



Guía Docente				
Datos Identificativos				2021/22
Asignatura (*)	Recuperación de Información		Código	614G02027
Titulación				
Descriptores				
Ciclo	Período	Curso	Tipo	Créditos
Grao	2º cuatrimestre	Terceiro	Obrigatoria	6
Idioma	Castelán			
Modalidade docente	Presencial			
Prerrequisitos				
Departamento	Ciencias da Computación e Tecnoloxías da Información			
Coordinación	Parapar López, Javier	Correo electrónico	javier.parapar@udc.es	
Profesorado	Hasan Romero, Ismael Landín Piñeiro, Alfonso Parapar López, Javier	Correo electrónico	ismael.hasan@udc.es alfonso.landin@udc.es javier.parapar@udc.es	
Web	www.dc.fi.udc.es/~parapar/			
Descripción xeral	Tradicionalmente, os documentais, bibliotecarios ou avogados utilizaron sistemas de recuperación de información para buscar rexistros. Hoxe en día a situación cambiou radicalmente, centos de millóns de persoas usan diariamente sistemas de recuperación de información: buscan na web, buscan na caixa de correo, buscan no ordenador ou reciben recomendacións para o consumo de contido. A recuperación de información converteuse na área dominante no acceso á información, superando as bases de datos tradicionais. Os sistemas de recuperación de información son capaces de resolver as necesidades do usuario en canto a textos non estruturados sen necesidade de que o usuario teña que facer explícita a súa consulta dun xeito estándar. Esta materia explorará os conceptos teóricos que soportan o acceso á información e os sistemas de recuperación, así como o software e as ferramentas para construír sistemas avanzados de busca e filtrado.			
Plan de continxencia	1. Modificacións nos contidos Sin modificacións. 2. Metodoloxías *Metodoloxías docentes que se mantienen Todas *Metodoloxías docentes que se modifican Ninguna 3. Mecanismos de atención personalizada ao alumnado Mail, teams e moodle. 4. Modificacións na avaliación Ninguna *Observacións de avaliación: Esta asignatura está planeada como híbrida xa que a proba final realizarase de forma presencial si a situación o permite. 5. Modificacións da bibliografía ou webgrafía Sin modificacións			

Competencias / Resultados do título		
Código	Competencias / Resultados do título	

Resultados da aprendizaxe			
Resultados de aprendizaxe			Competencias / Resultados do título
Conocer, comprender y analizar los distintos modelos Recuperación de Información, las técnicas para su implementación eficiente y la metodología de evaluación de los mismos.		A27 B4	B3 C1 C4
Conocer, comprender y analizar las plataformas software para la creación de estos sistemas.		A27	B2 B4 B9 B10
Planejar y realizar la evaluación de los sistemas Recuperación de Información . Analizar los resultados de la evaluación de los sistemas de RI para mejorarlos en su eficacia y eficiencia.			B7 B8 C1 C4
Ser capaces de un correcto tratamiento de los aspectos éticos, de privacidad, confidencialidad y de seguridad de los estos sistemas.		A27 B9	B4 C4

Contidos	
Temas	Subtemas



Arquitectura básica dun buscador	A arquitectura básica dun motor de busca
Análise e procesamento de textos	Do documento ós tokens do índice
Construcción de índices e procesamiento de consultas	Archivos invertidos e estrategias de procesamiento de consultas
Avaliación clásica da recuperación de información	Avaliación offline
Avaliación clásica da recuperación de información	Máis alá da avaliação de laboratorio: avaliação en liña, satisfacción do usuario
Modelos de espacio booleano e vectorial	Modelos básicos de busca
Modelos de linguaxe	Modelos estadísticos de linguaxe
Operacións de retroalimentación e consulta	Relevancia e reformulación de consultas
Análise de ligazóns	Análise de grafos web

Planificación

Metodoloxías / probas	Competencias / Resultados	Horas lectivas (presenciais e virtuais)	Horas trabalho autónomo	Horas totais
Prácticas de laboratorio	B2 B7 B9 B10 C1	14	42	56
Traballos tutelados	B4 B7 B9	5	7.5	12.5
Proba mixta	A27 B2 B4 B7 B8	2	13	15
Sesión maxistral	A27 B3 B4 B8 C4	19	47.5	66.5
Atención personalizada		0		0

*Os datos que aparecen na táboa de planificación son de carácter orientativo, considerando a heteroxeneidade do alumnado

Metodoloxías

Metodoloxías	Descripción
Prácticas de laboratorio	Prácticas de laboratorio sobre plataformas de desenvolvemento de amplio uso na industria, nas compañías de Search Engines e nos grupos de investigación
Traballos tutelados	Traballos e problemas realizados de forma autónoma polo estudiante e tutelados polo profesor
Proba mixta	Prueba que versará sobre os contidos fundamentais da materia.
Sesión maxistral	O estudiante asistirá ás explicacións dadas polo profesor sobre os distintos modelos, técnicas e algoritmos de Recuperación de Información. O profesor utilizará distintos niveis de abstracción-detalle e orientará ao estudiante nas lecturas fundamentais e complementarias.

Atención personalizada

Metodoloxías	Descripción
Prácticas de laboratorio	Prácticas laboratorio y trabajos tutelados: Además de evaluar el resultado de la práctica conforme a los requisitos exigidos se hace un seguimiento del desarrollo de las mismas. Debe respetarse la autonomía del estudiante para que adquiera mayor 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.
Traballos tutelados	

Avaliación

Metodoloxías	Competencias / Resultados	Descripción	Cualificación
Prácticas de laboratorio	B2 B7 B9 B10 C1	Seguimiento, defensa e avaliação dos resultados das prácticas realizados nas horas de clases prácticas de laboratorio. É obligatorio alcanzar un 40% da cualificación para superar a materia	40
Traballos tutelados	B4 B7 B9	Participación e resultados na realización dos traballos e/ou cuestións.	10
Proba mixta	A27 B2 B4 B7 B8	Preguntas sobre os coñecementos adquiridos nas sesións maxistrais, actividades prácticas e de problemas e traballos. É obligatorio alcanzar un 40% da cualificación para superar a materia	50



Observacións avaliación

Para a segunda oportunidade e as

convocatorias non ordinarias, tanto as prácticas e traballos como a teorías avaliaranse no exame mixto. Se non se acada a nota mínima nas distintas probas a nota máxima do alumno será 4.5

Na realización de traballos, o plaxio e a utilización de material non orixinal, incluído aquel obtido a través da internet, sen indicación expresa da súa procedencia e, se é o caso, o permiso do seu autor/a, poderá ser considerada causa de cualificación de suspenso. Todo iso sen prexuízo das responsabilidades disciplinarias ás que puidese haber lugar tras o correspondente procedemento

Fontes de información

Bibliografía básica	<ul style="list-style-type: none">- 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
Bibliografía complementaria	<ul style="list-style-type: none">- 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

Recomendacións

Materias que se recomenda ter cursado previamente

Materias que se recomienda cursar simultaneamente

Materias que continúan o temario

Observacións

(*)A Guía docente é o documento onde se visualiza a proposta académica da UDC. Este documento é público e non se pode modificar, salvo casos excepcionais baixo a revisión do órgano competente dacordo coa normativa vixente que establece o proceso de elaboración de guías