

# Recuperación de Información en Internet

P.O.P. *Língua e usos profesionais*

Miguel A. Alonso   Carlos Gómez   Jesús Vilares

Departamento de Computación  
Facultad de Informática  
Universidade da Coruña

- 1 Descriptores
- 2 Objetivos
- 3 Metodología
- 4 Evaluación
- 5 Bibliografía y software
- 6 Profesorado

# Descriptores

- Utilización de las tecnologías básicas de Internet.
- Conceptos básicos de redes de comunicación.
- Conceptos básicos de Internet y navegación en web.
- Buscadores web y recuperación de información.
- La web como base de datos documental.

# Objetivos de naturaleza teórica

- Conocer la estructura interna de una red de comunicaciones
  - considerando Internet como un caso práctico de especial relevancia.
- Estudiar en qué consiste la recuperación de información
  - cuál es su problemática específica
  - qué tipo de técnicas y modelos se utilizan para resolverla
  - cómo se evalúa el rendimiento de un sistema diseñado para intentar resolver esta tarea
- Considerar posibles técnicas de integración de las características
  - morfológicas
  - sintácticas
  - semánticas

de los idiomas, como método de mejora de los buscadores actuales.

# Objetivos de naturaleza aplicada

- Realizar prácticas de recuperación de información
  - en el dominio de toda la web
  - utilizando alguno de los muchos buscadores disponibles hoy en día, como puede ser Google
- Realizar prácticas de sintonización y evaluación
  - de un prototipo de buscador propio
  - en el dominio de una pequeña base de datos documental local

# Objetivos complementarios

- Sintetizar todos los conceptos estudiados en ideas concretas que permitan al alumno abordar tareas de perfeccionamiento de un sistema de recuperación de información particular, adaptándolo a la problemática específica de un determinado organismo o empresa.
- Aislarse momentáneamente de los métodos específicos de recuperación de información, para considerar otro tipo de procesos de nivel superior, como pueden ser las tareas de extracción de información.

# Temario

- 1 Internet
- 2 La web
- 3 Principios de Recuperación de Información
- 4 Búsqueda de información en la web
- 5 Búsqueda avanzada en la web
- 6 Integración del conocimiento lingüístico en las tareas de recuperación de información
- 7 Más allá de la búsqueda textual

# Metodología

- Clases teóricas
  - 1,5 créditos ECTS
  - 9 horas de clase
- Clases prácticas
  - 1,5 créditos ECTS
  - 9 horas de clase
- Tutorías

# Evaluación

- 40% examen escrito al final del curso
- 40% Prácticas, de carácter obligatorio
- 20% actividades individuales, que tendrán carácter obligatorio, que serán evaluadas de forma continua

# Bibliografía

- Baeza-Yates, R.; Ribeiro-Neto, B.; (1999).  
Modern information retrieval.  
Addison-Wesley and ACM Press, Harlow (England).
- Gralla, Preston; (2007).  
Cómo funciona Internet.  
Anaya Multimedia, Madrid.
- Calishain, Tara; Dornfest, Rael; (2005).  
Google. Los mejores trucos (2ª edición).  
Anaya Multimedia, Madrid.
- Martos Rubio, Ana; (2006).  
Aprende a buscar en Internet. Edición 2006.  
Anaya Multimedia, Madrid.
- Abaurrea Velarde, Jorge; (2005).  
Internet. Edición 2006.  
Anaya Multimedia, Madrid.

# Software

- Sistema operativo: Linux, Windows, ...
- Navegador: Firefox, Mozilla, Opera, Konqueror, IE, ...
- Búscador web: Google, Yahoo, ...
- Servicios de búsqueda avanzada: Google Scholar, Google Book Search, Google Maps, Google Video, ...
- Herramientas de búsqueda personal: Google Desktop, Yahoo! Desktop Search, Copernic Desktop Search, Windows Desktop Search, ...
- Motores de indexación y búsqueda libremente disponibles: Swish-e (Simple Web Indexing System for Humans — Enhanced).

# Los profesores

## Área de **Ciencias de la Computación e Inteligencia Artificial** Departamento de Computación, Facultad de Informática

- Miguel A. Alonso Pardo
  - Profesor Titular de Universidad
  - Despacho 2.17 de la FIC
  - e-mail: [alonso@udc.es](mailto:alonso@udc.es)
- Carlos Gómez Rodríguez
  - Profesor Ayudante
  - Edificio Xoana Capdevielle
  - e-mail: [cgomezr@udc.es](mailto:cgomezr@udc.es)
- Jesús Vilares Ferro
  - Profesor Contratado Doctor
  - Depacho S0.1 de la FIC
  - e-mail: [jvilares@udc.es](mailto:jvilares@udc.es)

Horarios de tutorías disponibles en <http://www.fic.udc.es>