Felipe Diogo

Data Engineer · Python Developer · AI & Analytics Engineering

Lisbon, Portugal | Open to relocation within EU or US

+351 928 381 366 felipe.g.diogo@gmail.com Linkedin Github Personal Website

Summary

Python developer and data engineer with 3+ years building end-to-end data systems in production: ETL/ELT pipelines, time-series platforms, ML inference at the edge, and AI-powered analytics. Strong SQL and Python (pandas, NumPy, scikit-learn) with a pragmatic track record of shipping real systems on Linux, Docker and CI/CD rather than notebooks that never leave a laptop. My control-and-automation engineering background is a concrete advantage for data work: I have spent years upstream of any dashboard, wiring the sensors, MQTT brokers, PLCs and databases that generate the kind of messy, high-volume, multi-source time-series data most teams only see after the fact. Daily user of AI tooling (Claude Code, OpenAI API, AI agents) as a genuine force multiplier across coding, data wrangling, schema design and debugging, the kind of setup that turns one engineer into a small team's worth of output. Autodidact, fast to pick up a new stack, and used to owning the full loop from raw ingestion to model, API or client-facing insight. Actively seeking a Data Engineer or Python Developer role; Movewer Technologies will continue in an advisory capacity to ensure full commitment to the new role.

Technical Skills

Languages & Core Data: Python (pandas, NumPy, scikit-learn), SQL, JavaScript/TypeScript, Bash, some R; PostgreSQL, SQLite, InfluxDB (time-series), Supabase.

Data Engineering & Pipelines: ETL/ELT design, data modeling, REST APIs, webhooks, web scraping, ingestion from databases and sensor streams (MQTT, OPC-UA, Modbus), Telegraf, scheduled jobs, systemd timers, data validation, awareness of orchestration tools (Airflow, Prefect, Dagster, n8n) and modern data-warehouse patterns.

ML & AI: scikit-learn, TensorFlow, TFLite (XNNPACK), Anomalib (PaDiM), OpenCV, computer vision, edge AI inference, anomaly detection, predictive analytics; OpenAI API, Claude Code, AI agents, prompt engineering, LLM-in-the-loop workflows, Retrieval-Augmented Generation patterns; familiar with PyTorch and MLOps basics.

Backend & APIs: FastAPI, Node.js, Express.js, REST and WebSocket APIs, Supabase Edge Functions, authentication and session design, API versioning, async I/O.

Cloud, DevOps & Infra: Docker and Docker Compose, Linux / VPS administration, Git/GitHub, GitHub Actions (CI/CD), Nginx reverse proxy, Let's Encrypt / TLS, Vercel, AWS (familiar), environment hardening, encrypted backups, systemd services.

Analytics & Visualization: Grafana, Power BI, Next.js dashboards, Jupyter notebooks, KPI dashboards, ad-hoc client-facing analysis, natural-language insight generation via LLM APIs.

Industrial Data (Differential): Large-scale exposure to real-world noisy time-series from industrial sensors, cameras, PLCs, SCADA historians; IT/OT convergence; edge computing on Raspberry Pi and Jetson.

Experience

Movewer Technologies **Jun. 2025 – Present**

Founder & Data/IT Engineer *Lisbon, Portugal*

- Designed and deployed a production IIoT data platform with a full ETL path (MQTT/TLS → Telegraf → InfluxDB → Python analytics → Grafana) running as six Docker services on a VPS; added AES-256 encrypted backups, Nginx with Let's Encrypt, UFW, non-root UIDs, and CI/CD via GitHub Actions. Commercial-grade data monitoring at zero recurring cost.
- Built a Python analytics layer that reads sensor time-series from InfluxDB, runs anomaly detection and trend analysis, and exposes AI-generated natural-language insights through a FastAPI REST service powered by an LLM (OpenAI-compatible API).
- Shipped an edge ML visual-inspection stack on Raspberry Pi 4 with PaDiM anomaly detection (Anomalib) and TensorFlow/TFLite, plus a SQLite history store, FastAPI + WebSocket live stream, and a SCADA-style Next.js dashboard; <500 ms end-to-end latency, fully offline.
- Engineered a real-time teleoperation data path for UGVs with 50 Hz MQTT/TLS telemetry, 720p WebRTC video at 30 fps, hardware fail-safe watchdog, and a 74-test automated suite across three repositories.

