SID AHMED REDJINI

Data Engineer — Pipelines, Warehousing & Automation

Paris, France · 07 68 18 31 21 · sidprod752@gmail.com · Portfolio: sredjini.dev · LinkedIn: linkedin.com/in/sid-ahmed-redjini-349001303 · GitHub: github.com/sredjini

PROFILE

Data engineer building **reliable ingestion, transformation and storage**. Design pipelines and data models over **PostgreSQL**, **ClickHouse** and cloud platforms, in **Python, SQL and Rust** — including **LLM-driven data-quality** automation, FinOps pipelines and observable production systems.

SKILLS

- **Data & Pipelines:** SQL, PostgreSQL, ClickHouse, pgvector, ETL/ELT, data modeling, time-series, idempotent ingestion, BigQuery
- **Languages:** Python, SQL, Rust, TypeScript, Bash
- **Cloud & Orchestration:** Azure (CLI, Monitor, KQL), Scaleway, Docker, CI/CD, scheduling (cron / APScheduler)
- **Observability:** OpenTelemetry (traces, metrics), SigNoz, pipeline monitoring, structured logging
- **AI / LLM:** LLM integration, RAG (embeddings, hybrid search, vector search), tool-calling, MCP, evaluation (hit-rate, MRR)
- **BI:** Power BI (DAX, Power Query), reporting, KPIs

EXPERIENCE

Software / Data Engineer — Free / dataX
03/2026 – 09/2026 · Paris

Rust · Python · PostgreSQL · ClickHouse · SQL · LLM · ETL

- Designed a **normalized data model** and persistence across **PostgreSQL** and **ClickHouse**; transactional writes, scoring, scheduled campaigns.
- Built a **data-quality audit engine** in **Rust**: asynchronous ingestion pipeline over the data catalog, **LLM**-based evaluation with **structured JSON output**, bounded concurrency.
- Fetched fresh metadata directly from **ClickHouse** system tables at run time (freshness over cache); exposed computed scores via API — **consumed by other teams' ETL pipelines**.
- Instrumented the pipeline with **OpenTelemetry** (traces & metrics) into **SigNoz** — monitored run latency, throughput and failures.
- Contributed to a real-time voice agent and its **MCP** tool integration (Python).

Cloud / FinOps Data Engineer — Innoha
06/2025 – 10/2025

Azure · Azure CLI · Azure Monitor · KQL · Python · SQL · Power BI

- Built an **automated ingestion pipeline** for Azure metrics (Azure CLI, Monitor, KQL): collection → **normalization** → **historization** → BI model, **multi-subscription**.
- Cost/usage analysis and oversizing detection; delivered a **€29K/yr** savings recommendation to IT & Finance.

Software Engineer, Internal R&D — Free
07/2024 – 01/2025 · Paris

Python · CI/CD · Docker · Agile

- Built internal **Python** automation (extraction, cleaning) for **data quality** and reliability, integrated into **CI/CD** pipelines.

PROJECTS

Ardoise — Retrieval + Accounting Workflow
deployed

Python · FastAPI · LangGraph · RAG · Docker · Render

Live demo: ardoise-1.onrender.com · **GitHub:** github.com/sredjini/ardoise

Expense workflow using a PCG referential: extraction → retrieval → deterministic accounting calculations → independent verification, with tested safeguards for critical fields.

RAG Retrieval Pipeline in Rust
GitHub

Rust · Postgres · pgvector · embeddings · ETL/ELT · Docker

github.com/sredjini/rag-demarches-fr

Ingestion → chunking → **embeddings** → idempotent upsert into **Postgres/pgvector** (HNSW). **Hybrid** dense + full-text retrieval, plus an **evaluation harness** (hit-rate@k, MRR) driving measured gains (73%→100%).

EDUCATION

- **42 Paris** — Architecte numérique (certifying track) · 2021 – 2023
- **Université Paris II Panthéon-Assas** — Bachelor of Laws (Public Law) · 2018 – 2021

ADDITIONAL

- **Certification:** Azure Fundamentals (AZ-900) · **Languages:** French (native), English (fluent)
- **Publications:** FinOps & Azure SQL — Dev.to, LinkedIn, Cloud Efficiency Hub (2025)