

Sidy BADJI

Data Engineer

📞 07 89 03 27 52

✉️ sidybadji935@gmail.com

🌐 linkedin.com/in/sidybadji

🐙 github.com/sidybadji

PROFIL PROFESSIONNEL

Data Engineer expérimenté en Python et Java, spécialisé dans la construction de pipelines de données robustes et le traitement de données massives avec Spark et GCP. Expertise avérée dans la conception d'architectures data, l'orchestration de workflows (Airflow), et la mise en place de pipelines ETL/ELT. Compétences solides en conteneurisation (Docker), développement d'APIs, gestion de bases de données et transformation de données brutes en insights business exploitables.

EXPÉRIENCE PROFESSIONNELLE

Développeur Back-End & Data Engineer

GoGP

Décembre 2023 – Juillet 2025

- Conception et implémentation d'API RESTful sécurisées avec Spring Boot (MVC, Spring Security, JWT)
- Modélisation et gestion de bases MySQL ; optimisation des requêtes pour la performance
- Scripts Python avancés pour : Web scraping (tarifs, disponibilités, données externes), Automatisation de flux de données, Monitoring et reporting
- Exploitation de GCP (BigQuery, Cloud Functions, Pub/Sub, Cloud Storage) pour le traitement et le stockage des données
- Déploiement automatisé via CI/CD GitLab, intégration continue et monitoring applicatif
- Documentation technique (Swagger, Confluence) et collaboration étroite avec l'équipe front-end

Résultats :

- Architecture stable, scalable et sécurisée mise en place
- Réduction significative des bugs grâce aux tests automatisés
- Exploitation des données via GCP pour enrichir l'application

Ingénieur Data & ML

Renault Group

Avril 2023 - Septembre 2023

- Développement Python & Spark pour l'ingestion et le traitement de données IoT (GPS, télémétrie véhicules)
- Mise en place de pipelines GCP (BigQuery, Dataflow, Dataproc) pour centraliser et fiabiliser les flux de données
- Machine Learning & Deep Learning : Détection d'attention du conducteur (images embarquées) via PyTorch & OpenCV
- Modèles de reconnaissance faciale et comportementale pour renforcer la sécurité
- Industrialisation des modèles ML en production (CI/CD, Docker, Airflow)
- Collaboration étroite avec Data Scientists et équipes métier pour transformer les POC en solutions IA industrialisées

Résultats :

- Réduction de 40% du temps de traitement grâce aux optimisations Spark & GCP
- Amélioration de 30% de la précision analytique via feature engineering et modèles ML
- Documentation complète facilitant la reprise et l'évolution du projet

Data & Cloud ODC (Opérations et Déploiement Cloud)

Claranet

Avril 2021 – Septembre 2022

- Développement de scripts Python pour l'automatisation (sauvegardes, monitoring PostgreSQL, collecte de logs)
- Participation à la migration d'applications et de bases de données vers Google Cloud Platform (BigQuery, Cloud Storage)
- Mise en place de pipelines d'ingestion de données pour centraliser et traiter les métriques de supervision
- Gestion réseau & systèmes (adressage IP, configuration serveurs en datacenter, surveillance des performances)
- Collaboration avec les équipes DevOps pour renforcer la sécurité et la haute disponibilité

Résultats :

- Fiabilisation des sauvegardes et réduction du risque de perte de données
- Automatisation des tâches récurrentes (gain de temps de 25%)
- Contribution à la migration cloud et meilleure exploitation des données via GCP

COMPÉTENCES TECHNIQUES

Data Engineering Python, SQL, Apache Spark, Apache Kafka
MLOps & Orchestration Airflow, Docker, Git, CI/CD, Kubernetes
Cloud (GCP/AWS) BigQuery, Dataflow, Pub/Sub, Cloud Functions, Cloud Storage
ML & IA PyTorch, OpenCV, NLP, Computer Vision
Backend & APIs Python (Django), Java (Spring Boot), Swagger, Postman
Bases de Données PostgreSQL, MySQL, Cassandra, Neo4j

FORMATION

Master Informatique Fondamentale et Appliquée Données Université Paris Cité Algorithme, Traitement et Analyse	2023
Licence Informatique Générale Université Paris Cité	2021

CERTIFICATIONS

ETL and Data Pipelines with Shell, Airflow and Kafka (Coursera)

LANGUES

Français	Natif
Anglais	B2

INFORMATIONS

Disponibilité :	Immédiate
Localisation :	Le Perreux-sur-Marne
Mobilité :	Région parisienne
Type de contrat :	Freelance / CDI