



National University of Engineering (UNI)

School of Artificial Intelligence

Syllabus 2024-I

1. COURSE

CS272. Databases II (Mandatory)

2. GENERAL INFORMATION

2.1 Course	:	CS272. Databases II
2.2 Semester	:	5 th Semester.
2.3 Credits	:	3
2.4 Horas	:	1 HT; 4 HP;
2.5 Duration of the period	:	16 weeks
2.6 Type of course	:	Mandatory
2.7 Learning modality	:	Face to face
2.8 Prerequisites	:	CS271. Data Management. (4 th Sem)

3. PROFESSORS

Meetings after coordination with the professor

4. INTRODUCTION TO THE COURSE

Information Management (IM) plays a leading role in almost every area where computers are used. This area includes the capture, digitization, representation, organization, transformation and presentation of information; Algorithms to improve the efficiency and effectiveness of access and update of stored information, data modeling and abstraction, and physical file storage techniques.

It also covers information security, privacy, integrity and protection in a shared environment. Students need to be able to develop conceptual and physical data models, determine which IM methods and techniques are appropriate for a given problem, and be able to select and implement an appropriate IM solution that reflects all applicable constraints, including scalability and Usability.

5. GOALS

- To make the student understand the different applications that the databases have, in the different areas of knowledge.
- Show appropriate ways of storing information based on their various approaches and their subsequent retrieval of information.

6. COMPETENCES

- 1) Analizar un problema computacional complejo y aplicar los principios computacionales y otras disciplinas relevantes para identificar soluciones. (**Evaluar**)
- 4) Reconocer las responsabilidades profesionales y hacer juicios informados en el campo profesional de computación con principios éticos. (**Evaluar**)
- 6) Aplicar la teoría de la computación y los fundamentos del desarrollo de software para producir soluciones basadas en computación. (**Evaluar**)
- 7) Desarrollar tecnología computacional buscando el bien común, aportando con formación humana, capacidades científicas, tecnológicas y profesionales para solucionar problemas sociales de nuestro entorno. (**Evaluar**)

7. TOPICS

Unit 1: Diseño Físico de Bases de Datos (10 hours)	
Competences Expected:	
Topics	Learning Outcomes
<ul style="list-style-type: none"> • Almacenamiento y estructura de archivos. • Archivos indexados. • Archivos Hash. • Archivos de Firma. • Árboles B. • Archivos con índice denso. • Archivos con registros de tamaño variable. • Eficiencia y Afinación de Bases de Datos. 	<ul style="list-style-type: none"> • Explica los conceptos de registro, tipos de registro, y archivos, así como las diversas técnicas para colocar registros de archivos en un disco [Usar] • Da ejemplos de la aplicación de índices primario, secundario y de agrupamiento [Usar] • Distingue entre un índice no denso y uno denso [Usar] • Implementa índices de multinivel dinámicos usando árboles-B [Usar] • Explica la teoría y la aplicación de técnicas de hash internas y externas [Usar] • Usa técnicas de hasp para facilitar la expansión de archivos dinámicos [Usar] • Describe las relaciones entre hashing, compresión, y búsquedas eficientes en bases de datos [Usar] • Evalúa el costo y beneficio de diversos esquemas de hashing [Usar] • Explica como el diseño físico de una base de datos afecta la eficiencia de las transacciones en ésta [Usar]
Readings : [Bur04], [Cel05]	

Unit 2: Procesamiento de Transacciones (12 hours)	
Competences Expected:	
Topics	Learning Outcomes
<ul style="list-style-type: none"> • Transacciones. • Fallo y recuperación. • Control concurrente. • Interacción de gestión de transacciones con el almacenamiento, especialmente en almacenamiento. 	<ul style="list-style-type: none"> • Crear una transacción mediante la incorporación de SQL en un programa de aplicación [Usar] • Explicar el concepto de confirmaciones implícitas [Usar] • Describir los problemas específicos para la ejecución de una transacción eficiente [Usar] • Explicar cuando y porqué se necesita un <i>rollback</i>, y cómo registrar todo asegura un <i>rollback</i> adecuado [Usar] • Explicar el efecto de diferentes niveles de aislamiento sobre los mecanismos de control de concurrencia [Usar] • Elejir el nivel de aislamiento adecuado para la aplicación de un protocolo de transacción especificado [Usar] • Identificar los límites apropiados de la transacción en programas de aplicación [Usar]
Readings : [Phi97], [Ram04]	

