

**Données & IA**

## Fondamentaux et Architecture Big Data

Maîtrisez les principes fondamentaux, l'écosystème technologique et l'architecture des systèmes Big Data, de la collecte des données à leur visualisation.

DURÉE <b>35 heures</b>	NIVEAU <b>Intermédiaire</b>	FORMAT <b>Présentiel &amp; en ligne</b>
LANGUE <b>Français</b>	CERTIFICAT —	FORMATEUR <b>Nizar Haddad</b>

### Présentation

Cette formation dispensée par Expert Team Tunisie, centre de formation IT agréé par l'État, vous initie aux fondamentaux du Big Data et à l'architecture des systèmes de gestion des données massives. Elle couvre l'ensemble de la chaîne de valeur de la donnée : enjeux stratégiques, collecte, stockage distribué, traitement, analyse et datavisualisation.

Le programme combine une approche conceptuelle (cycle de vie de la donnée, gouvernance, qualité, aspects juridiques et éthiques) et une découverte des technologies de référence de l'écosystème Big Data : Hadoop et son écosystème (HDFS, YARN, MapReduce, Hive, Pig, Spark, HBase), les bases NoSQL (Cassandra, MongoDB, Neo4j), les outils ETL comme Talend, ainsi que les outils de datavisualisation (Tableau, Qlik). Une ouverture vers la Data Science, le Machine Learning et le Deep Learning complète le parcours.

À l'issue de la formation, les participants disposent d'une vision d'ensemble structurée du domaine, leur permettant de comprendre les architectures Big Data, de dialoguer avec les équipes techniques et de contribuer à des projets de valorisation des données au sein de leur organisation.

### Objectifs pédagogiques

- ✓ Comprendre le rôle stratégique de la gestion des données et les enjeux du Big Data pour l'entreprise
- ✓ Identifier les caractéristiques, types et sources de données ainsi que les critères de qualité
- ✓ Maîtriser le cycle de vie de la donnée et les principes de gouvernance et de contrôle qualité
- ✓ Décrire l'architecture et l'écosystème Hadoop (HDFS, YARN, MapReduce, Hive, Pig, Spark)
- ✓ Mettre en œuvre les mécanismes de collecte, de stockage distribué et de traitement des données (ETL, NoSQL)
- ✓ Appréhender les méthodes d'analyse de données, de Machine Learning et de Deep Learning
- ✓ Concevoir des restitutions et tableaux de bord grâce aux outils de datavisualisation (Tableau, Qlik)
- ✓ Conduire une démarche méthodologique de projet Big Data, du besoin métier à la décision

### Introduction au Big Data

- Enjeux et perspectives du Big Data
- Chiffres clés français et internationaux
- La donnée comme matière première : types, structures et sources
- Écosystème Big Data et acteurs du marché
- Nouveaux métiers de la data

### La collecte des données Big Data

- Définitions clés : Big Data, Data Science, Machine Learning
- Caractéristiques des données massives
- Architectures orientées stockage, calcul et temps réel
- Enjeux juridiques et éthiques
- Cycle de vie et qualité de la donnée
- Création de valeur et processus Data Science

### Méthodologie d'implémentation d'un projet Big Data

- Définition des besoins métiers et objectifs
- Spécificités Big Data : incertitude, budget, délais, organisation
- Sélection des outils appropriés
- Exploration et prétraitement des données
- Analyse, modélisation et communication des résultats
- Transformation en décisions et actions

### Outils et technologies : l'écosystème Hadoop

- Écosystème Apache Hadoop et distributions
- Clusters Hadoop, HDFS, YARN, MapReduce
- Composants : HBase, Pig, Hive, Spark
- Cas d'utilisation et implémentations

### Collecte, stockage et traitement des données

- Sources et caractéristiques des jeux de données
- Principes ETL et Talend
- Gestion des données semi et non structurées
- Stockage distribué et HDFS/MapReduce
- Bases NoSQL : Cassandra, Neo4j, MongoDB
- Pig, Hive et Elasticsearch

### Analyse des données

- Méthodes et techniques d'analyse statistique et prédictive
- Apache Spark
- Convergence Big Data et Intelligence Artificielle

- Familles de modèles de Machine Learning : régression, classification, clustering, NLP
- Feature engineering, entraînement et test d'algorithmes
- Principes du Deep Learning et réseaux de neurones

## Datavisualisation

- Enjeux et objectifs de la communication des résultats
- Principes de représentation visuelle
- Outils : Tableau et Qlik
- Bonnes pratiques : couleurs, graphiques, tableaux de bord, infographies
- Visualisation interactive

## Prérequis

Aucune connaissance technique particulière n'est requise. Une connaissance de base du langage SQL est recommandée, ainsi qu'idéalement une sensibilisation préalable aux enjeux et perspectives du Big Data.

## Public visé

Professionnels de l'informatique, analystes, chefs de projet, décideurs et toute personne souhaitant acquérir une vision globale et des compétences en architecture et gestion des données massives (Big Data).

## Prochaines sessions

DATE	VILLE	HORAIRE	FORMATEUR
03/08/2026	Tunis	09h00 - 16h00	Nizar Haddad
14/09/2026	Sfax	09h00 - 16h00	Nizar Haddad

**1 250 TND** / personne

Acompte de 30% possible à la réservation · Financement sur demande.

Experteam — El Ghazala, Ariana (face au Technopôle La Ghazela)

+216 29 674 204 · contact@expertunisie.com · experteam.itt.tn