SIDI CAMARA

Lead / Senior Data Engineer

camarasidi12@gmail.com | Paris Area, France | +33 7 68 36 13 22

PROFESSIONAL SUMMARY

Senior Data Engineer with 7+ years of experience designing and optimizing large-scale data pipelines in the financial services industry. Deep expertise in Anti-Money Laundering (AML) and fraud detection systems. Proven track record in technical leadership: building teams from the ground up, mentoring engineers, and establishing engineering best practices. Proficient across the Big Data ecosystem (Scala, Spark, Kafka) and cloud platforms (AWS, Azure, GCP, Kubernetes).

TECHNICAL SKILLS

Languages	Scala, Python, Java, PySpark, SQL, Node.js, React / React Native
Big Data & Streaming	Spark, Kafka, Filebeat, Logstash, Flume, Hive, HBase
Databases	MongoDB, MySQL, ElasticSearch, HBase, S3
Cloud & Infra	AWS, Kubernetes, Docker, Terraform, Databricks, Hortonworks, Cloudera
CI/CD & Orchestration	Jenkins, GitHub Actions, Skaffold, Kustomize, Airflow, Maven, SBT
AI & Analytics	Quantexa (AI platform), H2O, Scikit-learn, Deep Java Library
Methodology	Agile / Scrum, Git, Ansible, Jira, Confluence

PROFESSIONAL EXPERIENCE

May 2022 — Present | **BNP Paribas CIB** — *Lead / Senior Data Engineer*
IT Trade Finance — AML & Fraud Detection Program, Paris

Leadership & Management
- Built and structured a new data engineering team from scratch: hiring, onboarding, resource planning, and upskilling
- Drove architectural decisions and conducted code reviews, ensuring maintainability and code quality across the entire scope
- Defined development standards, architectural patterns, and best practices adopted team-wide

Data Engineering
- Designed and developed end-to-end data pipelines (ingestion, transformation, scoring, alert generation, business exposure) processing high-volume transactional data
- Optimized Spark jobs through advanced tuning and data skew management, significantly reducing execution times
- Integrated the Quantexa AI platform for building entity-resolution graphs and enriching AML alerts with contextual intelligence
- Deployed and orchestrated services on a private Kubernetes cloud using Skaffold, Kustomize, and S3
- Implemented monitoring frameworks, performance benchmarks, and quality gates to ensure production reliability

Stack : Scala, Spark, Kubernetes, Jenkins, Git, Skaffold, S3, Deep Java Library, Quantexa

Jan. 2022 — May 2022 | **Bedrock Streaming** — *Data Engineer (Freelance)*
A/B Testing Team, Lyon

- Developed and maintained data pipelines for the A/B testing platform, processing large-scale audience data for France's leading streaming services

Stack : AWS, Terraform, Python, Scala, Spark, Databricks, Jenkins, GitHub Actions, Airflow, Docker

Nov. 2020 — Jan. 2022 | **Société Générale** — *Data Engineer*
MOSAIC Team — Technical Lead for Check Fraud Detection, Paris

- Built microservices-based data pipelines from raw data parsing to transaction scoring in both batch and streaming modes
- Developed and deployed a Machine Learning model (H2O / GBM) for automated check fraud detection into production
- Served as the technical point of contact for check fraud: daily stand-ups, weekly reporting, and cross-team coordination

Stack : Scala, Python, Java, Spark, Kafka, HBase, Hive, Jenkins, Cloudera

Oct. 2018 — Nov. 2020 | **Moobifun** — *Data Engineer & Scrum Master*
Data Platform Owner (7+ TB), Lyon
- Managed the full data platform built on the ELK stack: collection, storage, transformation, and data serving
- Led a POC and implemented a Lambda Architecture, improving both real-time and batch processing capabilities
- Built Machine Learning engines for customer segmentation and retention, deployed via REST API (Flask)
- Acted as Scrum Master for a 6-person team and mentored junior engineers and interns

Stack : MySQL, Python, Node.js, Spark, Hadoop, Kafka, ElasticSearch, Hive, Scikit-learn

Feb. 2018 — Aug. 2018 | **AXA France** — *Process Automation Intern*
Nanterre
- Automated business processes and functional testing using C#, Java, LeanFT, and Azure DevOps

Jan. 2017 — May 2017 | **LIP6 (Sorbonne University)** — *Research Intern*
Paris
- Analyzed and implemented (OCaml) a random graph generation algorithm derived from graph theory research

EDUCATION

2016 — 2018 — **M.Sc. Computer Science — Software Science & Technology**
Sorbonne University (formerly UPMC), Paris

2013 — 2016 — **B.Sc. Mathematics & Computer Science — Ranked 1st out of 300**
University of Hassiba Ben Bouali, Chlef, Algeria

CERTIFICATIONS

- Machine Learning — Stanford University (Coursera)
- Deep Learning Specialization — deeplearning.ai (Coursera)
- Hadoop Platform and Application Framework

INTERESTS

Volunteering • Martial Arts • Karate Champion