



## Mamadou DIAKHABY

### Data Engineer en recherche d'un CDI

06 03 70 08 96

diakhaby14@gmail.com

92360, MEUDON

@mamadou-diakhaby-879856156/

Permis B

@Diakhaby-data

## PROFIL

“Data Engineer formé à la conception et industrialisation de pipelines ETL/ELT robustes, scalables & automatisés. Expérience pratique en ingestion multi-sources, modélisation SQL/NoSQL et orchestration avec Airflow/dbt. Pipelines containerisés et testés via FastAPI & Docker. Recherche un poste de Data Engineer pour produire des pipelines scalables et mesurables qui créent de la valeur métier.”

Python | SQL | Airflow | dbt | Docker | FastAPI | PySpark | Kafka | GitLab CI/CD | AWS (S3) | Prometheus/Grafana

## FORMATIONS

### DATA ENGINEER

DataScientest, Paris, France - Juillet 2025 / Octobre 2025

Formation axée sur ingestion, transformation, automatisation et déploiement de données. Projet pipeline de données complet : SQL/NoSQL, Python, MongoDB, API, Docker, Airflow.

### MASTER 1 - RSE

Université de Versailles, Paris, France - 2018 / 2019

Une formation axée sur la RSE et le développement durable.

### MASTER - GÉOLOGIE

Université de Rennes 1, France - 2016 / 2018

Prospection des Ressources Minérales et Pétrolières - RMP.

## LANGUES

### FRANÇAIS

Courant

### ANGLAIS

Professionnel basique (en progression)

### DIAXHANKÉ

Langue maternelle

## INTÉRÊTS

### DATA & IA

Veille technique (NLP, pipelines ETL), participation à meetups Data/ML, lecture d'articles.

### LECTURE & RÉFLEXION

Essais sur technologie et Science, éthique de l'IA et innovation.

### SPORT

Fitness régulier pour le maintien de la forme, résilience mentale et performance.

## EXPÉRIENCES

### PROJET DATA ENGINEER

Pipeline complet d'analyse de satisfaction client : collecte web, modélisation ML, exposition API et monitoring production, Paris, France - 2025

- Ingestion multi-sources : scraping 20 259 avis (BeautifulSoup) + stockage MongoDB --> validation 98,7% d'intégrité des données.
- Pipeline ETL automatisé : orchestration Airflow (DAG 7 tâches), transformations MySQL + Random Forest (89% accuracy) --> réduction 75% du temps manuel.
- API & observabilité production : FastAPI (latence <200 ms), dashboard Streamlit interactif, monitoring Prometheus/Grafana avec alerting automatisé.
- Industrialisation & DevOps : microservices Docker Compose (4 conteneurs), CI/CD GitHub Actions, tests unitaires (87% couverture).
- Impact business : 20k+ avis traités automatiquement, 340 avis critiques détectés, 12 h/semaine libérées, architecture scalable 100k+ avis.
- Environnement : Python, SQL (MySQL), MongoDB, Airflow, FastAPI, Docker, Streamlit, Prometheus/Grafana, Git/GitHub.

### ANALYSTE MOLP/MET NIVEAU 3 & FORMATEUR (CDI)

ITGA Analytics, ARCUEIL - Depuis Février 2022

- Analyse de matériaux de construction et filtres d'air pour détection des fibres d'amiante via Microscopie Optique (MOLP) et Électronique (MET).
- Production & qualité : analyse de ~2 400 échantillons/an (200/mois) avec respect strict protocoles ISO, taux de conformité >99%.
- Gestion de données complexes : traçabilité exhaustive de milliers d'échantillons, détection de patterns anormaux, reporting précis sous contraintes réglementaires.
- Formation & transmission : montée en compétence des nouveaux analystes sur protocoles techniques et administratifs.
- Compétences transférables vers Data Engineering : rigueur méthodologique (protocoles stricts = code quality), analyse de données complexes (échantillons = datasets), gestion qualité (contrôles = tests unitaires), documentation exhaustive.

## COMPÉTENCES

### DATA ENGINEERING

- Langages & Systèmes : Python, SQL, Linux, Bash, Git/GitLab
- Bases de données : MySQL, Snowflake, MongoDB
- Big Data & Streaming : PySpark, Apache Kafka, Spark
- Streaming ETL & Orchestration : Apache Airflow, DBT, automatisation de pipelines de données.
- Cloud & DevOps : Docker, Kubernetes, CI/CD (GitLab), AWS (S3)
- API & Industrialisation : FastAPI, tests unitaires, conteneurisation et déploiement.
- Monitoring & Observabilité : Prometheus, Grafana