

GUÍA DOCENTE

ARQUITECTURA Y
ALMACENAMIENTO DE
DATOS ESCALABLES
/ SCALABLE DATA
ARCHITECTURE AND
STORAGE

Arquitectura Y Almacenamiento De Datos Escalables / Scalable Data Architecture And Storage

Número total de créditos ECTS	6	
Tipología	Obligatoria	
Organización temporal	Curso 1; Semestre 1	
Idioma	Castellano e inglés	
Contenidos	<p>Arquitecturas de sistemas de archivos distribuidos</p> <p>1.1. HDFS</p> <p>1.2. Ceph</p> <p>1.3. Google Colossus</p> <p>Almacenamiento en columnas vs. filas</p> <p>2.1. HBase</p> <p>2.2. Cassandra</p> <p>2.3. Amazon Redshift</p> <p>Soluciones cloud de Data Lake</p> <p>3.1. AWS Lake Formation</p> <p>3.2. Azure Data Lake Storage</p> <p>Sistemas de bases de datos NoSQL escalables</p> <p>4.1. MongoDB</p> <p>4.2. DynamoDB</p> <p>-----</p> <p>Distributed File System Architectures</p> <p>1.1. HDFS</p> <p>1.2. Ceph</p> <p>1.3. Google Colossus</p> <p>Column-Oriented vs. Row-Oriented Storage</p> <p>2.1. HBase</p> <p>2.2. Cassandra</p> <p>2.3. Amazon Redshift</p> <p>Cloud Data Lake Solutions</p> <p>3.1. AWS Lake Formation</p> <p>3.2. Azure Data Lake Storage</p> <p>Scalable NoSQL Database Systems</p> <p>4.1. MongoDB</p> <p>4.2. DynamoDB</p>	
Resultados de aprendizaje TÍTULO	Conocimientos y contenidos	<p>CC03 Comprender los principios de los modelos de regresión lineal y su aplicación para la predicción en contextos reales, evaluando sus supuestos, diagnóstico y bondad de ajuste. / Understand the principles of linear regression models and their application for prediction in real-world contexts, evaluating their assumptions, diagnostics, and goodness of fit.</p> <p>CC07 Comprender la arquitectura y principios de contenedores y sistemas de orquestación —Docker y Kubernetes— para la gestión de servicios en entornos de computación en la nube. / Understand the architecture and principles of containers and orchestration systems—Docker and Kubernetes—for service management in cloud computing environments.</p>
	Habilidades y destrezas	<p>HD03 Utilizar distintos modelos de almacenamiento y sistemas de gestión de bases de datos, así como el lenguaje de definición, consulta y manipulación de estos / Use different storage models and database management systems, as well as their definition, query, and manipulation languages.</p>

		HD04 Emplear herramientas estadísticas avanzadas y software estadístico especializado en la aplicación de las técnicas en Ciencia de Datos / Use advanced statistical tools and specialized statistical software in the application of techniques in Data Science.
	Competencias	CP01 Desarrollar e implementar soluciones de Ciencia de Datos en proyectos con incertidumbre técnica, asegurando que los resultados respondan a los objetivos de la organización. / Develop and implement Data Science solutions within projects characterized by technical uncertainty, ensuring that outcomes align with organizational objectives. CP03 Analizar nuevas metodologías y técnicas para la gestión y explotación de grandes bases de datos que apoyen la toma de decisiones en la dirección estratégica e inteligencia empresarial / Analyze new methodologies and techniques for managing and exploiting large databases to support decision-making in strategic management and business intelligence.

Resultados de aprendizaje ASIGNATURA

1. Diseñar arquitecturas escalables de almacenamiento de datos en entornos distribuidos.
 2. Gestionar clusters HDFS y soluciones de Data Lake en entornos cloud.
 3. Optimizar el rendimiento, la disponibilidad y la seguridad del almacenamiento masivo de datos.
 4. Integrar sistemas de almacenamiento en pipelines de ingesta y procesamiento de datos.
-
5. Design scalable data-storage architectures in distributed environments.
 6. Manage HDFS clusters and Data Lake solutions in cloud settings.
 7. Optimize performance, availability, and security of large-scale data storage.
 8. Integrate storage systems into data-ingestion and processing pipelines.

	Actividades formativas	Horas totales	Horas presenciales	Horas virtuales síncronas	Horas virtuales asíncronas	
Modalidad Presencial	Clases Expositivas / Participatory learning	22	22	0	0	
	Seminarios / Seminars	2	2	0	0	
	Clases prácticas / Practical lessons	10	10	0	0	
	Prácticas de Laboratorio / Laboratory practicals	12	12	0	0	
	Trabajo autónomo / Autonomous work	102	0	0	102	
	Prueba de evaluación final / Final evaluation test	2	2	0	0	
	Total	150	48	0	102	
	Sistemas de evaluación				MÍNIMO	MÁXIMO
	Evaluación Final: prueba o examen / Final evaluation test				50	50
	Resolución de problemas / Problem resolution				10	30
	Estudio de casos / Proyectos / Case studies / projects				10	30
	Total				70	110