

## PLAN DOCENTE DE LA ASIGNATURA

**Curso académico: 2025/2026**

Identificación y características de la asignatura					
Código	502285				
Denominación (español)	<b>RECUPERACIÓN DE LA INFORMACIÓN</b>				
Denominación (inglés)	INFORMATION RETRIEVAL				
Titulaciones	Grado en Información y Documentación Doble Grado en Comunicación Audiovisual/Información y Documentación Doble Grado en Periodismo/Información y Documentación				
Centro	Facultad de Ciencias de la Documentación y la Comunicación				
Módulo	Tecnologías y aplicación práctica de la información y documentación				
Materia	Representación y recuperación de la información				
Carácter	Obligatorio	ECTS	6	Semestre	1º
Profesorado					
Nombre		Despacho		Correo-e	
Cristina López Pujalte		22A		cloppuj@unex.es	
Área de conocimiento	Biblioteconomía y Documentación				
Departamento	Información y Comunicación				
Profesor/a coordinador/a (si hay más de uno)					
Competencias					
<p>Competencias básicas:</p> <p>CB1 - Que los estudiantes hayan demostrado poseer y comprender conocimientos en un área de estudio que parte de la base de la educación secundaria general, y se suele encontrar a un nivel que, si bien se apoya en libros de texto avanzados, incluye también algunos aspectos que implican conocimientos procedentes de la vanguardia de su campo de estudio.</p> <p>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.</p> <p>CB3 - Que los estudiantes tengan la capacidad de reunir e interpretar datos</p>					

relevantes (normalmente dentro de su área de estudio) para emitir juicios que incluyan una reflexión sobre temas relevantes de índole social, científica o ética.

CB4 - Que los estudiantes puedan transmitir información, ideas, problemas y soluciones a un público tanto especializado como no especializado.

CB5 - Que los estudiantes hayan desarrollado aquellas habilidades de aprendizaje necesarias para emprender estudios posteriores con un alto grado de autonomía.

#### Competencias generales:

CG2 - Conocimiento de los principios teóricos y metodológicos para la reunión, selección, organización, representación, preservación, recuperación, acceso, difusión e intercambio de la información.

CG3 - Conocimiento de las tecnologías de la información que se emplean en las unidades y servicios de información.

CG4 - Habilidades en el manejo de las tecnologías como medio indispensable en los procesos de tratamiento y transferencia de la información.

#### Competencias específicas:

CE1 - Conocimiento de los principios teóricos y metodológicos para la planificación, organización y evaluación de sistemas, unidades y servicios de información.

CE4 - Comprender y aplicar los principios y las técnicas para la planificación, organización y evaluación de sistemas, unidades y de servicios de información.

CE5 - Comprender y aplicar los principios y las técnicas para la reunión, selección, organización, representación, preservación, recuperación, acceso, difusión e intercambio de la información.

CE10 - Capacidad de usar y aplicar las técnicas, las normativas y otros instrumentos utilizados en la reunión, selección, organización, representación, preservación, recuperación, acceso, difusión e intercambio de la información.

#### Competencias transversales:

CT1 - Capacidad de análisis y de síntesis aplicadas a la gestión y organización de la información.

CT3 - Habilidades en el uso de Internet y software genérico (ofimática).

CT5 - Capacidad de organización y planificación del trabajo propio.

CT8 - Razonamiento crítico en el análisis y la valoración de alternativas.

CT10 - Capacidad para el aprendizaje autónomo.

## Contenidos

Descripción general del contenido: Historia, Evaluación y Leyes aplicadas a la Recuperación y Representación Documental. Estrategias Booleanas de Búsqueda.

### Temario

Denominación del tema 1: Introducción. Conceptos básicos.

Contenidos del tema 1:

- 1.1. Introducción. Acotación del ámbito abarcado por la asignatura
- 1.2. Concepto de Recuperación de Información (RI). Su papel dentro del Proceso Documental
- 1.3. Recuperación de información e indización
- 1.4. Sistemas de Recuperación de Información
- 1.5. Breve historia de la RI: *Astia, Cranfield, SMART...*

Denominación del tema 2: Evaluación de la Recuperación de Información.

Contenidos del tema 2:

- 2.1. Criterios de evaluación. Eficacia. Concepto de relevancia
- 2.2. Medidas de evaluación
- 2.3. Evaluación según la técnica de recuperación empleada
- 2.4. Test de diferencias de medias

Descripción de las actividades prácticas del tema 2: Batería de ejercicios prácticos correspondiente a los contenidos del tema 2.

Denominación del tema 3: Modelos de RI y Representación documental.

Contenidos del tema 3:

- 3.1. Sistemas basados en índices invertidos
- 3.2. Leyes de Zipf
- 3.3. Valor de información
- 3.4. Modelos de RI
- 3.5. Representación documental. Modelo del Espacio Vectorial
- 3.6. Técnicas de recuperación. Medidas de similitud
- 3.7. Selección de las componentes y ponderación
- 3.8. Componentes de un buscador

Descripción de las actividades prácticas del tema 3: Batería I y II de ejercicios prácticos correspondiente a los contenidos del tema 3.

