

SAMIR ANBRI

Senior Data & AI Engineer — 8 yrs Machine Learning | RAG | AI Agents | Evaluation & Guardrails

Bordeaux, France | +33 6 36 87 26 71 | samir.anbri@gmail.com

linkedin.com/in/samir-anbri | github.com/AnbriSamir | samir-anbri.com

Professional Summary

Eight years in machine learning, from research to revenue. Data engineering started for me at UC Berkeley and Lawrence Berkeley National Lab, then four years of R&D at a major French research institute.

Then I crossed to the business side: I own two modules of a traffic-analytics SaaS end to end, the neural networks, the product and the clients who pay for them (**1M EUR** Annual Recurring Revenue, **27%** of company revenue). For the last two years I have held Large Language Model (LLM), Retrieval-Augmented Generation (RAG) and AI Agent systems to that same bar of evaluation, testing and guardrails, and I lead AI adoption for my team.

I am looking for a tech company where AI is an explicit strategic bet, with a team that raises my level. If I quote a number, I measured it.

Skills

- **Core stack (daily):** Python, TensorFlow/Keras, deep neural networks, Large Language Models (LLMs), AI Agents, Retrieval-Augmented Generation (RAG), SQL
- **AI engineering:** LangGraph, LangChain, Model Context Protocol (MCP), Vector Databases (Qdrant), hybrid search (BM25 + dense, reranking), Human-in-the-Loop (HITL) gates; Anthropic Claude, OpenAI GPT-4o, Google Gemini, open-source bge embedding & reranking models
- **Evaluation & reliability:** golden-set eval harnesses, LLM-as-judge gated on Cohen's kappa, bootstrap 95% confidence intervals, deterministic replay, policy-as-code guardrails, prompt-injection defense
- **Machine learning:** scikit-learn, quantile neural networks, time-series forecasting, geospatial feature engineering, MLflow, statistical release gates
- **Data & platform:** PostgreSQL/PostGIS, geospatial pipelines, Docker, Kubernetes, FastAPI, Sentry, Prometheus

Professional Experience

Senior Machine Learning Engineer, AI Lead | Traffic analytics company owned by the Vinci Group | Bordeaux Jan 2024 - Present (2 yrs 8 mos)

Road-traffic analytics company (I keep its name confidential while I am employed there). Its SaaS helps metropolises and departments manage their road infrastructure. I own two modules end to end and lead AI adoption across the team.

- Grew the flagship traffic-flow module from **1 pilot to 23 territories**, owning models, geospatial pipeline and serving end to end: **1M EUR** Annual Recurring Revenue (ARR), **27%** of company revenue.
- Built the deep quantile neural networks that turn partial connected-vehicle data (Floating Car Data, FCD) into all-vehicle traffic volumes (TensorFlow/Keras), held to hard tolerance constraints against field measurements on every territory: **under 9% error at the 80th percentile, 98% of sensors within tolerance** on the best-calibrated territory.
- Automated ingestion and map-matching of **30,000+ client sensor feeds** and **172 GB** of Floating Car Data (STRtree, Hausdorff, Hungarian assignment): **99.93%** segment coverage, **zero** manual matching.
- Shipped the Intersection module to production alone in **2 weeks** against a 7-8 month lead time, replacing a 3-freelancer team: multi-agent AI orchestration with generated evaluation datasets and context engineering, held to the platform's standards, tested, deployed and sold.
- Named AI referent for the team: I set up the AI agents and the ground rules everyone codes with, review what they produce, and coach engineers on where AI helps and where it does not: **11 specialist sub-agents, 2 automated reviewers**, adopted team-wide.

Data & AI Engineer, Founder of the Sustainable Mobility unit | Nepsen (energy-transition consultancy) | Bordeaux

Jan 2023 - Dec 2023 (1 yr)

Unit created from zero, sales to software, and run as a business.

- Created the Sustainable Mobility unit from zero, sales to software: **0 to 4 engineers** hired, mentored and led, **650k EUR** revenue in year one across **~20 client projects**.
- Shipped two SaaS products solo: a carbon-audit prospecting tool (open-data ETL, LLM scoring) that became Glan, now paid for by Nepsen as a client, and the multi-tenant CRM running the unit's ~20 projects.

Machine Learning Research Engineer | IFP Energies Nouvelles (IFPEN) | Lyon area

Sep 2018 - Dec 2022 (4 yrs 4 mos)

- Led IFPEN's contribution to SAM, the **~110M EUR** national autonomous-vehicle programme, across an **18-partner** cross-functional consortium: managed **3 senior engineers** and co-designed the evaluation methodology the consortium published.
- Ran my own research line inside it: machine learning models predicting traffic flows and vehicle emissions from speed profiles (Python, scikit-learn, TensorFlow/Keras), **~10% median error** against field measurements.
- Shipped the transport-mode classifier behind GECO Air, a national mobility app, as both product manager and ML engineer: **~85% accuracy**, **~20,000 users**; drove UX/UI and front-end delivery with freelance designers and developers.

Selected Projects

Three MIT repos run as one loop: the agents produce the traces, the traces become the benchmark, and I ship only what passes it.

- [Multi-Agent Orchestrator](#) (2026): LangGraph orchestration with bounded execution proven by test, fail-closed tool registry and deterministic span-by-span replay; **255 offline tests**, **7 Architecture Decision Records (ADRs)**.
- [Eval Dataset Generator](#) (2026): production logs turned into provenance-stamped golden datasets, gated on measured human agreement; **Cohen's kappa 0.804** (double-blind, 95% confidence interval), **677 offline tests**.
- [Hybrid RAG Pipeline](#) (2026): hybrid retrieval (dense + BM25 + reranking) grounding my daily LLM answers in my own work corpus; **faithfulness 0.981**, **attribution 55/55**, **nDCG@10 0.983**, **370 offline tests**.
- **AI Recruiter (AI Salimi)** (2025): agentic screening product delivered solo for a paying client; **1,894 applications** screened, 950 scored down to a shortlist of 3, **3 hires** completed.

Education

- **IFP School (ENSPM)** — Engineering double degree, Energy & Mobility (2018-2019), on a work-study contract at IFPEN.
- **INSA Lyon** — Diplôme d'ingénieur, general engineering (2013-2018), with an exchange year at **UC Berkeley** (2016-2017) and a research internship at **Lawrence Berkeley National Laboratory**: statistical analysis in R of indoor-air-quality field data, contributed to the HENGH study (Chan et al., 2019). My first data work.

Languages

- **English** — bilingual (TOEIC 910/990, 1 year at UC Berkeley) | **French** — native