

Senior Data Engineer

☎ 0624010036 | ✉ relmaazouz@proton.me | 🌐 relmaazouz | 📍 Paris, France

EXPERIENCE

Candriam

09/2024 - present

Senior Data Engineer: Data Lake implementation

Paris, France

- Built metadata-driven PySpark ingestion frameworks, making pipelines self-configuring and reusable across data domains at scale.
- Engineered a data quality framework covering completeness, consistency, and reliability — embedded directly into production pipeline runs.
- Designed and optimized medallion-architecture ETL/ELT pipelines on Azure Synapse and Delta Lake for high-throughput financial data.
- Exposed analytical datasets via C# APIs, delivering low-latency endpoints consumed directly by risk analysts.
- Developed a multi-source data testing framework with automated validation and performance benchmarking to gate every pipeline release.
- Partnered with financial risk management teams to define and deliver data products powering analytics and execution strategies.

Informatique CDC

08/2023 - 09/2024

Senior Data Engineer: Data platform

Paris, France

- Designed and implemented diverse business use cases, including data structures, CI/CD pipelines, and PySpark jobs for scalable processing.
- Established and enforced best practices in development: CI/CD workflows, automated testing (unit tests, non-regression tests), and GitFlow.
- Modeled and implemented a Data Vault architecture across ingestion, pre-processing, and datamart layers to ensure structured and governed data flows.
- Developed ingestion, transformation, and exposure jobs in PySpark and Python for large-scale data pipelines.
- Integrated external market and risk data sources (e.g., Bloomberg, third-party risk, ratings) using Kafka batch ingestion.
- Worked within a distributed data environment (**Cloudera, Hive, Spark, Kafka**) with automation and orchestration tools (Jenkins, Control-M, Bitbucket).
- Delivered high-quality, reliable data products tailored to business and risk management needs, enabling advanced risk and analytics use cases.

Crédit Agricole

11/2021 - 06/2023

Data Engineer: **DataLake: run & migration**

Paris, France

- Ensured run and continuous operation of a large-scale Data Lake on MapR leveraging Hive, Hadoop, Sqoop, Tez, Oozie, and PySpark.
- Developed ingestion, transformation, and exposure jobs in PySpark to process and deliver high-volume datasets.
- Led the migration and redesign of a new Data Lake, including data zoning and ingestion pipelines to improve scalability and governance.
- Translated and optimized legacy Hive SQL scripts into Oracle SQL for compatibility and performance improvements.
- Automated the generation of ODI objects (interfaces, workflows, scenarios) via Python scripting to accelerate development cycles.
- Implemented automated data validation frameworks (schemas, columns, types, views) to ensure quality and completeness across pipelines.
- Containerized and deployed applications (Cash In Time) with Kubernetes and ArgoCD, enabling cloud-native scaling and reproducible deployments.
- Designed and maintained CI/CD pipelines to automate source code deployment and testing workflows.
- Worked with a modern data ecosystem: **PySpark, Hive, Sqoop, Pandera, Pydantic, Tez, Hue, Oozie, Jenkins, Docker, ArgoCD, Kubernetes, GitLab, Oracle ODI, Oracle 19c**.

Informatique CDC

11/2020 – 10/2021

Consultant Data: Migration eFront FIA, Mise en place d'un data hub exchange

Paris, France

- Participated in requirements gathering workshops to define screens, approval workflows, and additional data fields for business processes.
- Authored detailed technical specifications, including mapping of transactions to the new codification framework.
- Extracted and transformed data from FIC using Front Script (funds, investors, companies, instruments, investment links, benchmarks, contacts), including transaction mappings and enrichment.
- Implemented a data hub on PostgreSQL to exchange transactions, funds, instruments, and investment links, integrating via Kinesis streams into AWS S3 .
- Developed a custom web application in Spring Boot to support valuation workflows and specific business use cases.
- Worked in a hybrid environment leveraging **SQL Server, PostgreSQL, Apache Kafka, AWS S3, Spring Boot, eFront FIA**.

Consolis Group

07/2018 – 09/2020

Mise en place d'un cockpit de pilotage de la performance financière sous GCP

Paris, France

- Designed and developed a data lake architecture on BigQuery, including raw ingestion and curated transformation layers.
- Built and maintained data ingestion and transformation pipelines, ingesting subsidiary files via Pub/Sub and enriching data in transformation layers.
- Developed analytical and summary reports in Data Studio, enabling business stakeholders to access timely insights.
- Worked in the Google Cloud Platform ecosystem: BigQuery, DataProc, Composer, Data Studio, and Cloud Storage.

Inwi

08/2016 – 05/2018

Data engineer: CDR analytics platform

Casablanca-Morocco

- Built near real-time data ingestion pipelines for CDR data.
- Performed data cleaning, normalization, and harmonization, and designed business-oriented datamarts for network, equipment, voice/SMS/MMS, mobile, and customer datasets.
- Prepared enriched customer datasets (social networks, geolocation, call/SMS history) to support churn prediction and advanced analytics.
- Developed interactive reports and visualizations to track temporal metrics and geographical patterns.
- Worked with a distributed big data environment: **Apache Spark (PySpark), Flink, Sqoop, Hive, Hue, Oozie, Tez, Ambari, ORC, Oracle OBIEE**.

OCP

05/2014 – 08/2016

Oracle E-Business Suite implementation

Casablanca, Morocco

- Developed analytical and operational reports using SQL, PL/SQL, and XML Publisher.
- Customized and built data entry and reporting interfaces with Oracle OAF and Oracle Forms .
- Developed data APIs and web services to enable structured data exchange with legacy systems.

S2M

08/2013 – 04/2014

E-banking web app

Casablanca, Morocco

- Designed **data schemas and exchange interfaces** to integrate core banking data (SELECT SYSTEM V6) into downstream systems.
- Defined the **application and data architecture**, including modular web components and service-oriented APIs.
- Developed and deployed **data services and APIs** (authentication, fund transfers, account statements, salary files) to enable secure, real-time exchange of financial datasets.
- Worked across a **modern data ecosystem**: Hibernate, Spring Web, Apache CXF, Oracle Database, PostgreSQL, Jenkins, and HTML5.

CERTIFICATIONS

Google Cloud Certified Data Engineer

- Credential ID: HPWFGH

MongoDB Certified DBA

- Licence N.983323957

EDUCATION

CNAM

2023 - 2025

Masters in Financial Markets

Paris, France

- Master Thesis: *Spoofing Detection in the Order Book* – studied limit order book dynamics, high-frequency trading, and applied machine learning (DNN, RNN) for spoofing detection.
- Coursework: Portfolio Management & Markowitz Theory. Fixed Income Assets; Probability & Statistics for Finance. Interest Rate Products (bonds, swaps, structured products). Futures & Options.
- Skills: Quantitative Finance, Portfolio Optimization, Portfolio Performance Analysis, Credit Risk, Credit Derivatives, Equity Derivatives, Machine Learning (DNN, RNN).

Mohammadia School of Engineers

2010 - 2013

Engineer Degree in Software and Computer Engineering

Rabat, Morocco

CPGE AL QALAM

2008 - 2010

Preparatory classes in Mathematics and Physics.

Agadir, Morocco

Salah Eddine AL AYOUBI High School

2007 - 2008

Baccalaureat in mathematics

Tinghir, Morocco