

PROGRAMA DE LA ASIGNATURA
Curso académico 2008/2009

Identificación y características de la asignatura			
Denominación y código	TÉCNICAS AVANZADAS DE RECUPERACIÓN DOCUMENTAL	código	105288
Créditos (T+P)	2,5T+2P		
Titulación	LICENCIATURA EN DOCUMENTACIÓN		
Centro	FACULTAD DE BIBLIOTECONOMÍA Y DOCUMENTACIÓN		
Curso	2	Temporalidad	1er Cuatrimestre
Carácter	TRONCAL		
Descriptor (BOE)	Métodos y técnicas aplicadas al estudio y a las actividades propias de la recuperación del conocimiento humano		
Despacho y correo-e	3.01 correo-e: atmar@alcazaba.unex.es		
Profesor/es	Nombre	Ana Teresa García Martínez	
	Despacho	3.01	
	Correo-e	atmar@alcazaba.unex.es	
	Página web		
Área de conocimiento	Biblioteconomía y Documentación		
Departamento	Información y Documentación		
Profesor coordinador (si hay más de uno)			

Objetivos y/o competencias

- Conocer el contexto paradigmático en el que se originan y desarrollan las técnicas de recuperación de información
- Delimitar el entorno de funcionamiento de las técnicas de recuperación, radicado en los IRS
- Comprender el concepto IR. Diferenciación IR/DR.
- Establecer la taxonomía de las técnicas de recuperación resaltando las características esenciales de cada una de ellas
- Conocer la estructura subyacente del modelo booleano. Establecer las cuestiones clave que fundamentan la crítica del sistema booleano.
- Comprender las relaciones entre los elementos del sistema de recuperación booleano y las operaciones que sustenta
- Conocer los fundamentos teóricos de las técnicas de recuperación basadas en características
- Desarrollar la fundamentación teórica que sustenta el modelo vectorial de recuperación de información
- Aplicar las distintas funciones de similaridad documental
- Desarrollar supuestos prácticos que simulen la recuperación de información en un sistema de recuperación automatizado
- Aplicar el esquema de formalización de modelos de recuperación de información.

Temas y contenidos
(especificar prácticas, teoría y seminarios)

TEORÍA

Tema 0

Paradigma IR. Enfoques arquetípico y cognitivo de la Recuperación de Información

Tema 1

Contexto de las técnicas avanzadas IR. Entorno IRS

- 1.1. Origen de los IRS
- 1.2. Funcionamiento de los IRS
- 1.3. Operaciones de los IRS
- 1.4. Clasificación Facetada de los IRS
 - 1.4.1. Modelo conceptual
 - 1.4.2. Estructura de ficheros
 - 1.4.3. Operaciones sobre las consultas
 - 1.4.4. Operaciones sobre los términos
 - 1.4.5. Operaciones sobre los documentos
 - 1.4.6. Hardware
- 1.5. Visión general del paradigma IRS

Tema 2

Taxonomía de las técnicas avanzadas IR

- 2.1. Information Retrieval
- 2.2. Data Retrieval versus Information Retrieval
 - 2.2.1. Data Retrieval
 - 2.2.2. Information Retrieval
 - 2.2.3. Parámetros para la distinción
- 2.3. Taxonomía de las Técnicas de Recuperación Documental
 - 2.3.1. Modelos clásicos
 - 2.3.1.1. Booleano
 - 2.3.1.2. Espacio Vectorial
 - 2.3.2.3. Probabilístico
 - 2.3.2. Modelos alternativos
 - 2.3.2.1. Modelos de Teoría de Conjuntos
 - 2.3.2.2. Modelos Algebraicos
 - 2.3.2.3. Modelos Probabilísticos

TEMA 3

Modelos de recuperación documental tradicionales
Críticas al modelo booleano

- 3.1. Modelo Booleano
 - Estructura de ficheros
 - Operadores booleanos
 - Operadores de proximidad
- 3.2. Críticas al modelo booleano
 - Salton
 - Belkin, Bookstein y Cooper
 - Baeza-Yates y Ribeiro-Neto
- 3.3. Los sistemas booleanos en las bibliotecas

