



Marko Kolarek

Freelance Data Engineer



Berlin, DE



marko.kolarek@protonmail.com



<https://mkolarek.github.io/>



BACKGROUND



ABOUT

I am Marko Kolarek, a dedicated data engineer with over 9 years of experience in designing, building, and optimizing data architectures. My career has spanned across diverse industries, including e-commerce, travel, construction & real-estate, and legal where I have successfully implemented scalable data solutions and contributed to data-driven decision-making.

With a strong background in ETL processes, data warehousing, and big data technologies, I excel in transforming complex data into actionable insights. I am proficient in SQL, Python, Scala and various data pipeline tools such as Apache Spark and Kafka. My expertise also includes cloud platforms like AWS and Azure, enabling efficient and secure data storage and processing.

I am recognized for my analytical mindset, meticulous attention to detail, and ability to collaborate effectively with cross-functional teams. My passion lies in solving challenging data problems and continuously improving data infrastructure to support business growth and innovation.

I am eager to bring my skills to a forward-thinking organization where I can drive the development of robust data systems and contribute to impactful projects. My goal is to work in an environment that values continuous learning and the application of cutting-edge data technologies.

Outside of work, I am a tech enthusiast and proud owner of three cats. I also love running, hiking and exploring the great outdoors, which helps me stay energized and inspired.



WORK EXPERIENCE

- **Freelance Data Engineer**, Freelance consultant

Apr, 2024 - Present

- Data Engineering: Expertise in ETL pipelines, data warehousing, and data lakes.
- Big Data Technologies: Proficient in Hadoop, Spark, Kafka, SQL, and NoSQL databases.

- Cloud Platforms: Experienced with AWS and Azure for data storage, processing, and analytics.
- Data Modeling & Architecture: Strong ability to design logical and physical data models, ensuring data integrity and performance.
- Programming Languages: Proficient in Python, SQL, and Scala.
- Data Infrastructure: Experienced with Docker, Kubernetes, and Terraform.
- Data Governance: Ensuring compliance with data privacy regulations and best practices in data security.
- Data Architecture
- Data Engineering
- Amazon Web Services (AWS)
- Microsoft Azure
- Cloud Infrastructure
- Devops
- Python

Staff Data Engineer, GROPYUS

Mar, 2023 - Mar, 2024 ⓘ 1 year

- Overarching role in data department, participating in daily business of both data teams
- Designing and implementing data architecture, ensuring successful integration with overall IT architecture
- As interim Head of Data, collaborating with product management on product roadmap, aligning with other IT departments and providing higher-level management to data department
- Participated in broad IT infrastructure efforts with planning and implementation, such as service mesh migration
- Microsoft Azure
- Data Architecture
- Terraform
- Kubernetes
- Knowledge Graphs
- Python

Head of Data, VAAAM

Apr, 2022 - Feb, 2023 ⓘ 10 months

- Designing and implementing data architecture, ensuring successful integration with the overall IT architecture
- Managing data science, analytics and engineering efforts operationally and professionally
- Working with CTO on data and product strategy, investor and vendor relationships
- Working with CTO to establish VAAAM AG externally as a leader in engineering topics
- Amazon Web Services (AWS)
- Data Architecture
- Data Engineering
- Python
- Terraform
- Kubernetes
- Django

- Microservices

Senior Data Engineer, GROPYUS

Nov, 2020 - Mar, 2022 ⌚ 1 year 4 months

- Developed event-based streaming pipelines based on serverless solutions
- Implemented infrastructure for digital twin and related products
- Participated in interviewing, onboarding and mentoring of new and existing colleagues
- Microsoft Azure
- Data Engineering
- Streaming data
- Terraform
- Kubernetes
- Knowledge Graphs
- Resource Description Framework (RDF)
- Python

Senior Data Engineering Consultant, DAIN Studios

Mar, 2019 - Oct, 2020 ⌚ 1 year 7 months

- Developed and maintained existing data ingestion, cleaning and processing pipeline for large remote learning provider
- Designed and implemented data platform for large telecom provider
- Collaborated on various data architecture project proposals
- Amazon Web Services (AWS)
- Microsoft Azure
- Apache Kafka
- Scala
- Python
- Data Architecture

BE2 Cohort Member, Entrepreneur First

Oct, 2018 - Apr, 2019 ⌚ 6 months

- Researched and prototyped different start up ideas, with a focus on deep tech
- Market Research
- Entrepreneurship

Data Engineer, Flixbus

Feb, 2018 - Sep, 2018 ⌚ 7 months

- Implemented and maintained data ingestion, cleaning and processing pipelines
- Productionized prototype of machine learning API and deployed it to production
- Amazon Web Services (AWS)
- Python
- Apache Spark
- Flask
- Kubernetes

Research Engineer, Zalando

Jul, 2016 - Jan, 2018 ⌚ 1 year 6 months

- Performed A/B test analyses for diverse business and technical features
- Researched advanced statistical analysis methods for automating hypothesis tests
- Implemented and maintained various data ingestion, cleaning and processing pipelines for A/B test analysis automation
- Published and contributed to open-source statistical analysis library ExpAn
- Amazon Web Services (AWS)
- Python
- Scala
- Apache Spark
- Google BigQuery
- EXASOL
- Flask

Backend Engineer, Zalando

Sep, 2015 - Jun, 2016 ⌚ 9 months

- Implemented system, application and business level monitoring of the A/B testing platform
- Implemented various data ingestion and processing pipelines for A/B test analysis automation
- Amazon Web Services (AWS)
- Python
- Scala
- Flask
- EXASOL

Research Associate, Visage Technologies

Jul, 2014 - Jun, 2015 ⌚ 11 months

- Researched and prototyped various machine-learning based approaches to eye gaze detection
- Implemented eye gaze detection feature of the Visage SDK
- C++
- Python
- Computer Vision
- OpenCV

Software Engineering Intern, divIT

Jul, 2013 - Aug, 2013 ⌚ 1 month

- Implemented a module for the mash.me interactive motion-capture project (face and facial features detection)- Implemented a QR code reader with basic augmented reality features for a client event
- C#
- Computer Vision
- OpenCV
- Unity3D



PROJECTS EXPERIENCE

ExpAn

May, 2016 - Jan, 2018 ⌚ 1 year 8 months

A/B tests (a.k.a. Randomized Controlled Trials or Experiments) have been widely applied in different industries to optimize business processes and user experience. ExpAn (Experiment Analysis) is a Python library developed for the statistical analysis of such experiments and to standardise the data structures used.

- Python
- pandas
- Statistics
- A/B Testing



SKILLS

Data Architecture

Amazon Web Services (AWS) Microsoft Azure Streaming Architecture
Event-driven Architecture Serverless Knowledge Graph Data Platform OLTP OLAP

Data Engineering

Python Scala Apache Spark Apache Kafka Google BigQuery Cassandra EXASOL
MongoDB SQL Jupyter

Devops

Amazon Web Services (AWS) Microsoft Azure Kubernetes Docker Grafana
Terraform



EDUCATION

Information Processing and Multimedia Systems, Bachelors, Faculty of Electrical Engineering and Computing, University of Zagreb

Jan, 2008 - Jan, 2012

Information Processing, Masters, Faculty of Electrical Engineering and Computing, University of Zagreb

Jan, 2012 - Jan, 2014



INTERESTS

Cats

Running

Hiking