



ORGANISME DE FORMATION AUX TECHNOLOGIES ET METIERS DE L'INFORMATIQUE

Formation Big Data, état de l'art

N° ACTIVITÉ : 11 92 18558 92

TÉLÉPHONE : 01 85 77 07 07

E-MAIL : inscription@hubformation.com

L'accroissement continu des données numériques dans les organisations a conduit à l'émergence du Big Data. Ce concept recouvre les questions de stockage et de conservation de vastes quantités de données mais aussi celles liées au gisement potentiel de valeur que représentent ces masses de données. Ce séminaire présente les problèmes spécifiques du Big Data et les solutions techniques potentielles, de la gestion des données aux différents types de traitement.

Référence	BGA
Durée	2 jours (14h)
Tarif	1 990 €HT
Repas	repas inclus

Objectifs

- | Découvrir les principaux concepts du Big Data
- | Identifier les enjeux économiques
- | Evaluer les avantages et les inconvénients du Big Data
- | Identifier les principaux problèmes et les solutions potentielles
- | Identifier les principales méthodes et champs d'application du Big Data

Public

- | DSI
- | directeurs techniques
- | chefs de projet
- | architectes
- | responsables SI

Prérequis

- | connaissances de base des architectures techniques

Programme de la formation

Introduction

- | Les origines du Big Data : un monde de données numériques, l'e-santé, chronologie.
- | Une définition par les quatre V : la provenance des données.
- | Une rupture : changements de quantité, de qualité, d'habitudes.
- | La valeur de la donnée : un changement d'importance.
- | La donnée en tant que matière première.
- | Le quatrième paradigme de la découverte scientifique.

Big Data : traitements depuis l'acquisition jusqu'au résultat

- | L'enchaînement des opérations. L'acquisition.
- | Le recueil des données : crawling, scraping.
- | La gestion de flux événementiel (Complex Event Processing, CEP).
- | L'indexation du flux entrant.
- | L'intégration avec les anciennes données.
- | La qualité des données : un cinquième V ?
- | Les différents types de traitement : recherche, apprentissage (machine learning, transactionnel, data mining).
- | D'autres modèles d'enchaînement : Amazon, e-Santé.
- | Un ou plusieurs gisements de données ? De Hadoop à l'in-memory.
- | De l'analyse de tonalité à la découverte de connaissances.

SESSIONS PROGRAMMÉES

A DISTANCE (FRA)

- du 18 au 19 juin 2024
- du 1er au 2 octobre 2024
- du 26 au 27 novembre 2024

PARIS

- du 11 au 12 juin 2024
- du 24 au 25 septembre 2024
- du 19 au 20 novembre 2024

[VOIR TOUTES LES DATES](#)

Relations entre Cloud et Big Data

- | Le modèle d'architecture des Clouds publics et privés.
- | Les services XaaS.
- | Les objectifs et avantages des architectures Cloud.
- | Les infrastructures.
- | Les égalités et les différences entre Cloud et Big Data.
- | Les Clouds de stockage.
- | Classification, sécurité et confidentialité des données.
- | La structure comme critère de classification : non structurée, structurée, semi-structurée.
- | Classification selon le cycle de vie : données temporaires ou permanentes, archives actives.
- | Difficultés en matière de sécurité : augmentation des volumétries, la distribution.
- | Les solutions potentielles.

Introduction à l'Open Data

- | La philosophie des données ouvertes et les objectifs.
- | La libération des données publiques.
- | Les difficultés de la mise en oeuvre.
- | Les caractéristiques essentielles des données ouvertes.
- | Les domaines d'application. Les bénéfices escomptés.

Matériel pour les architectures de stockage

- | Les serveurs, disques, réseau et l'usage des disques SSD, l'importance de l'infrastructure réseau.
- | Les architectures Cloud et les architectures plus traditionnelles.
- | Les avantages et les difficultés.
- | Le TCO. La consommation électrique : serveurs (IPNM), disques (MAID).
- | Le stockage objet : principe et avantages.
- | Le stockage objet par rapport aux stockages traditionnels NAS et SAN.
- | L'architecture logicielle.
- | Niveaux d'implantation de la gestion du stockage.
- | Le "Software Defined Storage".
- | Architecture centralisée (Hadoop File System).
- | L'architecture Peer-to-Peer et l'architecture mixte.
- | Les interfaces et connecteurs : S3, CDMI, FUSE, etc.
- | Avenir des autres stockages (NAS, SAN) par rapport au stockage objet.

