



**UNIVERSIDAD AUTÓNOMA DE SINALOA  
COORDINACIÓN GENERAL DE INVESTIGACIÓN Y POSGRADO**

<b>Datos de Identificación</b>	
Nombre de la Asignatura	<b>Plataformas Computacionales para la Administración, Procesamiento y Análisis de Datos a Gran Escala</b>
Unidad Académica	Facultad de Informática
Programa	<b>Maestría en Ciencias de la Información</b>
Tipo	Optativa
Horas Clase	48
Horas Trabajo Independiente	16
Horas Laboratorio	16
Créditos	5
<b>Breve Descripción de la Asignatura:</b>	
Se Implementaran distintos sistemas de procesamiento y análisis de grandes volumen de datos.	
<b>Objetivo General:</b>	
El alumno conocerá e implementará distintos entornos de procesamiento, análisis y visualización de grandes volumen de datos.	
<b>Objetivos Específicos:</b>	
El alumno conocerá los fundamentos de la captura, almacenamiento, búsqueda, compartición, análisis y visualización de grandes volúmenes de datos.	
El alumno implementara distintas herramientas de software basadas en cómputo distribuido para el procesamiento de grandes volúmenes de datos.	
<b>Contenido Sintético:</b>	
<b>Unidad I. Introducción al Big Data</b>	
Definiciones	
Aplicaciones	
Ventajas y desventajas	
<b>Unidad II. Procesamiento de datos con HADOOP</b>	
Fundamentos	
Instalación, configuración y administración	
El ecosistema Hadoop (Hive, Pig, Snoop, etc).	
Elementos de desarrollo con R	
<b>Unidad III. Procesamiento de datos con SCIDB</b>	
Fundamentos	
Instalación, configuración y administración	
Tipos de datos	



Arreglos  
Operadores  
Funciones  
API HTTP  
Elementos de desarrollo con Java  
Elementos de desarrollo con Phyton  
Elementos de desarrollo con R

**Unidad IV. Procesamiento de datos con DISTRIBUTED R**

Fundamentos  
Instalación, configuración y administración  
Modelo de programación  
Paquetes de algoritmos

**Modalidades o Formas de Conducción de los Procesos de Enseñanza y Aprendizaje:**

Exposición oral por parte del profesor  
Planteamiento de problemas  
Presentación de propuestas de soluciones por parte de los alumnos.  
Exámenes  
Planteamiento de la problemática final  
Defensa de la solución del alumno de la problemática final

**Modalidades de Evaluación y Acreditación:**

Aspectos	Ponderación
Ejercicios de Programación	35 %
Proyecto de Curso	15 %
Examen de Medio Término	20 %
Examen Final	30 %

**Bibliografía, Documentación y Material de Apoyo Didáctico:**

Nathan Marz, Big Data, Manning Publications; 1 edition, 2015  
Tom White, Hadoop, The Definitive Guide, O'Reilly Media; 4 edition, 2015  
Jason Venner, Pro Apache Hadoop 2nd ed. 2014  
<http://hadoop.apache.org/docs/current/>  
[http://www.paradigm4.com/HTMLmanual/14.12/scidb\\_ug/](http://www.paradigm4.com/HTMLmanual/14.12/scidb_ug/)  
<https://www.vertica.com/hp-vertica-documentation/hp-vertica-distributed-r-product-documentation/>

**Responsable(s) de la Elaboración del Programa de la Asignatura:**

M.C. Gerardo Beltrán Gutiérrez