

# Data Engineer

**Djouldé BARRY**

📍 Saint-Denis (93210)

☎ 07 53 22 39 38

✉ [barrydjoulde15@gmail.com](mailto:barrydjoulde15@gmail.com)



## SAVOIR-FAIRE

---

### Compétences fonctionnelles :

- Modélisation et gestion de base de données (PostgreSQL, SQL Server, etc.).
- Développement et automatisation de pipeline ETL/ELT.
- Optimisation et sécurisation des flux de données.
- Reporting et aide à la décision.
- Collaboration en méthodologie Agile et DevOps.

**Secteurs d'intervention :** Tout type de secteur

## COMPÉTENCES TECHNIQUES

---

**OUTILS :** GitlabCI, GithubActions, Jira/Confluence, Jenkins, PowerBI, Looker, Airflow , Dagster, Nifi

**LANGAGES / AGL :** Python, PySpark, PL/SQL, NoSQL

**CLOUD:** GCP, AWS, DATABRICKS, DBT

**SYSTÈMES :** Windows, Linux/Mac

**SÉCURITÉ / RÉSEAUX :** ELK, Kafka, Grafana, Prometheus , Docker

**BASES DE DONNÉES / STOCKAGE :** PostgreSQL, SQL Server, Oracle, ClickHouse, InterSystems IRIS, MongoDB S3, Minio

**METHODES :** Agile (Scrum&Kanban), DevOps

## EXPÉRIENCES PROFESSIONNELLES

---

**SOCIÉTÉ – ONCOPELE CLAUDIS RIGAUD**

**10/2024 – En cours**

### DATA ENGINEER

**Projet (Oncolab) : Conception et mise en œuvre de solutions avancées de traitement de données pour optimiser la recherche et l'innovation en oncologie.**

- Développement et optimisation de pipelines de données scalables sur GCP pour l'intégration de données issues du DPI et de l'EDS.
- Traitement et harmonisation des données de santé selon les standards HL7, FHIR et le modèle OMOP pour assurer l'interopérabilité entre différentes sources cliniques et génomiques.
- Assurance qualité et gouvernance des données cliniques et biomédicales, en garantissant l'intégrité, la traçabilité et la conformité réglementaire (RGPD, Hébergement de Données de Santé - HDS).
- Automatisation et surveillance des flux de données en production.
- Collaboration avec l'équipe de recherche et les équipes analytiques.
- Documentation du projet.

**Environnement et méthodologie :**

- *Databricks, Pyspark, GCP (BigQuery, Cloud Storage, CloudRun, Composer, Dataflow, Artifact Registry), SQL, ClickHouse, InterSystems IRIS, GithubActions*
- *Interopérabilité des données : HL7, FHIR, CDA, OMOP, DICOM*
- *Méthode agile (Kanban)*

**SOCIÉTÉ – ARISTID RETAIL TECHNOLOGY**

**10/2023 – 09/2024**

**DATA ENGINEER**

**Projet 1 : Analyse de prix concurrentiels pour améliorer la compétitivité des produits clients sur le marché.**

- Mise en place de pipelines de données optimisés pour le traitement des flux concurrentiels sur GCP.
- Contrôle et validation de la qualité des données pour assurer des analyses précises.
- Surveillance et sécurisation des flux de données pour garantir leur performance.
- Support aux équipes business en adaptant les pipelines à leurs besoins.
- Amélioration continue de la gouvernance des données et gestion du datalake.
- Documentation du projet sur confluence.

**Environnement et méthodologie :**

- *Databricks, Pyspark, GCP (BigQuery, Cloud Storage, CloudRun, Cloud Composer, Dataflow) Airflow, GithubActions, Docker, Jenkins, Jira/Confluence*
- *Méthode Agile Scrum*

**Projet 2 : Refonte du système de facturation et industrialisation des modèles analytiques.**

- Analyse et migration de l'architecture de facturation vers une structure plus performante (GCP)
- Automatisation des pipelines et monitoring en continu pour assurer la fiabilité des données.
- Mise en place des meilleures pratiques de CI/CD pour le déploiement des solutions analytiques.
- Documentation du projet sur confluence.

**Environnement et méthodologie :**

- *Python, ELK, SFTP, Airflow, BigQuery, GraphQL, Cloud Storage, Cloud SQL, GithubActions, Jira/Confluence*
- *Agile scrum*

***Projet 1 : Automatisation des fichiers prix hebdomadaires/mensuels pour les clients pharmaciens.***

- Conception et mise en œuvre d'un pipeline de données automatisé sous Nifi.
- Optimisation des requêtes et validation des structures de données.
- Documentation du projet sur confluence.

***Environnement et méthodologie :***

- *Nifi, PostgreSQL, Oracle, GitlabCI, Jenkins, Jira/Confluence*
- *Agile Scrum*

***Projet 2 : Génération automatisée de feedbacks pour les médecins.***

- Conception et développement de pipelines ETL performants.
- Visualisation des données avec Power BI.
- Documentation du projet sur confluence.

***Environnement et méthodologie :***

- *Nifi, SQL Server, Oracle, PowerBI, GitlabCI, Jenkins, Jira/Confluence*
- *Agile Scrum*

## **FORMATIONS**

---

**2021** MASTER DATA ENGINEER – YNOV CAMPUS

**2019** Licence Informatique – UNIVERSITE DE MONTPELLIER

## **LANGUES**

---

**Anglais :** Professionnel

**Français :** Bilingue

**PEULH :** Langue Maternelle