DANIELSON RASOAVELONIRINA

Senior Data Engineer · Consultant (CDI)

Antananarivo, Madagascar · (+261) 34 71 147 49 · d.rasoavelonirina@gmail.com
Portfolio · GitHub · LinkedIn

PROFIL

Ingénieur de données **senior** avec **8+ ans d'expérience continue** chez eTech Consulting Madagascar, spécialisé en conception d'**architectures data robustes** (Data Lakes, Data Warehouses, event-driven GCP) et de **pipelines ETL/ELT performants** (Apache Spark, Kafka, Airflow, Nifi, Talend). Évolution récente vers le **leadership technique** : Tech Lead Backend & Cloud R&D, co-architecte du département **Data & AI Security** (RBAC/ABAC, RGPD, guardrails LLMs). Solide maîtrise du **Cloud (AWS, GCP)**, de l'écosystème **Big Data (Kafka, Hadoop, Delta Lake)** et des plateformes **BI (Power BI, Qlik Sense, Tableau, SAS)**. **Titulaire d'un permis de travail belge**, je recherche un poste de **Data Engineer / Lead BI en Europe**.

COMPÉTENCES TECHNIQUES

Data Engineering	Apache Spark · Kafka · Airflow · Nifi · Talend · Hadoop (HDFS) · Delta Lake · Architecture event-driven · Data Lake/Warehouse · Real-time streaming
Cloud Platforms	AWS (Lambda, EMR, RDS, S3, Terraform) · GCP (BigQuery, Dataflow, Cloud Run, Pub/Sub, Eventarc, Firestore, Vertex AI, Cloud Composer)
Business Intelligence	Power BI · Qlik Sense · Tableau · SAS (Base, Enterprise Guide, Visual Analytics) · SQL
Data & AI Security	RBAC/ABAC · RGPD/GDPR · Chiffrement · Anonymisation · Workload Identity · IAM granulaire · Guardrails LLMs · Tests adverses
Bases de données	PostgreSQL · MySQL · BigQuery · Firestore · DuckDB · NoSQL
Programmation & Frameworks	Python (orienté Data) · SQL · Shell · JavaScript/Node.js · Java (routines Talend) · FastAPI · Flask · Pandas · Polars · Dask · Pydantic v2
Machine Learning & IA	TensorFlow · PyTorch · Keras · scikit-learn · Fine-tuning BERT/LLMs (LLaMA, Qwen) avec LoRA · YOLO · Gemini function calling
DevOps & IaC	Git/GitLab · Jenkins · Cloud Build · Terraform · Docker · Artifact Registry · Secret Manager · VTOM · ADRs · CI/CD

EXPÉRIENCE PROFESSIONNELLE

eTech Consulting Madagascar

Antananarivo (remote clients EU)

Senior Data Engineer · Consultant (CDI) *2018 → présent*

8+ ans d'évolution continue : *Data Scientist (2018-2023) → Data Engineer & Quality Manager (2023+) → Tech Lead Backend & Data/AI Security (2026+)*. Missions internes (R&D, sécurité) et sous-traitance pour grands comptes : **Orange Madagascar TM, Orange Money, RAVATE, Argos Vétérinaire, Ghanty Groupe**.

▸ Tech Lead Backend & Cloud Infrastructure · R&D interne (PoC) *Avril 2026 → présent*

Lead technique Backend / Cloud sur un PoC interne de plateforme AI-powered événementielle combinant Speech-to-Text, traduction, NER, agent conversationnel et intégrations métier (architecture poly-repo, déploiement GCP).

- **Conception d'une architecture event-driven sur GCP** : Cloud Run + Pub/Sub + Eventarc + Firestore + BigQuery, déployée via Terraform (IaC).
- **Développement de l'API publique en FastAPI** (Pydantic v2, OpenAPI, JWT, observabilité) : ingestion sécurisée de fichiers via pre-signed URLs Cloud Storage.
- **Intégration des modèles IA** : pipelines orchestrés appelant **Whisper/MMS** (STT), **Gemini** (traduction + NER + agent function calling), **Vertex AI**.
- **Mise en place du DevOps / CI/CD** : Cloud Build + Artifact Registry + Secret Manager, environnements dev/staging/prod, monitoring Cloud Logging + Trace.
- **Sécurisation by-design** : Workload Identity Federation, IAM granulaire, secrets centralisés, conformité RGPD dès la conception.

