

Nwalozie Raphael

Data Engineer

Paris, France

nwalozie.raaphael@gmail.com

Tél : +33751404710

<https://www.linkedin.com/in/raphaelnwalozie/>

<https://github.com/eremah?tab=repositories/>

RÉSUMÉ

Data engineer avec expérience dans la conception, la mise en œuvre et la maintenance de pipelines ETL et des bases des données. Solide expérience de la mise en œuvre et de la gestion de l'infrastructure de données. Compétent avec une variété de technologies, y compris Docker, Spark, Azure et SQL, et Python, Airflow. Capacité avérée à travailler en collaboration avec les équipes de développement et d'exploitation pour obtenir des résultats de haute qualité.

COMPÉTENCES

- Langages: Python, Scala.
- DevOps: Git, Docker, Terraform, Ansible, Kubernetes.
- Data: Spark, Pyspark, Databricks, Airflow, kafka, Snowflake.
- Cloud : Azure.
- Data Modeling.
- Base de données : SQL, PostgreSQL.
- ETL: Talend.

NIVEAU D'INTERVENTIONS

- Développement des scripts de chargement vers PostgreSQL.
- Optimisations des traitements sur Pyspark.
- Utilisation de pyspark pour stocker les données en continu vers Snowflake et amélioration du traitement des données en utilisant Spark de 29%.
- Développements des Jobs Spark avec Python et préparation des data.
- Transformation des datas avec Pyspark.

CERTIFICATIONS

- Microsoft Certified: Azure Data Engineer Associate.
- Astronomer Certification for Apache Airflow Fundamentals.
- Data engineer: Datascientest.
- Microsoft Certified: Azure Fundamentals.

EXPÉRIENCE PROFESSIONNELLES

**Data Engineer pyspark/Azure
Buzzer**

Dec. 2021 – Oct. 2023
New York

Buzzer permet aux amateurs de sport de découvrir les moments en direct qui comptent le plus pour eux. Buzzer est le meilleur moyen d'attraper et de vivre des moments « éclair dans une bouteille » en direct de n'importe où.

Projet : Développement et maintenance de l'application Buzzer – Gestion flux des données et stockages.

- Développement des scripts de chargement vers PostgreSQL
- Optimisations des traitements sur Pyspark
- **Utilisation de pyspark pour stocker les données en continu vers Snowflake** et amélioration du traitement des données en utilisant Spark de 29%.
- **Développements des Jobs Spark avec Python et préparation des data.**
- **Transformation des datas avec Pyspark**
- Développement et maintenance de pipelines de données à l'aide des services de données Azure, entraînant une augmentation de 40 % de la vitesse de traitement des données.
- Expérience pratique de l'utilisation d'Azure Data Analytics pour analyser des **mégadonnées structurées, semi-structurées et non structurées.**
- Réduction de la charge de calcul de l'entrepôt d'environ 26% en refactorisant les structures de données et en implémentant une mise à jour sur le cache d'événements.
- Travailler avec des architectes de données pour construire un data-Warehouse dans Snowflake.
- **Configurer des alertes pour les processus d'ingestion (Kafka) de données en temps réel.**
- **Développement de code SQL Spark complexe dans le bloc-notes Azure Databricks** et invoqué dans les pipelines ADF pour le traitement des données et l'infrastructure et l'architecture des données.
- ETL processus en utilisant Postgres base de données, python et Fast Api Framework.

**Data Engineer
HubHaus**

Sept. 2018 – Nov. 2020
San Francisco Bay Area

HubHaus est une communauté d'habitation qui aide les professionnels à obtenir un logement partagé à long terme.

- **Faciliter le parcours des données de la source au hub cible** sur l'écosystème Azure en utilisant PySpark, SQL, Python.
- **Effectuer des analyses de données, des tests unitaires, nettoyer et résoudre les problèmes de données.**
- Développer des codes génériques de validation des données pour tester le processus ETL pour certaines sources de données.
- Orchestration des pipelines des données en utilisant Airflow.

- Création de pipelines et de déclencheurs Azure dans Azure Data Factory (ADF) pour automatiser le processus de transformation des données.
- DevOps utilisant Docker et Terraform.

FORMATIONS

Master Géophysique de proche surface
Sorbonne Université

July 2022
Paris, France

Master Géologie de l'exploration
Université Hassan II

July 2020
Casablanca, Morocco