		Guia d	locente			
	Datos Identificativos			2022/23		
Asignatura (*)	Recuperación de Información	Recuperación de Información Código		614G02027		
Titulación	Grao en Ciencia e Enxeñaría de	Datos				
		Descr	iptores			
Ciclo	Periodo	Cu	rso		Tipo	Créditos
Grado	2º cuatrimestre	Ter	cero		Obligatoria	6
Idioma	Castellano					
Modalidad docente	Presencial					
Prerrequisitos						
Departamento	Ciencias da Computación e Tecr	noloxías da Info	rmación			
Coordinador/a	Parapar López, Javier Correo electrónico javier.parapar@udc.es			udc.es		
Profesorado	Parapar López, Javier	Correo electrónico javier.parapar@udc.es			udc.es	
Web	www.dc.fi.udc.es/~parapar/					
Descripción general	Tradicionalmente los sistemas d	le recuperación	de información	venían s	siendo usados po	or documentalistas, bibliotecarios o
	abogados para la búsqueda de re	egistro. A día d	e hoy la situació	n ha car	mbiado radicalme	ente, cientos de millones de
	personas usan sistemas de recu	peración de info	ormación a diario	o: busca	n en la web, bus	can en su bandeja de correo,
	buscan dentro de su ordenador o reciben recomendaciones de consumo de contenido. La recuperación de información se			a recuperación de información se		
	ha convertido en el área dominante dentro del acceso a la información superando a las bases de datos tradicionales. Los					
	sistemas de recuperación de información son capaces de resolver necesidades del usuario sobre textos no estructurado			rio sobre textos no estructurados		
	sin necesidad de que este tenga que explicitar su consulta de una manera estándar. En esta asignatura se explorarán lo			esta asignatura se explorarán los		
	conceptos teóricos que dan sopo	orte a los sistem	nas de recupera	ición y a	cceso a la inform	ación, así como el software y
	herramientas para la construcció	n de sistemas a	avanzados de bu	úsqueda	y filtrado.	

	Competencias / Resultados del título
Código	Competencias / Resultados del título
A27	CE27 - Compresión y dominio de fundamentos y técnicas básicas para la búsqueda y el filtrado de información en grandes colecciones
	de datos.
B2	CB2 - Que los estudiantes sepan aplicar sus conocimientos a su trabajo o vocación de una forma profesional y posean las competencias
	que suelen demostrarse por medio de la elaboración y defensa de argumentos y la resolución de problemas dentro de su área de estudio
В3	CB3 - Que los estudiantes tengan la capacidad de reunir e interpretar datos relevantes (normalmente dentro de su área de estudio) para
	emitir juicios que incluyan una reflexión sobre temas relevantes de índole social, científica o ética
B4	CB4 - Que los estudiantes puedan transmitir información, ideas, problemas y soluciones a un público tanto especializado como no
	especializado
B7	CG2 - Elaborar adecuadamente y con cierta originalidad composiciones escritas o argumentos motivados, redactar planes, proyectos de
	trabajo, artículos científicos y formular hipótesis razonables.
B8	CG3 - Ser capaz de mantener y extender planteamientos teóricos fundados para permitir la introducción y explotación de tecnologías
	nuevas y avanzadas en el campo.
В9	CG4 - Capacidad para abordar con éxito todas las etapas de un proyecto de análisis de datos: exploración previa de los datos,
	preprocesado, análisis, visualización y comunicación de resultados.
B10	CG5 - Ser capaz de trabajar en equipo, especialmente de carácter multidisciplinar, y ser hábiles en la gestión del tiempo, personas y toma
	de decisiones.
C1	CT1 - Utilizar las herramientas básicas de las tecnologías de la información y las comunicaciones (TIC) necesarias para el ejercicio de su
	profesión y para el aprendizaje a lo largo de su vida.
C4	CT4 - Valorar la importancia que tiene la investigación, la innovación y el desarrollo tecnológico en el avance socioeconómico y cultural
	de la sociedad.

Resultados de aprendizaje	
Resultados de aprendizaje	Competencias /
	Resultados del título

Conocer, comprender y analizar los distintos modelos Recuperación de Información, las técnicas para su implementación	A27	В3	C1
eficiente y la metodología de evaluación de los mismos.		B4	C4
Conocer, comprender y analizar las plataformas software para la creación de estos sistemas.	A27	B2	
		B4	
		В9	
		B10	
Planear y realizar la evaluación de los sistemas Recuperación de Información . Analizar los resultados de la evaluación de los		В7	C1
sistemas de RI para mejorarlos en su eficacia y eficiencia.		В8	C4
Ser capaces de un correcto tratamiento de los aspectos éticos, de privacidad, confidencialidad y de seguridad de los estos	A27	B4	C4
sistemas.		В9	

