



Présentation Métier **Ingénieur Données / Data Engineer**

21/10/2021

Edouard LEICHT

Diplôme Universitaire (M2)



- Le **Big Data** (ou *données massives*) est né de l'**explosion quantitative** des données numériques souvent redondantes.
 - Problèmes associés aux “**3V**” : *Volume, Variété, Vélocité*
 - De l'information **accessible** et **exploitable** : *Cloudification*
 - La donnée au **coeur des décisions**
- De **nombreux métiers liés** au domaine:
Data Engineer, Data Analyst, Data Architect, Business Intelligence Engineer, et encore tant d'autres !



Formation initiale

Jeunesse - G33k du matériel informatique (*montage pc*)

2011 - Baccalauréat scientifique

2014 - CPGE PCSI / PSI (*de belles années !*)

2015 - Licence de Physique Appliquée (*Pythonista*)

2016 - Master 1 de Physique Appliquée

2017 - Master 2 d'Ingénierie des Systèmes Spatiaux

et ensuite...

- Des formations en autodidacte
- Des lectures, des rencontres entre les métiers, etc...

La **formation** est nécessaire **à tout âge**, même après le diplôme.



Stagiaire Data Scientist



- ❑ Évaluation d'un algorithme de Machine Learning en Python
- ❑ Utilisation des ressources du Worldwide LHC Computing Grid

Ingénieur Système / Ingénieur Données



- ❑ Support opérationnel du Flight Management System
- ❑ Utilisation de Skywise, plateforme de maintenance prédictive

Product Owner / Ingénieurs Données



- ❑ Digitalisation des procédures d'essais en vol
- ❑ Automatisation de l'ingestion des données sur Skywise

Ingénieur Données



- ❑ Développement de la plateforme des comportements d'achat en Scala
- ❑ Utilisation des services Amazon Web Services



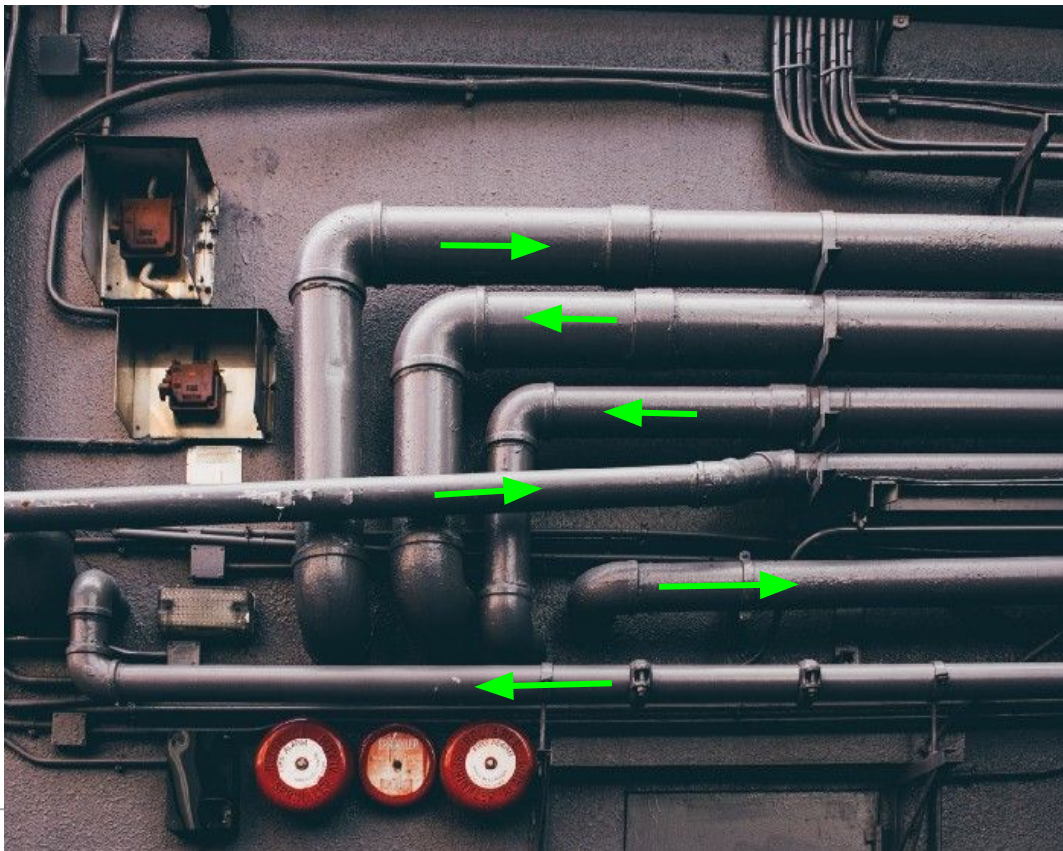
4 missions

- **Concevoir l'architecture technique** nécessaire pour la valorisation de données
- **Fournir l'expertise technologique** nécessaire pour développer les solutions data appropriées aux différents cas d'usage data émanant des unités métiers de l'entreprise
- **Effectuer les croisements de données** nécessaires ainsi que les travaux de validation, correction, qualité, nécessaires pour soutenir le travail des data scientists en aval.
- **Effectuer des analyses décisionnelles**

