

---

# ARQUITECTURA BIG DATA



## FORMACIÓN GRATUITA

### CONVOCATORIA FORMACIÓN OFERTA SECTORIAL ESTATAL

---

- **DURACIÓN:** 165 horas
- **MODALIDAD:** Teleformación
- **OBJETIVO GENERAL:**

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:

### 1. BATCH PROCESSING.

- 1.1. Hadoop.
- 1.2. Pig.
- 1.3. Hive.
- 1.4. Sqoop.
- 1.5. Flume.
- 1.6. Spark Core.
- 1.7. Spark 2.0.

### 2. STREAMING PROCESSING.

- 2.1. Fundamentos de Streaming Processing.
- 2.2. Spark Streaming.
- 2.3. Kafka.
- 2.4. Pulsar y Apache Apex.
- 2.5. Implementación de un sistema real-time.

### 3. SISTEMAS NOSQL.

- 3.1. Hbase.
- 3.2. Cassandra.
- 3.3. MongoDB.
- 3.4. Neo4J.
- 3.5. Redis.
- 3.6. Berkeley DB.

### 4. INTERACTIVE QUERY.

- 4.1. Lucene + Solr.

### 5. SISTEMAS DE COMPUTACIÓN HÍBRIDOS.

- 5.1. Arquitectura Lambda.
- 5.2. Arquitectura Kappa.

5.3. Apache Flink e implementaciones prácticas.

5.4. Druid.

5.5. ElasticSearch.

5.6. Logstash.

5.7. Kibana.

## **6. CLOUD COMPUTING.**

6.1. Amazon Web Services.

6.2. Google Cloud Platform.

## **7. ADMINISTRACIÓN DE SISTEMAS BIG.**

7.1. Administración e Instalación de clusters: Cloudera y Hortonworks.

7.2. Optimización y monitorización de servicios.

7.3. Seguridad: Apache Knox, Ranger y Sentry.

## **8. VISUALIZACIÓN DE DATOS.**

8.1. Herramientas de visualización: Tableau y CartoDB.

8.2. Librerías de Visualización: D3, Leaflet, Cytoscape.