

Amine SGHIR

Senior Data Engineer • Tech Lead • Freelance

Paris, France • sghir.ma@gmail.com • +33 6 95 70 26 00

linkedin.com/in/amine-sghir-29b598101 • github.com/pitchopp • pitchop.com



Profil

Data Engineer Senior avec **10+ ans d'expérience** Python, cloud (Azure / AWS) et pipelines à grande échelle. Spécialisé en **séries temporelles**, **ETL serverless** et **plateformes data internes** dans des contextes grands comptes (ENGIE x2, Crédit Agricole, EDF, TotalEnergies, Lyxor). Double casquette **AI/LLM** (Tech Lead Sodexo / Klaimy) et **fondateur** de 3 SaaS. Disponible immédiatement — **remote FR/EU/International** ou hybride 2j max.

Compétences

- **Data Engineering** — Pipelines ETL, TimeSeries, Kafka, Spark, PySpark, Pandas, NumPy, Dataiku, Artifactory, MapReduce
- **Python & Backend** — Python (10 ans), Django, FastAPI, Flask, Celery, Redis, REST APIs, Pytest
- **Cloud** — Azure (DevOps, ML, Databricks, OpenAI, Entra), AWS (S3, Lambda, Glue), C3.ai
- **DevOps & CI/CD** — Docker, Azure Pipelines, GitHub Actions, GitLab CI, Jenkins, Terraform, Ansible
- **Bases de données** — PostgreSQL, Cassandra, MS SQL Server, MySQL, CouchBase
- **Observabilité** — Application Insights, OpenTelemetry, Datadog, Grafana, Splunk, ELK
- **ML & GenAI** — Scikit-learn, text mining, clustering, LangChain, Azure OpenAI (GPT-5, GPT-4o), Claude, Gemini, LLM-as-judge

Expérience professionnelle

Senior Data Engineer — ENGIE GEMS (Trading énergie) • Freelance • 01/2023 — 05/2025

ENGIE GEMS achète et vend de l'énergie sur les marchés. Les analystes et traders s'appuient sur un large catalogue de séries temporelles — prix gaz et électricité (spot, forward, short-term), météo, flux de gaz liquide entre pays, etc. — pour leurs analyses et leurs décisions de marché.

- **Pipelines data cloud** (Azure, AWS) avec CI/CD : collecte, traitement, stockage et mise à disposition de la donnée aux équipes métier.
- **Outil de gestion timeseries** end-to-end : scrapers → traitement → stockage → catalogue + **système de formules avancé** pour combiner et dériver les séries.
- **Bibliothèques Python internes** déployées via Artifactory (doc + tests rigoureux), adoptées par plusieurs équipes.
- *Python • Azure • AWS • Azure DevOps • CI/CD • Artifactory • TimeSeries*

Tech Lead / Data Engineer — ENGIE (projet MASTERMIND) • CDI Visian • 06/2020 — 07/2022

Projet MASTERMIND (équipe de 11, SAFe) : permettre aux clients particuliers d'ENGIE de suivre et réduire leur consommation énergétique en centralisant les données des compteurs communicants Linky (électricité) et Gazpar (gaz).

- **Architecture ETL serverless** sur AWS (S3, Glue, Lambda) et plateforme **C3.ai** pour l'ingestion massive Linky / Gazpar.
- **APIs + jobs MapReduce** + modèles ML pour le traitement ; supervision prod (Splunk, Grafana) ; gestion d'incidents ; encadrement de développeurs.
- *Python • AWS Lambda/S3/Glue • C3.ai • PySpark • Pandas • Cassandra • PostgreSQL • Docker • Splunk • Grafana • MapReduce • Scikit-learn • Jenkins • Pytest*

Tech Lead & AI Engineer — Sodexo (projet Pegasus) • Freelance • 06/2025 — présent

Pegasus automatise l'enrichissement du catalogue produit Sodexo (food, services, équipements) à partir de documents fournisseurs multi-langues (PDF, fiches techniques, images).

Tech Lead unique du projet au sein de la Data Factory globale.

- **Architecture from scratch** : UI + API REST + workers asynchrones + **PostgreSQL**, abandon d'un workflow Azure Machine Learning Studio (AMLS).
- **Pipeline batch multi-étapes** avec retry, vote, scoring, comparaison à une base de référence — observabilité OpenTelemetry / Datadog.

- *Python · Azure OpenAI · LangChain · Django · Celery · Redis · PostgreSQL · Next.js · Docker · Azure Pipelines · OpenTelemetry · Datadog*

Data Engineer / Analyst / Data Scientist — Lyxor Asset Management · CDI Visian · 12/2017 — 06/2019

Lyxor (gestion d'actifs) : refondre un référentiel unique des données financières (transactions, clients, produits) éparpillées sur plusieurs SI non-réconciliés et industrialiser le reporting réglementaire AMF.

- **Triple casquette Data Eng / Analyst / Science** : architecture de déploiement, **reporting AMF automatisé**, alertes qualité ; réconciliation des positions clients, harmonisation des portefeuilles, dashboards **Tableau** ; text mining, clustering, classification clients et fonds.
- *Python · Pandas · Scikit-learn · API REST · Tableau · Google Analytics*

Senior AI & Backend Engineer — Klaimy (Insurtech IA) · Freelance · 01/2025 — 05/2025

Klaimy automatise les processus d'assurance par l'IA (souscription de polices, déclaration et instruction de sinistres) à destination des courtiers et compagnies.

- Conception **from scratch** de l'espace client web et de l'API ; **système de queue asynchrone** pour les traitements ML/LLM ; intégration des modèles prédictifs.
- *Python · Django · Next.js · React · Celery · Docker · Azure · OpenAI · Claude · Gemini*

Mentor Python & Data — OpenClassrooms · Indépendant · 04/2021 — présent

Accompagnement long-terme d'étudiants (Développeur Python, Data Analyst), évaluation des livrables et validation des soutenances. **50+ étudiants** passés et actuels.

Projets entrepreneuriaux

- **Locary** (*founder, depuis 2024*) — *locary.fr* — SaaS de gestion locative immobilière avec automatisation et IA : synchronisation bancaire, génération de quittances, relances automatiques.
- **Take My Keys** (*CTO & co-fondateur, depuis 2023*) — *takemykeys.fr* — Plateforme de conciergerie et gestion locative.
- **Judia** (*founder, 2024*) — *judia.pitchop.com* — Assistant email IA.
- **Open source** : `django-nextjs-tw-boilerplate` (template publié, GitHub).

Formation

- **Diplôme d'ingénieur** — École Centrale de Lille (2014–2017), spécialité **E-business & IoT**.
- **CPGE MPSI / MP*** — Lycée Michel Montaigne, Bordeaux (2012–2014), Mathématiques.
- **TOEIC** : 895 (avril 2017).

Certifications

- **AWS Certified Solutions Architect** — AWS (avril 2022)
- **Certified ScrumMaster** — Scrum Alliance (novembre 2019)
- **Machine Learning** — Columbia University (mars 2016)
- **Hadoop / Spark** — OpenClassrooms (juillet 2017)
- **Gestion de Projet** — MOOC (octobre 2014)

Langues

Français (natif) · Arabe (natif) · Anglais (courant — TOEIC 895) · Allemand (notions)