[Source](#)



Pipeline



Architecture

Surveillance

Ingestion



Stockage | Ingestion | Analyse | Machine Learning | Inférence (Viz.)

Opérationnel:

- **Acheminer** vers un “centre de dépôt” (Datalake, Warehouse, etc...)
- **Rendre accessible** aux consommateurs (utilisateurs)
- **Déployer** les modèles statistiques (dont Machine Learning)

Technique:

- ☐ Réaliser les flux en suivant l’architecture proposée (ex: ETL)
- ☐ Traiter les ensembles de données avec des **framework/libs** de calcul pour la plupart distribué (ex: Hadoop/Spark)
- ☐ Accélérer la mise en production grâce aux services **cloud** (ex: AWS)
- ☐ Automatiser les calculs, algo de ML et les déploiements (ex: CI/CD)

A garder constamment en tête:

Le **système** doit être **scalable**, **maintenable** et **évolutif**.

L’ingénieur données est le **garant** de ses **performances**!



Aptitudes comportementales pour le poste : Ingénieur Données / Data Engineer



**Aimer
ranger**

... à automatiser !

(sa chambre aussi!)



**Faire des compromis
... et des maths !**



(sa chambre aussi!)

Satisfaire les utilisateurs



... et bien sûr aimer coder !

Qu'apporte le métier: Ingénieur Données / Data Engineer ?

Satisfaction d'intégrer des projets variés

Capacité d'aider à l'échelle d'une organisation

Être en phase avec l'ère du temps

Data everywhere !



Quelles sont les perspectives d'évolution du métier: Ingénieur Données / Data Engineer?

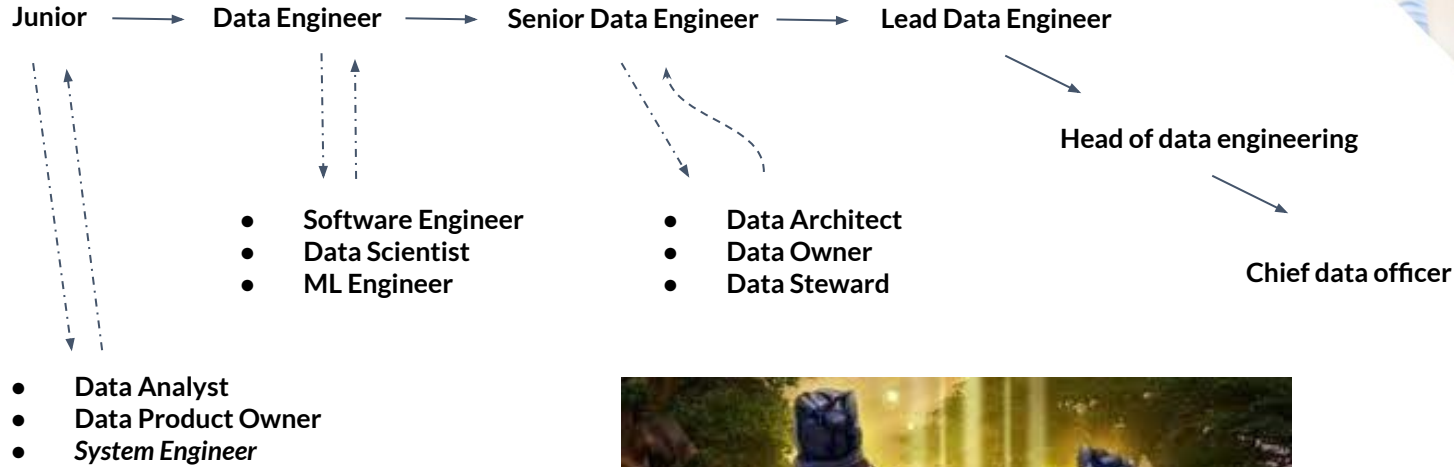


Photo by Blizzard World Of Warcraft



Questions/Réponses