Big Data - L'essentiel

De la technologie aux usages du Big Data

Présentiel ou en classe à distance

2 jours (14 h)

Prix inter : 1.990,00 € HT
Forfait intra : 3.850,00 € HT

Réf.: SEM34

Le **Big Data** est devenu un levier incontournable de transformation numérique pour les entreprises. Au-delà des usages emblématiques des géants du web, il s'impose désormais dans les organisations traditionnelles pour optimiser la **gestion de la relation client (CRM)**, renforcer la **sécurité informatique via l'analyse de logs** et développer une **vision 360° des données**.
Cette formation apporte une compréhension claire des **architectures distribuées** et des **technologies de référence comme Hadoop, HDFS, YARN, MapReduce ou les bases NoSQL**, ainsi que des nouveaux modèles **Data Lake et Lake House**. Elle explore également les enjeux de qualité et gouvernance des données (MDM, RGPD, CSA) et les apports de l'IA et du Machine Learning dans l'analyse avancée.

A qui s'adresse cette formation ?

- Responsables de la DSI s'interrogeant sur les apports et le déploiement du Big Data
- Chefs de projets, Responsables de métiers et consultants souhaitant aborder les projets
- Toute personne impliquée dans la réflexion et l'étude du Big Data

- Connaissances sommaires en informatique
- Disposez-vous des connaissances nécessaires pour suivre cette formation ? Testez-vous !

Programme

Pendant

1 - Exemples d'usage pertinent du Big Data

- Usages du Big Data chez les géants de l'Internet
- Usages du Big Data dans les entreprises traditionnelles
- Exemples de complexité technique et organisationnelle
- Gestion des clients (CRM) : Vue 360° des clients / Multicanal
- Sécurité informatiques (étude de logs) : identification des tentatives d'attaques
- Analyse des logs d'Internet (Web)
- Profiling d'individus
- Causes d'échec et d'insatisfaction des projets Big Data : confusion entre la BI évoluée, le Big Data adapté aux entreprises et le Big Data dans sa conception originelle (vélocité, véracité, valeur...)

2 - Définition commune du Big Data selon les grands acteurs du marché

- Concept et apports
- Architecture technique de l'écosystème de la chaine des valeurs d'une solution de Big Data
- Caractéristiques techniques des 3V de Gartner et les variantes d'usage (Véracité, Valeur, Validité...)
- Collecte et traitement des données structurées, semi-structurées et non-structurées
- Création et Transformation des données en informations

3 - Introduction aux architectures des solutions de calcul distribué du Big Data

- Principe
- Scalabilité horizontale et verticale
- Architecture de cluster et composants économiques
- Traitement parallèle des données
- Enjeux de sécurité des architectures distribuées, lors de l'intégration dans le système informatique des organisations

4 - Technologies de référence du Big Data à connaître

- Traitement des données par les clusters (Hadoop)
- Stockage et traitement des données dans le Data Lake, le Lake House : précautions à prendre
- Différents types de base des données NoSQL selon les usages
- Hadoop : un modèle de traitement distribué du Big Data adopté par les grands acteurs de l'informatique : HDFS, YARN, MapReduce...
- Ecosystème de Hadoop : Pig, Flume, Zookeeper, HBase, Oozie...
- Analyse de données : 4 types d'analytique
- Algorithmes et modèles d'IA et Machine Learning pour l'analyse avancée et automatisée

5 - Approches de déploiement du Big Data

- Déploiement sur site : définition des objectifs, choix des solutions d'analyse et d'intégration, présentation des informations / revue des fournisseurs de composants Big Data
- Déploiement sur site en version distribuée
- Déploiement dans les plates-formes Cloud Big Data
- Précautions à prendre (métriques de qualité, système fermé ...)
- Difficultés techniques à anticiper
- Validation de la pertinence d'une plate-forme

6 - Synthèse des plates-formes Cloud public du Big Data Analytiques

- IBM Analytics
- Amazon Web Services
- Google Cloud Platform
- Microsoft Azure
- Points communs et différents entre les plates-formes Big Data

7 - Qualité des données

- Données : la matière première pour construire le Big Data
- Création des données de qualité
- Définition et consolidation de la valeur des données
- Enrichissement des données brutes avec l'Open Data
- Gouvernance des données pour le maintient de la pertinence des données et Rôle primordial du MDM (Master Data Management)

8 - Sécurité des données et confidentialité du Big Data

- Règlementation RGPD à respecter
- Recommandation des bonnes pratiques de l'organisme international CSA (Cloud Security Alliance) pour le Big Data
- Panorama des moyens techniques de sécurité des données et d'accès au Datacenter (Cryptage, IDA et DLP...)
- Recommandations de sécurité dans la manipulation des données
- Recours aux technologies d'IA/ML et Cybersécurité pour faire face aux fonctionnalités spécifiques des plateformes de traitement Big Data

9 - Impacts du Big Data à anticiper

- Adaptation des bonnes pratiques actuelles à l'environnement Big Data
- Impacts sur les compétences des équipes informatiques en place
- Nouveaux métiers (Data Scientist, Data Steward...) et compétences nouvelles à acquérir
- Rôle de la DSI et de la Direction Numérique face à la montée du Big Data
- Progression des activités numériques de l'entreprise



Les objectifs de la formation

- Comprendre le concept du Big Data et mesurer ses apports
- Connaître l'écosystème du Big Data et appréhender les technologies associées
- Savoir anticiper son intégration dans les activités informatiques de l'entité
- Être en mesure de l'exploiter dans le respect des règles de sécurité et de confidentialité

Evaluation

Pendant la formation, le formateur évalue la progression pédagogique des participants via des QCM, des mises en situation et des travaux pratiques. Les participants passent un test de positionnement avant et après la formation pour valider leurs compétences acquises.

Les points forts de la formation

- Un panorama détaillé des acteurs et solutions du marché du Big Data
- L'expertise d'un consultant en avance de phase sur le sujet et fréquemment impliqué dans la définition et la mise en oeuvre de projets Big Data
- Le partage d'une vision objective des apports et limites du Big Data
- 80% des participants à cette formation se sont déclarés satisfaits ou très satisfaits au cours des 12 derniers mois.

Evaluation

Pendant la formation, le formateur évalue la progression pédagogique des participants via des QCM, des mises en situation et des travaux pratiques. Les participants passent un test de positionnement avant et après la formation pour valider leurs compétences acquises.

Dates et villes 2027 - Référence SEM34

Dernières places disponibles Session garantie

A distance

du 4 mars au 5 mars	du 30 sept. au 1 oct.
du 3 juin au 4 juin	du 22 déc. au 23 déc.

Paris

du 4 mars au 5 mars	du 30 sept. au 1 oct.
du 3 juin au 4 juin	du 22 déc. au 23 déc.