Datos de identificación			
Nombre del EE: Almacén de Datos		Área Formativa: Vocacional	
Departamento que da el servicio: Departamento de Ingeniería Industrial			
Clave:	Modalidad: Presencial	Idiomas: Español	
Horas totales al semestre: 64	Valor en créditos: 4	Semestre en que se cursa: N/A	
Carácter: Optativa	EE Antecedente:	EE subsecuente: N/A	
Opciones de promoción: Calificación		Mecanismos alternativos de promoción: Equivalencia	
Presentación			
El espacio educativo de almacén de datos ofrece una formación completa sobre los conceptos y técnicas fundamentales para el almacenamiento y procesamiento de grandes volúmenes de datos. Es una herramienta valiosa para que los estudiantes desarrollen sus habilidades en el ámbito del Big Data. Se abordará el diseño de procesos y su implementación en diferentes plataformas.			
Desempeños			
<i>Competencias genéricas que se ejercitan</i>		<i>Unidades de competencia profesionales</i>	
<ul style="list-style-type: none"> Utiliza con eficiencia las tecnologías digitales para la comunicación y la gestión de información académica y profesional, en un entorno de trabajo colaborativo. 			
Resultados de Aprendizaje			
<ul style="list-style-type: none"> Crear un almacén de datos para un sistema. Implementar un almacén de datos utilizando Hadoop / Spark. Implementar sistema con Map Reduce para procesamiento de datos. Analizar los requisitos para establecer un almacén de datos. Analizar la seguridad para un Almacén de datos. 			
Orientación didáctica			
El estudiante asistirá 64 horas semestrales (4 horas semanales) a clases presenciales dirigidas por el profesor donde se desarrollarán los conceptos y prácticas propias de un curso de almacén de datos y se evidenciará la flexibilidad y utilidad de este para modelar y resolver problemas de desarrollo de sistemas con esta metodología.			
<i>Actividades del estudiante</i>		<i>Actividades del profesor</i>	
<i>Horas/ semestre</i>	<i>Actividades</i>	<i>Horas/ semestre</i>	<i>Actividades</i>
64	<ul style="list-style-type: none"> Asistencia y participación en clase. 	64	<ul style="list-style-type: none"> Impartir clases presenciales Definir las reglas de uso de algoritmos generativos de IA (GitHub, Copilot, ChatGPT, Gemini, etc.) en actividades de la materia que permitan utilizarlas como complemento para incrementar la productividad, más no como la

			solución total de los problemas a resolver o tareas por entregar.
<i>Evaluación del aprendizaje</i>			
<i>Criterios de cumplimiento</i>	<i>Evidencias de desempeño</i>	<i>Evidencias de conocimiento</i>	
<ol style="list-style-type: none"> Cumplir con la asistencia, puntualidad (Presencial o Virtual). Entrega de trabajos (investigación, tareas, exámenes) y/o prácticas a tiempo y siguiendo las especificaciones descritas. Cumplir con los criterios acordados de desarrollo con la vinculación. 	<ol style="list-style-type: none"> Realización de exámenes (en línea y/o en papel). Entrega de tareas y trabajos en plataforma electrónica. Realización de exposiciones en inglés sobre el tema. Desarrollo de un proyecto que integre todos los conceptos y tecnologías vistos. 	<ol style="list-style-type: none"> Proyecto final que refleje los conocimientos aprendidos durante el curso. Material y/o diapositivas de las exposiciones. El alumno responderá con ideas, conocimiento y aprendizaje a preguntas del profesor. Entrega de las actividades desarrolladas. 	
<i>Técnicas e instrumentos de evaluación</i>	Rúbricas para los exámenes, para tareas, prácticas e investigación y el proyecto final.		
<i>Recursos para la formación</i>			
<i>Contenidos básicos</i>		<i>Materiales</i>	
<ul style="list-style-type: none"> Introducción a Almacenes de Datos y Big Data. MapReduce y su papel en el procesamiento distribuido. Hadoop / Spark. Procesamiento de datos. 		<ul style="list-style-type: none"> Bibliografía física. Documentos electrónicos. Material audiovisual. Equipo de proyección. Plumones y pintarrón. Computadora. Conexión a Internet. Recursos en la nube. 	
<i>Bibliografía</i>			
<ul style="list-style-type: none"> Damji, J.S., Wenig, B., Das, T. & Lee D. (2020). Learning spark: lightning-fast big data analysis (2nd ed.). O'Reilly Media, Inc. Eri, T., Barceló-Monroy, E. (2023). Cloud Computing: Concepts, Technology, Security, and Architecture (2nd ed.). Pearson Parsian, M. (2022). Data Algorithms with Spark. O'Reilly Media, Inc. Gowda, P. (2024). Apache Spark for Machine Learning. Packt Publishing. Tandon, A., Ryza, S., Laserson, U., Owen, S., Wills, J. (2022). Advanced Analytics with PySpark. O'Reilly Media, Inc. Nudurupati, S. (2021). Essential PySpark for Scalable Data Analytics. Packt Publishing. Kleppmann, M. (2022). Diseño de aplicaciones mediante el uso intensivo de datos. Marcombo. 			
Perfil deseable del profesor que lo conduce o lo coordina			
Grado académico: Licenciatura. De preferencia con estudios de Maestría o Doctorado		Área de formación: Ing. en Sistemas de Información, Ing. en Desarrollo de Software, Maestría en Desarrollo de Software, Ciencias Computacionales Se recomienda un nivel intermedio-alto del idioma inglés para impartir esta materia.	

Experiencia docente: 1 año.	Experiencia profesional en el campo: 1 año.
Elaboró: Dra. Raquel Torres Peralta.	Fecha: 24 de octubre de 2024