

01/. Acquérir les bases du Big Data

Introduction au Big Data : de quoi s'agit-il ?
Les technologies concernées
Les outils
Les langages : Hortonworks, MapR, Cloudera, IBM Watson
Démystification du Big Data
Les acteurs principaux
Les différents métiers du Big Data

02/. Lister les solutions Big Data

Projets, applications, plateformes
Solutions de stockage des données
Considérations Hardware : Stockage, CPU, Mémoire, Réseau, Systèmes distribués
La science et l'art de l'analyse prédictive

03/. Expérimenter l'architecture applicative

Les différentes étapes de gestion des données
Les outils du marché
L'ingestion
La transformation
La restitution / visualisation

04/. Expérimenter l'architecture technique

La capture des données
La mise en qualité des données (Data Quality) : Utilisation d'une exploitation d'un flux Open Data avec Talend Data Quality
Le stockage des données : Hadoop, HDFS, NoSQL (Cassandra, MongoDB)
La diffusion des données : Streaming avec Apache Kafka, Amazon AWS
Traitements en temps réel ou différé : Apache Spark
Les architectures réparties : Clustering Hadoop
Supervision des plateformes : App Dynamix, Ambari

05/. Appliquer au métier

Text analytics
Détection de fraude
Ciblage clientèle

REF. BD-001

PUBLIC

Administrateurs systèmes, développeurs, architectes et chefs de projets techniques, consultants Business Intelligence. **Si vous êtes en situation de handicap, merci de nous contacter.**

OBJECTIFS

Exploiter les architectures Big Data et mettre en place des socles techniques complets pour des projets Big Data.

PRÉREQUIS

Maîtriser Shell Linux, des bases du réseau, des bases de données relationnelles et des architectures réparties.

MOYENS PÉDAGOGIQUES

Exposés alternant théorie et pratique, études de cas, correction travaux pratiques et ateliers.

1 support de cours par personne.