Personal Projects & Independent Work **2024 – Present**

Python Developer & Data Engineer *Lisbon, Portugal*

- Built Julius, an AI-powered personal-finance PWA (Next.js 15, Supabase/PostgreSQL, OpenAI API): an AI agent extracts vendor, amount, category and date from natural-language user input, writes structured rows to Postgres via Edge Functions, and drives interactive dashboards filtered by category and period.
- Developed Python ETL pipelines that pull from REST APIs, transform with pandas, and persist to PostgreSQL on scheduled runs; applied the same pattern to a LinkedIn content pipeline using n8n/Make plus Telegram human-in-the-loop approval.
- Built internal tooling around Claude Code and the OpenAI API (custom agents, prompt templates, code-review loops, data-exploration assistants), using LLMs as a daily multiplier on real engineering output rather than a demo toy.

Psyche AeroSpace **Jun. 2024 – Jun. 2025**

Automation & UGV Engineering Manager *São Paulo, Brazil*

- Led a 5-engineer cross-functional team across two product lines (agricultural UGVs and drone-support infrastructure), running structured Scrum sprints on Jira and cutting time-to-prototype by ˜30%.
- Defined data-collection and test-analysis workflows for prototype telemetry: field-test data capture, structured logging, and post-run analysis supporting engineering decisions and investor-facing technical proof points.
- Translated operational requirements into technical user stories, data schemas and automation logic, bridging business stakeholders and engineering delivery.

Controvale **Jun. 2018 – Jun. 2024**

Automation Engineer (Industrial Data & Systems Integration) *São Paulo, Brazil*

- Integrated 20+ industrial systems with continuous sensor streams over Modbus, Profinet and EtherCAT into SCADA historians and client databases; designed the data models behind operator dashboards and KPI reporting used across 35+ client facilities.
- Built reporting and monitoring pipelines (SQL, spreadsheets, Grafana, Power BI) to surface production metrics, process drift and maintenance signals from years of messy, real-world industrial time-series.
- Worked directly with clients on requirements gathering, data needs, and post-deployment analytics, sharpening an ability to translate vague business asks into concrete queries and dashboards.

Projects

AI Data Platform | *Python, MQTT, InfluxDB, Telegraf, Grafana, Docker, FastAPI, LLM API* **Jan. 2026**

- Production IIoT data platform on a VPS: six containerized services (Mosquitto, Telegraf, InfluxDB, Grafana, Python analytics, encrypted backup), six-layer security (TLS 1.2+, AES-256 backups via Rclone, UFW, non-root UIDs, OpenSSL-generated credentials, provisioned Grafana), Docker Compose dev/prod overlays, Nginx with Let's Encrypt, and AI-generated natural-language insights via a custom REST API. GDPR-compliant; a zero-cost alternative to commercial SaaS (US$50 to 200/month). ***Click to check it***

Julius — AI Personal Finance PWA | *Next.js 15, TypeScript, Supabase (PostgreSQL), OpenAI API, PWA* **Feb. 2026**

- Mobile-first PWA where users describe expenses in natural language; an AI agent extracts vendor, amount, category and date, then writes structured entries to Supabase (PostgreSQL with Edge Functions). Includes per-user accounts, dashboards with interactive charts filtered by category and period, and app settings. End-to-end example of data app design: ingestion, normalization, storage, query and presentation. ***Click to test it***

MLAI — Edge ML Visual Inspection | *TensorFlow/TFLite, Anomalib (PaDiM), Python, FastAPI, SQLite, Next.js* **2026**

- Self-contained edge ML inspection station on a Raspberry Pi 4 with two inference modules (PaDiM anomaly detection on MVTec AD for industrial QC, and SSD MobileNet V2 for fruit detection and quality grading), unified under a SCADA-style Next.js dashboard with FastAPI + WebSocket live stream, camera calibration, dimensional measurement (px→mm), and SQLite history. Runs fully offline via TFLite (XNNPACK); three auto-restarting systemd services; <500 ms end-to-end latency. ***Click to check it***

Education

Federal Institute of São Paulo (IFSP) **Jan. 2019 – Dec. 2023**
Bachelor of Control and Automation Engineering *São José dos Campos, Brazil*

Self-Directed Software, Data & AI Engineering **2022 – Present**
freeCodeCamp, One Bit Code, hands-on project work with LLM tooling and the Python data stack *1000+ hours*

Languages

Portuguese — Native **English** — Advanced (C1) **Dutch / French / German** — open to learning for role