

**Universidad de San Carlos de Guatemala**

**Facultad de Ingeniería**

**Escuela de Ciencias y Sistemas**



## **PROGRAMA DEL CURSO**

### **I. Información General**

Nombre: Seminario de Sistemas 2

Código: 0798

Créditos: 3

Prerrequisitos: Seminario de Sistemas 1 + 190 créditos

### **II. Descripción**

En las últimas décadas se ha presenciado como empresas han construido su posición competitiva a partir del buen análisis de sus datos. Para poder hacer de esto una estrategia sustentable y diferenciadora, muchas empresas han extendido el concepto analítico desde sus bases de datos corporativas, a bases de datos nuevas, más grandes y de más rápido crecimiento, que han surgido producto de las redes sociales, smartphones y nuevos sensores, que en combinación con las bases de datos corporativas han demostrado tener altísimo valor en el descubrimiento de nuevo conocimiento. En su conjunto, el fenómeno de Big Data engloba una serie de conceptos del Business Intelligence tradicional, pero re-escritos desde la perspectiva de las nuevas posibilidades que abren tecnologías como las bases de datos no relacionales, procesamiento paralelo, visualización, inteligencia artificial y procesamiento en tiempo real.

### **III. Objetivo General**

EL objetivo del curso es desarrollar en los estudiantes las competencias necesarias para entender los aspectos tecnológicos que permiten la toma de decisiones basadas en datos. Esto significa ser capaces de interpretar todos los diferentes tipos de datos y los procesos detrás de ellos para llevarlos de la manera correcta y oportuna a quienes deben analizarlos y tomar decisiones.

### **IV. Objetivos Específicos**

- Saber las diferencias entre sistema operacional y sistema decisional

- Comprender la importancia de lograr diferenciación analítica en los mercados en los que compiten
- Conocer las tecnologías disponibles para desarrollar emprendimientos escalables y sustentables basados en Big Data y Analítica avanzada
- Desarrollar metodologías ágiles aplicadas a ciencia de datos

#### **IV. Metodología**

Se impartirán clases magistrales, con la participación de los estudiantes en foros de discusión. Autoaprendizaje y lectura constante de información relacionada al curso. Foros de discusión, exposiciones y evaluaciones parciales, para el control del progreso. Investigaciones sobre los diferentes productos disponibles en cada una de las áreas del curso.

#### **V. Evaluación**

- Clase teórica, 50 puntos (integrada por tareas, revisiones de las lecturas, exámenes y asistencia)
- Laboratorio, 25 puntos
- Examen Final, 25 puntos
- Es necesario aprobar el laboratorio para tener derecho a examen final.

#### **VI. Contenido**

##### 1. Inteligencia de negocios

- 1.1. Introducción al Business Intelligence y Analítica Avanzada
- 1.2. Datos para la toma de decisiones
- 1.3. Niveles de madurez analítica
- 1.4. Tipos de información
- 1.5. Proceso de toma de decisiones basadas en dato

##### 2. Arquitectura de sistemas de BI

- 2.1. Procesos ETL
- 2.2. Introducción al Data Warehouse / Datamart
- 2.3. Diseño multidimensional
- 2.4. Datamarts vrs. cubos multidimensionales y nuevas estructuras de datos

2.5. Arquitecturas OLAP

2.6. Arquitectura moderna de un DataWarehouse

### 3. Nuevas arquitecturas de bases de datos

3.1. Base de datos en memoria

3.2. Base de datos Columnares

3.3. Bases de datos multidimensionales

3.4. Appliance Analíticos y Bases de datos MPP

3.5. Bases de datos en la nube (Cloud computing)

3.5.1. Virtualización y Federación de datos

3.5.2. Centros de datos para Cloud Computing

3.5.3. Ventajas y desventajas del Cloud Computing

3.5.4. Análisis FODA de la implantación del Cloud Computing

3.5.5. Aspectos Legales y Seguridad del Cloud Computing

3.5.6. Implementación de la seguridad en el Cloud Computing

3.5.7. Nubes híbridas

3.5.8. IAAS, PAAS, SAAS

### 4. Análisis de datos presente y futuro

4.1. Metodologías de indicadores y KPI

4.2. Diferencia entre análisis histórico e indicadores

4.3. Balance Scorecard

4.4. Análisis de datos para presupuestos

4.5. Rolling Forecast y cubos multidimensionales

4.6. Proyecciones de datos y analítica avanzada

### 5. Fundamentos de Big Data & Analytics

4.1. Introducción Fundamentos de Big Data & Analytics

4.2. ¿Qué es Big Data?

4.3. Las 4 "V" de Big Data

4.4. Arquitectura de Hadoop y componentes principales

4.5. Lago de datos y aplicaciones sobre Hadoop

4.6. Spark

4.7. Hadoop sobre nube privada (Cloudera Private Cloud)

4.8. Bases de datos NoSQL

4.9 Internet de las cosas

5. Data Science

5.1. Modelos predictivos (Análítica avanzada)

5.2. Roles principales: Científico de Datos e Ingeniero de Datos

5.3. Metodología de Ciencia de Datos, metodología DMCRISP

5.4. Herramientas para Ciencia de datos

6. Gobierno de Datos

6.1. Componentes principales de Gobierno de datos

6.2. Beneficios del Gobierno de dato

6.3. Master Data Management

6.4. Gobierno de datos y Datawarehouse

## **VIII. Bibliografía**

- Documentos, videos y lecturas disponibles en internet referentes a cada tema.
- Descripciones de productos proporcionados por las compañías fabricantes de Software.