

# SISTEMAS DE RECUPERACIÓN DE LA INFORMACIÓN E INTERNET

---

Prof. Gonzalo Marco Cuenca  
Prof. Rodrigo Sánchez Jiménez  
Curso 2015-2016  
Créditos: 6

## **Objetivos docentes**

- Conocimiento de las características y especificidades de la Recuperación de Información en Internet.
- Introducción a la administración de Sistemas de Recuperación de Información sobre colecciones Web.
- Adquisición de conocimientos y competencias para mejorar la difusión de información de información en Internet.

## **Metodología**

Las clases tendrán un carácter teórico práctico, de forma que en la mayor parte de las ocasiones se combinarán ambos aspectos dentro de una lección. Sin embargo también se impartirán lecciones magistrales y algunas de las clases tendrán un carácter completamente práctico y orientado fundamentalmente al trabajo en grupo.

## **Evaluación**

Se tendrá en cuenta la asistencia, participación, pruebas y prácticas realizadas durante el curso y, además, se realizará: una prueba escrita parcial al finalizar los módulos I y II (50%) y una prueba escrita parcial al finalizar los módulos III y IV (50%).

## **PRIMERA PARTE**

### **Módulo I. Conceptos fundamentales**

**Tema 1.** Introducción a la Recuperación de la Información (RI) y a los Sistemas de Recuperación de la Información. Modelos tradicionales de RI.

**Tema 2.** Introducción a la Estructura y Funcionamiento de la Web. Motores de búsqueda.

**Tema 3.** RI en Internet VS RI Tradicional. Calidad de la información en Internet.

### **Módulo II. Administración de Servicios de Búsqueda Propios**

**Tema 4.** Optimización de la Recuperación. Filtros. Búsqueda por campos. Operadores avanzados.

**Tema 5.** Utilización de metadatos. Estándares de metadatos.

**Tema 6.** Recuperación de la información en la empresa. Tipos de datos. Formatos. Portales de datos.

## **SEGUNDA PARTE**

### **Módulo III. SEO - Detección de problemas y planteamiento de una estrategia**

**Tema 7:** Aspectos básicos del SEO. RI con adversario - SPAM y buscadores - Confianza, Autoridad y Popularidad - Limitación de la aplicación de metadatos - Componentes del Ranking de un buscador.

**Tema 8:** Problemas de accesibilidad y relevancia. Defectos en el código - Defectos estructurales - Malas prácticas de nombrado - Utilización de Site Maps - Buenas prácticas de redacción y contenidos - Mitos y verdades sobre la relevancia de un sitio Web.

**Tema 9:** Problemas de reputación, visibilidad y popularidad. Enlaces - visitas - PageRank y TrustRank.

**Tema 10:** Planteamiento de una estrategia. Auditoría de Sitios Web - Tipos de campañas - Establecimiento de objetivos - Establecimiento del tipo de trabajo a realizar - Rutinas de mantenimiento.

## Módulo IV: SEO - Ejecución de una campaña

**Tema 11:** Keyword research. Long Tail/Broad Head - Recopilación de datos - Selección de palabras clave - Mitos y verdades sobre la densidad de palabras clave - Creación de contenido SEF.

**Tema 12:** Análisis de la competencia. Selección de criterios de búsqueda - Delimitación de la competencia

**Tema 13:** Análisis de enlaces. Aplicación de la Webmetría al análisis competitivo de enlaces - Herramientas para el análisis - Técnicas de análisis.

**Tema 14:** Construcción de enlaces. Enlaces internos - Pago por inclusión - Redes Sociales - Malas prácticas.

**Tema 15:** Medición del éxito. Tráfico vs Conversiones - Software para el análisis.

## Bibliografía

---

Baeza-Yates, R. & de Araújo Neto Ribeiro, B. (2010), *Modern information retrieval*. Second edition., Pearson Higher Education, New Jersey.

Chowdhury, G. (2010), *Introduction to modern information retrieval*, Facet publishing.

Croft, W Bruce; Metzler, Donald y Strohman Trevor (2010). *Search Engines. Information Retrieval in practice*. Pearson Higher Education, New Jersey

Manning, Christopher; Schütze, Hinrich y Raghavan, Prabhakar (2008) *Introduction to Information Retrieval*. Cambridge, Cambridge University Press.  
<http://www-csli.stanford.edu/~hinrich/information-retrieval-book.html>

Baeza-Yates, Ricardo, y Berthier Ribeiro-Neto (1999). *Modern Information Retrieval*. Addison Wesley & ACM Press, New York.

Korfhage, R. R. (1997), *Information storage and retrieval*, John Wiley & Sons, Inc., New York, NY, USA.

Salvador Oliván, José Antonio (2008). *Recuperación de la información*. Alfagrama, Buenos Aires.

van Rijsbergen, Keith (1979). *Information Retrieval*. Rutterworths, London.  
<http://www.dcs.gla.ac.uk/Keith/Chapter.2/Ch.2.html>

## **RI en la empresa**

Bailey, P; Hawking D. y Matson B. (2006) Secure search in enterprise webs: tradeoffs in efficient implementation for document level security. *Proceedings of the 15th ACM international conference on Information and knowledge management*.

Hawking, D. (2004) "Challenges in Enterprise Search". *ACM International Conference Proceeding Series*; Vol. 52.

## **SEO**

(Accesibles a texto completo desde la UCM)

Bailyn, E. y Bailyn, B. (2011), *Outsmarting Google: SEO Secrets to Winning New Business*, Que Publishing.

Levene, M. (2011), *An introduction to search engines and web navigation*, Wiley.

Safko, L. (2010), *The social media bible: tactics, tools, and strategies for business success*, John Wiley & Sons.