



Zakaria TOZY

Data Engineer | CDI

☎ 0617407077 | ✉ zakaria.tozy@icloud.com | in zakaria-tozy | 📍 Paris | 🌐 zack242

Ingénieur diplômé d'un double cursus en systèmes d'information et data science, avec une forte passion pour l'IA et les nouvelles technologies. Expérience en tant que Data Engineer chez AXA IM, je suis à la recherche d'un CDI pour innover dans le domaine de la data et de l'intelligence artificielle.

COMPETENCES

- Programmation** : Python, SQL, Java, JavaScript
- Ingénierie Données Cloud** : (PySpark,Hadoop), (Databricks,Snowflake), (Azure,GCP,S3), dbt
- Machine Learning** : Scikit-Learn, NLP, LLM
- Devops** : Docker, CI/CD, Git
- Certifications** : HackerRank SQL (Expert) - 2024, IBM Introduction to data engineering - 2024, GCP Professional Data Engineer (en cours)
- Soft Skills** : Autonomie, Rigueur, Esprit d'analyse, Apprentissage rapide
- Langues** : Français (Bilingue), Arabe (Bilingue), Anglais (Courant - TOEIC 875)

FORMATIONS

- École Centrale d'Électronique (ECE Paris)** Janvier 2024
Ingénieur en Systèmes d'Information et Big Data Analytics Top 5% de la promotion
- Spécialisation : Programmation, Bases de données, Électronique
- Institut Polytechnique de Paris (École Polytechnique)** Janvier 2024
Master 2 en Data Sciences Mention Très bien
- Spécialisation : Machine Learning, Big Data, Cloud Infrastructure, Deep Learning, NLP

EXPERIENCES

- Projet Personnel** Février 2024 — Juin 2024
Ingénieur Blockchain Paris
- Monitoring de tous les tokens créés sur la blockchain Solana en Rust
 - Mise en place d'un système pour acheter ces tokens dès leur création (moins de 1s)
 - Vente automatique des tokens après réalisation d'une plus-value
 - Déploiement sur DigitalOcean avec Docker
- AXA Investment Managers - AXA IM** Août 2023 — Janvier 2024
Stage de Fin d'Études - Data Engineer Paris
- Conception et mise en place de pipelines de données pour l'ingestion et la gestion des données métiers
 - Optimisation du modèle de données et migration des pipelines de distribution pour le reporting financier de 844 milliards d'euros d'actifs
 - Tests de nouveaux clusters Databricks et création d'un PoC pour l'anonymisation des données sensibles
 - Création de tableaux de bord et calcul des KPIs pour la vision distribuée (ME, OV, CIS, FX)
- Kalima Blockchain et IoT** Avril 2022 — Août 2022
Stage - Software Engineer Paris
- Développement d'un explorateur de blockchain propriétaire Kalima en Java/LevelDB (+1000 transactions par seconde)
 - Implémentation de nouvelles fonctionnalités d'administration pour la blockchain propriétaire Kalima en Node.js
 - Conception et développement d'un système d'authentification multisignature pour les DApps (Decentralized Applications)
- Le Crédit Lyonnais - LCL** Janvier 2020 — Février 2020
Stage - IT Support Paris
- Étude de migration Microsoft Office 2010 vers 2016 et analyse de son impact sur les processus métiers
 - Développement et exécution de procédures de tests sur un parc de 23 000 postes de travail

PROJETS

Bitcoin Analysis

Août 2023

- Création d'un pipeline de données avec Python, SQL, Airflow, dbt et Snowflake pour l'analyse des transactions Bitcoin, traitant plus de 1 million de transactions par jour.

Twitter Real-time Analysis

Juin 2022

- Développement d'un système de visualisation en temps réel des flux Twitter par sujets à l'aide de Kafka et de l'algorithme KNN (K-Nearest Neighbors).

Energy Disaggregation avec Caggemini

Mars 2021

- Extraction de la composante thermosensible de la consommation énergétique pour prévoir la demande future en France.

Analyse des Communications FED avec Natixis

Novembre 2023

- Développement d'un modèle NLP basé sur FinBERT pour l'analyse de sentiment des communications de la FED
- Mise en place d'un pipeline de traitement des données textuelles pour l'extraction d'insights financiers

Data Quality Monitoring System

Septembre 2022

- Conception et implémentation d'un système de monitoring de la qualité des données avec Great Expectations
- Intégration avec Airflow pour l'automatisation des contrôles de qualité et la génération de rapports

Agent AI Autonome

Décembre 2023

- Développement d'un agent AI autonome utilisant le framework a16z pour l'automatisation de tâches complexes
- Implémentation d'une architecture de prise de décision basée sur les LLMs avec gestion de contexte et mémoire

AI de Persuasion

Octobre 2023

- Développement d'un système d'IA conversationnel spécialisé dans la persuasion et l'argumentation
- Implémentation de techniques avancées de NLP pour l'analyse et la génération de discours persuasif