

Franck Mboutou

Tel : +33 695767339 Mail : mboutoufranck@icloud.com GitHub: <https://github.com/FranckRuisseau> LinkedIn: www.linkedin.com/in/franck-mboutou

Profil

Étudiant en Master Data Engineering à l'EFREI Paris, je me spécialise dans la conception d'architectures de données robustes et auditables. Fort d'une expertise en **Big Data (Spark/Kafka)** et en **Monitoring (ELK)**, je cherche à rejoindre une équipe pour relever des défis liés à la qualité des données et au traitement temps réel.

COMPÉTENCES

- **Big Data** : Apache Spark (Structured Streaming), Kafka, Architecture Kappa (Faible latence).
- **Observabilité & Audit** : Suite ELK (**Elasticsearch, Logstash, Kibana**), Grafana, Tableau.
- **Data Quality & Gouvernance** : Gestion des doublons (Deduplication), Conformité des données, Data Lineage.
- **Bases de Données** : PostgreSQL, NoSQL (MongoDB), ElasticSearch, MinIO(S3).
- **DevOps & Industrialisation** : Docker, CI/CD, Développement d'API (FastAPI/Flask)

PROJETS TECHNIQUES

Real-Time Financial Security Engine (Architecture Kappa) | Janv. 2026 Stack: Python, Spark Streaming, Kafka, Delta Lake, MinIO, Docker. **Code source :** github.com/FranckRuisseau/DATA-STREAM-PLATFORM

- **Conception d'une Architecture Kappa** pour l'ingestion et le traitement unifié de flux transactionnels massifs (>10k events/sec), supprimant la latence des couches batch traditionnelles.
- **Implémentation de logiques Stateful** complexes (agrégations par fenêtres temporelles, jointures de flux) pour enrichir les données en temps réel.
- **Fiabilisation du pipeline** via un pattern "Router" isolant les messages invalides (Dead Letter Queue) pour garantir une qualité de donnée stricte sans bloquer le flux.
- **Gestion du stockage ACID** et activation du "Time Travel" via **Delta Lake** pour assurer la cohérence et l'auditabilité des transactions historiques.

Pipeline d'Audit & Monitoring de Logs (Stack ELK) | 2025 Solution de centralisation des logs pour la traçabilité et la sécurité.

Code source: <https://github.com/FranckRuisseau/eltk-kafka-pipeline.git>

- Ingestion temps réel de logs systèmes et applicatifs critiques via **Logstash**.
- Structuration des données dans **Elasticsearch** pour permettre des recherches rapides lors d'audits.
- Création de dashboards de contrôle sous **Kibana** pour visualiser les anomalies et incidents de production.
- *Stack : Elasticsearch, Logstash, Kibana, Kafka, Docker*

EXPÉRIENCES PROFESSIONNELLES

Ingénieur Logiciel Junior (Data & Backend) | 3HCIG Juin 2024 – Sept. 2024

- Conception d'APIs REST Flask pour la **normalisation de données** hétérogènes .
- Participation à la modélisation des processus de collecte et de transformation
- Mise en place d'un système d'alerte critique pour la surveillance de la production.

Développeur Junior Data & Cloud | ETech For Self-Worth Juin 2023 – Sept. 2023

- Création d'un pipeline pour la collecte de séries temporelles environnementales.
- Analyse de données (Python/Pandas) pour optimiser les cycles d'irrigation.
- Déploiement et administration d'un cluster **MongoDB Atlas**

LICENCES ET CERTIFICATIONS

- **Établir une gouvernance des données durable et efficace** (LinkedIn, Déc. 2025)
- **L'essentiel d'Apache Spark** (LinkedIn, Déc. 2025)

ÉDUCATION

Master en Data Engineering & AI – EFREI PARIS (Ecole d'ingénieurs – Paris)

2025 - *en cours*

Diplôme d'Ingénieur de Travaux option Génie Logiciel, IAI (Ecole d'ingénieurs)

2022 - 2024