Protection des données

- | La conservation dans le temps face aux accroissements de volumétrie.
- | La sauvegarde, en ligne ou locale ?
- | L'archive traditionnelle et l'archive active.
- | Les liens avec la gestion de hiérarchie de stockage : avenir des bandes magnétiques.
- | La réplication multisites.
- | La dégradation des supports de stockage.

Méthodes de traitement et champs d'application

- | Classification des méthodes d'analyse selon le volume des données et la puissance des traitements.
- | Hadoop : le modèle de traitement Map Reduce.
- | L'écosystème Hadoop : Hive, Pig. Les difficultés d'Hadoop.
- | OpenStack et le gestionnaire de données Ceph.
- | Le Complex Event Processing : un exemple ? Storm.
- | Du BI au Big Data.
- | Le décisionnel et le transactionnel renouvelés : les bases de données NoSQL. Typologie et exemples.
- | L'ingestion de données et l'indexation. Deux exemples : splunk et Logstash.
- | Les crawlers Open Source.
- | Recherche et analyse : Elasticsearch.
- | L'apprentissage : Mahout. In-memory.
- | Visualisation : temps réel ou non, sur le Cloud (Bime), comparaison QlikView, Tibco Spotfire, Tableau.
- | Une architecture générale du data mining via le Big Data.

Cas d'usage à travers des exemples et conclusion

- | L'anticipation : besoins des utilisateurs dans les entreprises, maintenance des équipements.
- | La sécurité : des personnes, détection de fraude (postale, taxes), le réseau.
- | La recommandation. Analyses marketing et analyses d'impact.
- | Analyses de parcours. Distribution de contenu vidéo.
- | Big Data pour l'industrie automobile ? Pour l'industrie pétrolière ?

- | Faut-il se lancer dans un projet Big Data ?
- | Quel avenir pour les données ?
- | Gouvernance du stockage des données : rôle et recommandations, le Data Scientist, les compétences d'un projet Big Data.

Méthode pédagogique

Chaque participant travaille sur un poste informatique qui lui est dédié. Un support de cours lui est remis soit en début soit en fin de cours. La théorie est complétée par des cas pratiques ou exercices corrigés et discutés avec le formateur. Le formateur projette une présentation pour animer la formation et reste disponible pour répondre à toutes les questions.

Méthode d'évaluation

Tout au long de la formation, les exercices et mises en situation permettent de valider et contrôler les acquis du stagiaire. En fin de formation, le stagiaire complète un QCM d'auto-évaluation.

Suivre cette formation à distance

Voici les prérequis techniques pour pouvoir suivre le cours à distance :

- | Un ordinateur avec webcam, micro, haut-parleur et un navigateur (de préférence Chrome ou Firefox). Un casque n'est pas nécessaire suivant l'environnement.
- | Une connexion Internet de type ADSL ou supérieure. Attention, une connexion Internet ne permettant pas, par exemple, de recevoir la télévision par Internet, ne sera pas suffisante, cela engendrera des déconnexions intempestives du stagiaire et dérangera toute la classe.
- | Privilégier une connexion filaire plutôt que le Wifi.
- | Avoir accès au poste depuis lequel vous suivrez le cours à distance au moins 2 jours avant la formation pour effectuer les tests de connexion préalables.
- | Votre numéro de téléphone portable (pour l'envoi du mot de passe d'accès aux supports de cours et pour une messagerie instantanée autre que celle intégrée à la classe virtuelle).
- | Selon la formation, une configuration spécifique de votre machine peut être attendue, merci de nous contacter.
- | Pour les formations incluant le passage d'une certification la dernière journée, un voucher vous est fourni pour passer l'examen en ligne.
- | Pour les formations logiciel (Adobe, Microsoft Office...), il est nécessaire d'avoir le logiciel installé sur votre machine, nous ne fournissons pas de licence ou de version test.
- | Horaires identiques au présentiel.

Accessibilité

Les sessions de formation se déroulent sur des sites différents selon les villes ou les dates, merci de nous contacter pour vérifier l'accessibilité aux personnes à mobilité réduite.

Pour tout besoin spécifique (vue, audition...), veuillez nous contacter au 01 85 77 07 07.