BIG DATA : LES FONDAMENTAUX

Public	toute personne souhaitant profiter des nombreux avantages liés aux technologies dédiées au Big Data.
Durée	3 jours - 21 heures
Pré-requis	Vous devez avoir des connaissances pratiques de la plateforme Microsoft Windows. Des notions de programmation sont utiles sans toutefois être obligatoires.
Objectifs	acquérir des connaissances pratiques fondamentales des différents concepts, techniques et produits liés au Big Data apprendre à stocker, gérer, traiter et analyser des ensembles volumineux de données non structurées
Méthodes pédagogiques	Pour bien préparer la formation, le stagiaire remplit une évaluation de positionnement et fixe ses objectifs à travers un questionnaire. La formation est délivrée en présentiel ou distanciel (e-learning, classe virtuelle, présentiel et à distance). Le formateur alterne entre méthodes démonstratives, interrogatives et actives (via des travaux pratiques et/ou des mises en situation). La validation des acquis peut se faire via des études de cas, des quiz et/ou une certification. Cette formation est animée par un consultant-formateur dont les compétences techniques, professionnelles et pédagogiques ont été validées par des diplômes et/ou testées et approuvées par l'éditeur et/ou par Audit Conseil Formation.
Moyens techniques	1 poste de travail complet par personne De nombreux exercices d'application Mise en place d'ateliers pratiques Remise d'un support de cours Passage de certification(s) dans le cadre du CPF Remise d'une attestation de stage
Modalité d'évaluation des acquis	Evaluation des besoins et objectifs en pré et post formation Evaluation technique des connaissances en pré et post formation Evaluation générale du stage
Délai d'accès	L'inscription à cette formation est possible jusqu'à 5 jours ouvrés avant le début de la session
Accessibilité handicapés	Au centre d'affaires ELITE partenaire d'ACF à 20 m. Guide d'accessibilité à l'accueil.

1. DÉFINITION

- Les quatre dimensions du Big Data : volume, vélocité, variété, véracité
- Présentation de l'ensemble MapReduce, stockage et requêtes

2. AMÉLIORER LES RÉSULTATS DE L'ENTREPRISE GRÂCE AU BIG DATA

- Mesurer l'importance du Big Data au sein d'une entreprise
- Réussir à extraire des données utiles
- Intégrer le Big Data aux données traditionnelles

3. ANALYSER LES CARACTÉRISTIQUES DE VOS DONNÉES

- Sélectionner les sources de données à analyser
- Supprimer les doublons
- Définir le rôle de NoSQL

4. PRÉSENTATION DES ENTREPÔTS DE BIG DATA

- Modèles de données : valeur clé, graphique, document, famille de colonnes
- Système de fichiers distribué Hadoop (HDFS)
- HBase
- Hive
- Cassandra
- Hypertable
- Amazon S3
- BigTable
- DynamoDB
- MongoDB
- Redis
- Riak
- Neo4J

5. CHOISIR UN ENTREPÔT DE BIG DATA

- Choisir un entrepôt de données en fonction des caractéristiques de vos données
- Injecter du code dans les données, mettre en œuvre des solutions de stockage des données multilingues
- Choisir un entrepôt de données capable de s'aligner avec les objectifs de l'entreprise

6. INTÉGRER DIFFÉRENTS ENTREPÔTS DE DONNÉES

- Mapper les données avec le framework de programmation, se connecter aux données et les extraire de l'entrepôt de stockage, transformer les données à traiter
- Fractionner les données pour Hadoop MapReduce

7. UTILISER HADOOP MAPREDUCE

- Créer les composants des tâches Hadoop MapReduce
- Distribuer le traitement des données entre plusieurs fermes de serveurs, exécuter les tâches Hadoop MapReduce
- Surveiller l'avancement des flux de tâches

8. PRINCIPES FONDAMENTAUX DE HADOOP MAPREDUCE

- Identifier les démons Hadoop
- Examiner le système de fichiers distribué Hadoop (HDFS)
- Choisir le mode d'exécution : local, pseudo-distribué, entièrement distribué

9. SYNTHÉTISER LES TÂCHES HADOOP MAPREDUCE AVEC PIG

- Communiquer avec Hadoop en Pig Latin
- Exécuter des commandes avec le shell Grunt
- Rationaliser les traitements de haut niveau

10. LANCER DES REQUÊTES AD HOC SUR LE BIG DATA AVEC HIVE

- Assurer la persistance des données dans le Hive MegaStore
- Lancer des requêtes avec HiveQL
- Examiner le format des fichiers Hive

11. EXTRAIRE DES DONNÉES DONNANT DE LA VALEUR À L'ENTREPRISE

- Analyser les données avec Mahout, utiliser des outils de génération d'états pour afficher le résultat du traitement

12. ÉLABORER UNE STRATÉGIE DÉDIÉE AU BIG DATA

- Définir les besoins en matière de Big Data
- Atteindre les objectifs grâce à la pertinence des données
- Évaluer les différents outils du marché dédiés au Big Data
- Répondre aux attentes du personnel de l'entreprise

13. UNE MÉTHODE ANALYTIQUE INNOVANTE

- Identifier l'importance des traitements métier
- Cerner le problème
- Choisir les bons outils
- Obtenir des résultats exploitables

14. ANALYSE STATISTIQUE DU BIG DATA

- Exploiter la fonctionnalité RHadoop
- Générer des états statistiques avec RHadoop
- Utiliser la visualisation RHadoop
- Exploiter les résultats des analyses

15. METTRE EN œUVRE UNE SOLUTION BIG DATA

- Bien choisir les fournisseurs et options d'hébergement
- Trouver le juste équilibre entre les coûts engendrés et la valeur apportée à l'entreprise
- Garder une longueur d'avance

NOUS CONTACTER

Siège social
16, ALLÉE FRANÇOIS VILLON
38130 ÉCHIROLLES

Centre de formation
87, RUE GÉNÉRAL MANGIN
38000 GRENOBLE

Téléphone
04 76 23 20 50 - 06 81 73 19 35

E-mail
contact@audit-conseil-formation.com

Suivez-nous sur les réseaux sociaux, rejoignez la communauté !

ACF Audit Conseil Formation @ACF_Formation ACFauditconseilformation