



Zaki Ilias

Data Engineer

Strasbourg, mobile partout en France
+33 7 44 71 79 53
zakii.iliaas@gmail.com
iliaszaki.github.io
github.com/iliaszaki
in/ilias-zaki-b5841a226

Data Engineer en fin d'alternance, je conçois et automatise des plateformes de données de bout en bout : 6 sources API intégrées, 18 tables Delta modélisées en étoile et 5 pipelines orchestrés dans Microsoft Fabric, aujourd'hui en production. Diplômé d'un M2 Big Data en 2026, je recherche un **CDI de Data Engineer dès septembre 2026**.

EXPÉRIENCE PROFESSIONNELLE

SEPT. 2025 – PRÉSENT

CTDI Europe (CTDI Group), Strasbourg
Data Engineer & Data Analyst (Alternance)

- Optimisation du modèle SOC : refonte de l'architecture, gain de performance et de fiabilité des flux de cybersécurité
- Architecture Microsoft Fabric Lakehouse : stockage Delta Lake, notebooks PySpark, pipelines Data Factory orchestrés
- Plateforme livrée en production : 6 sources, 18 tables Delta Lake, 5 pipelines de bout en bout
- Modèles sémantiques et rapports Power BI pour l'équipe SOC : mesures DAX, KPIs sécurité, alertes temps réel
- Automatisation IA du reporting : classification de risque (Power Automate, AI Builder), rapports d'investigation d'alertes générés automatiquement, agent de triage basé sur MCP

MAI – AOÛT 2025

CTDI GmbH (CTDI Group), Malsch, Allemagne
Data Engineer (Stage, 4 mois)

- Conception from scratch d'un pipeline de données SOC multi-sources, consolidant plusieurs APIs cybersécurité
- Livraison en production : 7 tables Delta Lake dans Azure Fabric Lakehouse, archivage automatisé

PROJETS

Agora Python, Apache Airflow, dbt, PostgreSQL

- Pipeline e-commerce : 1 500 à 3 000 commandes/jour, orchestration Airflow, transformations dbt

Cloud Transcriber Azure Container Apps, Terraform, Docker, FastAPI, Whisper, GitHub Actions

- Transcription audio sur Azure : infrastructure as code, CI/CD complet

ML Padel Project Python, scikit-learn, XGBoost, FFT, MFCC, Hilbert

- Classification zone d'impact, type et âge de raquette depuis signaux audio et vibration ; 5 features custom ; jusqu'à 100% d'accuracy (KNN, Random Forest, SVM, XGBoost)

QuickQCM Python, PHP, OpenCV, EasyOCR, MySQL

- Création, gestion et correction automatique de QCM par reconnaissance d'image ; suivi des résultats

COMPÉTENCES

DATA & BI Python, pandas, SQL, REST APIs, GraphQL, Delta Lake, dbt, Microsoft Fabric, Power BI, Power Automate, Azure Data Lake, Terraform
BIG DATA Apache Airflow, Spark (PySpark), Hadoop, Docker, Docker Compose, Azure Container Apps
ML & IA scikit-learn, XGBoost, CNN (MobileNetV2), Vision Transformer, Transfer Learning, R, Jupyter
DEVOPS FastAPI, Django, PHP, MySQL, GitHub Actions, CI/CD, OAuth2, OpenCV, EasyOCR

FORMATION

2025 – 2026 **Master 2 Big Data**

EN COURS

JUNIA ISEN, Lille

2024 – 2025 **Master 1 Big Data**

JUNIA ISEN, Lille

2023 – 2024 **1^{re} année cycle ingénieur**

MAJOR DE PROMOTION

ISEN, Rabat, Maroc

2021 – 2023 **Cycle préparatoire**

MAJOR DE PROMOTION

ISEN, Rabat, Maroc

CERTIFICATIONS

MICROSOFT **Power BI** : Harnessing Data, ETL, Data Modeling, Preparing Data for Analysis

STANFORD **Supervised Machine Learning** : Regression and Classification

IMPERIAL **Mathematics for ML** : Linear Algebra

MICHIGAN **Python** : Programming for Everybody, Data Structures

LANGUES

Français bilingue natif · Anglais bilingue (TOEIC 850/990) · Espagnol débutant