



## Sociedad Peruana de Computación (SPC)

Programa Profesional de  
Ciencia de la Computación  
Sílabo 2022-I

### 1. CURSO

CS272. Gerenciamiento de Datos II (Obligatorio)

### 2. INFORMACIÓN GENERAL

2.1 Créditos	:	3
2.2 Horas de teoría	:	1 (Semanal)
2.3 Horas de práctica	:	4 (Semanal)
2.4 Duración del periodo	:	16 semanas
2.5 Condición	:	Obligatorio
2.6 Modalidad	:	Presencial
2.7 Prerrequisitos	:	CS271. Gerenciamiento de Datos I. (4 <sup>to</sup> Sem)

### 3. PROFESORES

Atención previa coordinación con el profesor

### 4. INTRODUCCIÓN AL 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.

### 5. OBJETIVOS

- 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. COMPETENCIAS

- a) Aplicar conocimientos de computación y de matemáticas apropiadas para la disciplina. (**Usar**)
- b) Analizar problemas e identificar y definir los requerimientos computacionales apropiados para su solución. (**Evaluar**)
- d) Trabajar efectivamente en equipos para cumplir con un objetivo común. (**Evaluar**)
- i) Utilizar técnicas y herramientas actuales necesarias para la práctica de la computación. (**Usar**)

### 7. COMPETENCIAS ESPECÍFICAS

- a14) Utilizar de forma apropiada archivos para el almacenamiento y recuperación de información.
- b4) Identificar y aplicar de forma eficiente diversas estrategias algorítmicas y estructuras de datos para la solución de un problema dadas ciertas restricciones de espacio y tiempo.
- b5) Aplicar de forma eficiente diversas estrategias algorítmicas y estructuras de datos para la solución de un problema en ambientes paralelos y distribuidos.

d2) Desarrollar presentaciones grupales e informes sobre tópicos específicos.

d3) Desarrollar trabajo en grupo en cada tópico del curso.

i3) Utilizar de forma apropiada los módulos de optimización de consultas, desempeño, indexación y fragmentación de tablas para BD distribuídas utilizando un motor de bases de datos de código abierto.

## 8. TEMAS

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

Unidad 2: Procesamiento de Transacciones (12)	
Competencias esperadas: b,j	
Temas	Objetivos de Aprendizaje
<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> [Phi97], [Ram04]	

**Unidad 3: Almacenamiento y Recuperación de Información (10)****Competencias esperadas: b,j**

Temas	Objetivos de Aprendizaje
<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> [Pet98], [Ram04]	

Unidad 4: Bases de Datos Distribuidas (36)	
Competencias esperadas: b,j	
Temas	Objetivos de Aprendizaje
<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>
Lecturas : [M T99]	

## 9. PLAN DE TRABAJO

### 9.1 Metodología

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.

### 9.2 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.

### 9.3 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.

## 10. SISTEMA DE EVALUACIÓN

\*\*\*\*\* EVALUATION MISSING \*\*\*\*\*

## 11. BIBLIOGRAFÍA BÁSICA

- [Bur04] Donald K. Burleson. *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 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.