Denominación del tema 4: Mejoras de la RI

Contenidos del tema 4:

- 4.1. Concepto de unitérminos. Limitaciones
- 4.2. Mejoras a los unitérminos

- 4.3. La importancia de las consultas. Estrategias de búsqueda. Técnicas de modificación automática de consultas: Relevance Feedback
- 4.4. Tendencias de investigación relacionadas: Aplicaciones de Inteligencia Artificial en RI

Descripción de las actividades prácticas del tema 4: Batería de ejercicios prácticos correspondiente a los contenidos del tema 4.

### Actividades formativas

TEMA	TOTAL	Actividades Presenciales (AP)					Actividades Virtuales (AV)				TP	TA
		GG	CH	L	O	S	CST	CSP	CAT	CAP		
1	15								1			14
2	39,5									2	0,5	37
3	39,5									2	0,5	37
4	34									1	0,5	32,5
<b>Evaluación</b>	22	2										20
<b>Totales</b>	150	2							1	5	1,5	140,5
		1,33% Presencialidad					98,67% Virtualidad					

#### Actividades Presenciales (AP)

Actividades que se desarrollan en un único espacio físico y que implican interacción física entre estudiante y docente:

- GG: Grupo Grande (85 estudiantes).
- CH: Actividades de prácticas clínicas hospitalarias (7 estudiantes)
- L: Actividades de laboratorio o prácticas de campo (15 estudiantes)
- O: Actividades en sala de ordenadores o laboratorio de idiomas (20 estudiantes)
- S: Actividades de seminario o de problemas en clase (40 estudiantes).

#### Actividades Virtuales (AV)

Actividades que no se desarrollan en un espacio físico común. Pueden ser síncronas (implican interacción estudiante / docente) o asíncronas:

- CST: Clase síncrona teórica.
- CSP: Clase síncrona práctica.
- CAT: Clase asíncrona teórica.
- CAP: Clase asíncrona práctica.

- TP: Tutorías Programadas (seguimiento docente, tutorías ECTS).
- TA: Trabajo autónomo del estudiante.

### Metodologías docentes

- Utilización de material docente en diferentes tipos y formatos.
- Discusión de los contenidos.
- Análisis y resolución de problemas prácticos propuestos.

- Actividades de seguimiento del aprendizaje.
- Autoevaluaciones.

## Sistemas de evaluación

Según la normativa de evaluación vigente, el estudiante comunicará al profesor por el Campus Virtual de la Universidad (CVUEx) el tipo de evaluación elegido (modalidad de evaluación continua o modalidad de evaluación global) al principio de cada semestre. Cuando un estudiante no realice esta comunicación, se entenderá que opta por la modalidad de evaluación continua.

El sistema de evaluación por defecto, es decir, la **modalidad de evaluación continua** se realizará según los siguientes parámetros:

La evaluación de esta asignatura se realizará sobre los conocimientos teóricos y prácticos adquiridos por el alumno, y las prácticas realizadas. En todo caso, el examen teórico/práctico deberá aprobarse para superar la asignatura.

La calificación final será la suma de la evaluación continua y del examen atendiendo a los siguientes criterios:

- Evaluación continua (60%): realización de prácticas tuteladas relacionadas con la materia desarrolladas durante la impartición de la asignatura que se realizarán y entregarán a través de la plataforma del aula virtual. La evaluación continua **no es recuperable en ningún caso**, y deberá realizarse en los plazos establecidos por el profesor durante el semestre correspondiente (1º semestre). La calificación obtenida en la evaluación continua solo estará vigente para las convocatorias del curso académico correspondiente.
- Examen (40%): realización de una prueba escrita sobre conocimientos teóricos y prácticos de la materia. En todo caso **el examen deberá aprobarse para poder superar la asignatura**, y solo en este caso se sumará a la nota obtenida en el examen la nota de la evaluación continua.

En cuanto a la **modalidad de evaluación global**, para aquellos alumnos que lo hayan solicitado, consistirá en una prueba final de carácter global, de manera que la superación de esta suponga la superación de la asignatura.

**IMPORTANTE:** En ambas modalidades, **el examen será exactamente el mismo**, y la calificación final mínima para superar la asignatura será de Aprobado (5).

En cualquier caso, en las convocatorias extraordinarias, aunque la elección haya sido evaluación continua, el alumno tendrá la posibilidad de superar la

asignatura en una prueba final que englobe todos los contenidos de la misma.

Para los aspectos no contemplados en estos criterios se seguirá lo dispuesto en la normativa de evaluación vigente en la Universidad de Extremadura.

### Resultados de aprendizaje

- Encuadrar el concepto de recuperación de información en del proceso documental.
- Comprender la importancia de la representación documental (indización) y su relación con la recuperación de información.
- Conocer la historia de los sistemas de recuperación de información y cómo han evolucionado a lo largo del tiempo.
- Comprender la importancia de una correcta evaluación de la recuperación de la información, y aplicar las técnicas necesarias para llevarla a cabo.
- Conocer y emplear los distintos modelos de recuperación de información.
- Utilizar las metodologías de automatización de la representación documental y determinar su repercusión en la posterior recuperación a partir de sus ventajas e inconvenientes.
- Aplicar las técnicas básicas de recuperación y de representación documental, tanto estadísticas como lingüísticas.
- Comprender la importancia de la indización de las consultas en la mejora de los resultados de la recuperación, así como de las estrategias de búsquedas.

