

Nathan Leclercq

Data Engineer · ML Engineer

✉ nathan.leclercq9@protonmail.com  [LinkedIn](#)  [GitHub](#)  [Blog](#)

Experience — DataKhi, Tourcoing (2023 - present)

Data Engineer — Nyukom · Full-time · Oct 2025 - present

- End-to-end telecom data platform: 3CX scraping, Centreon API, MinIO data lake, PostgreSQL star schema warehouse, Power BI reporting
- Infrastructure: K3s, Airflow, Ansible, private Docker registry
- Multi-tenant, partitioning, idempotency, historical backfill

ML Engineer — Hall U Need · Full-time (continued from apprenticeship) · 2023 - present

- Restaurant demand forecasting: XGBoost quantile, multi-restaurant prediction models, feature engineering (weather, calendar, reservations)
- Custom Huber loss, confidence interval calibration, non-regression tests
- Full pipeline: Microsoft Fabric collection → training → prediction

Data Engineer & ML Engineer — Tossée · Apprenticeship · 2023 - 2025

- Eco-fashion data ecosystem, multi-brand
- Scraping (Playwright, Scrapy, custom YAML rules), normalization, environmental impact (Ecobalyse)
- FastAPI/pgvector API (semantic search), Flutter app (virtual try-on, scan, OAuth)
- React/TypeScript browser extension, AI agent (OpenAI Agents SDK), hybrid Azure deployment

FullStack Developer — Internship · 2023 · 4 months — PowerBI versioning: C++ backend, React frontend, Electron

Skills

Data Engineering — Airflow · PostgreSQL (star schema, partitioning) · MinIO (S3) · Parquet/PyArrow · Microsoft Fabric · End-to-end ETL pipelines

Machine Learning — XGBoost · Feature engineering · Temporal cross-validation · pgvector/embeddings · CamemBERT/Transformers · MLflow · Custom loss functions

DevOps / Infra — K3s · Docker · Ansible (IaC, vault) · Proxmox · Prometheus/Grafana · Azure (Fabric, Functions, DevOps)

Development — Python (FastAPI, Pandas, scikit-learn) · SQL · TypeScript (React) · Dart (Flutter) · Playwright/Scrapy · Julia · Familiar with: Go, Rust, C++

Personal projects

MLOps Homelab Platform (2024 - present) — Proxmox, GPU servers, ML services, crewAI/RAG, Prometheus/Grafana monitoring, Ansible deployment, Docker registry, Gitea · [Articles](#)

Book Recommendation System (2023 - 2025) — Large book catalogue scraping, TF-IDF + CamemBERT embeddings, FastAPI API, PostgreSQL/pgvector, MLflow, Vue.js · [Articles](#)

Competitive Programming (2020 - 2024) — Optimized solutions in Julia · Google Hash Code, Reply Challenge, Cloudflight

Education

Master's in Machine Learning · University of Lille · 2023 - 2025 — Deep Learning, NLP, MLOps

Bachelor's in Computer Science · University of Lille · 2020 - 2023 — Advanced algorithms, distributed architecture

Mathematics Studies (3 years) · University of Lille · 2017 - 2020 — Numerical analysis, probability, linear algebra

Languages · Publications · Interests

French native · English professional (TOEIC 885)

[Articles on DataKhi blog](#) (2026) · [Articles on personal blog](#) (2024-2025)

Jazz/soul saxophone · Cycling · Badminton · Tabletop RPGs