Unit 3: Almacenamiento y Recuperación de Información (10 hours)	
Competences Expected:	
Topics	Learning Outcomes
<ul style="list-style-type: none"> • Documentos, publicación electrónica, markup, y lenguajes markup. • Tries, archivos invertidos, Árboles PAT, archivos de firma, indexación. • Análisis Morfológico, stemming, frases, stop lists. • Distribuciones de frecuencia de términos, incertidumbre, fuzificación (fuzzyness), ponderación. • Espacio vectorial, probabilidad, lógica, y modelos avanzados. • Necesidad de Información , Relevancia, evaluación, efectividad. • Thesauri, ontologías, clasificación y categorización, metadata. • Información bibliográfica, bibliometría, citasiones. • Enrutamiento y filtrado. • Búsqueda multimedia. • Información de resumen y visualización. • Búsqueda por facetas (por ejemplo, el uso de citas, palabras clave, esquemas de clasificación). • Librerías digitales. • Digitalización, almacenamiento, intercambio, objetos digitales, composición y paquetes. • Metadata y catalogación. • Nombramiento, repositorios, archivos • Archivamiento y preservación, integrdad • Espacios (Conceptual, geográfico, 2/3D, Realidad virtual) • Arquitecturas (agentes, autobuses, envolturas / mediadores), de interoperabilidad. • Servicios (búsqueda, de unión, de navegación, y así sucesivamente). • Gestión de derechos de propiedad intelectual, la privacidad y la protección (marcas de agua). 	<ul style="list-style-type: none"> • Explica los conceptos básicos de almacenamiento y recuperación de la información [Usar] • Describe que temas son específicos para una recuperación de la información eficiente [Usar] • Da aplicaciones de estrategias alternativas de búsqueda y explica porqué una estrategia en particular es apropiada para una aplicación [Usar] • Diseña e implementa un sistema de almacenamiento y recuperación de la información o librería digital de tamaño pequeño a mediano [Usar] • Describe algunas de las soluciones técnicas a los problemas relacionados al archivamiento y preservación de la información en una librería digital [Usar]
Readings : [Pet98], [Ram04]	

Unit 4: Bases de Datos Distribuidas (36 hours)	
Competences Expected:	
Topics	Learning Outcomes
<ul style="list-style-type: none"> • DBMS Distribuidas <ul style="list-style-type: none"> – Almacenamiento de datos distribuido – Procesamiento de consultas distribuido – Modelo de transacciones distribuidas – Soluciones homogéneas y heterogéneas – Bases de datos distribuidas cliente-servidor • Parallel DBMS <ul style="list-style-type: none"> – Arquitecturas paralelas DBMS: memoria compartida, disco compartido, nada compartido; – Aceleración y ampliación, por ejemplo, el uso del modelo de procesamiento MapReduce – Replicación de información y modelos de consistencia débil 	<ul style="list-style-type: none"> • Explicar las técnicas usadas para la fragmentación de datos, replicación, y la asignación durante el proceso de diseño de base de datos distribuida [Usar] • Evaluar estrategias simples para la ejecución de una consulta distribuida para seleccionar una estrategia que minimice la cantidad de transferencia de datos [Usar] • Explicar como el protocolo de dos fases de <i>commit</i> es usado para resolver problemas de transacciones que acceden a bases de datos almacenadas en múltiples nodos [Usar] • Describir el control concurrente distribuido basados en técnicas de copia distinguibles y el método de votación. [Usar] • Describir los tres niveles del software en el modelo cliente servidor [Usar]
Readings : [M T99]	

8. WORKPLAN

8.1 Methodology

Individual and team participation is encouraged to present their ideas, motivating them with additional points in the different stages of the course evaluation.

8.2 Theory Sessions

The theory sessions are held in master classes with activities including active learning and roleplay to allow students to internalize the concepts.

8.3 Practical Sessions

The practical sessions are held in class where a series of exercises and/or practical concepts are developed through problem solving, problem solving, specific exercises and/or in application contexts.

9. EVALUATION SYSTEM

***** EVALUATION MISSING *****

10. BASIC BIBLIOGRAPHY

- [Bur04] Donald K. Bursleson. *Physical Database Design Using Oracle*. CRC Press, 2004.
- [Cel05] Joe Celko. *Joe Celko's SQL Programming Style*. Elsevier, 2005.
- [M T99] Patrick Valduriez M. Tamer Ozsu. *Principles of Distributed Database Systems, Second Edition*. Prentice Hall, 1999.
- [Pet98] Julita Vassileva Peter Brusilovsky Alfred Kobsa. *Adaptive Hypertext and Hypermedia, First Edition*. Springer, 1998.
- [Phi97] Eric Newcomer Philip A. Bernstein. *Principles of Transaction Processing, First Edition*. Morgan Kaufmann, 1997.
- [Ram04] Shamkant B. Navathe Ramez Elmasri. *Fundamentals of Database Systems, Fourth Edition*. Addison Wesley, 2004.