

# Paweł Żak

Senior Data Engineer

✉ zakpaw36@gmail.com    📞 +48 693 937 144    🌐 linkedin.com/in/pawel-zakpaw    📍 Remote - Gdańsk, Poland

## SUMMARY

Led **11 engineers** – fraud detection for **EU carbon-credit trading**, org-wide data strategy for **global aviation**. Previously **Amazon**: fulfillment analytics behind **millions in maintenance savings**. Outside of work, I build **MyElib**, cross-platform SaaS for memory retention.

## SELECTED PROFESSIONAL EXPERIENCE

### ■ Addepto – Big Data & AI

Warsaw (Remote)

Tech Lead, Data Engineering

Feb 2025 – Nov 2025

- Led **11 engineers** across two flagship projects; part of **10-person leadership team** steering company direction. Owned architecture, hiring, performance reviews, and client relationships.
- Architected **fraud detection for EU carbon-credit trading** – **15s end-to-end latency** on TB-scale XML with sequential validation constraints. Built multi-tenant Databricks platform with strict row/column security for **100+ analysts**.
- Led **data strategy for global aviation provider** (5k+ employees); unified fragmented architecture across autonomous teams. Built first data product – live baggage tracking – **securing executive approval to scale org-wide**.

Senior Data Engineer

Feb 2024 – Feb 2025

- Built **Nix-based** CI/CD framework – abstracted pipelines to four commands, making GitHub Actions, GitLab CI, and Azure DevOps interchangeable. Used by **~40% of company projects**.
- Built **enterprise Databricks platform template** – three automated environments with Terraform, private networking, Unity Catalog governance, DAB, and data contracts.

Data Engineering Consultant (part-time)

Apr 2023 – Feb 2024

- Replaced on-prem SQL Server with BigQuery, DuckDB, dbt, and Terraform, **cutting warehouse costs ~70%** with **zero-downtime migration**.

### ■ Amazon

Luxembourg (Remote)

Data Engineer

Jun 2022 – Jan 2024

- Defined and owned Cost-per-Asset and Lost Productivity Hours metrics approved by VP and adopted by operations leaders; contributed to **millions in annual maintenance savings**.
- Cut reporting latency **from 24 hours to 15 minutes** – CDC (Kafka, Redshift) on TB-scale fulfillment data.
- **Killed 1,000-line SQL monoliths** in QuickSight, rebuilt as dbt semantic layer – **established self-service analytics** across Global RME.

Business Intelligence Engineer

Feb 2021 – Jun 2022

- Built serverless data pipelines (Glue, Lambda) for near-real-time conveyor-jam detection, **reducing average conveyor downtime by 10%**.

## ENTREPRENEURIAL & COMMUNITY EXPERIENCE

### ■ MyElib – Knowledge Retention SaaS

Founder

Dec 2025 – Present

- Shipped SaaS to **five platforms solo** (Web, Mac, Windows, iOS, Android) – turns e-reader highlights into LLM-powered spaced-repetition flashcards. **30 paying customers**.
- **Own product end-to-end**: idea, roadmap, pricing, customer acquisition, releases, and support.

### ■ Data Science Student Research Group

University

President

Nov 2020 – Nov 2024

- Scaled **from 20 to 300+ members**; hosted 30+ workshops and a local data conference.
- Secured a **€30k grant** and **led a research project with 15 students** - the Polish real-estate market recommendation engine.

## SKILLS

**Deep:** Python, Apache Spark, SQL & dbt, Databricks, Kafka/Redpanda/Event Hub, Azure, Terraform, Golang, CI/CD, Bash, Nix, GenAI

**Proficient:** AWS, GCP, Kubernetes, Redshift, BigQuery, Airflow, DLT

**Project-specific:** ClickHouse, TypeScript/React/Astro, Dagster, Prefect, SvelteKit, Rust

## EDUCATION

**Gdańsk University of Technology** BEng, Data Engineering

Nov 2019 – Feb 2023

## INTERESTS

• Robotics

• Kitesurfing

• Snowboarding

• Traveling

• Reading

• Cooking