

# I Convocatoria de acciones para la adaptación de UEX al EEES

## Equipo

Director: M<sup>a</sup> Victoria Nuño Moral  
Felipe Zapico Alonso  
Vicente P. Guerrero Bote  
Cristina López Pujalte  
María J. Reyes Barragán  
Jorge Caldera Serrano

## I Convocatoria de acciones para la adaptación de UEx al EEES

<i>Datos del Proyecto</i>	
<i>Título del Proyecto</i>	<b>“3 Titulaciones con Europa”</b>
<i>Director</i>	María Victoria Nuño Moral
<i>Titulación/es implicada/s</i>	Diplomado en Biblioteconomía y Documentación. Licenciado en Documentación Licenciado en Comunicación Audiovisual

Plan Docente de la materia

**“Indización y Condensación Documental Automatizadas”.**

I. Descripción y contextualización

<i>Identificación y características de la materia</i>				
<i>Denominación</i>	Indización y condensación documental automatizadas			
<i>Curso y Titulación</i>	1º Licenciatura en Documentación			
<i>Profesor</i>	Cristina López Pujalte			
<i>Área</i>	Biblioteconomía y Documentación			
<i>Departamento</i>	<i>Informática</i>			
<i>Tipo y ctos. LRU</i>	Materia Troncal (3,5 + 1 créditos LRU)		Formación Inicial	
<i>Coefficientes</i>	Practicidad: 2 (Medio-alto, profesional)		Agrupamiento: 3 (Medio-alto)	
<i>Duración ECTS (créditos)</i>	Primer Cuatrimestre		4,15 ECTS (104 h.)	
<i>Distribución ECTS (rangos)</i>	Grupo Grande:35 %	Seminario-Lab.: 10%	Tutoría ECTS:5 %	No presenciales: 50 %
	36-37 horas	10-11 horas	5 horas	52 horas
<i>Descriptor (según BOE)</i>	Sistemas de recuperación de la información. Compatibilidad de idiomas y sistemas. Condensación de contenidos documentales			

<i>Competencias Específicas de la Titulación (CET)</i> Diplomado en Biblioteconomía y Documentación Licenciado en Documentación	Nº perfil/es
1. INTERACCIÓN CON LOS PRODUCTORES, LOS USUARIOS Y LOS CLIENTES DE LA INFORMACIÓN	E01
2. CONOCIMIENTO DEL ENTORNO PROFESIONAL DE LA INFORMACIÓN Y LA DOCUMENTACIÓN	E02
3. CONOCIMIENTO DEL MARCO JURÍDICO Y ADMINISTRATIVO NACIONAL E INTERNACIONAL DE LA GESTIÓN DE LA INFORMACIÓN	E03
4. IDENTIFICACIÓN, AUTENTIFICACIÓN Y EVALUACIÓN DE FUENTES Y RECURSOS DE INFORMACIÓN	E04
5. GESTIÓN DE COLECCIONES Y FONDOS	E05
6. PRESERVACIÓN, CONSERVACIÓN, Y TRATAMIENTO FÍSICO DE DOCUMENTOS	E06
7. ANÁLISIS Y REPRESENTACIÓN DE LA INFORMACIÓN.	E07
8. ORGANIZACIÓN Y ALMACENAMIENTO DE LA INFORMACIÓN	E08
9. BÚSQUEDA Y RECUPERACIÓN DE LA INFORMACIÓN	E09
10. ELABORACIÓN Y DIFUSIÓN DE LA INFORMACIÓN	E10
11. TECNOLOGÍAS DE LA INFORMACIÓN: INFORMÁTICA	E11
12. TECNOLOGÍAS DE LA INFORMACIÓN: TELECOMUNICACIONES	E12
13. TÉCNICAS DE PRODUCCIÓN Y EDICIÓN	E13
14. TÉCNICAS DE GESTIÓN ADMINISTRATIVA.	E14
15. TÉCNICAS DE MARKETING	E15
16. TÉCNICAS COMERCIALES	E16
17. TÉCNICAS DE ADQUISICIÓN	E17
18. TÉCNICAS DE GESTIÓN MICRO ECONÓMICA	E18
19. TÉCNICAS DE INSTALACIÓN, ACONDICIONAMIENTO Y EQUIPAMIENTO	E19
20. TÉCNICAS DE PLANIFICACIÓN Y GESTIÓN DE PROYECTOS	E20
21. TÉCNICAS DE DIAGNÓSTICO Y EVALUACIÓN	E21
22. TÉCNICAS DE GESTIÓN DE RECURSOS HUMANOS	E22
23. TÉCNICAS DE FORMACIÓN	E23

