

Nathan Leclercq

Data Engineer · ML Engineer

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

Expérience — DataKhi, Tourcoing (2023 - présent)

Data Engineer — Mission Nyukom · CDI · oct 2025 - présent

- Plateforme data télécom end-to-end : scraping 3CX, API Centreon, data lake MinIO, entrepôt PostgreSQL (schéma étoile), restitution Power BI
- Infrastructure : K3s, Airflow, Ansible, registre Docker privé
- Multi-client, partitionnement, idempotence, backfill historique

ML Engineer — Mission Hall U Need · CDI (continuité alternance) · 2023 - présent

- Prédiction d'affluence restaurant : XGBoost quantile, modèles de prédiction multi-restaurant, feature engineering (météo, calendrier, réservations)
- Loss custom Huber, calibration des intervalles de confiance, tests de non-régression
- Pipeline complet : collecte Microsoft Fabric → entraînement → prédiction

Data Engineer & ML Engineer — Mission Tossée · Alternance · 2023 - 2025

- Écosystème data mode éco-responsable multi-marques
- Scraping (Playwright, Scrapy, règles YAML custom), normalisation, impacts environnementaux (Ecobal-yse)
- API FastAPI/pgvector (recherche sémantique), app Flutter (virtual try-on, scan, OAuth)
- Extension navigateur React/TypeScript, agent IA (OpenAI Agents SDK), déploiement hybride Azure

Développeur FullStack — Stage · 2023 · 4 mois — Versioning PowerBI : backend C++, frontend React, Electron

Compétences techniques

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

Machine Learning — XGBoost · Feature engineering · Cross-validation temporelle · pgvector/embeddings · CamemBERT/Transformers · MLflow · Loss functions custom

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

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

Projets personnels

Plateforme MLOps Homelab (2024 - présent) — Proxmox, serveurs GPU, services ML, crewAI/RAG, monitoring Prometheus/Grafana, déploiement Ansible, registre Docker, Gitea · [Articles publiés](#)

Système de recommandation de livres (2023 - 2025) — Scraping d'un large catalogue de livres, embeddings TF-IDF + CamemBERT, API FastAPI, PostgreSQL/pgvector, MLflow, Vue.js · [Articles publiés](#)

Compétitions algorithmiques (2020 - 2024) — Solutions optimisées en Julia · Google Hash Code, Reply Challenge, Cloudflight

Formation

Master Machine Learning · Université de Lille · 2023 - 2025 — Deep Learning, NLP, MLOps

Licence Informatique · Université de Lille · 2020 - 2023 — Algorithmique avancée, architecture distribuée

Cursus Mathématiques (L1-L3) · Université de Lille · 2017 - 2020 — Analyse numérique, probabilités, algèbre linéaire

Langues · Publications · Centres d'intérêt

Français natif · Anglais professionnel (TOEIC 885)

[Articles sur le blog DataKhi](#) (2026) · [Articles sur mon blog personnel](#) (2024-2025)

Saxophone jazz/soul · Vélo · Badminton · Jeux de rôle sur table