

**Universidad Católica San Pablo (UCSP)**  
**Escuela Profesional de**  
**Ciencia de la Computación**  
**SILABO**



**CS272. Bases de Datos II (Obligatorio)**

**1. Información general**

1.1 Escuela	:	Ciencia de la Computación
1.2 Curso	:	CS272. Bases de Datos II
1.3 Semestre	:	5 <sup>to</sup> Semestre.
1.4 Prerrequisitos	:	CS271. Bases de Datos I. (4 <sup>to</sup> Sem)
1.5 Condición	:	Obligatorio
1.6 Modalidad de aprendizaje	:	Presencial
1.7 horas	:	1 HT; 4 HP;
1.8 Créditos	:	3
1.9 Plan	:	Plan Curricular 2016

**2. Profesores**

**Titular**

- Yessenia Deysi Yari Ramos <ydyari@ucsp.edu.pe>  
– Master en Ciencias de la Computación, UFRGS, Brasil, 2011.

**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. Resumen**

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

**5. Objetivos Generales**

- 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 los resultados (*Outcomes*)

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

- 1) S.O. Analizar un problema computacional complejo y aplicar los principios computacionales y otras disciplinas relevantes para identificar soluciones. (**Evaluar**)
- 4) S.O. Reconocer las responsabilidades profesionales y hacer juicios informados en el campo profesional de computación con principios éticos. (**Evaluar**)
- 6) S.O. Aplicar la teoría de la computación y los fundamentos del desarrollo de software para producir soluciones basadas en computación. . (**Evaluar**)
- 7) S.O. 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. Contenido

### UNIDAD 1: Diseño Físico de Bases de Datos (10)

Competencias: 1,6

#### Contenido

- 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.

#### Objetivos Generales

- 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:** Burleson (2004), Celko (2005)

UNIDAD 2: Procesamiento de Transacciones (12)	
Competencias: 1,6	
Contenido	Objetivos Generales
<ul style="list-style-type: none"> <li>• Transacciones.</li> <li>• Fallo y recuperación.</li> <li>• Control concurente.</li> <li>• Interacción de gestión de transacciones con el almacenamiento, especialmente en almacenamiento.</li> </ul>	<ul style="list-style-type: none"> <li>• Crear una transacción mediante la incorporación de SQL en un programa de aplicación [Usar]</li> <li>• Explicar el concepto de confirmaciones implícitas [Usar]</li> <li>• Describir los problemas específicos para la ejecución de una transacción eficiente [Usar]</li> <li>• Explicar cuando y porqué se necesita un <i>rollback</i>, y cómo registrar todo asegura un <i>rollback</i> adecuado [Usar]</li> <li>• Explicar el efecto de diferentes niveles de aislamiento sobre los mecanismos de control de concurrencia [Usar]</li> <li>• Elejir el nivel de aislamiento adecuado para la aplicación de un protocolo de transacción especificado [Usar]</li> <li>• Identificar los límites apropiados de la transacción en programas de aplicación [Usar]</li> </ul>
<b>Lecturas:</b> Philip A. Bernstein (1997), Ramez Elmasri (2004)	