<i>Competencias específicas de la Materia 2:</i> Indización y condensación documental automatizadas	CET
1. INTERACCIÓN CON LOS PRODUCTORES, LOS USUARIOS Y LOS CLIENTES DE LA INFORMACIÓN	E01
2. CONOCIMIENTO DEL ENTORNO PROFESIONAL DE LA INFORMACIÓN Y LA DOCUMENTACIÓN	E02
3. GESTIÓN DE COLECCIONES Y FONDOS	E05
4. ANÁLISIS Y REPRESENTACIÓN DE LA INFORMACIÓN.	E07
5. ORGANIZACIÓN Y ALMACENAMIENTO DE LA INFORMACIÓN	E08
6. BÚSQUEDA Y RECUPERACIÓN DE LA INFORMACIÓN	E09
7. ELABORACIÓN Y DIFUSIÓN DE LA INFORMACIÓN	E10
8. TECNOLOGÍAS DE LA INFORMACIÓN: INFORMÁTICA	E11
9. TECNOLOGÍAS DE LA INFORMACIÓN: TELECOMUNICACIONES	E12
10. TÉCNICAS DE DIAGNÓSTICO Y EVALUACIÓN	E21

## II. Objetivos

<i>Relacionados con competencias académicas y disciplinares</i>	<i>Vinculación</i>
<i>Descripción</i>	<i>CET<sup>i</sup></i>
1. Recordar el concepto de indización y condensación documental, y las funciones que realizan dentro del Proceso Documental, así como el resto de conceptos básicos (documentales, estadísticos y tecnológicos) necesarios para poder acometer la docencia de la asignatura.	1, 2, 3, 10
2. Comprender la importancia de la indización y cómo está íntimamente relacionada con la Recuperación de Información. Asimismo, asimilar el papel fundamental del resumen en estos tiempos de sobrecarga de información.	4, 5, 6, 7
3. Conocer las distintas metodologías de automatización de la indización y condensación documental de manera general, cómo surgen, sus orígenes, características, y las bases científicas en las que se apoyan.	4, 6, 7, 8, 9
4. Profundizar posteriormente en cada una de ellas, analizando sus características específicas, sus técnicas básicas, sus principales ventajas y limitaciones, así como las principales investigaciones en torno a ellas.	4, 6, 7, 8, 9
5. Ahondar en una de estas metodologías, el Procesamiento del Lenguaje Natural, al ser ésta la que encierra mayor complejidad y a la vez mayores expectativas.	2, 8, 9
6. Comprender la importancia de la indización de las consultas, así como de las técnicas automáticas de formulación de las mismas.	4, 6
7. Manejar con soltura las técnicas básicas de automatización, tanto estadísticas como lingüísticas	4, 7, 8, 9
8. Conocer las últimas tendencias en sistemas automáticos de indización y resumen, y sus perspectivas de futuro en la sociedad de la información.	1, 2, 3, 8, 9

