

# Aliaksandr Zaparozhtsau

aliaksandr.zaparozhtsau@gmail.com

Poland, Warsaw | linkedin.com/in/zap-al | +48 (451) 246 926

## OBJECTIVE

Senior Data Engineer with 5+ years of experience in *clouds, data processing, orchestration, storage design, E2E pipelines*. Seeking full-time *Data Engineering, Solution Architect and DevOps-related* roles.

## PROFESSIONAL EXPERIENCE

### Capgemini – Senior Data Engineer – Poland

Aug 2023 – Present

#### Data Processing

- Codegeneration for Azure Databricks using Python and Jinja2
- Data processing using Apache Spark
- Implementing Azure Pipelines for orchestration of processing
- Implementing Azure Data factory pipelines for data processing

### Mimesis Inc – Senior Data Engineer & Co-Founder – Armenia

Jan 2021 – May 2023

#### AI-based mobile image editor (Jan 2021 – May 2023)

- Built end-to-end data scrapping for ML to AWS S3, AWS RDS using Spark (PySpark), NumPy, OpenCV
- Established Business to development communication processes
- Architected and managed backend REST API development from scratch using Python and FastAPI
- Flexibility improvemnt with YAML and Jinja templating
- Build docker containers for ML, Backend production, ETL pipelines
- Designed and implemented:
  - Data scrapping (Python BS4, Selenium) and aggregation workflows
  - Database (AWS RDS) structure
  - CI/CD pipelines (Jenkins)
  - REST API for data manipulation
  - Social bots (Selenium, REST API)
  - Clusters orchestration and monitoring (AWS ECS, AWS EKS)
  - Automation of large number of tasks by Jenkins and BASH scripts

### EPAM Systems – Data Engineer – Belarus

Jul 2019 – Sep 2021

#### eCommerce sales optimization by ML (Sep 2020 – Sep 2021)

- Built Airflow end-to-end data pipelines including custom operators for storage monitoring
- Implemented Apache Spark jobs wiht Kafka as datasource and Azure Blob Storage as target
- Done bunch of Azure Jenkins CI/CD pipelines to manage artifacts

#### PowerBI Analytical platform (Jan 2020 – Jul 2020)

- Developed Azure Data Factory ETL pipelines to ingest and deliver data
- Implemented Azure Pipelines for CI/CD
- Architected and implemented STAR database data model
- Data processing using Azure Data Factory
- Implemented data analytics engine for PowerBI

#### Marketing analytics platform (Oct 2019 – Jan 2020)

- Built Azure Databricks notebooks based on python wheel to make solution platform agnostic
- Implemented CI/CD for python wheel package from Github to Azure Databricks via Jenkins
- Architected and developed data analytical python wheel using Apache Spark
- Implemented multi-layer multi-zone optimal structure for Azure Blob Storage
- Implemented storage monitoring system for tables/columns metadata to monitor data updates

## EDUCATION/CERTIFICATIONS

### Azure Data Fundamentals – Microsoft

Feb – 2024

- Certificate link: microsoft website

### Azure Fundamentals – Microsoft

Dec – 2023

- Certificate link: microsoft website

### English B2 Language School – Streamline Language School

Sep – 2020

- Successfully passed exam with writing and speaking parts
- Certificate link: streamline website

### M.S. in Physics & Programming – BSU

Sep 2016 — Jun 2021

- Done 3 coursework and 1 diploma related to Computer Science

## KEY COMPETENCES

**Languages:** Python, NodeJS, SQL. **Clouds:** AWS, Azure, Terraform, Kubernetes. **Databases:** PostgreSQL, Microsoft SQL, MongoDB. **DevOps:** Jenkins, Airflow, Azure DevOps, Azure Pipelines. **Other:** Apache Hadoop, Apache Spark, Linux, Databricks, Lambdas.