

ACCIÓN FORMATIVA

ARQUITECTURA BIG DATA

OBJETIVO

Conocer en profundidad los diferentes paradigmas de procesamiento en sistemas Big Datas y dominar las principales tecnologías y su utilización para el diseño de arquitecturas escalables adaptadas a cada proyecto

CONTENIDOS**UNIDAD DIDÁCTICA 1. BATCH PROCESSING.**

- Hadoop.
- Pig.
- Hive.
- Sqoop.
- Flume.
- Spark Core.
- Spark 2.0.

UNIDAD DIDÁCTICA 2. STREAMING PROCESSING.

- Fundamentos de Streaming Processing.
- Spark Streaming.
- Kafka.
- Pulsar y Apache Apex.
- Implementación de un sistema real-time.

UNIDAD DIDÁCTICA 3. SISTEMAS NOSQL.

- Hbase.
- Cassandra.
- MongoDB.
- Neo4J.
- Redis.
- Berkeley DB.

UNIDAD DIDÁCTICA 4. INTERACTIVE QUERY.

- Lucene + Solr.

UNIDAD DIDÁCTICA 5. SISTEMAS DE COMPUTACIÓN HÍBRIDOS.

- Arquitectura Lambda.
- Arquitectura Kappa.
- Apache Flink e implementaciones prácticas.
- Druid.
- Elasticsearch.
- Logstash.
- Kibana.

UNIDAD DIDÁCTICA 6. CLOUD COMPUTING.

- Amazon Web Services.

- Google Cloud Platform.

UNIDAD DIDÁCTICA 7. ADMINISTRACIÓN DE SISTEMAS BIG.

- Administración e Instalación de clusters: Cloudera y Hortonworks.
- Optimización y monitorización de servicios.
- Seguridad: Apache Knox, Ranger y Sentry.

UNIDAD DIDÁCTICA 8. VISUALIZACIÓN DE DATOS.

- Herramientas de visualización: Tableau y CartoDB.
- Librerías de Visualización: D3, Leaflet, Cytoscape.