

Guía docente de la asignatura

Recuperación de Información (496114J)



Fecha de aprobación: 23/06/2023

Grado	Grado en Ingeniería Informática (Ceuta)	Rama	Ingeniería y Arquitectura				
Módulo	Formación de Especialidad 4: Sistemas de Información	Materia	Sistemas de Información en Web				
Curso	4º	Semestre	1º	Créditos	6	Tipo	Obligatoria

PRERREQUISITOS Y/O RECOMENDACIONES

No es necesario que los estudiantes tengan aprobadas asignaturas, materias o módulos previos como requisito indispensable para cursar este módulo. No obstante, se recomienda la superación de los contenidos y adquisición de competencias de las materias de formación básica y de rama, en especial las relacionadas con la programación.

BREVE DESCRIPCIÓN DE CONTENIDOS (Según memoria de verificación del Grado)

- Modelos de Recuperación de Información.
- Indexación.
- Recuperación de Información XML.
- Recuperación de Información Web.

COMPETENCIAS ASOCIADAS A MATERIA/ASIGNATURA

COMPETENCIAS GENERALES

- CG03 - Capacidad para diseñar, desarrollar, evaluar y asegurar la accesibilidad, ergonomía, usabilidad y seguridad de los sistemas, servicios y aplicaciones informáticas, así como de la información que gestionan.
- CG04 - Capacidad para definir, evaluar y seleccionar plataformas hardware y software para el desarrollo y la ejecución de sistemas, servicios y aplicaciones informáticas.

COMPETENCIAS TRANSVERSALES

- CT03 - Capacidad para el uso y aplicación de las TIC en el ámbito académico y profesional.



RESULTADOS DE APRENDIZAJE (Objetivos)

Objetivos formativos particulares

- Adquirir las destrezas, conocimientos y técnicas básicas para buscar información textual.
- Entender el concepto de modelo de recuperación de información.
- Adquirir una visión general del proceso de recuperación de información, así como de sus objetivos.
- Conocer los diferentes componentes de un sistema de recuperación de información, su funcionamiento y relaciones entre ellos.
- Comprender las peculiaridades de la recuperación de información estructurada con un lenguaje de marcado y las similitudes y diferencias con la recuperación de información clásica.
- Identificar los elementos que conforman la Web, así como conocer su estructura.
- Conocer las técnicas específicas para la recuperación de información en la Web.
- Asumir la importancia de la recuperación de información en el diseño y desarrollo de sistemas de información.
- Analizar problemas de acceso de información en el marco de los sistemas de información y diseñar e implementar un sistema de recuperación de información que les dé solución.
- Ser capaz de integrar un sistema de recuperación de información en un sistema de información.

Objetivos formativos de carácter general.

- Tener la capacidad de integrar soluciones de Tecnologías de la Información y las Comunicaciones y procesos empresariales para satisfacer las necesidades de información de las organizaciones, permitiéndoles alcanzar sus objetivos de forma efectiva y eficiente, dándoles así ventajas competitivas.
- Ser capaz de participar activamente en la especificación, diseño, implementación y mantenimiento de los sistemas de información y comunicación.

PROGRAMA DE CONTENIDOS TEÓRICOS Y PRÁCTICOS

TEÓRICO

Tema 1. Introducción a la recuperación de información.

- Definición de recuperación de información. Componentes de un sistema de recuperación de información y sus relaciones.

Tema 2. Procesamiento de Lenguaje Natural.

- Fuentes de datos textuales. Procesamiento de documentos y términos. Segmentación y lematización. Pesado de términos.

Tema 3. Modelos de recuperación.

- Concepto de modelo de recuperación. Modelo booleano. Modelo vectorial. BM25. Modelado de Lenguaje.

Tema 4. Evaluación de la recuperación.

- Conceptos básicos. Eficacia, eficiencia. Relevancia. Satisfacción de usuario. Precisión y exhaustividad. Medidas de la eficacia a partir de la curva precisión-exhaustividad. Medidas que consideran el orden.

Tema 5. Indexación de documentos.

- El proceso de indexación de documentos. Índices Invertidos. Indexación distribuida.

Tema 6. Recuperación de información Web.

- Características de la recuperación de información en la Web. Arañas. Reordenación de resultados basado en el análisis de enlaces. Posicionamiento Web.

Tema 7. Clasificación y Agrupamiento documental.

- Concepto de clasificación y agrupamiento documental, algoritmos básicos. Clasificación



de sentimientos.

PRÁCTICO

- Unidad P0: Búsqueda de información en la Web.
- Unidad P1: Indexación y Búsqueda con un motor de código abierto: Lucene
 - Práctica 2. Procesamiento de Textos.
 - Práctica 3. Diseño y desarrollo de un sistema de recuperación de información e implementación de un buscador utilizando Lucene.
 - P3.1 Indexación
 - P3.2 Búsqueda
- Unidad P2: Introducción a Solr.

Seminarios

- Seminario 1: Estrategias de búsqueda de información en la Web.
- Seminario 2: Sistemas de recuperación de información actuales.
- Seminario 3: Sistemas de Recomendación: Una aplicación sobre grandes datos.

