



Sociedad Peruana de Computación (SPC)

Programa Profesional de
Ciencia de la Computación
Sílabo 2021-I

1. CURSO

CS374. Procesamiento de Texto para Ciencia de Datos (Electivo)

2. INFORMACIÓN GENERAL

- 2.1 Créditos : 3
- 2.2 Horas de teoría : 1 (Semanal)
- 2.3 Horas de práctica : 2 (Semanal)
- 2.4 Duración del periodo : 16 semanas
- 2.5 Condición : Electivo
- 2.6 Modalidad : Presencial
- 2.7 Prerrequisitos : CS272. Gerenciamiento de Datos II. (5^{to} 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"> • ... • ... • ... 	<ul style="list-style-type: none"> • ... [Usar] • ... [Usar] • ... [Usar]
Lecturas : [Bur04], [Cel05]	

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.