

Universidad Católica San Pablo
Escuela Profesional de
Ciencia de la Computación
SILABO



CS272. Bases de Datos II (Obligatorio)

1. DATOS GENERALES

1.1 CARRERA PROFESIONAL	:	Ciencia de la Computación
1.2 ASIGNATURA	:	CS272. Bases de Datos II
1.3 SEMESTRE ACADÉMICO	:	5 ^{to} Semestre.
1.4 PREREQUISITO(S)	:	CS271. Bases de Datos I. (4 ^{to} Sem)
1.5 CARÁCTER	:	Obligatorio
1.6 HORAS	:	1 HT; 2 HP; 2 HL;
1.7 CRÉDITOS	:	3

2. DOCENTE

Dr. Javier Leandro Tejada Cárcamo

- Dr. Ciencia de la Computación, IPN, México, 2009.

3. FUNDAMENTACIÓN DEL CURSO

La Gestión de la Información (*IM-Information Management*) juega un rol principal en casi todas las áreas donde los computadores son usados. Esta área incluye la captura, digitalización, representación, organización, transformación y presentación de información; algoritmos para mejorar la eficiencia y efectividad del acceso y actualización de información almacenada, modelamiento de datos y abstracción, y técnicas de almacenamiento de archivos físicos.

Este también abarca la seguridad de la información, privacidad, integridad y protección en un ambiente compartido. Los estudiantes necesitan ser capaces de desarrollar modelos de datos conceptuales y físicos, determinar que métodos de IM y técnicas son apropiados para un problema dado, y ser capaces de seleccionar e implementar una apropiada solución de IM que refleje todas las restricciones aplicables, incluyendo escalabilidad y usabilidad.

4. SUMILLA

1. Diseño Físico de Bases de Datos
2. Procesamiento de Transacciones
3. Almacenamiento y Recuperación de Información
4. Bases de Datos Distribuidas

5. OBJETIVO GENERAL

- Hacer que el alumno entienda las diferentes aplicaciones que tienen las bases de datos, en las diversas áreas de conocimiento.
- Mostrar las formas adecuadas de almacenamiento de información basada en sus diversos enfoques y su posterior recuperación de información.

6. CONTRIBUCIÓN A LA FORMACIÓN PROFESIONAL Y FORMACIÓN GENERAL

Esta disciplina contribuye al logro de los siguientes resultados de la carrera:

- b) Analizar problemas e identificar y definir los requerimientos computacionales apropiados para su solución. ()
- e) Entender correctamente las implicancias profesionales, éticas, legales, de seguridad y sociales de la profesión. (**Evaluar**)
- i) Utilizar técnicas y herramientas actuales necesarias para la práctica de la computación. (**Evaluar**)
- j) Aplicar la base matemática, principios de algoritmos y la teoría de la Ciencia de la Computación en el modelamiento y diseño de sistemas computacionales de tal manera que demuestre comprensión de los puntos de equilibrio involucrados en la opción escogida. (**Evaluar**)

7. COMPETENCIAS ESPECÍFICAS DE COMPUTACIÓN

Esta disciplina contribuye a la formación de las siguientes competencias del área de computación (IEEE):

- C1.** La comprensión intelectual y la capacidad de aplicar las bases matemáticas y la teoría de la informática (computer science).⇒ **Outcome b**
- C7.** Ser capaz de aplicar los principios y tecnologías de ingeniería de software para asegurar que las implementaciones de software son robustos, fiables y apropiados para su público objetivo.⇒ **Outcome e**
- CS4.** Implementar la teoría apropiada, prácticas y herramientas para la especificación, diseño, implementación y mantenimiento, así como la evaluación de los sistemas basados en computadoras.⇒ **Outcome i**
- CS6.** Evaluar los sistemas en términos de atributos de calidad en general y las posibles ventajas y desventajas que se presentan en el problema dado.⇒ **Outcome j**

8. CONTENIDOS

UNIDAD 1: Diseño Físico de Bases de Datos(10)	
Competencias: C1	
CONTENIDO	OBJETIVO GENERAL
<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]
Lecturas: [Burlson, 2004], [Date, 2005], [Celko, 2005]	

UNIDAD 2: Procesamiento de Transacciones(12)	
Competencias: C1	
CONTENIDO	OBJETIVO GENERAL
<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]
Lecturas: [Philip A. Bernstein, 1997], [Ramez Elmasri, 2004]	

UNIDAD 3: Almacenamiento y Recuperación de Información(10)	
Competencias: C1	
CONTENIDO	OBJETIVO GENERAL
<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]
Lecturas: [Peter Brusilovsky, 1998], [Ramez Elmasri, 2004]	

UNIDAD 4: Bases de Datos Distribuidas(36)	
Competencias: C1	
CONTENIDO	OBJETIVO GENERAL
<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 distinguidos y el método de votación. [Usar] ▪ Describir los tres niveles del software en el modelo cliente servidor[Usar]
Lecturas: [M. Tamer Ozsü, 1999], [Date, 2005]	

9. METODOLOGÍA
<p>El profesor del curso presentará clases teóricas de los temas señalados en el programa propiciando la intervención de los alumnos.</p> <p>El profesor del curso presentará demostraciones para fundamentar clases teóricas.</p> <p>El profesor y los alumnos realizarán prácticas.</p> <p>Los alumnos deberán asistir a clase habiendo leído lo que el profesor va a presentar. De esta manera se facilitará la comprensión y los estudiantes estarán en mejores condiciones de hacer consultas en clase.</p>

10. EVALUACIONES
<p>Evaluación Permanente 1 : 20 %</p> <p>Examen Parcial : 30 %</p> <p>Evaluación Permanente 2 : 20 %</p> <p>Examen Final : 30 %</p> <p style="padding-left: 40px;">Trabajo Final : 50 %</p> <p style="padding-left: 40px;">Examen Final : 50 %</p>

Referencias

- [Burlson, 2004] Burlson, D. K. (2004). *Physical Database Design Using Oracle*. CRC Press.
- [Celko, 2005] Celko, J. (2005). *Joe Celko's SQL Programming Style*. Elsevier.

- [Date, 2005] Date, C. (2005). *Data Mining: Practical Machine Learning Tools and Techniques, Second Edition*. Elsevier.
- [M. Tamer Ozsü, 1999] M. Tamer Ozsü, P. V. (1999). *Principles of Distributed Database Systems, Second Edition*. Prentice Hall.
- [Peter Brusilovsky, 1998] Peter Brusilovsky, Alfred Kobsa, J. V. (1998). *Adaptive Hypertext and Hypermedia, First Edition*. Springer.
- [Philip A. Bernstein, 1997] Philip A. Bernstein, E. N. (1997). *Principles of Transaction Processing, First Edition*. Morgan Kaufmann.
- [Ramez Elmasri, 2004] Ramez Elmasri, S. B. N. (2004). *Fundamentals of Database Systems, Fourth Edition*. Addison Wesley.