



Unidad de Enseñanza Aprendizaje
E Impartición de Cátedra

Recuperación de Información

Reforma Curricular UAT 2023

Dependencia Académica: Facultad de Ingeniería y Ciencias

Programa Académico: Maestría en Ciencias e Ingeniería de Datos

Tipo: Único

Unidad de Enseñanza Aprendizaje e Impartición de Cátedra

PERFIL DOCENTE DE LA UNIDAD DE ENSEÑANZA-APRENDIZAJE		
FORMACIÓN PROFESIONAL	COMPETENCIAS PEDAGÓGICAS	EXPERIENCIA LABORAL
Posgrado en área afín a la ciencia de datos	Diseña, organiza y propone actividades didácticas y problemas para su solución por parte de los educandos Conoce y aplica distintas estrategias y formas de evaluación sobre el proceso educativo Establece un clima de relación en el grupo que favorece actitudes de confianza, respeto, disciplina y creatividad por el estudio Conoce los materiales de enseñanza y los recursos didácticos disponibles	Mínimo 1 año de experiencia en impartición de cátedra a nivel posgrado en el desarrollo de habilidades para aprender

UNIDAD DE ENSEÑANZA-APRENDIZAJE						
Datos Generales						
ASIGNATURA	Recuperación de Información				TIPO DE ASIGNATURA	PERIODO ESCOLAR
					Disciplinar (optativa)	2° y 3°
CLAVE	HTC	HTI	TH	TC	ASIGNATURA ANTECEDENTE	ASIGNATURA CONSECUENTE
MCID29	4 (56)	3(42)	7 (98)	6	Ninguna	ninguna
CONTRIBUCIÓN DE LA ASIGNATURA EN EL PERFIL DE EGRESO INSTITUCIONAL		El estudiante desarrollará capacidades de análisis para la solución de problemas relacionados con la recuperación de información digital presentada en distintos formatos				
CONTRIBUCIÓN DE LA ASIGNATURA EN EL PERFIL DE EGRESO DEL PROGRAMA ACADÉMICO		Capacidad para trabajar en grupo, capacidad para proponer y desarrollar soluciones a través de un programa de cómputo				
OBJETIVO GENERAL DE LA UEA		Al término del curso el alumno conocerá y utilizará herramientas de cómputo para hacer frente a los nuevos problemas relacionados con la recuperación de información				

ATRIBUTOS DE LA UEA			
SABER	SABER HACER	SABER SER	SABER CONVIVIR
Conocimientos Conoce los conceptos y herramientas del área de recuperación de información. Analiza problemas y reconocer las diferentes aplicaciones de la recuperación de información.	Habilidades y aptitudes Desarrolla sistemas capaces de resolver problemas del mundo real haciendo uso de herramientas de cómputo para la Recuperación de la Información digital.	Actitudes y valores Crea e integra soluciones basadas en las distintas técnicas para la Recuperación de la Información	Interrelación pacífica, justa y respetuosa Practicar la comunicación efectiva y respetuosa con los integrantes del grupo

CRITERIOS DE EVALUACIÓN	VALOR EN PORCENTAJE
PARTICIPACIÓN INDIVIDUAL	25%
EVALUACIONES PARCIALES	25%
ACTIVIDADES REALIZADAS	25%
PRODUCTO INTEGRADOR	25%
TOTAL	100%

Nota: los criterios de evaluación que se enlistan son probables, el profesor decidirá cuáles incluir, debiendo considerar al menos 4 y los porcentajes para cada uno de ellos dando una suma total de 100%

PRODUCTO INTEGRADOR DE LA UEA		
DESCRIPCIÓN		Evaluación de contenidos, evaluación de prácticas y proyecto final donde haga uso de técnicas de Recuperación de Información
NIVELES DE DESEMPEÑO		CRITERIOS DE DESEMPEÑO
10	Excelente	La aplicación cumple con el 100 % de requerimientos solicitados y presentación del proyecto ante el grupo.
9	Bueno	Asistencia al 90%, cumple con el 90% de la ponderación de los instrumentos de evaluación.
8	Regular	Asistencia al 80%, cumple con el 80% de la ponderación de los instrumentos de evaluación.
7	Básico	No alcanza el mínimo aprobatorio.
6	Elemental	No alcanza el mínimo aprobatorio.
NA	Aún no competente	No alcanza el mínimo aprobatorio.

