# CYCLE COMPLET BIG DATA

*Lien : https://innov-systems.com/formation/cycle-complet-big-data*

DURÉE
**10 jours (70h)**

RÉFÉRENCE
**BSI591**

CATÉGORIE
**Big Data : développement, langages et NOSQL**

Innov Systems

## OBJECTIFS DE LA FORMATION

À l'issue de cette formation, vous serez capable de :

- Comprendre les concepts et l’écosystème Big Data
- Maîtriser les architectures distribuées et leurs outils
- Développer des compétences sur Hadoop et Spark
- Utiliser des bases NoSQL pour des cas adaptés
- Automatiser la collecte et l’ingestion des données massives
- Construire et gérer un Data Lake
- Réaliser des traitements massifs avec Spark et Hive
- Appliquer des techniques de Machine Learning à grande échelle
- Intégrer les solutions Big Data dans le Cloud
- Assurer la sécurité et la conformité réglementaire des données
- Concevoir des projets Big Data adaptés à différents secteurs
- Développer des capacités de communication et de présentation des résultats

## POUR QUI ?

- Ingénieurs et développeurs désirant acquérir des compétences en traitement massif de données
- Professionnels de la data (Data Analyst, Data Scientist) voulant élargir leur champ d’expertise
- Responsables IT et architectes systèmes cherchant à comprendre les solutions Big Data
- Chefs de projets et managers souhaitant piloter des projets Big Data
- Toute personne en reconversion professionnelle vers les métiers de la Data

Innov Systems

## Programme détaillé

### 1 / Introduction au Big Data

- Comprendre les concepts fondamentaux du Big Data
- Identifier les enjeux et cas d’usage
- Découvrir l’écosystème des technologies Big Data

### 2 / Architecture des systèmes Big Data

- Comprendre le fonctionnement distribué
- Étudier les clusters et systèmes distribués
- Découvrir les approches Batch et Streaming

### 3 / Hadoop et son écosystème

- HDFS pour le stockage distribué
- MapReduce pour le traitement parallèle
- Outils associés : Hive, Pig, HBase

### 4 / Apache Spark

- Principes de Spark et RDD
- Spark SQL et DataFrames
- Spark Streaming pour le temps réel

### 5 / Bases de données NoSQL

- Découvrir MongoDB et Cassandra
- Comparer NoSQL vs SQL
- Cas d’usage des bases orientées documents et colonnes

Innov Systems

### 6 / Collecte et ingestion de données

- Apache Kafka pour la gestion de flux
- Apache Flume et Sqoop
- Connecteurs temps réel et batch

### 7 / Stockage et Data Lake

- Différences entre Data Warehouse et Data Lake
- Structuration des données brutes et raffinées
- Outils de gestion : AWS S3, Azure Data Lake

### 8 / Data Processing et ETL Big Data

- Transformation avec Spark et Hive
- Workflows avec Airflow
- Bonnes pratiques d’optimisation

### 9 / Analyse et exploration

- Utiliser Python et R pour analyser les données
- Visualisation avec Tableau ou Power BI
- Statistiques et modèles de base

### 10 / Machine Learning avec Big Data

- Spark MLlib
- Intégration avec Scikit-learn et TensorFlow
- Cas pratiques de classification et prédiction

### 11 / Cloud et Big Data

- Introduction à AWS, Azure, GCP
- Services managés Big Data

Innov Systems

- Scalabilité et coûts

### 12 / Sécurité et gouvernance

- Gestion des accès et confidentialité
- Règlementation RGPD
- Catalogues de données et Data Lineage

### 13 / Cas pratiques industriels

- Big Data pour la finance
- Big Data pour la demande électrique
- Big Data pour la logistique

### 14 / Projet de fin de formation

- Collecte, traitement et analyse d’un dataset réel
- Implémentation d’une architecture Big Data complète
- Présentation et soutenance du projet
- Ingénieurs et développeurs désirant acquérir des compétences en traitement massif de données
- Professionnels de la data (Data Analyst, Data Scientist) voulant élargir leur champ d’expertise
- Responsables IT et architectes systèmes cherchant à comprendre les solutions Big Data
- Chefs de projets et managers souhaitant piloter des projets Big Data
- Toute personne en reconversion professionnelle vers les métiers de la Data

Innov Systems

## Approche pédagogique

- Support Ecrit et Projection
- Exposés Interactifs, Podcasts et Vidéos
- Brainstorming et Jeux de Rôle
- Cas Pratiques et Labs inclus pour leur impact opérationnel
- Test de Validation des Acquis des Connaissances

## Prochaines dates programmées

| | |
|---|---|
| 23 Nov. au 04 Déc. 2026 | Présentiel - Casablanca |
| 18 au 29 Jan. 2027 | Présentiel - Casablanca |
| 15 au 26 Mars 2027 | Présentiel - Casablanca |
| 10 au 21 Mai 2027 | Présentiel - Casablanca |
| 05 au 16 Juil. 2027 | Présentiel - Casablanca |

Autres dates possibles sur demande. Contactez-nous pour organiser une session intra-entreprise.

## Réservation & Renseignements

**Téléphone :** +212 522 247 210
**Email :** contact@innov-systems.com
**Web :** https//www.innov-systems.com


Innov Systems — Tous droits réservés

Innov Systems