



INSTITUTO POLITECNICO NACIONAL
COORDINACION GENERAL DE POSGRADO E INVESTIGACION
DIVISION DE ESTUDIOS DE POSGRADO

FORMATO GUIA PARA REGISTRO DE ASIGNATURAS

Hoja 1 de 3

I. DATOS DEL PROGRAMA Y LA ASIGNATURA

1.1 NOMBRE DEL PROGRAMA: Maestría en Informática

1.2 COORDINADOR DEL PROGRAMA: Dr. ERIC MANUEL ROSALES PEÑA ALFARO

1.3 NOMBRE DE LA ASIGNATURA: Tecnologías de Búsqueda de Información No-estructurada

1.4 CLAVE: _____ (Para ser llenado por la CGPI)

1.5 TIPO DE ASIGNATURA:

	OBLIGATORIA	<input type="checkbox"/>	OPTATIVA	<input checked="" type="checkbox"/>
	SEMINARIO	<input type="checkbox"/>	ESTANCIA	<input type="checkbox"/>

1.6 NUMERO DE HORAS: 54

	TEORIA	<input checked="" type="checkbox"/>	PRACTICA	<input type="checkbox"/>	T-P	<input type="checkbox"/>
--	--------	-------------------------------------	----------	--------------------------	-----	--------------------------

1.7 UNIDADES DE CREDITO: 6

1.8 FECHA DE LA ELABORACION DEL PROGRAMA DE LA ASIGNATURA:

	16	10	2000
	d	m	a

1.9 SESION DEL COLEGIO DE PROFESORES EN QUE SE ACORDO LA IMPLANTACION DE LA ASIGNATURA:

	SESION No.	12	FECHA:	22	11	2000
				d	m	a

1.10 FECHA DE REGISTRO EN CGPI: (Para ser llenado por la CGPI)

d m A

II. DATOS DEL PERSONAL ACADEMICO

2.1 PROFESOR TITULAR: Mauricio Jorge Procel Moreno CLAVE: 6939-EE-10

2.2 PROFESORES ADJUNTOS: _____ CLAVE: _____

_____ CLAVE: _____

III. DESCRIPCION DEL CONTENIDO DEL PROGRAMA DE LA ASIGNATURA

III.1 OBJETIVO GENERAL:

Al concluir el curso, experimenta máquinas de búsqueda sobre bases de datos textuales con base en los principios, elementos y herramientas de los sistemas de recuperación de información ("Information Retrieval Systems –IRS-)

III.2 DESCRIPCION DEL CONTENIDO

TEMAS Y SUBTEMAS	TIEMPO
1. Conceptos Iniciales.	9:30 hs
1.1 Elementos del problema y de las soluciones propuestas.	
1.2 Antecedentes históricos.	
1.3 Tipos de búsquedas.	
1.4 Definiciones de elementos usualmente empleados en la búsqueda: Colecciones, documentos, ventanas, frases, palabras, lemas, etc.	
1.5 La ley de Zipf, la frecuencia de ocurrencia de letras, palabras, etc.	
2. Operaciones de manipulación de textos.	6:30 hs
2.1 Técnicas de compresión de textos.	
2.2 Técnicas de procesamiento de texto previo a su clasificación.	
2.3 "Inverted Files" y estructuras de datos eficientes en Information Retrieval.	
3. Modelos y estrategias de Index y de Búsqueda.	9:30 hs
3.1 Medidas de Similitud.	
3.1 El modelo Booleano.	
3.2 Modelos vectoriales (el Vector Space Model).	
3.4 El modelo probabilístico.	
3.5 El modelo basado en redes inferenciales.	
4. Medidas de evaluación de la respuesta de modelos de búsqueda.	6:30 hs
4.1 Cobertura (recall) y precisión (precision).	
4.2 Medidas integradas.	
4.3 Otras medidas.	
5. Realimentación (relevance feedback).	9:30 hs
5.1 Realimentación por parte del usuario:	
5.2 Análisis global y la independencia de los términos	
5.3 Pseudo Relevance Feedback.	
5.4 Análisis local sin ayuda del usuario.	

5.5 Perfiles de usuario y otros elementos que se pueden utilizar para mejorar el desempeño de un modelo de búsqueda.	
6. Estándares y búsqueda en la WEB	9:30 hs
6.1 Hipertexto: SGML, XML, HTML.	
6.2 Información básica sobre la WEB.	
6.3 Arquitectura de máquinas de búsqueda en la WEB.	
6.4 Browsing (navegación, exploración).	
6.5 Metabuscaadores.	

III.3 BIBLIOGRAFIA UTILIZADA EN LA ASIGNATURA

Journal of Information Storage and Retrieval, Elsevier, 2010. url:
<http://www.sciencedirect.com/science/journal/00200271>

Baeza-Yates, R. A., and Ribeiro-Neto, B. (2010). *Modern Information Retrieval* (2nd ed.).
 Reading, Massachusetts: Addison-Wesley. url:
<http://people.ischool.berkeley.edu/~hearst/irbook/>

Van Rijsbergen, J.V. *Information Retrieval*. url: <http://www.dcs.gla.ac.uk/Keith/Preface.html>

III.4 PROCEDIMIENTOS O INSTRUMENTOS DE EVALUACION A UTILIZAR

- Trabajos
- Exposición por parte de los alumnos
- Exámenes