



Universidad Nacional Mayor de San Marcos
Escuela Profesional de
Ciencia de la Computación
Silabo del curso
Periodo Académico 2018-II

1. **Código del curso y nombre:** CS2702. Bases de Datos II (Obligatorio)
2. **Créditos:** 3
3. **Horas de Teoría y Laboratorio:** 1 HT; 4 HL; (15 semanas)
4. **Docente(s)**

Atención previa coordinación con el profesor

5. Bibliografía

- [Bur04] Donald K. Bursleson. *Physical Database Design Using Oracle*. CRC Press, 2004.
- [Cel05] Joe Celko. *Joe Celko's SQL Programming Style*. Elsevier, 2005.
- [Dat05] C.J. Date. *Data Mining: Practical Machine Learning Tools and Techniques, Second Edition*. Elsevier, 2005.
- [M T99] Patrick Valduriez M. Tamer Ozsü. *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.

6. Información del curso

- (a) **Breve descripción 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.

- (b) **Prerrequisitos:** CS2701. Bases de Datos I. (4^{to} Sem)
- (c) **Tipo de Curso:** Obligatorio
- (d) **Modalidad:** Presencial

7. Objetivos específicos del curso.

- 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.

8. Contribución a los resultados (Outcomes)

- b) Analizar problemas e identificar y definir los requerimientos computacionales apropiados para su solución. (**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. (**Usar**)
- k) Aplicar los principios de desarrollo y diseño en la construcción de sistemas de software de complejidad variable. (**Usar**)

9. Competencias (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 j**
- 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 k**

10. Lista de temas a estudiar en el curso

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

11. Metodología y Evaluación

Metodología:

Sesiones Teóricas:

Las sesiones de teoría se llevan a cabo en clases magistrales donde se realizarán actividades que propicien un aprendizaje activo, con dinámicas que permitan a los estudiantes interiorizar los conceptos.

Sesiones de Laboratorio:

Para verificar que los alumnos hayan alcanzado el logro planteado para cada una de las unidades de aprendizaje, realizarán actividades que les permita aplicar los conocimientos adquiridos durante las sesiones de teoría y se les propondrá retos que permitan evaluar el desempeño de los alumnos.

Exposiciones individuales o grupales:

Se fomenta la participación individual y en equipo para exponer sus ideas, motivándolos con puntos adicionales en las diferentes etapas de la evaluación del curso.

Lecturas:

A lo largo del curso se proporcionan diferentes lecturas, las cuales son evaluadas. El promedio de las notas de las lecturas es considerado como la nota de una práctica calificada. El uso del campus virtual UTEC Online permite a cada estudiante acceder a la información del curso, e interactuar fuera de aula con el profesor y con los otros estudiantes.

Sistema de Evaluación:

12. Contenido

Unidad 1: Diseño Físico de Bases de Datos (10)	
Competencias esperadas: C1	
Objetivos de Aprendizaje	Tópicos
<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] 	<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.
Lecturas : [Bur04], [Dat05], [Cel05]	

Unidad 2: Procesamiento de Transacciones (12)	
Competencias esperadas: C1	
Objetivos de Aprendizaje	Tópicos
<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] 	<ul style="list-style-type: none"> • Transacciones. • Fallo y recuperación. • Control concurrente. • Interacción de gestión de transacciones con el almacenamiento, especialmente en almacenamiento.
Lecturas : [Phi97], [Ram04]	

Unidad 3: Almacenamiento y Recuperación de Información (10)	
Competencias esperadas: C1	
Objetivos de Aprendizaje	Tópicos
<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] 	<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).
Lecturas : [Pet98], [Ram04]	

Unidad 4: Bases de Datos Distribuidas (36)	
Competencias esperadas: C1	
Objetivos de Aprendizaje	Tópicos
<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 minimise 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] 	<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
Lecturas : [M T99], [Dat05]	