<i>Relacionados con otras competencias personales y profesionales</i>	<i>Vinculación</i>
<i>Descripción</i>	<i>CET</i>
9. Incentivar actitudes intelectuales positivas y métodos de trabajo científico	2, 10
10. Desarrollar el hábito de lectura de literatura científica	10
11. Despertar en el alumno cierta “inquietud” y curiosidad, así como aprender los rudimentos del método científico	1, 2, 10
12. Desarrollar habilidades de razonamiento y argumentación	2, 10
13. Acostumbrarse a manejar el lenguaje técnico especializado de la disciplina	2

### III. Contenidos

<b>Secuenciación de bloques temáticos y temas</b>	
<b>Tema 1: Introducción. Conceptos básicos.</b>	
1.1.	Introducción. Acotación del ámbito abarcado por la asignatura
1.2.	Conceptos de indización y condensación documental: su papel dentro de la Cadena Documental
1.3.	Recuperación de información (RI). Relación con la indización y el resumen
1.4.	Breve historia de la RI: <i>Astia, Cranfield, SMART...</i>
1.5.	Necesidad y ventajas de la automatización de la indización y condensación documental
1.6.	Ventajas del lenguaje natural en la sociedad de la información
1.7.	Sistemas basados en índices invertidos
1.8.	Modelos de representación documental. Modelo del espacio vectorial
1.9.	Evaluación de los sistemas de recuperación de información
<b>Tema 2: Metodología estadística</b>	
2.1.	Introducción
2.2.	Leyes de Zipf
2.3.	Crecimiento del vocabulario
2.4.	Teoría matemática de la comunicación. Concepto de entropía
2.5.	Primeras investigaciones de Luhn
2.6.	Esquemas de ponderación de términos
2.7.	Modelo probabilístico
2.8.	<i>Stemming</i>
2.9.	Proceso de indización automática básico
2.10.	Limitaciones de los métodos estadísticos. Conclusiones
<b>Tema 3: Metodología pseudolingüística: a medio camino entre los métodos estadísticos y los lingüísticos</b>	
3.1.	Problemas del empleo exclusivo de unitérminos
3.2.	Relaciones entre términos: tesauros y formación de frases
3.3.	Indización mediante citas
3.4.	Ejemplos de aplicaciones
3.5.	Limitaciones de los métodos pseudolingüísticos. Conclusiones
<b>Tema 4: Procesamiento del Lenguaje Natural (PLN)</b>	

4.1.	Introducción: Definición de PLN y contextualización
4.2.	Clasificación del conocimiento de la Lengua
4.3.	Principales aplicaciones del PLN
4.4.	Modelos computacionales del lenguaje
4.5.	Fases o niveles de análisis del PLN desde la perspectiva del análisis textual
4.6.	Ejemplos de aplicaciones
4.7.	Conclusiones
<b>Tema 5: Mejoras en la indización automática de consultas: Retroalimentación por relevancia</b>	
5.1.	Introducción: Importancia de la indización de las consultas
5.2.	Técnicas de mejora en la definición de consultas: Retroalimentación por relevancia
5.3.	Retroalimentación por relevancia implementada por métodos clásicos
5.4.	Retroalimentación por relevancia implementada por herramientas de la Computación Evolutiva: Algoritmos Genéticos
5.5.	Conclusiones
<b>Tema 6: Introducción a la condensación automática</b>	
6.1.	Particularidades del resumen desde la perspectiva de su automatización
6.2.	Clasificaciones de las estrategias de automatización de la condensación
<b>Tema 7: Métodos estadísticos de resumen automático</b>	
7.1.	Introducción
7.2.	El trabajo pionero de Luhn: principio básico del <i>sentence-scoring</i>
7.3.	Sistema de extracción de frases de Salton
7.4.	Investigaciones de Edmundson: Métodos Cue, Clave, Título y Localización
7.5.	Trabajos de Rush, Salvador y Zamora: Métodos eliminación-frases
7.6.	Sistema desarrollado por Paice basado en indicaciones contextuales
7.7.	Sistemas basados en párrafos
7.8.	Limitaciones de estos sistemas. Conclusiones
<b>Tema 8: Métodos basados en PLN de resumen automático</b>	
8.1.	Introducción
8.2.	Características más comunes de los sistemas de condensación automáticos basados en técnicas de PLN
8.3.	Ejemplos de aplicaciones basadas en PLN
8.4.	Limitaciones actuales de estos sistemas. Conclusiones