### Bibliografía (básica y complementaria)

Bibliografía básica:

Baeza-Yates, R.; Ribeiro-Neto, B. (1999). *Modern information retrieval*. Essex, UK: Addison-Wesley.

[Büttcher, S.; Clarke, C.L.A.; Cormack, G.V. \(2016\). \*Information Retrieval: Implementing and Evaluating Search Engines\*. Cambridge: MIT Press. Disponible en: <https://mitmecsept.wordpress.com/wp-content/uploads/2018/05/stefan-bc3bcttcher-charles-l-a-clarke-gordon-v-cormack-information-retrieval-implementing-and-evaluating-search-engines-2010-mit.pdf>](https://mitmecsept.wordpress.com/wp-content/uploads/2018/05/stefan-bc3bcttcher-charles-l-a-clarke-gordon-v-cormack-information-retrieval-implementing-and-evaluating-search-engines-2010-mit.pdf)

Cacheda Sijo, F.; Manuel Fernández Luna, J.; Huete Guadix, J.F. (coord.) (2011). *Recuperación de información: un enfoque práctico y multidisciplinar*. Madrid: RA-MA, cop.

Cheng, L., Yang, Y., Zhao, K., & Gao, Z. (2020). Research and improvement of tf-idf algorithm based on information theory. *Advances in Intelligent Systems and Computing*, 905, 608-616.

Chu, H. (2007). *Information Representation and Retrieval in the Digital Age*. Medford, New Jersey: Information Today.

Ellis, D. (1996). *Progress and problems in information retrieval*. London: The Library Association.

Frakes, W. B. & Baeza-Yates, R. (1992). *Information retrieval: data structures & algorithms*. New Jersey: Upper Saddle River.

Harman, D.K. (1992). Relevance feedback and other query modification techniques. In: Frakes, W.B.; Baeza-Yates, R. (Eds.) *Information retrieval: data structures and algorithms*. NJ: Prentice Hall, 241-263.

Hutter, M., Quarel, D., & Catt, E. (2024). *An Introduction to Universal Artificial Intelligence*. CRC Press.

Manning, C.D.; Raghavan, P.; Schütze, H. (2008). *Introduction to Information Retrieval*. UK: Cambridge University Press.

Moya Anegón, F. de. (1995). *Los Sistemas Integrados de Gestión Bibliotecarias*. Madrid: ANABAD.

Salton, G. (1989). *Automatic text procesing: the transformation, analysis, and retrieval of information by computer*. Reading, Massachusetts: Addison-Wesley.

Salton, G. & McGill, M.J. (1983). *Introduction to modern information retrieval*. New York: McGraw-Hill.

Zhai, C., & Massung, S. (2016). *Text Data Management and Analysis: A Practical Introduction to Information Retrieval and Text Mining*. New York: Association for Computing Machinery and Morgan & Claypool Publishers.

#### Bibliografía complementaria:

Cheng, L., Yang, Y., Zhao, K., & Gao, Z. (2020). Research and improvement of tf-idf algorithm based on information theory. *Advances in Intelligent Systems and Computing*, 905, 608-616.

Keyser, P. de. (2007). *Indexing From Thesauri to the Semantic Web*. Oxford, UK: Chandos Publishing.

Langville, A.N. & Meyer, C.D. (2011). *Google's PageRank and Beyond: The Cience of Search Engine Rankings*. Princeton, N.J.: Princeton University Press.

Mayer-Schönberger, V. & Cukier, K. (2013). *Big data. La revolución de los datos masivos*. Madrid: Turner Publicaciones S.L.

Parga, N. (2025). Premio Nobel de Física 2024. Descubrimientos e inventos fundamentales que impulsan el aprendizaje automático con redes neuronales artificiales. *Revista Española De Física*, 39(1), 47-53. Disponible en: <https://revistadefisica.es/index.php/ref/article/view/3093/2453>

Pestana Caldes, A.I.; López-Pujalte, C. (2009). Algoritmos Genéticos aplicados a la Recuperación de Información. Em: Borges, M.M. & Sanz Casado, E. (ed.). *A Ciência da Informação Criadora de Conhecimento*. Vol. II, 405-408. Coimbra: Coimbra University Press.

Strzalkowski, T. (ed.) (1999). *Natural Language Information Retrieval*. Dordrecht,

Netherlands: Kluwer Academic Publishers.

### Otros recursos y materiales docentes complementarios

- Espacio virtual de la asignatura en el campus virtual de la UEx: <http://campusvirtual.unex.es>
- Se utilizarán colecciones de test específicas de evaluación de recuperación de información (Cranfield, CISI, Medline, etc.) accesibles on-line: [http://ir.dcs.gla.ac.uk/resources/test\\_collections/](http://ir.dcs.gla.ac.uk/resources/test_collections/)