TEMA 4

Modelos de recuperación avanzada de información
Formalización de modelos

- 4.1. Técnicas Basadas en Características
- 4.2. Visión lógica de los Documentos
 - Fundamentos teóricos de la indización automática: Lhun y Zipf
- 4.3. Factor de Significado
- 4.4. Medidas de Frecuencia Absoluta
- 4.5. Medidas de Frecuencia Relativa
 - Inverse Document Frequency (IDF)
 - Teoría Matemática de la Comunicación – Shannon
 - Ratio Ruido-Señal
 - Valor de discriminación
- 4.6. Formalización de modelos

PRÁCTICAS

1. Operaciones sobre la consulta: Parsing, utilizando un Analizador Morfosintáctico
2. Operaciones sobre los términos: stemming, mediante algorithmo
3. Operaciones sobre los términos: Stop List
4. Cálculo de Inverse Document Frequency
5. Ejercicios sobre la reducción de incertidumbre de los documentos
6. Ejercicios de cálculo de relevancia documental a una consulta
7. Cálculo de la similaridad media de una colección documental
8. Cálculo del Valor de Discriminación mediante procedimiento real
9. Cálculo del Valor de Discriminación mediante procedimiento aproximado

Criterios de evaluación

La evaluación de la asignatura se llevará a cabo mediante la realización de una prueba escrita, en la que el alumno deberá demostrar los conocimientos teóricos adquiridos sobre la materia, así mismo deberá evidenciar su aplicación práctica mediante la resolución de los presupuestos prácticos o ejercicios que se planteen.

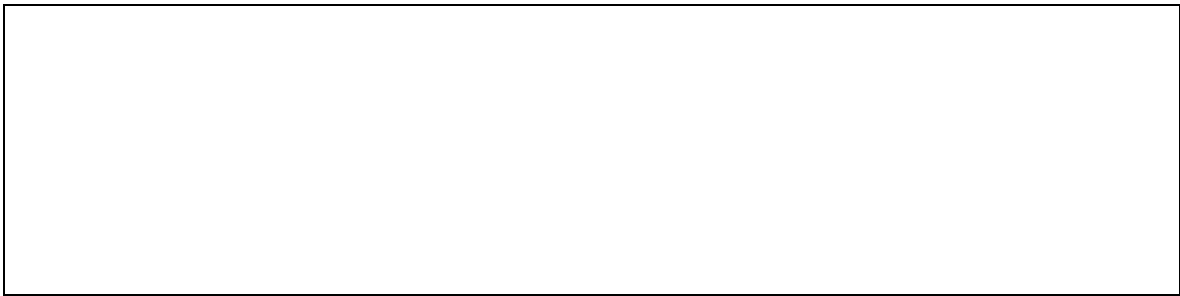
Otro aspecto a contemplar en la evaluación será la formación complementaria a través de las lecturas obligatorias y recomendadas, cuya cuantificación podrá completar la calificación final.

El examen consistirá en responder a las cuestiones que se planteen en las preguntas teórico – prácticas, que oscilarán entre 8 y 15 preguntas, siendo necesario obtener una puntuación mínima de 5 puntos para superar el examen.

Valoración de las respuestas:

Cada respuesta acertada tendrá un valor diferente dependiendo del tipo de cuestión, resultando en 10 puntos la suma total.

Para aprobar la asignatura será necesario haber entregado todas las prácticas propuestas en el transcurso del desarrollo de la asignatura. Por tanto aunque se haya aprobado el examen estaría suspensa la asignatura si no se han entregado las prácticas.