<i>Interrelación</i>					
Requisitos (Rq) y redundancias (Rd)		Tema		Procedencia	
Indización y condensación intelectual	Rq	1.2		Análisis de contenido y Lenguajes documentales: sistemas postcoordinados (2º de ByD)	
Conceptos básicos de informática	Rq	La asignatura en su totalidad		Introducción a la tecnología de la información (1º de ByD)	
Conocimientos sobre recuperación de información y modelos de representación documental	Rq	1.3, 1.7, 1.8		Sistemas de almacenamiento y recuperación de la información (2º de ByD)	
Conceptos básicos de estadística	Rq	1.9, tema 2		Métodos estadísticos en documentación (1º de Documentación)	
Aplicaciones de Algoritmos genéticos a la RI	Rd	5.4		Procesamiento automático del conocimiento (1º de Documentación)	

#### IV. Metodología docente y plan de trabajo del estudiante

Actividades de enseñanza-aprendizaje				Vinculación	
Descripción y secuenciación de actividades	Tipo <sup>i</sup>		Duración <sup>i</sup>	Tema	Objetivos
	GG	C-E			
<b>1. Presentación de la asignatura</b>	GG	C-E	1	1-8	--
<b>2. Lectura recomendada de repaso/introducción al tema</b>	NP	T	1	1.2, 1.3, 1.7, 1.8	1
<b>3. Exposición general del tema y discusión</b>	GG	T	2	1	1, 2
<b>4. Estudio de los contenidos explicados</b>	NP	T	1	1	1, 2
<b>5. Resolución de problemas tipo sobre evaluación de sistemas de RI</b>	GG	T-P	2	1.9	1, 7, 9, 12
<b>6. Realización de ejercicios</b>	NP	P	3	1.9	1, 7, 9, 12
<b>7. Dudas sobre los ejercicios realizados</b>	GG	T-P	1	1.9	1, 7, 9, 12
<b>8. Explicación y discusión del tema</b>	GG	T	3	2	3, 4
<b>9. Estudio de los contenidos explicados</b>	NP	T	1	2	3, 4



