Marko Kolarek

Data Engineer

Berlin, DE | marko.kolarek@proton.me | markokolarek.com

SUMMARY

As a data engineer, I have extensive experience designing and implementing scalable data architectures on AWS and Azure. My expertise includes building robust data pipelines, managing large-scale migrations, and optimizing cloud infrastructure for e-commerce, fintech, and legaltech companies. I excel at integrating big data technologies such as Apache Spark and Kafka, and have led cross-functional teams to deliver innovative solutions in fast-paced environments.

PROFESSIONAL EXPERIENCE

Data Engineering Consultant (Self-employed)

April 2024 – Present

- Designing and implementing data platform architecture
- Prototyping and productionizing machine learning systems
- Mentoring and career advice

Selected projects:

- **Data engineer, VAAAM AG (legaltech)** - Critical data migration (multiple TBs, ~10M images), infrastructure optimization
- **Data & Infra engineer, Stealth (AI-based customer service)** - Data modeling, architecture design, web app development
- **Fractional CTO, ATMO GmbH (fintech)** - Led technology strategy and execution to build a scalable, AI-enhanced platform, aligning product development with business goals and driving innovation through ML and LLM integration

Skills: Data Architecture | Data Engineering | Data Modeling | Cloud Infrastructure | AWS | Azure | Python | SQL

Staff Data Engineer & Interim Head of Data, GROPYUS Technologies GmbH

March 2023 – March 2024

- Designed and implemented data architecture and data modeling to ensure integration with IT systems.
- Led daily operations of data teams and collaborated with product management on technical roadmaps.
- Contributed to service mesh integration planning and implementation.

Skills: Azure | DevOps | Kubernetes | Terraform | Data Architecture | SQL | Python

Head of Data, VAAAM AG

April 2022 – February 2023

- Managed data and backend engineering teams.
- Designed and implemented data architecture and models while collaborating on product strategy.
- Maintained investor and vendor relationships.

Skills: AWS | Data Strategy | Data Engineering | SQL | Python | Business Intelligence | Terraform

Senior Data Engineer, GROPYUS Technologies GmbH

November 2020 – March 2022

- Developed event-based streaming pipelines using serverless solutions.
- Implemented digital twin infrastructure based on semantic web stack.

Skills: Azure | Data Pipelines | Knowledge Graphs | RDF | Streaming Data

Senior Data Engineering Consultant, DAIN Studios GmbH

March 2019 – October 2020

- Developed data ingestion and processing pipelines in Spark (Python and Scala) and Hive.
- Designed and implemented customer data platforms.

Skills: AWS | Apache Spark | Data Modeling | SQL | Python | Azure | Hive

Previous Experience (2016 & Earlier)

- **Data Engineer, FlixBus SE (2018)** – Implemented and maintained data pipelines for feature engineering in PySpark.
- **Research Engineer, Zalando SE (2016-2018)** – Developed and automated A/B test analyses using Apache Spark.
- **Backend Engineer, Zalando SE (2015-2016)** – Built data ingestion and processing pipelines.
- **Research Associate, Visage Technologies AB (2014-2015)** – Researched and implemented eye gaze detection models.

SKILLS / TECH STACK

- **Data Engineering & Architecture**: ETL/ELT Development, Data Warehousing, Streaming Data Solutions
- **Cloud Technologies**: AWS (S3, EMR, EC2), Azure (Event Hub, AKS)
- **Programming**: Python (PySpark, Flask), SQL, Terraform, Kubernetes
- **Big Data Technologies**: Apache Spark, Kafka, Hive
- **DevOps & Infrastructure**: Terraform, Kubernetes, CI/CD Pipelines
- **Statistical Analysis & Machine Learning**: A/B Testing, Hypothesis Testing, ML Prototyping

EDUCATION

- **M.Sc. in Information Processing**, FER, University of Zagreb (2012 – 2014)
- **B.Sc. in Information Processing and Multimedia Systems**, FER, University of Zagreb (2008 – 2012)

INTERESTS

- Running, Hiking, Cats