Bibliografía

BIBLIOGRAFIA BÁSICA

BAEZA-YATES, Ricardo; RIBEIRO-NETO, Berthier. Modern information retrieval. England: Addison Wesley, 1999 {002BAEmod}

BELKIN, N.J. y CROFT, W.B. "Retrieval techniques". ARIST, 22, 1987

ELLIS, David. Progress and problems in information retrieval. 2ª ed. London: Library Association Publishing, 1996

FRAKES, W. B.; BAEZA-YATES. Information retrieval: data structures and algorithms. Englewood Cliffs, 1992

Information retrieval experiment / edited by Karen Soparck Jones. London: Butterworths, 1981

KANTOR, P.B. Information retrieval techniques. ARIST, 29, 1994, 53-90

KORFHAGE, R.R. Information storage and retrieval. New York: Wiley Computer Publishing, 1997

LANCASTER, F.W. Information retrieval systems: characteristics, testing and evaluation. 2ª ed. New York: Wiley – Interscience, 1979

MARCHIONINI, Gary. Information seeking in electronic environments. New York: Cambridge University Press, 1997

MEADOW, Charles T.; BOYCE, Bert R.; KRAFT, Donald H. Text information retrieval systems. 2ª ed. San Diego: Academic Press, 2000

MOYA ANEGÓN, Félix de. Los sistemas integrados de gestión bibliotecaria. Madrid: ANABAD, 1995

MOYA ANEGÓN, Félix de. Técnicas avanzadas de recuperación documental. En: Manual de Ciencias de la Documentación. Coord. López Yepes. Madrid: Pirámide, 2002

Natural language information retrieval. . Edited by Tomek Strzalkowski. Dordrecht: Kluwer Academic Publisher, 1999

Capítulos. 1, 4, 13, 14

PAO, Miranda Lee. Concepts of information retrieval. Englewood, Colorado: ; Libraries unlimited, 1989

PEÑA, Rosalía; BAEZA-YATES, Ricardo; RODRIGUEZ, José V. Gestión digital de la información: de bits a bibliotecas digitales y la web. Madrid: RAMA, 2002

Procesamiento de la información científica. Coordinadores Wilfrid Lancaster y María Pinto. Madrid: Arco/Libros, 2001

Readings in information retrieval. Edited by Karen Sparck Jones y Peter Millett. San Francisco, California: Morgan Kaufmann Publishers, 1997

Van RIJSBERGEN, C.J. Information retrieval. 2nd ed. London: Butterworths, 1979

SALTON,, Gerard. Introduction to modern information retrieval. New York. McGraw –Hill, 1983

SALTON, Gerard. Automatic text processing: the transformation, analysis, and retrieval of information by computer. EEUU; Canada: Addison-Wesley Publishing Company, 1989

BIBLIOGRAFÍA COMPLEMENTARIA

CODINA, Luis. Modelo conceptual de un sistema de información documental. Revista Española de Documentación Científica, 1994, v.17, n. 4, p. 440-449

CODINA. Luis. Recuperación de información e hipertextos: sus bases lógicas y su aplicación a la documentación periodística. En: Manual de documentación periodística. Ed. M^a Eulalia Fuentes i Pujol. Madrid: Síntesis, 1995, p. 213-230

CODINA, Luis. Teoría de la recuperación de información: modelos fundamentales y aplicaciones a la gestión documental. Information World en Español, 1995, n. 38, p. 18-22

FERNÁNDEZ MOLINA, J. C.; MOYA ANEGÓN, FÉLIX DE. Los catálogos de acceso público en línea. el futuro de la recuperación bibliográfica. Málaga: Asociación Andaluza de Bibliotecarios, 1998

LARSON, Ray R. Evaluation of advanced retrieval techniques in an

experimental online catalog. Journal of the American Society for Information Science, 43(1) 1992

LARSON, Ray R. The decline of subject searching: long-term trends and patterns of index use in an online catalog. Journal of the American Society for Information Science, 42(3) 1991

MARTINEZ MONTALVO, E.; MARTINEZ COMECHE, J. A.. Adecuación de modelos matemáticos a la ciencia de la documentación. Documentación de las Ciencias de la Información, n.16, 1993

REVISTAS

- * Annual Review for Information Science and Technology
- * Journal of American Society for Information Science
- * Journal of Documentation
- * Scientometrics

Horario de tutorías

Periodo lectivo

Lunes y miércoles, de 11:30 a 13:45

Jueves, de 18 a 19:30

Periodo no lectivo

Lunes y miércoles, de 11:30 a 13:45