<b>UNIDAD 3: Almacenamiento y Recuperación de Información (10)</b>	
<b>Competencias: 1,6</b>	
<b>Contenido</b>	<b>Objetivos Generales</b>
<ul style="list-style-type: none"> <li>• Documentos, publicación electrónica, markup, y lenguajes markup.</li> <li>• Tries, archivos invertidos, Árboles PAT, archivos de firma, indexación.</li> <li>• Análisis Morfológico, stemming, frases, stop lists.</li> <li>• Distribuciones de frecuencia de términos, incertidumbre, fuzificación (fuzzyness), ponderación.</li> <li>• Espacio vectorial, probabilidad, lógica, y modelos avanzados.</li> <li>• Necesidad de Información , Relevancia, evaluación, efectividad.</li> <li>• Thesauri, ontologías, clasificación y categorización, metadata.</li> <li>• Información bibliográfica, bibliometría, citasiones.</li> <li>• Enrutamiento y filtrado.</li> <li>• Búsqueda multimedia.</li> <li>• Información de resumen y visualización.</li> <li>• Búsqueda por facetas (por ejemplo, el uso de citas, palabras clave, esquemas de clasificación).</li> <li>• Librerías digitales.</li> <li>• Digitalización, almacenamiento, intercambio, objetos digitales, composición y paquetes.</li> <li>• Metadata y catalogación.</li> <li>• Nombramiento, repositorios, archivos</li> <li>• Archivamiento y preservación, integrdad</li> <li>• Espacios (Conceptual, geográfico, 2/3D, Realidad virtual)</li> <li>• Arquitecturas (agentes, autobuses, envolturas / mediadores), de interoperabilidad.</li> <li>• Servicios (búsqueda, de unión, de navegación, y así sucesivamente).</li> <li>• Gestión de derechos de propiedad intelectual, la privacidad y la protección (marcas de agua).</li> </ul>	<ul style="list-style-type: none"> <li>• Explica los conceptos básicos de almacenamiento y recuperación de la información [Usar]</li> <li>• Describe que temas son específicos para una recuperación de la información eficiente [Usar]</li> <li>• Da aplicaciones de estrategias alternativas de búsqueda y explica porqué una estrategia en particular es apropiada para una aplicación [Usar]</li> <li>• Diseña e implementa un sistema de almacenamiento y recuperación de la información o librería digital de tamaño pequeño a mediano [Usar]</li> <li>• Describe algunas de las soluciones técnicas a los problemas relacionados al archivamiento y preservación de la información en una librería digital [Usar]</li> </ul>
<b>Lecturas:</b> Peter Brusilovsky (1998), Ramez Elmasri (2004)	

UNIDAD 4: Bases de Datos Distribuidas (36)	
Competencias: 1,6	
Contenido	Objetivos Generales
<ul style="list-style-type: none"> <li>• DBMS Distribuidas <ul style="list-style-type: none"> <li>– Almacenamiento de datos distribuido</li> <li>– Procesamiento de consultas distribuido</li> <li>– Modelo de transacciones distribuidas</li> <li>– Soluciones homogéneas y heterogéneas</li> <li>– Bases de datos distribuidas cliente-servidor</li> </ul> </li> <li>• Parallel DBMS <ul style="list-style-type: none"> <li>– Arquitecturas paralelas DBMS: memoria compartida, disco compartido, nada compartido;</li> <li>– Aceleración y ampliación, por ejemplo, el uso del modelo de procesamiento MapReduce</li> <li>– Replicación de información y modelos de consistencia débil</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• Explicar las técnicas usadas para la fragmentación de datos, replicación, y la asignación durante el proceso de diseño de base de datos distribuida [Usar]</li> <li>• Evaluar estrategias simples para la ejecución de una consulta distribuida para seleccionar una estrategia que minimice la cantidad de transferencia de datos [Usar]</li> <li>• 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]</li> <li>• Describir el control concurrente distribuido basados en técnicas de copia distinguibles y el método de votación. [Usar]</li> <li>• Describir los tres niveles del software en el modelo cliente servidor [Usar]</li> </ul>
<b>Lecturas:</b> M. Tamer Ozsu (1999)	

#### 8. Metodología

1. El profesor del curso presentará clases teóricas de los temas señalados en el programa propiciando la intervención de los alumnos.
2. El profesor del curso presentará demostraciones para fundamentar clases teóricas.
3. El profesor y los alumnos realizarán prácticas
4. 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.

#### 9. Evaluar 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 Prácticas:

Las sesiones prácticas se llevan en clase donde se desarrollan una serie de ejercicios y/o conceptos prácticos mediante planteamiento de problemas, la resolución de problemas, ejercicios puntuales y/o en contextos aplicativos.

#### Sistema de Evaluación:

La nota final se obtiene a través de:

EVALUACIONES PERMANENTES	EVALUACIONES
<b>Evaluación Permanente 1</b> : 20 % <b>Evaluación Permanente 2</b> : 20 %	<b>Evaluación Parcial</b> : 30 % <b>Evaluación Final</b> : 30 % <b>Trabajo Final</b> : 50 % <b>Examen Final</b> : 50 %
40%	60%

Donde:

Evaluación Permanente: Comprende trabajos grupales, participación activa en clase, test de ejercicios.

- Permanente 1 (Semanas 1 - 9)
- Permanente 2 (Semanas 10 - 17)

Para aprobar el curso, el alumno debe obtener 11.5 o más en la nota final.

## References

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