



Formation professionnalisante de niveau 7
(demande d'enregistrement de la certification au RNCP en cours)

Data Engineer



Au cœur de la stratégie Big Data, votre mission est de collecter, stocker, et traiter les données, afin de les mettre à disposition des analystes. Conscient des enjeux de la data pour l'entreprise, vous développez et mettez en place l'infrastructure technique et les solutions IT adaptées aux modèles de traitement de la data et répondant aux besoins des métiers.

Vous êtes le maillon indispensable entre l'architecte data et le data scientist, et permettez ainsi aux analystes d'accéder à des données fiables, sécurisées et de qualité.

Objectifs de la formation

- Alimenter et préparer les données pour les projets Data,
- Manipuler/exploiter les données présentes dans les différents types de bases de données,
- Automatiser les traitements de données sur une infrastructure Big Data,
- Industrialiser les traitements de données.

A l'issue de cette formation, vous pourrez intégrer des postes de « Data Engineer », « Ingénieur de données », ou « Ingénieur Big Data » dans les moyennes et grandes entreprises de tout secteur.



Durée
445 H
prévues



Lieu
Distanciel

Contacts : contact@academiedunumerique.net

DATA ENGINEER

Compétences visées

Analyse des besoins métier et mise œuvre d'une stratégie Big Data

- Veille technologique et stratégique : prise en compte des besoins métiers pour choisir l'environnement technologique adapté
- Analyse des besoins et cadrage de projet : conception des spécifications de la plateforme data
- Gouvernance des données : mise en œuvre de la politique de données, en matière de qualité, de sécurité et de traçabilité des données.

Collecte et stockage des données

- Mise en œuvre d'une méthodologie de collecte adaptée et développement des data pipelines pour la collecte et la transformation des données : développement d'une solution technique, manipulation des données collectées, sécurisation des données
- Choix et déploiement d'une solution de stockage, sécurisation du stockage.

Déploiement d'une solution Big Data et maintenance des développements

- Déploiement d'une plateforme Big Data avec une approche DevOps et la méthode Agile : application des concepts agiles dans le développement, automatisation des phases d'intégration et de livraison selon la méthode DevOps, conteneurisation et orchestration des conteneurs
- Maintenance évolutive et corrective

Traitement et exposition des données

- Préparation de données de qualité, fiables et sécurisées, à destination des data scientists, data analysts et ingénieurs IA : développement des traitements en temps réels, automatisation des flux...
- Mise à disposition des données pertinentes afin d'aider à la prise de décision des directions métiers : organisation de l'exposition des données avec les outils adéquats.





Programme détaillé

Module	Durée par module (en h)
Accueil : 1er jour de formation	7
Maîtrise du domaine du Big Data	21
Maîtrise de l'IA et la Dataviz	14
Maîtrise des fondamentaux du Cloud	14
Maîtrise du système Linux	30
Conteneurisation (Docker, Kubernetes)	8
Maîtrise de la Data gouvernance	7
Cybersécurité et protection des données (RGPD)	7
Manipulation de BD relationnelles	33
Développer en langage Python	30
Traitement de données avec le langage Scala et Spark	41
Collecte et traitement de données (Apache NIFI)	15
Collecte et traitement de données (DBT)	15
Manipulation de BD non relationnelles	16
Le moteur de recherche Elasticsearch	20
Manipulation de BD orientées objet	15
Manipulation de donnée en temps réel (Spark Streaming)	17
Manipulation de donnée en temps réel (Kafka)	25
Data Ops (Déployer une plateforme Data...)	26
Les fondamentaux des méthodes Agiles	14
Préparation du projet	70

DATA ENGINEER

Evaluations & validation des compétences

- Mise en pratique des compétences lors de travaux pratiques évalués tout au long du parcours, et projet final évalué en fin de formation,
- Dossier technique présentant le projet final permettant de valider l'ensemble des compétences visées par le titre,
- Soutenance devant un jury de professionnels, afin de démontrer la maîtrise de l'ensemble des compétences visées par le titre.

Public cible

- Vous disposez d'un niveau VI (bac+3/4) scientifique ou technique, avec une première expérience professionnelle
- Vous disposez d'un niveau VII (bac+5) scientifique ou technique
- Vous avez un autre profil, que nous étudierons sur dossier

Pré-requis

- Vous avez une première expérience du développement informatique (tout langage), et des notions sur Linux
- Vous avez une appétence pour le développement dans le domaine de la data
- Vous avez une bonne compréhension de l'anglais écrit

Conditions d'inscriptions

- Candidature en ligne avec CV
- Tests de positionnement en anglais, logique (algorithmie), et programmation
- Entretien de motivation

.....
Cette formation est accessible aux personnes en situation de handicap.



Méthodes pédagogiques

- Sessions collectives en présentiel ou en distanciel
- Modules e-learning
- Travail en groupe et coaching sur projets, travaux pratiques et prise en main de solutions technologiques



Moyens pédagogiques

- Supports de formation accessibles sur la plateforme de e-learning
- Cours en présentiel ou en distanciel avec des outils collaboratifs

De nombreux aménagements pédagogiques sont possibles pour accueillir les candidats concernés. Pour toute question, se rapprocher du référent handicap de l'Académie du Numérique :

*Isabelle Belzanne-Blanche
isabelle.belzanne@atos.net*





Dates des sessions



- Se reporter au site internet de l'Académie du Numérique pour les dates de formation : www.academiedunumerique.net

Lieux de formation



- Formation en distanciel,
- Nécessité d'une bonne connexion internet.

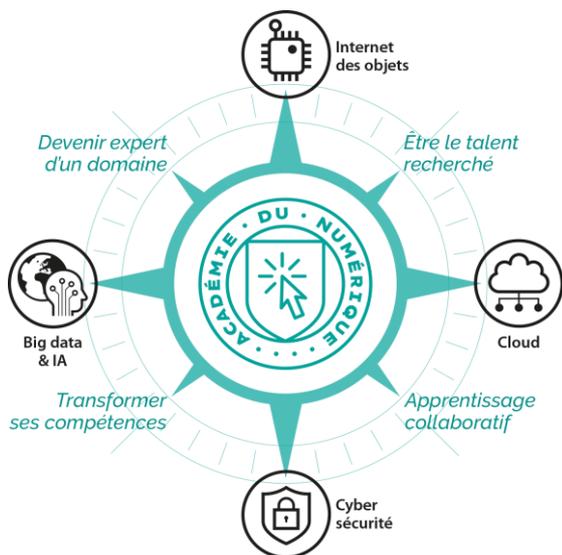
Contacts

Pour toute candidature, vous pouvez nous faire parvenir votre CV À l'adresse suivante : contact@academiedunumerique.net



Prix public de la formation

- 12 500 euros HT (tarif 2023, hors frais de restauration) ,
- Financement de votre formation : nous contacter.



Les 4 points forts de votre formation

1. Une formation au plus près des besoins métiers du domaine de la Data et de l'IA
2. Un parcours à l'état de l'art conçu par des experts de la Data
3. Une grande diversité de technologies abordées
4. Une formation centrée sur la pratique et l'entraînement, avec des cas d'usages concrets et récents, pour être immédiatement opérationnel



Projet soutenu par l'Etat dans le cadre du dispositif « Ingénierie de formations professionnelles et d'offres d'accompagnement innovantes » de France 2030 opéré par la Caisse des Dépôts (Banque des Territoires).

