Big Data : état de l’art

Lien : https://innov-systems.com/formation/big-data-etat-de-lart

DURÉE
3 jours (21h)

RÉFÉRENCE
BSI03

CATÉGORIE
Big Data : fondamentaux et architecture

Innov Systems

OBJECTIFS DE LA FORMATION

À l'issue de cette formation, vous serez capable de :

- Connaître l'ensemble des concepts du Big Data
- Appréhender les enjeux économiques
- Connaître les avantages, les limites et les inconvénients du Big Data
- Identifiez les nouvelles problématiques liées au Big Data
- Identifier l’écosystème et comprendre les technologies associées
- Anticiper son intégration dans les activités informatiques de l’entreprise
- Exploiter dans le respect des règles de sécurité et de confidentialité

POUR QUI ?

- DSI
- Directeurs techniques
- Chefs de projet
- Architectes
- Responsables SI

Programme détaillé

1 / Comprendre le Big Data

- Big Data, l'Analytics, la Data Science : définitions
- Concepts fondamentaux et caractéristiques clés du Big Data
- Big Data : origines
- La source des données
- 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
- 3V de Gartner (Vélocité, Variété et Volume) et les variantes (Véracité, Valeur, Validité...)
- Collecte et traitement des données structurées, semi-structurées et non-déstructurées
- Transformation des données en informations
- Création de la valeur à partir des données
- Différences et apports par rapport aux technologies traditionnelles
- Exemples d'application et usages des technologies Big Data

2 / Big Data : exemples d'usage

- Réseaux sociaux : Google, Twitter, Youtube
- Gestion des clients (CRM) : Vue 360° des clients / Multicanal
- Sécurité informatiques : identification des tentatives d'attaques
- Analyse des logs d'Internet (Web)
- Profiling d'individus : ADN numérique
- Les projets Big Data : critères de succès et causes d'échec

3 / Architectures des solutions de calcul distribué du Big Data : introduction

- Principe

Inov Systems

- Scalabilité horizontale et verticale
- Architecture de cluster et composants économiques
- Traitement parallèle des données
- Enjeux des architectures distribuées selon l'organisme CSA

4 / Stockage

- Sources de données publiques, privées, la philosophie Open Data
- Classifier les données selon leur structure
- Vue d'ensemble des modes et architectures de stockage
- Stockage des données Data Lake
- Différents types de base des données NoSQL : MongoDB, Cassandra, Neo4j…

5 / Méthodes de traitement du Big Data

- Les types de traitement : recherche, apprentissage (Machine Learning, transactionnel, data mining, Mahout)
- Ecosystème de Hadoop : Pig, Flume, Zookeeper, HBase, Oozie …
- Principes de fonctionnement du modèle MapReduce
- OpenStack et le gestionnaire de données Ceph
- Le Complex Event Processing : un exemple, Storm
- Du Business Intelligent (BI) au Big Data
- Décisionnel et le transactionnel renouvelés : les bases de données NoSQ, typologie et exemples
- Ingestion de données et l'indexation (exemples : Splunk et Logstash)
- Crawlers Open Source
- Recherche et analyse : Elasticsearch
- Une architecture générale du data mining via le Big Data

6 / Visualisation des données

- Principes fondamentaux et objectifs de la dataviz
- Limites des outils et méthodes traditionnels (Excel, BusinessObjects, etc)
- Spécificités du Big Data

Innov Systems

- Panorama des solutions de datavisualisation : Tableau, Qlik, Microstrategy, Microsoft Power BI…
- Quels types de visualisation pour quels usages ? comparaison QlikView, Tibco Spotfire, Tableau
- La visualisation interactive pour représenter les données complexes
- Raconter une histoire avec les données : Data Storytelling

7 / Outils et techniques pour l'analyse de données massives

- Analyse statistique au Deep Learning
- Caractéristiques de l'Analytics Big Data, spécificités
- Qu'est-ce que le Data Mining ? Relations avec la Business Intelligence
- Sélectionner les bons outils suivant la structure de données
- Analyse de données avec Pig et Hive, Apache Storm, Apache Spark
- ElasticSearch pour la recherche de données
- Big Data et intelligence artificielle, une convergence révolutionnaire
- Machine Learning : vue d'ensemble des technologies et algorithmes d'apprentissage automatique
- Spécificités du Machine Learning pour le Big Data et domaines d'application
- Introduction aux réseaux neuronaux (deep learning)

8 / Cloud et Big Data

- Clouds publics et privés
- Services XaaS
- Objectifs et avantages des architectures Cloud
- 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

Innov Systems

9 / Open Data

- Philosophie des données ouvertes et les objectifs
- Libération des données publiques
- La mise en œuvre : les difficultés
- Caractéristiques essentielles des données ouvertes
- Domaines d'application
- Les bénéfices escomptés

10 / Sécurité et Big Data

- Les données personnelles
- Les informations sensibles, interdites de collecte
- La CNIL régule les données numériques
- Les accords intra-pays

11 / Cas d'usage

Approche pédagogique

- Support Ecrit et Projection
- Exposés Interactifs, Podcasts et Vidéos
- Brainstorming et Jeux de Rôle
- Cas Pratiques et Labs inclus pour leur impact opérationnel
- Test de Validation des Acquis des Connaissances

Innov Systems

Prochaines dates programmées

06 au 08 Jan. 2027	Présentiel - Casablanca
10 au 12 Mars 2027	Présentiel - Casablanca
05 au 07 Mai 2027	Présentiel - Casablanca
30 Juin au 02 Juil. 2027	Présentiel - Casablanca
01 au 03 Sep. 2027	Présentiel - Casablanca

Autres dates possibles sur demande. Contactez-nous pour organiser une session intra-entreprise.

Réservation & Renseignements

Téléphone : +212 522 247 210
Email : contact@innov-systems.com
Web : https//www.innov-systems.com

Innov Systems — Tous droits réservés

Innov Systems