BIBLIOGRAFÍA

BIBLIOGRAFÍA FUNDAMENTAL

- Cacheda F. Fernández-Luna, JM and Huete, J. Recuperación de Información: Un enfoque práctico y multidisciplinar. Ed. Rama 2011.
- M. McCandless, E. Hatcher and O. Gospodnetic. Lucene in Action (2d Edition) Manning 2011.
- Manning, P. Raghavan, and H. Schütze. Introduction to Information Retrieval, C. Cambridge University Press, 2008.
- Bruce Croft, Donald Metzler, Trevor Strohman, "Search Engines: Information Retrieval in Practice:" Addison Wesley/Prentice Hall, 2010.
- ChengXiang Zhai t Sean Massung. Text Data Management and Analysis. A practical introduction to Information Retrieval and text mining. ACM & Morgan Claypool Pubs. 2016
- A. Sharma. Practical Apache Lucene 8. Uncover the Search Capabilities of Your Application. Apress, 2020

BIBLIOGRAFÍA COMPLEMENTARIA

- Goker, J. Davies. Searching in the 21st century. Wiley. 2009.
- Feldman, R. and Sanger, J. The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data. Cambridge U. Press, 2006.
- D. A. Grossman, O. Frieder, Information Retrieval: Algorithms and Heuristics, Springer. 2004.
- Belew Richard K., "Finding Out About. A Cognitive perspective on Search Engine Technology and the WWW", Cambridge University Press, 2000.
- Salton, Automatic Text Processing, Addison-Wesley, 1989.
- Van Rijsbergen, Information Retrieval, Second Edition, Butterworth, 1979.
- Butcher S, Clarke C. Cornack G Information Retrieval: Implementing and Evaluating Search Engines (The MIT Press) 2016



ENLACES RECOMENDADOS

- [SIGIR](#)
- [Information Retrieval specialist group](#)
- [Introduction to Information Retrieval](#)

METODOLOGÍA DOCENTE

- MD01 - Lección Magistral (Clases Teóricas-Expositivas)
- MD02 - Actividades Prácticas (Resolución de Problemas, Resolución de Casos Prácticos, Desarrollo de Proyectos, Prácticas en Laboratorio, Taller de Programación, Aula de Informática, Prácticas de Campo).
- MD03 - Seminarios (Debates, Demos, Exposición de Trabajos Tutelados, Conferencias, Visitas Guiadas, Monografías).
- MD04 - Actividades no presenciales Individuales.
- MD05 - Actividades no presenciales Grupales.
- MD06 - Tutorías Académicas.

EVALUACIÓN (instrumentos de evaluación, criterios de evaluación y porcentaje sobre la calificación final)

EVALUACIÓN ORDINARIA

Todo lo relativo a la evaluación se regirá por la [Normativa de evaluación y calificación de los estudiantes vigente en la Universidad de Granada](#).

Régimen de Asistencia

Para considerar el sistema de evaluación continua, el estudiante deberá asistir al menos a un 80% de las clases teóricas y prácticas. Será necesario asistir a las clases prácticas en las que se realicen la evaluación de trabajos.

La evaluación seguirá las siguientes pautas:

Evaluación Continua

- Para la parte teórica se realizará mediante pruebas escritas cuya ponderación total es del 50% de la asignatura.
- Para la parte práctica se realizarán prácticas de laboratorio, resolución de problemas y desarrollo de proyectos, y se valorarán las entregas/presentaciones de las memorias realizadas por los alumnos. La ponderación de este bloque es del 40%.
- La parte de trabajo autónomo y los seminarios se evaluarán teniendo en cuenta la asistencia, los problemas propuestos que hayan sido resueltos y entregados por los alumnos y la presentación oral de los trabajos desarrollados. La ponderación es del 10%.

La calificación global será una calificación numérica obtenida mediante la suma ponderada de las calificaciones obtenidas en cada parte

EVALUACIÓN EXTRAORDINARIA

Todo lo relativo a la evaluación se regirá por la [Normativa de evaluación y calificación de los estudiantes vigente en la Universidad de Granada](#).

La evaluación en convocatorias extraordinarias se realizará en una prueba única que se valorará en el rango 0 a 10, donde se incluirá tanto la parte teórica como práctica. En el caso en que el alumno haya podido seguir el sistema de evaluación continua en el año en curso se podrá



considerar dicho esquema (ponderando la calificación de la parte teórica de la prueba al 50% de la nota final y respetando las calificaciones obtenidas en los otros apartados por el alumno en la evaluación continua). La calificación del alumno será el máximo entre las calificaciones obtenidas por cada sistema de evaluación, continua (si procede) y prueba única.

EVALUACIÓN ÚNICA FINAL

La evaluación única final se regirá por lo establecido en la [Normativa de evaluación y de calificación de los estudiantes de la Universidad de Granada](#).

Se realizará en un único acto académico el día de la convocatoria oficial de examen para la asignatura. Dicha prueba (evaluada de 0 a 10) incluirá preguntas tanto de tipo teórico como práctico que garanticen que el alumno ha adquirido la totalidad de las competencias descritas en esta misma guía docente, incluyendo aquellas herramientas software utilizadas en la parte práctica de la asignatura. La ponderación entre teoría/práctica será la misma que la indicada en la evaluación extraordinaria.

INFORMACIÓN ADICIONAL

No es necesario que los estudiantes tengan aprobadas asignaturas, materias o módulos previos como requisito indispensable para cursar este curso. No obstante, se recomienda la superación de los contenidos y adquisición de competencias de las materias de formación básica y de rama, en especial las relacionadas con la programación.

