



## **Big Data Technology Training**

### **Course description**

The objective of the course is to learn the basic concepts and technical principles of big data.

Also it will allow us to understand the concept and ecosystem architecture of big data. Be familiar with typical applications of big data in telecom services. Understand the principles of Hadoop core components (HDFS, MapReduce, and Spark). Understand the basic NoSQL and typical products.

### **Target :**

All employees in the ICT industry

### **Content of the Training**

## **1) Big Data Concepts and Ecosystems**

- 01 Big Data Era
- 02 Typical Application of Big Data in the Telecom Industry
- 03 Big Data Ecosystem and Common Tools

## **2) Hadoop Big Data Platform and Technology**

- 01 Introduction to the Hadoop Big Data Platform
- 02 HDFS - Distributed File System
- 03 Distributed Computing Framework Based on MapReduce
- 04 YARN Resource Scheduling Implementation Principle
- 05 Spark - An In-Memory Distributed Computing Engine
- 06 HBase - Distributed NoSQL Database
- 07 Elasticsearch - Distributed Search Engine
- 08 Hive - Distributed Data Warehouse
- 09 ClickHouse - A Column-based OLAP Data Warehouse
- 10 Sqoop Data Conversion
- 11 Flink - Stream and Batch Processing in a Single Engine
- 12 Flume - Massive Log Aggregation
- 13 Kafka Distributed Publish-Subscribe Messaging System
- 14 Distributed coordination service of the ZooKeeper cluster

## **3) NoSQL Basic Knowledge**

- 01 NoSQL Development History and Trend
- 02 CAP Theory
- 03 Classification and Representative Works of NoSQL
- 04 Introduction to Redis
- 05 Introduction to MongoDB
- 06 Introduction to Cassandra
- 07 Introduction to Neo4J