- **Tech leadership** : revue d'architecture, ADRs (Architecture Decision Records), mentorat de l'équipe poly-repo, conventions Conventional Commits + protected branches.

▸ Data & AI Security · Création du département Cybersécurité *Janvier 2026 → présent*

- **Membre fondateur** de l'équipe transverse de **5 personnes** pilotant la création du nouveau département Cybersécurité d'eTech Consulting (référentiels **RGPD + OWASP LLM Top 10**).

SecData — Sécurisation de la donnée :

- **Conception et déploiement** de politiques **RBAC/ABAC** uniformes sur **3 environnements** (dev/staging/prod) couvrant Data Lakes et projets Cloud (AWS/GCP).
- **Implémentation** des protocoles de **chiffrement (au repos / en transit), tokenisation et anonymisation** des données sensibles en conformité RGPD.
- **Audit et monitoring continu** de 10 pipelines de tests via **Cloud Audit Logs GCP + Wiz** : détection d'accès anormaux, intégrité des flux ETL/ELT.

SecIA — Sécurisation de l'IA (PoC R&D interne) :

- **Conception de Security Guardrails** pour LLMs : prévention des injections de prompts, jailbreaking et fuites de données d'entraînement.
- **Définition de tests de robustesse adverse** (adversarial testing) et de pipelines de **nettoyage anti-PII** sur les jeux d'entraînement.
- **Évaluation des risques** liés aux outils IA tiers et définition de **standards de déploiement sécurisé** pour les modèles en production.
- **Projet phare** : conception et industrialisation d'une **politique IAM uniforme** garantissant le moindre privilège et la traçabilité sur les 3 environnements eTech.

▸ Data Engineer & Data Quality Manager · Mission Orange Madagascar TM *Avril 2023 → Décembre 2025*

- **Architecture et industrialisation** de **170 DAGs Apache Airflow + 184 jobs Talend** consolidant **25 sources** multi-systèmes (DATA GSM, CRM, billing, réseau) à hauteur de **350 millions de lignes/semaine**.
- **Migration et optimisation** de l'architecture Data Lake / Data Warehouse vers **Delta Lake sur Hadoop** : gain de **×3** sur la latence des requêtes critiques, garanties ACID natives et time-travel.
- **Orchestration et supervision** des workflows sous **Apache Airflow + VTOM**, atteignant un **SLA de livraison de 90 %** sur les livrables business journaliers.
- **Conception et industrialisation de 35 tableaux de bord décisionnels** (Power BI, Qlik Sense, QlikView, Tableau) déployés auprès des équipes **Marketing, Finance, Risque et Direction** pour pilotage en quasi temps réel.
- **Mise en place d'un framework Data Quality automatisé** (45 règles métier + ML pour la détection d'anomalies) : **18 anomalies bloquantes détectées/mois**, fiabilisant les indicateurs transmis à la Direction et aux régulateurs (ARTEC).
- **Projet phare** : refonte du pipeline DATA GSM en **Spark Structured Streaming**, divisant le temps de calcul par 3 et débloquant l'analyse temps-réel sur les événements réseau.
- **Mentorat de 6 juniors** sur les bonnes pratiques Apache Airflow et Delta Lake : revue de code, design de DAGs, gestion d'erreurs, ACID patterns.

▸ Data Analyst · Mission Orange Money Madagascar (SAS) *Avril 2022 → Avril 2023*

- **Conception et automatisation de 15 tableaux de bord stratégiques** sous **SAS Base, Enterprise Guide et Visual Analytics** couvrant les KPIs Orange Money (**~180 000 transactions/jour** : dépôts, retraits, commissions, fraude), déployés auprès de **12 utilisateurs des équipes Marketing, Finance et Risque**.
- **Segmentation comportementale et géographique de 725 000 abonnés** en 5 segments, socle des stratégies de rétention et de ciblage Marketing.
- **Conception et déploiement en production d'un modèle de prédiction du churn**, ayant permis **35 % de réduction du churn** sur la population ciblée par les campagnes.
- **Conception et déploiement d'un modèle de scoring fraude en production** en collaboration avec les équipes IT, Conformité et Systèmes de paiement (gouvernance des données).
- **Préparation et nettoyage** des jeux de données en **SQL + Python (pandas)** en amont des analyses SAS.
- **Animation de 5 sessions de formation** auprès de 12 utilisateurs métiers sur les outils SAS et l'interprétation des résultats.