<b>10. Resolución de problemas tipo sobre técnicas estadísticas</b>	<b>GG</b>	<b>T-P</b>	<b>2</b>	<b>2.2, 2.3, 2.4, 2.5, 2.6, 2.9</b>	<b>4, 7, 9, 12</b>
<b>11. Realización de ejercicios</b>	<b>NP</b>	<b>P</b>	<b>3</b>	<b>2.2, 2.3, 2.4, 2.5, 2.6, 2.9</b>	<b>4, 7, 9, 12</b>
<b>12. Dudas sobre los ejercicios realizados</b>	<b>GG</b>	<b>T-P</b>	<b>1</b>	<b>2.2, 2.3, 2.4, 2.5, 2.6, 2.9</b>	<b>4, 7, 9, 12</b>
<b>13. Exposición general del tema y discusión</b>	<b>GG</b>	<b>T</b>	<b>3</b>	<b>3</b>	<b>3, 4</b>
<b>14. Estudio de los contenidos explicados</b>	<b>NP</b>	<b>T</b>	<b>1</b>	<b>3</b>	<b>3, 4</b>
<b>15. Resolución de problemas tipo sobre técnicas pseudolingüísticas</b>	<b>GG</b>	<b>T-P</b>	<b>2</b>	<b>3.1, 3.2</b>	<b>4, 7, 9, 12</b>
<b>16. Realización de ejercicios</b>	<b>NP</b>	<b>P</b>	<b>3</b>	<b>3.1, 3.2</b>	<b>4, 7, 9, 12</b>
<b>17. Dudas sobre los ejercicios realizados</b>	<b>GG</b>	<b>T-P</b>	<b>1</b>	<b>3.1, 3.2</b>	<b>4, 7, 9, 12</b>
<b>18. Exposición general del tema y discusión</b>	<b>GG</b>	<b>T</b>	<b>3</b>	<b>4</b>	<b>3, 4, 5</b>
<b>19. Estudio de los contenidos explicados</b>	<b>NP</b>	<b>T</b>	<b>1</b>	<b>4</b>	<b>3, 4, 5</b>
<b>20. Resolución de problemas tipo sobre técnicas de PLN</b>	<b>GG</b>	<b>T-P</b>	<b>1</b>	<b>4.5</b>	<b>5, 7, 9, 12</b>
<b>21. Realización de ejercicios</b>	<b>NP</b>	<b>P</b>	<b>3</b>	<b>4.5</b>	<b>5, 7, 9, 12</b>
<b>22. Dudas sobre los ejercicios realizados</b>	<b>GG</b>	<b>T-P</b>	<b>1</b>	<b>4.5</b>	<b>5, 7, 9, 12</b>
<b>23. Exposición general del tema</b>	<b>GG</b>	<b>T</b>	<b>2</b>	<b>5</b>	<b>6</b>
<b>24. Estudio de los contenidos explicados</b>	<b>NP</b>	<b>T</b>	<b>1</b>	<b>5</b>	<b>6</b>
<b>25. Lectura recomendada sobre investigaciones actuales de Algoritmos Genéticos</b>	<b>NP</b>	<b>T</b>	<b>1</b>	<b>5.4</b>	<b>6, 10, 11</b>
<b>26. Comentario y discusión del artículo</b>	<b>GG</b>	<b>T-P</b>	<b>1</b>	<b>5.4</b>	<b>6, 11, 13</b>
<b>27. Exposición general del tema y discusión</b>	<b>GG</b>	<b>T</b>	<b>1</b>	<b>6</b>	<b>2, 8</b>
<b>28. Estudio de los contenidos explicados</b>	<b>NP</b>	<b>T</b>	<b>1</b>	<b>6</b>	<b>2, 8</b>
<b>29. Exposición general del tema y discusión</b>	<b>GG</b>	<b>T</b>	<b>3</b>	<b>7</b>	<b>7, 8</b>
<b>30. Estudio de los contenidos explicados</b>	<b>NP</b>	<b>T</b>	<b>1</b>	<b>7</b>	<b>7, 8</b>
<b>31. Lectura recomendada sobre aplicaciones estadísticas de resumen automático</b>	<b>NP</b>	<b>T</b>	<b>1</b>	<b>7</b>	<b>8, 10, 11</b>
<b>32. Comentario y discusión del artículo</b>	<b>GG</b>	<b>T-P</b>	<b>1</b>	<b>7</b>	<b>8, 11, 13</b>

<b>33. Exposición general del tema y discusión</b>	<b>GG</b>	<b>T</b>	<b>3</b>	<b>8</b>	<b>7, 8</b>
<b>34. Estudio de los contenidos explicados</b>	<b>NP</b>	<b>T</b>	<b>1</b>	<b>8</b>	<b>7, 8</b>
<b>35. Lectura recomendada sobre aplicaciones de PLN de resumen automático</b>	<b>NP</b>	<b>T</b>	<b>1</b>	<b>8.3</b>	<b>8, 10, 11</b>
<b>36. Comentario y discusión del artículo</b>	<b>GG</b>	<b>T-P</b>	<b>1</b>	<b>8.3</b>	<b>8, 11, 13</b>
<b>37. Manejo de programas de indización y resumen automáticos</b>	<b>S</b>	<b>P</b>	<b>10</b>	<b>3, 4, 5, 7, 8</b>	<b>7, 8</b>
<b>*<sup>i</sup> Preparación de un trabajo sobre investigaciones recientes de indización/resumen automático</b>	<b>NP</b>	<b>T-P</b>	<b>8</b>	<b>3, 4, 5, 7, 8</b>	<b>8, 9, 10, 11, 12, 13</b>
<b>* Tutorización y evaluación de la actividad anterior</b>	<b>Tut</b>	<b>P</b>	<b>5</b>	<b>3, 4, 5, 7, 8</b>	<b>8, 9, 10, 11, 12, 13</b>
<b>38. Estudio y preparación del examen final</b>	<b>NP</b>	<b>T-P</b>	<b>20</b>	<b>1-8</b>	<b>Todos</b>
<b>39. Examen final</b>	<b>GG</b>	<b>C-E</b>	<b>2</b>	<b>1-8</b>	<b>Todos</b>