UNIDAD O BLOQUE TEMÁTICO		
NÚMERO Y NOMBRE	Unidad 1. Introducción a la recuperación de información	
OBJETIVO DE LA UNIDAD O BLOQUE TEMÁTICO	El objetivo de esta unidad es que el alumno se familiarice con los conceptos generales de la recuperación de información: sistema de recuperación de información, motor de búsqueda, crawler, relevancia, entre otros.	
TIEMPO/DURACIÓN	Horas: 8 Semanas: 2	
RECURSOS EDUCATIVOS:	Apuntes de clase	
DESGLOSE DE CONTENIDOS ESPECÍFICOS		
1.1 Conceptos básicos de RI, búsqueda, recuperación y relevancia 1.2 Problemas de la RI 1.3 Arquitectura de un sistema RI 1.4 Interfaces de búsqueda 1.5 Vocabulario y diccionarios		

IMPARTICIÓN DE CÁTEDRA

Semana No. 1-2

NÚMERO Y NOMBRE DE LA UNIDAD O BLOQUE TEMÁTICO:		Unidad 1. Introducción a la recuperación de información		
OBJETIVO DE LA UNIDAD O BLOQUE TEMÁTICO:		El objetivo de esta unidad es que el alumno se familiarice con los conceptos generales de la recuperación de información: sistema de recuperación de información, motor de búsqueda, crawler, relevancia, entre otros.		
CONTENIDO TEMÁTICO	MOMENTOS DE LA CÁTEDRA			
	Apertura	Desarrollo		Cierre
1.1 Conceptos básicos de RI, búsqueda, recuperación y relevancia 1.2 Problemas de la RI 1.3 Arquitectura de un sistema RI 1.4 Interfaces de búsqueda 1.5 Vocabulario y diccionarios	Presentación y repaso de conceptos generales de los componentes de un sistema de recuperación de información	Presentación y desarrollo de ideas relacionadas a escenarios reales donde se requieran técnicas de recuperación de información		Retroalimentación del tema visto en clase mediante cuestionarios o exposición de ideas sobre la recuperación de información

UNIDAD O BLOQUE TEMÁTICO

NÚMERO Y NOMBRE	<i>Unidad 2. Modelos de Recuperación de Información</i>
OBJETIVO DE LA UNIDAD O BLOQUE TEMÁTICO	<i>El objetivo de esta unidad es conocer los esquemas de representación de información que los sistemas de recuperación manejan para realizar consultas en lenguaje natural y obtener los documentos adecuados a partir de dichas consultas.</i>
TIEMPO/DURACIÓN	Horas: 8 Semanas: 2
RECURSOS EDUCATIVOS:	Documentos de consulta y reportes de practicas
DESGLOSE DE CONTENIDOS ESPECÍFICOS	

2. Modelos de Recuperación de Información
 2.1 Taxonomía de los modelos de RI
 2.2 Modelo de recuperación booleano
 2.3 Modelo espacio vector
 2.4 Modelo Esquema de pesado de términos
 2.5 Ponderación TF-IDF
 2.6 Modelo probabilístico

IMPARTICIÓN DE CÁTEDRA

Semana No. 3-4

NÚMERO Y NOMBRE DE LA UNIDAD O BLOQUE TEMÁTICO:		Unidad 2. Modelos de Recuperación de Información		
OBJETIVO DE LA UNIDAD O BLOQUE TEMÁTICO:		El objetivo de esta unidad es conocer los esquemas de representación de información que los sistemas de recuperación manejan para realizar consultas en lenguaje natural y obtener los documentos adecuados a partir de dichas consultas.		
CONTENIDO TEMÁTICO	MOMENTOS DE LA CÁTEDRA			
	Apertura	Desarrollo		Cierre
2. Modelos de Recuperación de Información 2.1 Taxonomía de los modelos de RI 2.2 Modelo de recuperación booleano 2.3 Modelo espacio vector 2.4 Modelo Esquema de pesado de términos 2.5 Ponderación TF-IDF 2.6 Modelo probabilístico	Presentación y repaso de las características de los modelos de representación de la información	Estudio de casos donde se requieran los distintos modelos de representación de información Exposición de las características de los modelos		Retroalimentación del tema visto en clase mediante cuestionarios o exposición de ideas sobre los modelos de recuperación de información Auto reflexión de lo que se aprendió en clase Lluvia de ideas de los modelos de representación vistos

UNIDAD O BLOQUE TEMÁTICO

NÚMERO Y NOMBRE	Unidad 3. Retroalimentación y expansión de consulta
OBJETIVO DE LA UNIDAD O BLOQUE TEMÁTICO	El alumno conocerá mecanismos para mejorar la representación de consultas basado en modelos de expansión, retroalimentación y/o reformulación de la consulta para expresa la necesidad de información del usuario.
TIEMPO/DURACIÓN	Horas: 12

