

Curso Académico: ( 2016 / 2017 )

Fecha de revisión: 24/03/2015 13:13:32

Departamento asignado a la asignatura:

Coordinador/a:

Tipo: Optativa Créditos ECTS : 6.0

Curso : Cuatrimestre :

**REQUISITOS (ASIGNATURAS O MATERIAS CUYO CONOCIMIENTO SE PRESUPONE)**

Como complemento y con carácter opcional, es aconsejable haber cursado la asignatura denominada Ingeniería de la Información, ya que dicha materia complementa algunos conceptos tratados a lo largo de la asignatura

**OBJETIVOS**

1. Adquisición de los conocimientos relevantes en el contexto de las modernas técnicas de Recuperación de Información, centrándose especialmente en el entorno Web y las Intranets.
2. Conocimiento de las metodologías utilizadas en los sistemas actuales (técnicas de procesamiento del lenguaje natural, minería de textos y aprendizaje automático, extracción de información, utilización de redes conceptuales,... ).
3. Utilización de aplicaciones existentes para la recuperación de información

**DESCRIPCIÓN DE CONTENIDOS: PROGRAMA**

Recuperación y Acceso a la Información

- Algoritmos de Recuperación
- Asignación de Pesos
- Modelos de Recuperación

Recuperación en Internet

- Tipos de buscadores
- Algoritmos de Posicionamiento
- El futuro de la Recuperación de Información: Ontologías y Web Semántica

Medida de la calidad en la Recuperación de documentos

- Tipos de medidas
- Problemática
- Colecciones de Prueba

Corpora

- Corpus Documentales
- Corpus como Recursos Lingüísticos

Indización Automatizada

- Tipos de Indización
- Metodología clásica: Tokens y Filtradores

Análisis basado en el PLN: Analizadores, Recursos y Problemática

Lenguajes para recuperación de la Información

Optimización de sistemas de recuperación

- Reconocimiento de entidades y extracción de información
- Asignación de pesos mediante minería de textos
- Utilización de sistemas de representación del conocimiento

**ACTIVIDADES FORMATIVAS, METODOLOGÍA A UTILIZAR Y RÉGIMEN DE TUTORÍAS**

Práctica I. Creación de Páginas Web para mejora de su posicionamiento en la recuperación

Práctica II. Detección de Plagios

Práctica III. Creación de un motor de recuperación desarrollado contra una colección de prueba:

- a) Motor de recuperación implementado mediante técnicas lingüísticas y documentalistas
- b) Motor de recuperación con extracción de entidades y minería de textos
- c) Motor de recuperación con incorporación de redes léxicas y conceptuales

## SISTEMA DE EVALUACIÓN

Valoración de los conocimientos teórico prácticos sobre el temario de la asignatura.

Calidad y características de las prácticas realizadas. Las prácticas serán evaluadas mediante el cuaderno de prácticas y su exposición oral en su caso (60%). Además se realizará un examen final para evaluar la adquisición de los conocimientos teóricos compartidos (40%).

## BIBLIOGRAFÍA BÁSICA

- Antoniou, Grigoris & Van Harmelen, Frank A semantic Web Primer, Massachussets: MIT, 2004.
- Baeza-Yates, Ricardo and Berthier Ribeiro-Neto Modern information retrieval, New York : ACM Press ; Harlow ; Madrid [etc.]: Addison-Wesley , 1999.
- Manning, Christopher D; Raghavan, Prabhakar; Schütze, Hinrich Introduction to Information Retrieval, Cambridge University Press, 2008 ISBN 978-0-521-86571-5.
- Moens Marie-Francine Information Extraction: algorithms and prospects in a retrieval context. Capítulos 1, 2 y 4, Springer, 2006.
- [Distintos Autores] Posicionamiento Web, Revista El Profesional de la Información, Vol. 14, números 1 y 2. 2005.

## BIBLIOGRAFÍA COMPLEMENTARIA

- C. J. Van Rijsbergen Information Retrieval, <http://www.dcs.gla.ac.uk/Keith/Preface.html>.
- Codina, Lluís Teoría de recuperación de información: modelos fundamentales y aplicaciones a la gestión documental, Information World en Español, nº 38, octubre 1995.
- Dale R Handbook of Natural Language Processing, New York, 2000.
- Karen Sparck Jones Readings in information retrieval, Peter Willett. San Francisco : Morgan Kaufmann , cop. 1997.
- Méndez Rodríguez, Eva M<sup>a</sup> Metadatos y recuperación de información : estándares, problemas y aplicabilidad en bibliotecas digitales, Gijón : Trea.
- Nadeau D. and Sekine S. A survey of named entity recognition and classification, Linguisticae Investigationes vol. 30 nº1. 2007.
- Park J and Hunting S XML Topic Maps. Creating and Using Topics Maps for the Web, Boston, 2003.