

Mariusz Strzelecki

DATA ENGINEER · PYTHON DEVELOPER

Śliwkowa 13 Street, 87-123 Brzozówka, Poland

☎ (+48) 603-093-358 | ✉ szczeles@gmail.com | 📱 [szczeles](#) | 🌐 [mariusz-strzelecki](#)

"The impossible is often the untried."

Work Experience

GetInData

Poland, Remotely

DATA ENGINEER

08.2019 - ...

- Created and maintained ETL system based on AWS using Airflow, EMR, Spark and Glue.
- Implemented data anomalies detection module to alert about abnormal traffic.
- Designed and implemented system for Machine Learning support (training, serving, scheduling) based on Kubeflow and MLflow and Seldon.

Self-employed

Remotely

FREELANCER

05.2018 - ...

- Implemented Airflow-oriented data warehousing systems.
- Deployed and maintained Hadoop on-premise Hadoop distributions.
- Implemented Kafka-based systems for streams ingestion, processing and visualisation using Kafka Streams, Connect, Druid and Superset.
- Optimized big data ETLs in PySpark and SparkR.
- Data science applications, including text mining.

Fandom Inc.

Poland, Remotely

DATA ENGINEER

08.2018 - 08.2019

- Maintained AWS-based data pipelines based on batch and streaming processes.
- Co-created Open Source projects related to data processing and querying.

DXC Technology

Poland, Remotely

BIG DATA PLATFORM EXPERT

06.2017 - 07.2018

- Designed and created Big Data platforms system with Apache Kafka, Apache Hadoop and Apache Spark.

Allegro Group

Toruń, Poland

SENIOR DATA ENGINEER

04.2015 - 05.2017

- Co-created and maintained Big Data ecosystem based on Apache Hadoop and ETL between Apache Kafka and Hadoop.
- Conducted workshops on Apache Spark.
- Maintained recommendation and relevancy-targeting projects based on big data analysis.

Allegro Group

Toruń, Poland

SENIOR JAVA/PHP DEVELOPER

07.2010 - 03.2015

- Implemented and maintained microservices for storing and serving files, operating on thumbnails and content validation.
- Co-created mobile sites.

Education

Nicolaus Copernicus University

Toruń, Poland

M.S. IN COMPUTER SCIENCE

2006 - 2011

Skills

DEVELOPMENT

- very good knowledge: Python, Java, scripting languages
- good knowledge: R, PHP, Scala

TECHNOLOGIES

- **Big Data:** Spark, Hadoop, Hive, Presto, Oozie, Airflow, Druid, Elasticsearch, Kubernetes
- **Data Science/ML Ops:** Kubeflow, MLflow, Seldon, Jupyter, Amundsen
- **AWS:** S3, Athena, EMR, Fargate Glue, DynamoDB, CF, CDK
- **Streaming:** Kafka (Streams, Connect, KSQL), Debezium

SOFT SKILLS

- Experience in conducting pySpark, Hadoop and Kafka workshops for developers and analysts.
- A few public presentations on Kafka, Hadoop, Big Data and Spark.
- Interns mentoring during 3 internship programs.
- Contributions in several big open source projects: Apache Hive, Apache Spark, Hue, Apache Airflow.
- 3-month experience in being Product Owner of own project.

LANGUAGES

- Polish: native.
- English: fluent.

INTERESTS

- DIY
- Running
- 3d printing