<b>Distribución del tiempo (ECTS)</b>		<i>Dedicación del alumno</i>		<i>Dedicación del profesor</i>		
<i>Distribución de actividades</i>	<i>Nº alumnos</i>	<i>H. presenc.</i>	<i>H. no presenc.</i>	<i>H. presenc.</i>	<i>H. no presenc.</i>	
GG	C-E	40	3	--	3	15
	T	40	20	12	20	7
	P	40	--	12	--	--
	T-P	40	14	--	14	5
	Subtotal	40	37	24	37	27
S	C-E	20	--	--	--	10
	T	20	--	--	--	--
	P	20	10	--	20	5
	T-P	20	--	--	--	--
	Subtotal	20	10	0	20	15
Tut.	C-E	5	--	---	--	5
	T	5	--	---	--	--
	P	5	5	---	40	8
	T-P	5	--	8	--	--
	Subtotal	5	5	8	40	13
Tutoría comp. y preparación de ex.		1	--	18-20	8	2
Totales			52 (2,075 ECTS)	52 (2,075 ECTS)	105	57

## V. Evaluación

<i>Criterios de evaluación</i>	<i>Vinculación</i>	
	<i>Objetivo</i>	<i>CC<sup>i</sup></i>
1. Explicar razonadamente los conceptos básicos y las técnicas más utilizadas en la automatización de la indización y resumen	1-6	75 %
2. Analizar de manera crítica los resultados obtenidos con estas técnicas, ventajas, inconvenientes y limitaciones	1-6	
3. Aplicar las metodologías de automatización estudiadas a casos prácticos y evaluar los rendimientos obtenidos por ellas aplicando las técnicas de diagnóstico explicadas	7, 9, 12	
4. Participar activamente en la resolución de problemas realizados en clase	7, 9, 12	25 % (N.R.)
5. Participar activamente en los comentarios/discusiones de artículos realizados en clase, analizándolos con rigor y desde el punto de vista del método científico	8, 10, 11, 13	
6. Manejar con soltura los programas informáticos de indización y resumen automático	7, 8	
* 7. Realizar un análisis crítico o una revisión sobre investigaciones recientes de indización/resumen automático (siguiendo las normas de elaboración de trabajos para revistas científicas)	8-13	10 %

<i>Actividades e instrumentos de evaluación</i>		
Seminarios y Tutorías ECTS	<ul style="list-style-type: none"> <li>• Valoración de la participación del alumno en clase mediante evaluación continua, tanto en la resolución de problemas como en la discusión de artículos científicos.</li> <li>• Será necesario haber realizado positivamente las prácticas informáticas para poder aprobar la asignatura.</li> </ul>	25 %
	<ul style="list-style-type: none"> <li>• La realización del trabajo voluntario, siempre que éste se realice satisfactoriamente, aumentará en un punto la nota final</li> </ul>	10 %
Examen final	<ul style="list-style-type: none"> <li>• La evaluación final consistirá en una prueba escrita que contendrá preguntas teóricas y varios supuestos prácticos similares a los resueltos en clase</li> </ul>	75 %