

Best Technology Talent For Companies Around The World



Martin

Senior Data Engineer

Summary

Senior Data Engineer with a track record of building high-trust, high-ownership teams and designing scalable data platforms. Currently leading the data engineering function. Balances pragmatism, cost, and long-term platform health while keeping engineering effort aligned with business impact.

Led department-wide Spark adoption, implemented a modern Apache Iceberg lakehouse on AWS that improved query performance by 60%, and designed end-to-end batch and streaming architectures making explicit trade-offs between latency, cost, and complexity.

Earlier career includes growing an engineering organisation from 2 to 20 developers at AEOS/Adscanner, building Spark/Scala ETL pipelines processing 20M+ records per day, and designing distributed PostgreSQL systems using Citus for parallel query execution.

TECHNICAL SKILLS:

- Python, Scala, SQL
- Apache Spark (Scala, PySpark), Apache Kafka, RabbitMQ, Apache Airflow, ETL/ELT pipelines, batch and streaming architectures
- Apache Iceberg, AWS S3, AWS Glue, Delta Lake patterns, ACID transactions, metadata management
- AWS (S3, EMR, Glue, Step Functions, Lambda, DynamoDB, Athena)
- PostgreSQL (Citus, PGQ, Partman, PgAgent), SQL query optimisation, execution plans
- Terraform, Redis caching
- Team coaching, ownership frameworks, decision logs, delivery predictability, cross-functional stakeholder alignment

Professional Career

Lead Software Engineer

January 2023 – Present | Zagreb, Croatia

Project 1 Info:

B2B sales intelligence company. Engineering Manager for the data department, partnering with Product, Data Science, and Analytics while leading a distributed team of data engineers.

Responsibilities:

- Led a Terraform monorepo standardisation for the data department, decoupling teams from ad-hoc DevOps and security dependencies and significantly reducing infrastructure provisioning and deployment lead time
- Coached and upskilled engineers by systematically increasing ownership, architectural responsibility, and decision-making scope, enabling multiple team members to operate at senior level and lead initiatives independently
- Partnered closely with Product, Data Science, and Analytics to translate product goals and SLAs into clear technical requirements, ensuring engineering effort was aligned with business impact and data reliability expectations
- Established clear communication and collaboration standards across teams — ownership, decision logs, escalation paths — reducing misalignment and execution friction in multi-team initiatives
- Improved delivery predictability by clarifying scope boundaries, defining success criteria upfront, and enforcing incremental delivery, reducing late-stage rework and last-minute escalations

Software and technologies:

Terraform, AWS, Spark, Python, Kafka, Airflow, Iceberg, cross-team leadership

Project 2 Info:

Senior Data Engineer progressing to technical lead on department-wide platform initiatives, including Spark standardisation, Iceberg lakehouse implementation, and AWS data architecture.

Responsibilities:

- Led department-wide adoption of Apache Spark (Scala and PySpark) as the standard data processing tool, enabling reliable incremental processing while reducing reprocessing and infrastructure waste
- Implemented a modern data lakehouse architecture using Apache Iceberg on AWS S3 and Glue, enabling ACID transactions, pipeline visibility, and improving query performance by 60% through efficient metadata management
- Breaking down requirements into executable designs, coaching engineers on distributed-systems trade-offs, and pair-implementing critical components

- Designed end-to-end batch and streaming architectures on AWS (S3, EMR, Glue, Step Functions, Lambda, DynamoDB, Athena), making explicit trade-offs between batch vs streaming, latency vs cost, and complexity vs operability early in project lifecycles
- Implemented event-driven pipelines using Kafka and RabbitMQ, and introduced Redis caching for high-read workloads, enabling low-latency use cases while avoiding unnecessary streaming complexity where batch was sufficient

Software and technologies:

Apache Spark (Scala, PySpark), Apache Iceberg, AWS (S3, EMR, Glue, Step Functions, Lambda, DynamoDB, Athena), Kafka, RabbitMQ, Redis, Airflow, Terraform, Python, SQL

Database Developer

July 2019 – December 2022 | Zagreb, Croatia

Project Info:

Ad intelligence platform. Lead Database Developer through a high-growth phase that scaled the engineering organisation from 2 to 20 developers.

Responsibilities:

- Built Spark-based ETL pipelines in Scala processing 20M+ records per day
- Orchestrated complex workflows using Apache Airflow
- Designed and operated distributed PostgreSQL systems using Citus for parallel query execution
- Leveraged PostgreSQL extensions (PGQ, Partman, PgAgent) to improve performance and maintainability
- Optimised complex SQL queries, functions, and procedures with a focus on execution plans and latency

Software and technologies:

Apache Spark, Scala, Apache Airflow, PostgreSQL, Citus, PGQ, Partman, PgAgent, SQL

Education

Zagreb University of Technology

Bachelor's Degree in Computer programming

Foreign Languages Skills

English: Advanced

Croatian: Native