

# Zakaria TOZY

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

Jeune ingénieur en Data Engineering, avec une double diplomation en Systèmes d'Information et Data Science. Passionné par l'optimisation des données et leur transformation en solutions pratiques.

## EXPERIENCE

### AXA Investment Managers

Août 2023 – Janvier 2024

Data Engineer / Data Analyst

Paris

- Conception et mise en place de **10+ pipelines ETL** pour l'ingestion de données métiers dans un environnement Cloud
- Optimisation des modèles de données, migration vers Python 3.9 et Unity Catalog, gérant **844 milliards d'euros d'actifs**
- Participation aux tests **Databricks** et création d'un **PoC** pour l'anonymisation des données clients
- Création de tableaux de bord et **calcul des KPIs** (ME, OV, CIS, FX) pour la vision distribuée/managériale

### Kalima Blockchain & IoT

Avril 2022 – Août 2022

Software Engineer

Paris

- Développement d'un **explorateur de blockchain** avec Java et LevelDB, traitant **1000+ transactions/seconde**
- Maintenance et ajout de fonctionnalités pour l'outil d'administration blockchain sur **Node.js**
- Conception d'un algorithme d'**authentification multisignature** pour les DApps

### Le Crédit Lyonnais - LCL

Janvier 2020 – Février 2020

IT

Paris

- Réalisation d'une étude sur la **migration d'Office 2010 vers Office 2016**, analyse des fonctionnalités
- Exécution de tests sur **23 000 postes de travail** et développement de procédures d'installation

## PROJECTS

### Bitcoin Analysis (Python, SQL, Airflow, dbt, Snowflake)

- Développement d'un **pipeline de données** pour l'analyse des transactions Bitcoin
- Optimisation des processus d'analyse de données massives et **automatisation** du traitement

### Twitter Real-time Analysis (Kafka, NLP, KNN)

- Création d'une solution de **visualisation et classification en temps réel** des flux Twitter
- Application de modèles de **machine learning** pour la classification des données non structurées

### Energy Disaggregation @Capgemini (Numpy, Pandas, Feature Engineering)

- Extraction de la composante **thermosensible** de la consommation énergétique en France
- Amélioration des **prévisions énergétiques** avec des techniques avancées de traitement de données

## EDUCATION

### École Centrale d'Électronique (ECE Paris)

Janvier 2024

Ingénieur en Systèmes d'information et Big data analytics

Paris

- **Top 5%** de la promotion
- Cours : Programmation, Systèmes d'exploitation, Bases de données avancées, DevOps, Sécurité

### Institut Polytechnique de Paris

Janvier 2024

Master 2 en Data Sciences

Paris

- **Mention Très bien**
- Cours : Machine Learning, Big Data Streaming, Cloud Infrastructure, Deep Learning, NLP

## SKILLS

---

**Programmation** : Python, SQL, Java, C++, C#

**Data Engineering** : PySpark, BigQuery, ETL, Data Pipelines

**Cloud & Databases** : Databricks, Snowflake, GCP, dbt

**Machine Learning** : Scikit-Learn, Deep Learning, NLP, Computer Vision

**DevOps** : Docker, CI/CD, Git, Scrum methodology

## LANGUAGES

---

**Français** : Langue maternelle    **Anglais** : Courant (TOEIC : 875)    **Arabe** : Langue maternelle