|  |  |  |  |
| --- | --- | --- | --- |
|  | |  | | --- | | **Université Moulay Ismail**  **Faculté Des Sciences et Techniques Errachidia** | |  |

Une image contenant texte

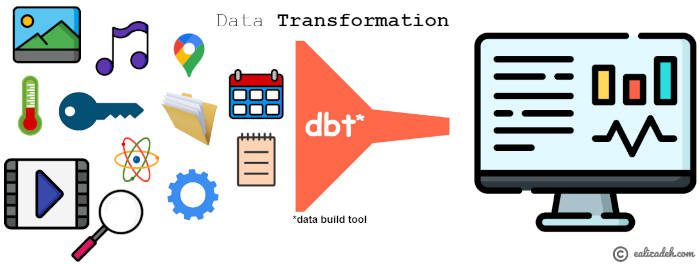
Description générée automatiquement

***Master : Système d’information décisionnel et imagerie***

***Module : Big Data***

**Rapport de Mini projet intitulé :**

***DATA TRANSFORMATION DBT***



|  |  |
| --- | --- |
| **Réalisé par :**   * **SABIRI MOHAMED** | **Encadré par :**   * **Prof. AMINE BENHAMZA** |

**Année Universitaire 2021/2022**

Table des matières

[**I. Introduction 3**](#_Toc94952878)

[**II. Définitions 3**](#_Toc94952879)

[**III. Informations général 3**](#_Toc94952880)

[**IV. Fonctionnalités Principales 3**](#_Toc94952881)

[**V. Connectivité 3**](#_Toc94952882)

[**VI. Déploiement 3**](#_Toc94952883)

[**VII. Les Pratiques 3**](#_Toc94952884)

[**VIII. Feuille Route 3**](#_Toc94952885)

## **Introduction**

De nos jour la recherche et l’exploitation des outils d’analyse des base données NoSQL ou Big Data constitue un point très importante pour les analystes dans le processus ETL ou ELT (Extraction, Transformation, Chargement de données), alors que dans chaque un de ces opération nous somme dans l’obligation d’intégration des outils pour garder la fiabilité des résultats ainsi que la rapidité de traitement de données. Dans ce sujet je vais aborder l’outils BDT (Data Build Tool) outils utilisé dans l’étape de transformation des données dans les entrepôts de données.

