

## FORMATION BIG DATA

Notre formation de deux jours vous permettra de maîtriser les concepts du Big Data et savoir quelles sont les technologies implémentées. A l'issue de cette formation, vous serez en mesure d'analyser les difficultés propres à un projet Big Data tels que les freins, les apports..

### PROGRAMME

#### 1/ Introduction

- L'essentiel du BigData.
- Besoins fonctionnels et caractéristiques techniques des projets.
- La valorisation des données.
- Le positionnement respectif des technologies de cloud, BigData et noSQL, et les liens, implications.
- Eléments d'architecture.
- L'écosystème du BigData
- Cycle de vie des projets BigData.
- Emergence de nouveaux métiers : Datascientists, Data labs,...

#### 2/ Stockage

- Caractéristiques NoSQL.
- Structure de données proches des utilisateurs, développeurs.
- Les types de bases de données : clé/valeur, document, colonne, graphe.
- Données structurées et non structurées, documents, images, fichiers XML, JSON, CSV, ...
- Les différents modes et formats de stockage.
- Stockage réparti : réplication, sharding, gossip protocol, hachage,
- Systèmes de fichiers distribués : GFS, HDFS,
- Quelques exemples de produits et leurs caractéristiques : Cassandra, MongoDB, CouchDB, DynamoDB, Riak, Hadoop, HBase, BigTable, ...
- Qualité des données, gouvernance de données.

#### 3/ Indexation et recherche

- Moteurs de recherche.Principe de fonctionnement.
- Méthodes d'indexation.
- Mise en oeuvre avec elasticsearch.
- Exemple de Lucene/solr.
- Recherche dans les bases de volumes importants.
- Exemples de produits et comparaison : Dremel, Drill, ElasticSearch, MapReduce,

#### 4/ Calcul et restitution, intégration

- Différentes solutions : calculs en mode batch, ou en temps réel, sur des flux de données ou des données statiques.
- Les produits : langage de calculs statistiques, R Statistics Language, sas, RStudio.
- Ponts entre les outils statistiques et les bases BigDatas
- Outils de calcul sur des volumes importants : storm en temps réel, hadoop en mode batch.
- Zoom sur Hadoop : complémentarité de HDFS et MapReduce.
- Restitution et analyse : logstash, kibana, elk, pentaho
- Présentation de pig pour la conception de tâches MapReduce sur une grappe Hadoop.

## NOUS CONTACTER

**Openska**  
21 rue Louise Weiss  
75013 Paris

Tel : 01 84 25 05 10  
Tel : 01 84 17 44 76  
[www.openska.com](http://www.openska.com)

**PRIX (INTER-ENTREPRISE) : 1490 € euros**

 **14h jours**

 **BD01**

Vous souhaitez organiser  
cette formation dans vos locaux ?

**Demandez Houilly au**  
**01 84 25 05 10**

### OBJECTIFS

- Connaître les concepts essentiels du BigData, et les technologies implémentées
- Être en mesure d'analyser les difficultés propres à un projet BigData

### PRE-REQUIS

- Une bonne culture générale sur les systèmes d'information.

### PUBLIC CONCERNE

- Chefs de projets
- Architectes
- Développeurs
- Data-scientists

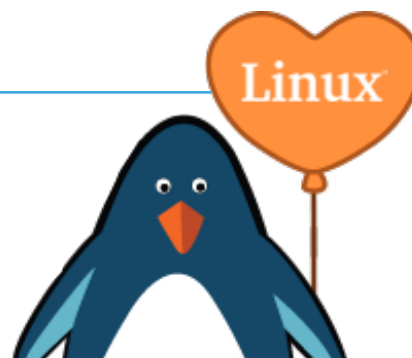
### DATES INTER-ENTREPRISES

Delais d'entree : sans

**Paris et à distance**

06/02/2023, 09/05/2023, 11/09/2023, 27/11/2023,

Si vous souhaitez organiser cette formation a une autre date contactez-nous.



## ELEMENTS COMPLÉMENTAIRES À TRANSMETTRE IMPÉRATIVEMENT À VOTRE OPCO POUR VOTRE DEMANDE DE PRISE EN CHARGE

### HORAIRES

- 9h30-13h
- 14h-17h30

### ACCÈS HANDICAP

Oui, Consultez notre référent interne.

### MODALITÉS D'ÉVALUATION

L'évaluation des acquis se fait tout au long de la formation au travers de multiples exercices et mise en situation

### MODALITÉS DE SUIVI DE L'EXÉCUTION

- Contrôle systématique des présences par demi-journée
- Attestation de présence remis à chaque participant

### COMPÉTENCES DU FORMATEUR

Les experts qui animent la formation sont des spécialistes des matières abordées. Ils ont été validés par nos équipes pédagogiques tant sur le plan des connaissances métiers que sur celui de la

pédagogie, et ce pour chaque formation qu'ils animent. Ils ont majoritairement cinq à dix années d'expérience dans leur domaine de compétences et ont une grande pratique en entreprise

### MOYENS PÉDAGOGIQUES

Les supports pédagogiques sont imprimés par nos soins et transmis à chacun des stagiaires. Les supports sont aussi remis au format électronique aux participants. Ce support est projeté via un vidéo projecteur afin d'animer la formation. Des exercices d'application ou études de cas sont prévus afin de valider les acquis des stagiaires.

### MOYENS D'ACCOMPAGNEMENT

8 personnes maximum, formation animée par un formateur expert sur le sujet.  
Mise en situation : Pédagogie active et participative.  
Apport théorique et méthodologique.  
Etude de cas. Atelier pratique.

### NATURE DES TRAVAUX DEMANDÉS AU STAGIAIRE

Un exercice est réalisé par le participant à la fin de chaque chapitre. Le participant dispose de 20 min avant de passer à la correction avec le formateur.

### MODALITÉS TECHNIQUES EN CAS DE PROBLÈMES

Les connexions et installations d'outil peuvent être testées en amont de la formation avec le formateur et les stagiaires ou le donneur d'ordre. Ce test permet d'éviter tout accident technique lié aux outils de connexion à distance.

Le formateur prend la main sur le poste du participant en cas de difficulté durant la formation.



ORGANISME DE FORMATION RÉFÉRENCÉ SOUS LE NUMÉRO : 117 555 432 75