

CONCEVOIR ET PILOTER UN PROJET BIG DATA

3 jours en présentiel (21 heures)

Objectifs pédagogiques

Cette formation Concevoir et piloter un projet Big Data vous permettra de : Comprendre et traiter les spécificités d'un projet Big Data au sens organisationnel, méthodologique, technologique, économique, juridique et humain, composer et piloter une équipe Big Data, manager le processus de mise en place d'un projet Big Data, identifier les besoins et le type de données à traiter avec les métiers (use cases), organiser la collecte et le stockage des données, déterminer l'exploitation des données, mettre en œuvre l'analyse des données, sélectionner une méthode de visualisation des données, piloter et maîtriser les risques des projets Big Data, réaliser des tests et analyses de performances, consolider ses connaissances à travers un cas d'usage.

Population visée

Cette formation Concevoir et piloter un projet Big Data est destinée aux concepteurs Big Data, Data Stewards.

Pré-requis

Cette formation Concevoir et piloter un projet Big Data nécessite une expérience de pilotage de projets BI et de la gouvernance de données.

Méthodes pédagogiques

1 poste et 1 support par stagiaire 8 à 10 stagiaires par salle 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

Formateur

Formateur consultant expert en BIG DATA

Modalités de validation des acquis

Auto-évaluation des acquis par le stagiaire via un questionnaire en ligne Attestation de fin de stage remise au stagiaire

Contenu

COMPRENDRE ET TRAITER LES SPÉCIFICITÉS D'UN PROJET BIG DATA AU SENS ORGANISATIONNEL, MÉTHODOLOGIQUE, TECHNOLOGIQUE, ÉCONOMIQUE, JURIDIQUE ET HUMAIN

- Bâtir une vision Data Centric pour l'entreprise
- Etudier l'environnement concurrentiel de l'entreprise
- Comment créer de la valeur ou apporter de la valeur complémentaire aux données
- Comment utiliser les Big Data qui doivent être un levier technologique pour accompagner les enjeux métiers et non l'inverse
- Comprendre les acteurs du Big Data et leur positionnement
- Quelle Gouvernance mettre en place
- Quel impact dans la relation DSI / Métiers
- L'émergence de la culture DevOPS
- Les nouvelles méthodologies de développement
- Découvrir le Visual Thinking
- Les impacts organisationnels
- Comment positionner le Big Data face à l'existant ?
- Quelles sont les possibilités offertes par le Big Data ?
- La gestion des données personnelles
- Surveillance et sanctions de la CNIL
- Les nouveaux profils du Big Data

COMPOSER ET PILOTER UNE EQUIPE BIG DATA

- Définition des profils en fonction des typologies de projet
- Les nouveaux profils du Big Data
- L'architecte des données
- Le Data Steward
- L'administrateur Hadoop
- Le Data Scientist
- Le Data Analyst
- Les développeurs Java Hadoop
- Les développeurs apache Spark

MANAGER LE PROCESSUS DE MISE EN PLACE D'UN PROJET BIG DATA

- Le cadrage du projet
- Cadrage métier identification des cas d'usages
- Cadrage technique identification des cas d'usage
- Identification des sources de données
- Evaluation des besoins technologiques
- Estimation budgétaire et planification projet
- Management de projet Big Data
- Définition de la méthodologie de gestion de projet
- Mise en place de la structure de gouvernance du projet
- Définition de la liste des livrables
- Traitement des obligations juridiques: CNIL, propriétés intellectuelles
- Compétences & organisation des projets Big Data
- Identification des compétences nécessaires
- Mise en place des différentes équipes projet (profils DSI, experts métiers, Data Scientists et Data Steward)
- Mise en place des outils de pilotage
- Déploiement des outils de data preparation
- Mise en place des outils d'alimentation
- Modélisation de la structure d'accueil des données au sein du Data Lake(landing area)
- Mise en place des outils de développements (externe à la plateforme Big data si nécessaire)
- Mise en place des outils d'exposition des données

PILOTER ET MAITRISER LES RISQUES DES PROJETS BIG DATA

- Les enjeux règlementaires et la conformité : la dimension "Gouvernance" du projet
- La disponibilité effective des données (en regard des différents contrats souscrits auprès des fournisseurs de données)
- La maîtrise du format des données

CONCEVOIR ET PILOTER UN PROJET BIG DATA

- La disponibilité des acteurs en charge de vous fournir les données
- Anticiper la volumétrie attendue
- Savoir correctement modéliser la structure d'accueil des données
- Avoir une bonne vision des traitements à effectuer et leur performance en regard de la volumétrie attendue
- Choisir avec soin son outil de modélisation prédictif (capable de gérer une forte volumétrie)
- Maîtriser sa solution de datavisualisation et s'assurer de sa compatibilité avec le Big Data

TESTS ET ANALYSES DE PERFORMANCES

- Tests de performances. Objectif : Tests sur un ou plusieurs scénarios sous une charge modérée du système complet et/ou d'un sous-système nécessitant un point d'attention
- Exemple : La souscription est testée pour 1 utilisateur et, pour chaque étape du use case, on mesure le temps passé dans les différents composants de l'application
- Test de vieillissement. Objectif : Déterminer la capacité de l'application à fonctionner sur une période étendue
- Exemple : On simule l'utilisation de l'application pendant 48h, avec une charge constante et égale à la charge moyenne
- Test de charge. Objectif : Mesurer la tenue en charge de l'application sur la population cible
- Exemple : On simule l'utilisation de l'application par 200 utilisateurs (avec des scénarios différents) en parallèle pendant 2h
- Test de rupture / Stress test. Objectif : Déterminer les limites de l'application
- Exemple : On augmente le nombre d'utilisateurs en parallèle sur l'application jusqu'à ce que le taux d'erreurs / les temps de réponse ne soient plus acceptables
- Choix des scénarios
- Choix des métriques
- Choix de l'outillage

ETUDES DE CAS / MISES EN SITUATION

- Mission de choix d'outil et montée en compétence
- Déploiement de distributions Hadoop différente et réalisation d'un test de performance