

retengr

Formation Big Data, une journée pour comprendre

Durée : 1 jour – Réf : PRESENTATION-BIGDATA



Présentation

Limité au monde des GAFAs il y a quelques années encore, la gestion des données volumineuses est une préoccupation pour toutes les entreprises, quel que soit leur domaine d'activité. Le Big Data est un vaste sujet : stratégique, organisationnel, et technique, il requiert des adaptations importantes à la fois dans les technologies mais aussi dans les esprits. Ces arbitrages demeurent complexes, surtout face à un sujet dont l'étendue et la complexité en rendent une lecture difficile et souvent pleine de fausses idées.

Cette conférence se décompose en deux parties distinctes : la matinée est consacrée à une présentation générale du Big Data, et l'après-midi à un tour d'horizon des technologies du Big Data.

Objectifs

Première partie : le matin

- Connaître les origines du BigData
- Découvrir les concepts qui se cachent derrière le terme BigData
- Déterminer les usages pertinents au travers d'exemples
- Définir les étapes qui jalonnent la mise en œuvre d'une approche BigData
- Mesurer l'impact du BigData sur notre quotidien
- Évaluer la relation entre BigData et Intelligence Artificielle

Deuxième partie : l'après-midi

- Dégrossir le panorama technologique des outils autour du BigData
- Utiliser nifi et/ou node-red pour automatiser vos workflows de collecte
- Pourquoi utiliser Hadoop comme infrastructure de stockage et de traitement ?
- Comprendre l'usage des frameworks Spark et Dask
- Appréhender les différentes approches en Machine Learning et Deep Learning



- 40 rue des Filatiers 31000
- Organisme de formation : 73 31 08023 31
- SIRET : 812 673 903 00017

Audience

DSI, Architectes, Développeurs

Pré-requis

La connaissance d'un langage de programmation et du monde des bases de données est nécessaire.

Le formateur

Le formateur est un expert du domaine qui intervient sur le sujet depuis plusieurs années en formation mais aussi en conseil.

Doté d'une grande qualité d'écoute, sa pédagogie et sa compétence technique vous permettront d'acquérir une vue d'ensemble sur le Big Data.

Tarif

790 euros HT



Programme

Première partie : le matin

Cette présentation utilise un fil rouge ludique et dont la compréhension est à la portée de l'ensemble de l'audience (aucun prérequis technique). Les différentes étapes d'un projet Big Data ainsi que les outils y afférent seront expliqués en s'appuyant sur ce cas concret.

Origine des données

- Définir les différentes sources de données disponibles
- Données structurées / non structurées
- Comprendre les différents moyens d'y accéder
- Confidentialité des données, RGPD

Collecte

- Réflexion sur les besoins : origine, fiabilité, fréquence, coût
- Les différents formats de données
- Nettoyage et consolidation
- Problématiques techniques

Stockage

- Les limites des modèles actuels
- Le principe du datalake
- Les nouveaux modèles NoSql et NewSql
- Problématiques techniques du stockage

Traitement

- Comment traiter de gros volumes de données
- Paralléliser pour aller plus vite



Deuxième partie : l'après-midi

Collecte de données

- Nifi
- Node-red

Les nouveaux moyens de stockages

- Les nouveaux types de bases de données : key / value, document, wide-column
- MongoDB : une base de données de type document
- Scalability & Sharding
- Hadoop : Un framework de stockage et de traitement en parallèle

Hadoop

- Le framework BigData
- Les piliers : HDFS, Yarn et Mapreduce
- HDFS : architecture, découpage en block, réplication
- YARN : le chef d'orchestre des tâches Hadoop
- Mapreduce : un template de programmation pour les traitements en parallèles
- L'écosystème Hadoop : Hive, impala, ...

Spark

- Un autre framework pour paralléliser les traitements
- Plusieurs langages : Python, Scala, Java
- RDD et DAG : les principes fondamentaux de Spark
- Dataframe et Dataset : structuration et typage des données
- Infrastructure d'exécution : Spark sur Yarn, mesos

*

Intelligence Artificielle

- Machine learning et Deep Learning
- Les librairies : Scikit learn et sparkml



- 40 rue des Filatiers 31000
- Organisme de formation : 73 31 08023 31
- SIRET : 812 673 903 00017

- Régression | Classification | Clustering
- Analyse, nettoyage et préparation des données
- Apprentissage et scoring

The background is a vibrant, abstract composition of overlapping shapes and patterns. It features a large yellow shape on the left, a blue shape at the top, and a red shape on the right. A pattern of yellow dots is visible on the right side. In the bottom right, there are overlapping shapes in light pink, teal, and purple. A dashed blue line with a circular loop is located in the bottom left area.

Vous allez nous adorer si
comme nous vous pensez que...

Une formation doit être au service de la performance du collaborateur et de l'entreprise

Ceci nécessite une quête constante d'excellence de la part de l'organisme formateur avec une adaptation systématique aux enjeux de l'entreprise, la mise à jour régulière des supports de cours et une veille technologique indispensables pour toujours être à la pointe du domaine.



L'expertise technique est aussi importante que les qualités pédagogiques



Nos formateurs sont tous des experts de leur domaine. Mais qu'ont-ils de plus que les autres ? Nous les sélectionnons en plus pour leurs qualités de pédagogue et leurs méthodes d'enseignements. Nous plaçons les qualités pédagogiques au même niveau que l'expertise afin que nos stagiaires tirent le meilleur de leurs formations.

L'excellence naît de l'excellence

Beaucoup de nos clients se classent parmi les leaders de leurs industries respectives ou parmi les start-ups les plus prometteuses. Nous savons que former les collaborateurs de telles entreprises nécessite de prêter attention à chaque détail en prodiguant un accompagnement à la hauteur de l'ambition de nos stagiaires. C'est pourquoi nous savons faire des leaders d'aujourd'hui les champions de demain !



The logo for 'retengr' is displayed in a white, lowercase, sans-serif font. The letter 'r' is stylized with a circular dot at its top and a horizontal bar at its base. The background features abstract, overlapping shapes in various shades of red and pink, along with a pattern of small white dots on the right side.

retengr

Faire du leader
d'aujourd'hui, le champion
de demain

www.retengr.com

05.82.95.70.19

36 rue d'Alsace-Lorraine, 31000 Toulouse