

Séminaire BigData : Concepts et enjeux

Référence : PYCB000

Durée : 1 jour

Certification : **Aucune**

CONNAISSANCES PREALABLES

- Avoir une bonne culture générale sur les systèmes d'information.

PROFIL DES STAGIAIRES

- Chefs de projets, architectes, et toute personne souhaitant connaître les impacts du BigData sur l'entreprise au niveau de l'organisation et des architectures.

OBJECTIFS

- Comprendre les concepts et les apports du BigData, les impacts sur l'organisation de l'entreprise.

CERTIFICATION PREPAREE

Aucune

METHODES PEDAGOGIQUES

- Mise à disposition d'un poste de travail par stagiaire
- Remise d'une documentation pédagogique papier ou numérique pendant le stage
- La formation est constituée d'apports théoriques, d'exercices pratiques, de réflexions et de retours d'expérience
- Le suivi de cette formation donne lieu à la signature d'une feuille d'émargement

FORMATEUR

Consultant-Formateur expert Bigdata

METHODE D'EVALUATION DES ACQUIS

- Auto-évaluation des acquis par le stagiaire via un questionnaire
- Attestation de fin de stage adressée avec la facture

CONTENU DU COURS

Introduction

- Le besoin : volumes importants de données, traitements optimisés de flux de données au fil de l'eau, liés aux nouvelles technologies et aux nouveaux usages
- Domaines concernés : recherche scientifique, médical, e-commerce, sécurité, prédictif, ...
- Exemples : lutte contre la criminalité, fraude, santé, ressources énergétiques
- Développement des techniques sur différents aspects : stockage, indexation/recherche, calcul
- Définitions : ETL, Extract Transform Load, CAP, 3V, 4V, ...
- Les acteurs
- Le positionnement des technologies de cloud, BigData et noSQL
- Eléments d'architecture

Stockage

- Caractéristiques NoSQL

- Les différents modes et formats de stockage- Besoin de distribution- Elasticité
- Stockage réparti : réplication, sharding, gossip protocol, hachage
- Systèmes de fichiers distribués : GFS, HDFS, Ceph
- Les bases de données : Cassandra, HBase, MongoDB, CouchBase, Riak, BigTable, ..

Calcul et restitution

- Apport des outils de calculs statistiques
- Langages adaptés aux statistiques, liens avec les outils BigData

Evolutions

- Liens vers les nouveaux métiers : Hadoop scientists, Data scientists, CDO
- Analyse des données au service de l'entreprise
- Rôle de la DSI dans la démarche BigData
- Ouverture sur l'OpenData : principe, la démarche publique, les licences

- Exemple : portail data.gouv.fr
- Les offres Saas BigData comme Google BigQuery
- Les limites. Les nouveautés annoncées