Contenidos		
Tema	Subtema	
Arquitectura básica de un buscador	La arquitectura básica de un motor de búsqueda	
Análisis y procesamiento de textos	Del documento a los tokens del índice	
Índice invertido y procesamiento de consultas	Inverted files y estrategias de procesamiento de consultas	
Evaluación en recuperación de información	Métricas y métodos	
Modelos booleanos y de espacio vectorial	Modelos básicos de búsqueda	
Modelos de lenguaje	Modelos estadísticos de lenguaje	
Retroalimentación y operaciones de consulta	Retroalimentación de relevancia y reformulación de consultas	
Análisis de enlaces	Análisis de grafos web	

	Planificac	ión		
Metodologías / pruebas	Competencias /	Horas lectivas	Horas trabajo	Horas totales
	Resultados	(presenciales y	autónomo	
		virtuales)		
Prácticas de laboratorio	B2 B7 B9 B10 C1	14	42	56
Trabajos tutelados	B4 B7 B9	5	7.5	12.5
Prueba mixta	A27 B2 B4 B7 B8	2	13	15
Sesión magistral	A27 B3 B4 B8 C4	19	47.5	66.5
Atención personalizada		0		0
(*)Los datos que aparecen en la tabla de plani	ficación són de carácter orie	entativo, considerando	la heterogeneidad de	los alumnos

	Metodologías
Metodologías	Descripción
Prácticas de	Prácticas de laboratorio en plataformas de desarrollo ampliamente utilizadas en la industria, empresas de motores de
laboratorio	búsqueda y grupos de investigación.
Trabajos tutelados	Trabajos y problemas realizados de forma autónoma por el alumno y supervisados ??por el profesor
Prueba mixta	Prueba que girará sobre los contenidos fundamentales de la asignatura.
Sesión magistral	El alumno asistirá a las explicaciones del profesor sobre los diferentes modelos, técnicas y algoritmos de Recuperación de
	Información. El profesor utilizará diferentes niveles de abstracción-detalle y guiará al alumno en las lecturas fundamentales y
	complementarias.

	Atención personalizada
Metodologías	Descripción
Prácticas de	Prácticas laboratorio y trabajos tutleados: Además de evaluar el resultado de la práctica conforme a los requisitos exigidos se
laboratorio	hace un seguimiento del desarrollo de las mismas. Debe respetarse la autonomía del estudiante para que adquiera mayor
Trabajos tutelados	destreza con las plataformas software empleados pero el profesor podrá resolver ciertas dificultades que puedan bloquear al
	estudiante un tiempo excesivo dada la planificación de la asignatura.

		Evaluación	
Metodologías	Competencias /	Descripción	
	Resultados		
Prácticas de	B2 B7 B9 B10 C1	Seguimiento, defensa y evaluación de los resultados de las prácticas realizadas en	40
laboratorio		las horas de clases prácticas de laboratorio.	
		Es obligatorio alcanzar el 40% de la calificación para aprobar la asignatura 40	
Trabajos tutelados	B4 B7 B9	Participación y resultados en la realización del trabajo y / o preguntas.	10
Prueba mixta	A27 B2 B4 B7 B8	Preguntas sobre los conocimientos adquiridos en las sesiones magistrales,	50
		actividades prácticas y problemas y trabajos.	
		Es obligatorio alcanzar el 40% de la nota para aprobar la asignatura	

Observaciones evaluación

Para la segunda oportunidad y convocatorias no ordinarias, se evaluarán en el examen mixto tanto las prácticas como los trabajos y las teorías. Si no se alcanza la nota mínima en las distintas pruebas, la nota máxima del alumno será de 4,5

En la realización de la obra, el plagio y el uso de material no original, incluido el obtenido a través de Internet, sin indicación expresa de su origen y, en su caso, permiso de su autor, podrá ser considerado motivo de salvedad. Todo ello sin perjuicio de las responsabilidades disciplinarias que pudieran producirse con posterioridad al correspondiente trámite.

	Fuentes de información
Básica	 W.B. Croft, D. Metzler, T. Strohman (2009). Search Engines. Information Retrieval in Practice. Pearson Education C.D. Manning, P. Raghavan, H. Schutze (2008). Introduction to Information Retrieval. Cambridge University Press Baeza-Yates and B. Ribeiro-Neto (2011). Modern Information Retrieval (second edition). Addison Wesley/Pearson Education F. Cacheda, J.M. Fernández, J. Huete (editores) (2011). Recuperación de Información. Un enfoque práctico y multidisciplinar. Ra-Ma
Complementária	 Ian H. Witten (1999). Managing Gigabytes: Compressing and Indexing Documents and Images. Morgan Kaufmann Amy N. Langville, Carl D. D. Meyer (2011). Google's PageRank and Beyond: The Science of Search Engine Rankings. Princeton University Press

Recomendaciones
Asignaturas que se recomienda haber cursado previamente
Asignaturas que se recomienda cursar simultáneamente
Asignaturas que se resonnenta sur sin attantente
Asignaturas que continúan el temario
Total
Otros comentarios

(*) La Guía Docente es el documento donde se visualiza la propuesta académica de la UDC. Este documento es público y no se puede modificar, salvo cosas excepcionales bajo la revisión del órgano competente de acuerdo a la normativa vigente que establece el proceso de elaboración de guías