ftp/ftp.cs.cornell.edu/pub/smart-0.html>
- Se utilizarán documentos hechos con una hoja de cálculo Excel (.excel) para la aplicación de test de diferencias de medias.

Horario de tutorías

Tutorías Programadas: Las tutorías programadas serán en el mes de diciembre y se comunicarán al alumnado con antelación suficiente.

Tutorías de libre acceso: El horario de tutorías de libre acceso se establecerá para cada semestre dentro de los plazos previstos por la Universidad y podrá ser consultado en la Web de la Facultad, y en el tablón de anuncios del profesor.

Recomendaciones

Se recomienda al alumno estudiar con suficiente tiempo la asignatura, dado que los contenidos de la misma tienen cierta complejidad y son totalmente nuevos para él. Además, es importante ir realizando durante el semestre los ejercicios prácticos para una correcta comprensión de los contenidos teóricos.

Se estiman aproximadamente entre 70/90 horas de estudio necesarias para asimilar adecuadamente la asignatura (incluida la preparación y elaboración del trabajo obligatorio).

Es recomendable tener aprobadas previamente las siguientes asignaturas del grado:

- Introducción a la Tecnología de la Información y la Comunicación
- Introducción a la Estadística Documental
- Edición Digital