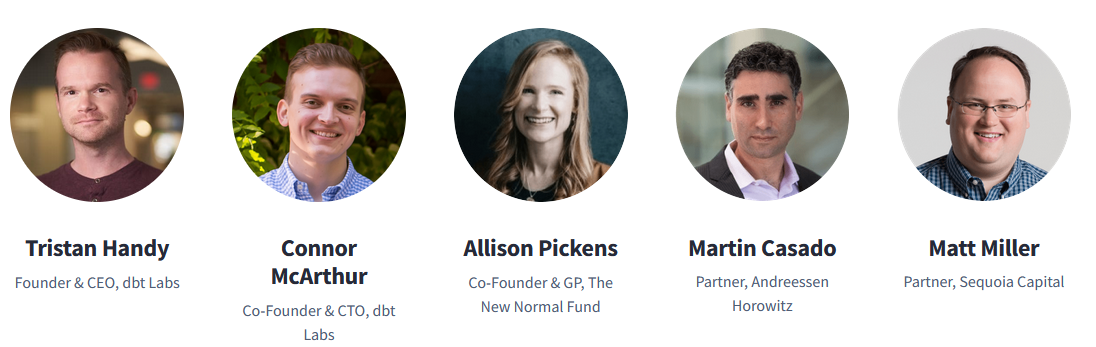
## **Définitions**

* Data Build Tool

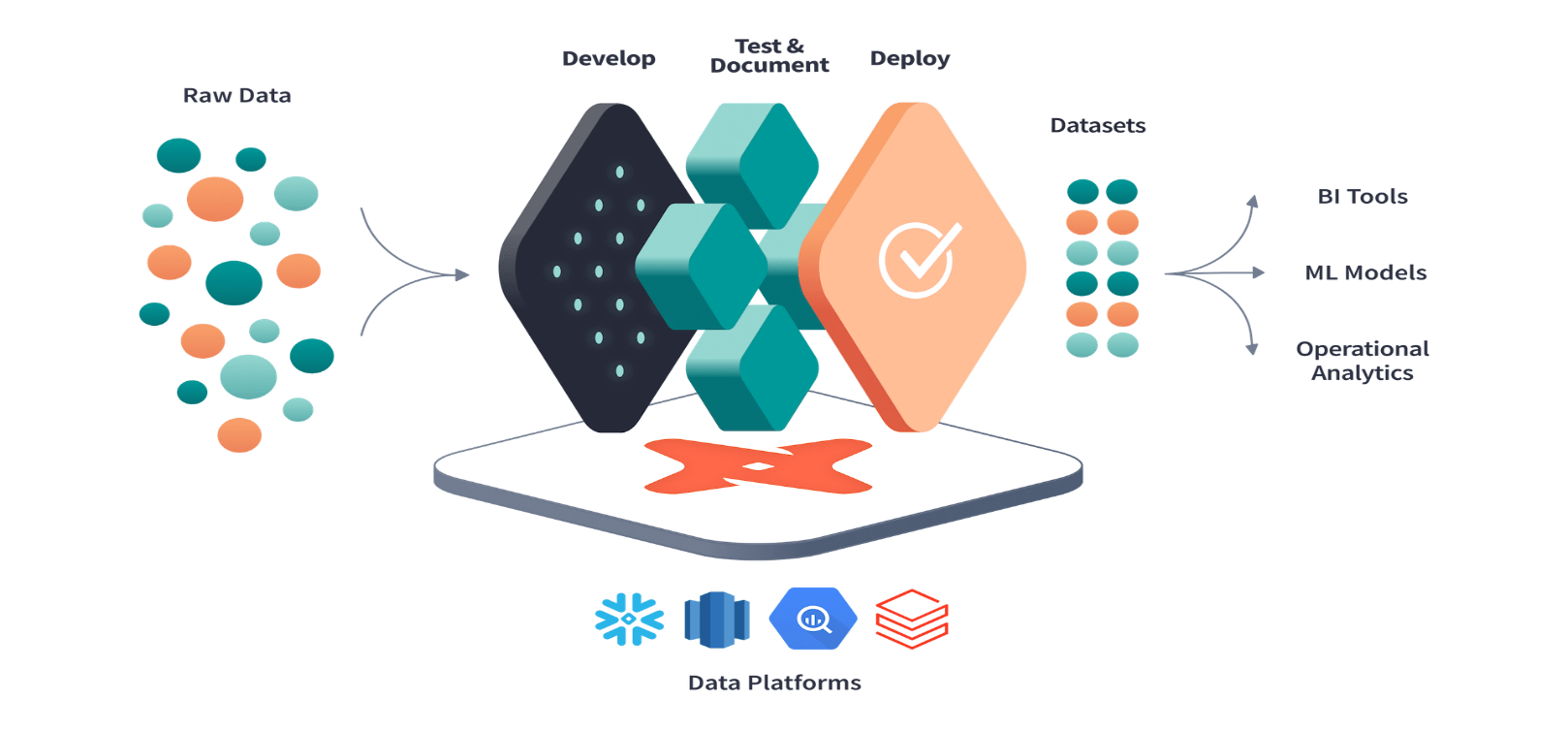
DBT outils permettant les Analytics et les ingénieurs d'analyse de faire la transformation les données dans leurs entrepôts en écrivant simplement des instructions de sélection. il gère la transformation de ces instructions select en tables et vues, aussi effectue les processus T dans ELT(Extract, Load, Transform) par contre il n'extrait ni ne charge de données, mais il est extrêmement efficace pour transformer des données déjà chargées dans les entrepôt.

## **Informations générales**

* **Historique**
* **2016 : commencé chez** [**RJMetrics**](https://en.wikipedia.org/wiki/RJMetrics)
* **2018 : acquis par Talend, avec lancement du 1ere produit commercial**
* **2020 : En avril 2020, dbt Labs a annoncé sa série A dirigée par Andreessen Horowitz, En novembre, dbt Labs a annoncé sa série B dirigée par Andreessen Horowitz et Sequoia.**
* **2021 : dbt Labs a relevé sa série C dirigée par Altimeter, Sequoia et Andreessen Horowitz.**
* **Site Internet**
* **https://docs.getdbt.com/**
* **Dépôt GitHub**
* **https://github.com/dbt-labs/**
* **Taille de la communauté**
* **Octobre 2018 : mise à jour du nombre d'utilisateurs dbt au début du post de 100 à 280**
* **Septembre 2019 : mise à jour du nombre d'utilisateurs dbt au début de la publication de 280 à 850**
* **Mai 2021 : nombre d'utilisateurs dbt mis à jour au début de la publication à 5000**
* **Fondateurs**



## **Fonctionnalités Principales**



Dans le processus ETL le DBT est utilisé pour objectif de faire la transformation de données, et comme illustrer dans le workflow, on distingue les fonctionnalité suivantes qui sont déduite de trois couches constituant le model DBT :

* **Develop : développer**

Écrivez des modèles SQL modulaires avec des instructions SELECT et la fonction ref() - dbt gère la corvée de gestion des dépendances.

* **Test & Document : Tester et documenter**

Testez chaque modèle avant la production et partagez la documentation générée dynamiquement avec toutes les parties prenantes des données.

* **Deploye : Contrôle de version et CI/CD**

Déployez en toute sécurité à l'aide d'environnements de développement. Le contrôle de version activé par Git permet la collaboration et un retour aux états précédents.

## **Connectivité et integration**

## **Déploiement**

## **Les Pratiques**

## **Feuille Route**

## **Références**

https://docs.getdbt.com/docs/introduction