	Semanas: 3
RECURSOS EDUCATIVOS:	Documentos de consulta, reportes de practicas
DESGLOSE DE CONTENIDOS ESPECÍFICOS	
3.1 Retroalimentación de la consulta 3.1.1 Expansión de la consulta con términos extraídos de un tesoro 3.1.2 Construcción automática de un tesoro 3.2 Retroalimentación por relevancia 3.2.1 Método Rocchio 3.3 Retroalimentación probabilística 3.4 Retroalimentación a través de análisis local	

IMPARTICIÓN DE CÁTEDRA			
Semana No. 5-7			
NÚMERO Y NOMBRE DE LA UNIDAD O BLOQUE TEMÁTICO:	<i>Unidad 3. Retroalimentación y expansión de consulta</i>		
OBJETIVO DE LA UNIDAD O BLOQUE TEMÁTICO:	<i>El alumno conocerá mecanismos para mejorar la representación de consultas basado en modelos de expansión, retroalimentación y/o reformulación de la consulta para expresa la necesidad de información del usuario.</i>		
CONTENIDO TEMÁTICO	MOMENTOS DE LA CÁTEDRA		
	Apertura	Desarrollo	Cierre
3.1 Retroalimentación de la consulta 3.1.1 Expansión de la consulta con términos extraídos de un tesoro 3.1.2 Construcción automática de un tesoro 3.2 Retroalimentación por relevancia 3.2.1 Método Rocchio 3.3 Retroalimentación probabilística 3.4 Retroalimentación a través de análisis local	Presentación y repaso de los mecanismos de expansión y retroalimentación de consultas a partir de elementos de retroalimentación	Estudio de casos del uso de la retroalimentación y reformulación de consultas Exposición de los mecanismos de retroalimentación	Retroalimentación del tema visto en clase mediante cuestionarios o exposición de ideas Auto reflexión de lo que se aprendió en clase Lluvia de ideas

UNIDAD O BLOQUE TEMÁTICO

NÚMERO Y NOMBRE	<i>Unidad 4. Características de los documentos</i>
OBJETIVO DE LA UNIDAD O BLOQUE TEMÁTICO	<i>El alumno conocerá los diferentes tipos de documentos con los que puede trabajar un sistema de recuperación de información.</i>
TIEMPO/DURACIÓN	Horas: 8 Semanas: 2
RECURSOS EDUCATIVOS:	Documentos de consulta, reportes de practicas
DESGLOSE DE CONTENIDOS ESPECÍFICOS	
4.1 Metadatos 4.2 Descriptores 4.3 Indización 4.4 Formatos de los documentos 4.5 Pre-procesamiento de los documentos 4.6 Propiedades del texto 4.7 Vocabularios controlados 4.8 Organización de documentos	

IMPARTICIÓN DE CÁTEDRA			
Semana No. 8-9			
NÚMERO Y NOMBRE DE LA UNIDAD O BLOQUE TEMÁTICO:	<i>Unidad 4. Características de los documentos</i>		
OBJETIVO DE LA UNIDAD O BLOQUE TEMÁTICO:	<i>El alumno conocerá los diferentes tipos de documentos con los que puede trabajar un sistema de recuperación de información.</i>		
CONTENIDO TEMÁTICO	MOMENTOS DE LA CÁTEDRA		
	Apertura	Desarrollo	Cierre
4.1 Metadatos 4.2 Formatos de los documentos 4.3 Pre-procesamiento de los documentos 4.4 Propiedades del texto 4.5 Organización de documentos	Presentación y repaso de los distintos formatos de los documentos y su preprocesamiento	Estudio de casos del uso Exposición de los tipos de formatos y propiedades	Retroalimentación del tema visto en clase mediante cuestionarios o exposición de ideas Auto reflexión de lo que se aprendió en clase Lluvia de ideas

UNIDAD O BLOQUE TEMÁTICO	
NÚMERO Y NOMBRE	Unidad 5. Indexamiento y búsqueda
OBJETIVO DE LA UNIDAD O BLOQUE TEMÁTICO	El alumno obtendrá conocimientos relacionados con la puntuación de términos y modelos que permitan representar los índices para obtener mejor desempeño de los sistemas de recuperación de información.
TIEMPO/DURACIÓN	Horas: 12 Semanas: 3
RECURSOS EDUCATIVOS:	Documentos de consulta, reportes de practicas
DESGLOSE DE CONTENIDOS ESPECÍFICOS	
5.1 Índices invertidos 5.2 Estructuras de archivos indizados 5.3 Árboles B 5.4 Búsqueda secuencial 5.5 Indexamiento multidimensional 5.6 Estructura de la web 5.7 Motores de búsqueda 5.7.1 Ranking de los motores de búsqueda 5.8 Manejo de información web 5.9 Crawlers web 5.10 Taxonomía de los crawlers	

