

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

# Aliaksandr Zaparozhtsau

aliaksandr.zaparozhtsau@gmail.com

## Education/Certifications

---

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

*Sep 2016 — Jun 2021*

- Done 3 coursework related to Computer Science
- Done diploma related to Computer Science

### English B2 Language School – *Streamline Language School*

*Sep – 2020*

- Learned language for B2 level
- Successfully passed exam with writing and speaking parts
- Certificate link: streamline website

### Azure Fundamentals – *Microsoft*

*Dec – 2023*

- Structurized basic knowledge about Azure Cloud
- Successfully passed exam
- Certificate link: microsoft website

## 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*)

- Builded end-to-end pipelines of data scrapping from web, processing it and storing for following using with ML to AWS S3 and AWS RDS using Spark (PySpark), Pandas, 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*)

- Builded 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*)

- Builded 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

## 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.