

# Sami Moustachir

Senior Data & ML Platform Engineer (AWS)

France / UAE / Remote

Email: [moustachir.sami@gmail.com](mailto:moustachir.sami@gmail.com)

Website: [sammous.dev](http://sammous.dev)

LinkedIn: [linkedin.com/in/samimoustachir](https://www.linkedin.com/in/samimoustachir)

GitHub: [github.com/sammous](https://github.com/sammous)

Builds and scales AWS data and ML platforms with DevOps/SRE rigor, IaC, CI/CD, and deep observability for product teams.

## Skills

Data & ML Platforms: mlops, model serving, feature pipelines, data platform, datalake, experimentation  
DevOps & Cloud: aws, terraform, cloudformation, kubernetes, docker, ci/cd, linux, iam, vpc, observability  
Data & Platforms: python, fastapi, dbt, airflow, spark, kafka, sql, databricks  
Product & Delivery: architecture reviews, stakeholder alignment, mentoring, documentation, roadmap design

## Experience

### Staff Software Engineer - Xendit

2021 - Present · Singapore/Remote

Building payment services with one of the biggest payment gateway in SEA

- Built internal developer CLI tooling in Golang to streamline local workflows and service operations across product teams and infrastructure.
- Maintaining core services and implementing monitoring at scale (internal tools, Datadog, Kubernetes deployments).
- Backend development for payment services, ewallet integrations, card issuing, core banking and crypto integrations (Golang, Typescript, Node).
- Data infrastructure deployment and developer tools for internal teams (Python, Scala, Databricks, AWS).
- Managed 2-3 developers across projects, delivering developer tools that enabled 50+ engineers to safely migrate microservices with 0 downtime.

### Consultant Data Scientist - Thales Digital Factory

2019 — 2021 · Singapore

Shipping ML models and streaming data products for an optimization engine in the defense sector.

- Built time-series models to improve operational reliability for a defense platform.
- Worked with the Digital Factory to ship a ML platform to internal teams through Dataiku Software and Azure.

### Founder in Residence & Lead Instructor - Entrepreneur First / General Assembly

2019 · Singapore

Built an early-stage product in construction intelligence while teaching data science full time.

- Prototyped defect detection workflows for construction stakeholders.
- Taught a 12-week data science immersive; refreshed materials on ML foundations.
- Coached teams on shipping lightweight experiments with Python and cloud notebooks.

### Innovation Fellow, Data Scientist - French Ministry for Higher Education & Research

2018 · Paris

Built NLP and knowledge graph pipelines to power the national research search engine (scanr.fr).

- Delivered an ETL that ingested research documents into a graph database with Airflow.
- Built a BiDAF-based model to classify documents by research concepts and surface them in search.
- Partnered with policy teams to prioritize signals and ship improvements without downtime.

### Data Scientist - AXA Data Innovation Lab

2017 · Paris

Improved insurance fraud detection models on top of Spark and Cloudera.

- Trained fraud scoring models in Scala/Spark; tuned pipelines for large datasets.
- Helped analysts productionize notebooks and align on evaluation metrics.

## **Data Scientist - Mention**

2016 — 2017 · Paris

Extended multilingual sentiment analysis with custom embedding work.

- Extended Gensim to train on aligned multilingual corpora to boost sentiment accuracy.
- Explored clustering strategies to automate tagging across languages.

## **Data Engineer - Stratagem Technologies**

2015 — 2016 · London

Delivered data pipelines for a trading team with real-time execution needs.

- Built asynchronous data ingestion and real-time views for execution algorithms.
- Added Spark-based parallel processing for backtests and analytics.

## **Projects**

### **BiasImpact - Founder**

2020 — Present

Tooling to surface bias in news feeds with lightweight NLP.

- Started with NLP prototypes for news bias detection; pivoted to a consulting practice shipping ML and data infrastructure.
- Delivered production MLOps pipelines, cloud migrations, and real-time data systems for multiple clients (from startup to MNCs).

## **Education**

### **Masters of Engineering Science - Ecole Nationale Supérieure des Mines de Nancy**

2013 — 2016 · Nancy, France

Focus on statistics, probability, and applied mathematics; exchange semester at TUM on machine learning.

### **Preparatory Classes (Maths/Physics) - Lycée Saint Louis**

2010 — 2013 · Paris, France

Intensive mathematics and computer science curriculum to prepare for Grandes Ecoles.