IMPARTICIÓN DE CÁTEDRA

Semana No. 10-12

NÚMERO Y NOMBRE DE LA UNIDAD O BLOQUE TEMÁTICO:	<i>Unidad 5. Indexamiento y búsqueda</i>		
OBJETIVO DE LA UNIDAD O BLOQUE TEMÁTICO:	<i>El alumno obtendrá conocimientos relacionados con la puntuación de términos y modelos que permitan representar los índices para obtener mejor desempeño de los sistemas de recuperación de información.</i>		
CONTENIDO TEMÁTICO	MOMENTOS DE LA CÁTEDRA		
	Apertura	Desarrollo	Cierre
5.1 Índices invertidos 5.2 Estructuras de archivos indizados 5.3 Árboles B 5.4 Búsqueda secuencial 5.5 Indexamiento multidimensional 5.6 Estructura de la web 5.7 Motores de búsqueda 5.7.1 Ranking de los motores de búsqueda 5.8 Manejo de información web 5.9 Crawlers web 5.10 Taxonomía de los crawlers	Presentación y repaso de los algoritmos de indexamiento y búsqueda	Estudio de casos del uso Exposición de las características de los distintos algoritmos de indexamiento y búsqueda Análisis de ventajas y desventajas de los algoritmos	Retroalimentación del tema visto en clase mediante cuestionarios o exposición de ideas Auto reflexión de lo que se aprendió en clase Lluvia de ideas

UNIDAD O BLOQUE TEMÁTICO

NÚMERO Y NOMBRE	<i>Unidad 6. Evaluación de los modelos de RI</i>
OBJETIVO DE LA UNIDAD O BLOQUE TEMÁTICO	<i>El alumno conocerá e implementará los mecanismos de evaluación para un sistema de recuperación de información.</i>
TIEMPO/DURACIÓN	Horas: 8 Semanas: 2
RECURSOS EDUCATIVOS:	Documentos de consulta, reportes de practicas
DESGLOSE DE CONTENIDOS ESPECÍFICOS	

6.1 Métricas de evaluación
6.2 Relevancia
6.3 Colecciones públicas de referencia

IMPARTICIÓN DE CÁTEDRA

Semana No. 13-14

NÚMERO Y NOMBRE DE LA UNIDAD O BLOQUE TEMÁTICO:	Unidad 6. Evaluación de los modelos de RI
OBJETIVO DE LA UNIDAD O BLOQUE TEMÁTICO:	El alumno conocerá e implementará los mecanismos de evaluación para un sistema de recuperación de información.

CONTENIDO TEMÁTICO	MOMENTOS DE LA CÁTEDRA		
	Apertura	Desarrollo	Cierre
6.1 Métricas de evaluación 6.2 Relevancia 6.3 Colecciones públicas de referencia	Presentación y repaso de las métricas de evaluación de desempeño	Estudio de casos del uso Exposición de las métricas de evaluación del desempeño Búsqueda y estudio de las principales colecciones de documentos utilizadas por los sistemas de recuperación de información	Retroalimentación del tema visto en clase mediante cuestionarios o exposición de ideas Auto reflexión de lo que se aprendió en clase Lluvia de ideas

REFERENCIAS BIBLIOGRÁFICA Y ELECTRÓNICA (APA)

Básica	T. Roelleke, <i>Information Retrieval Models: Foundations and Relationships</i>, 1era. Ed. Morgan & Claypool Publishers, 2013, pp. 164. Christopher D. Manning, Prabhakar Raghavan and Hinrich Schütze, <i>Introduction to Information Retrieval</i>, Cambridge University Press. 2008.
--------	---

	<i>Dr. L.M. Palanivelu, N. Jayathi, S. Sadesh. Information Retrieval. Charulatha Publications, 2013.</i>
Complementaria	<i>R. Baeza-Yates, B. Ribeiro-Neto, Modern Information Retrieval: The Concepts and Technology behind Search, 2da. Ed. Addison-Wesley Professional, 2011, pp. 944</i> <i>Grossman D. Information Retrieval, Algorithms and Heuristics. 2a Edición, Springer, 2014</i>

Nota: las referencias bibliográficas (básicas y complementarias) se colocan solo al final de la UEA e IC. Se sugiere colocarlas de acuerdo con los criterios de la APA.

CRÉDITOS	
Elaborado por:	<i>Dra. Heidy Marisol Marin Castro</i>
Fecha:	<i>11-04-2022</i>
Actualizado por:	
Fecha:	
Autorizado por:	
Fecha:	