▸ Data Manager & Backend Developer · Mission RAVATE (Réunion) *Mars 2021 → Février 2022*

- **Conception et déploiement** d'une **infrastructure cloud hybride AWS + on-premise** via **Terraform (IaC)** : **35 services AWS** sur 2 comptes (dev/prod) — VPC, EC2, RDS, S3, Lambda, ALB.

- **Architecture de synchronisation continue** entre l'**ERP legacy on-premise** et le **frontal e-commerce AWS** (50 000 références produit synchronisées).
- **Industrialisation du CI/CD** : **7 pipelines AWS CodePipeline** pour le data layer + **GitLab CI** pour le backend Node.js + **Terraform** pour l'infra.
- **Conception et optimisation** d'une base **PostgreSQL (AWS RDS)** : ~25 tables, ~10 GB, schéma e-commerce + gestion de stock.
- **Développement backend Node.js + AWS Lambda** : ~30 endpoints REST + 10 Lambda functions pour les logiques d'accès et de manipulation des données.
- **Projet phare** : **migration complète de l'infrastructure legacy vers AWS via Terraform** avec architecture hybride zero-downtime.

▸ Data Engineer · Missions Argos Vétérinaire & Ghanty Groupe (Réunion / Montpellier) *2020 → 2021*

Deux missions menées en parallèle : Argos Vétérinaire (réseau de cliniques vétérinaires, FR/Réunion) et Ghanty Groupe (groupe diversifié, La Réunion).

- **Développement et maintenance de 35 jobs Talend ETL/ESB** pour l'intégration et la transformation de données métier (commandes, stock, facturation, CRM, reporting).
- **Conception de 6 DAGs Apache Airflow** pour l'orchestration des workflows inter-applicatifs.
- **Architecture de synchronisation quasi temps-réel** entre **5 systèmes clés** (Odoo, SFTP, ERP legacy, CRM, plateforme de paiement) supportant **~30 000 transactions/jour**.
- **Projet phare 1** : **migration zero-downtime de l'ERP legacy vers Odoo** pour **12 cliniques** Argos Vétérinaire.
- **Projet phare 2** : **architecture de synchronisation temps-réel SFTP ↔ Odoo** avec gestion automatisée des conflits de données.

▸ Data Scientist · Projets IA internes eTech (R&D) *2018 → 2023*

Contributions à ~10 projets de PoC R&D internes (dont 3 majeurs) en parallèle des missions clients, sur les domaines NLP, Computer Vision et systèmes de recommandation.

- **Conception et entraînement** d'un **moteur de recommandation produits e-commerce** (collaborative filtering + content-based) pour PoC interne.
- **Développement d'un classifieur topologique de documents PDF** (analyse de structure + classification de pages) sur corpus métier.
- **PoC Computer Vision** : détection d'objets sur **images urbaines** via **YOLO** (préparation dataset, fine-tuning, évaluation).
- **Fine-tuning BERT** pour extraction d'entités spécialisées (NER pédagogique) sur corpus métier.
- **Extension 2023** : **PoC chatbot conversationnel basé sur LLaMA-7B fine-tuné avec LoRA** sur GPU consumer (assistant interne), exploration Qwen-T en parallèle.

FORMATION

École Supérieure Polytechnique d'Antananarivo (ESPA) *2018 - 2019*
Parcours Master 2 — Système de Traitement d'Information (STI)

École Supérieure Polytechnique d'Antananarivo (ESPA) *2016 - 2017*
Licence — Télécommunications

CERTIFICATIONS

Data Engineering, Big Data & Machine Learning on GCP — Coursera (Google) *2018*
BigQuery, Dataflow, Cloud Storage, Pub/Sub, Cloud Composer, Vertex AI, BigQuery ML.

Neural Networks and Deep Learning — Coursera (DeepLearning.AI) *2020*
Architectures Deep Learning, structuration et évaluation de projets ML.

AI Programming with Python — Udacity *2020*
Python orienté Data, Pandas, NumPy, fondamentaux IA.

LANGUES

Malagasy : natif · **Français** : C2 (langue de travail) · **Anglais** : C1

MOBILITÉ & DISPONIBILITÉ

Permis de travail belge en main · Mobilité Europe (Belgique en priorité) · Onsite avec relocation, hybride ou remote · Disponibilité selon préavis (négociable).