



# Murat Cakir

Senior Data Engineer

## About me

Murat is a data engineer possessing broad industry expertise. Enthusiastic about crafting seamless data pipelines and distributed systems, he has overseen petabyte-scale data initiatives. As a collaborative professional with a "can-do" attitude, excels in communication and tackling challenges.

## Education

2010 - 2014

BSc Information Technologies

## Biography

Murat is a data-minded developer with a wide range of industry experience. He is passionate about building flawless data pipelines.

Murat has a deep interest in distributed systems. He helped organizations store, and process Petabyte scale data during his previous experiences using distributed systems.

He is a complete team player with getting things done approach. Effective communication and addressing problems are among his most important soft skills.

## Roles

### Senior Data/Platform Engineer

Developing new features and optimizing data platform.

### Senior Data Engineer

Built a DWH ELT solution from scratch using dbt, Snowflake, AWS.

## Certifications

AWS certified Solutions Architect - Associate

## Compentences

### Programming

Python, Scala, Java

### Cloud

AWS, Azure

### Tech

Spark, Docker, Git, Kubernetes, Bash

### Languages

English, Turkish

## Recent work experience

### Ahold-Delhaize – Jan 2023 – March 2023 - Senior Data/Platform Engineer

As a Data/Platform Engineer, I contributed to the enhancement and optimization of the data platform utilized by all entities within the Ahold Group. This platform was built upon Azure Databricks. My responsibilities included developing new features, boosting performance, and streamlining CI/CD deployments by transitioning to Terraform. Additionally, I implemented features for easy and robust development environments, while providing support to various teams and organizations.

### Didomi – March 2022 – December 2022 - Senior Data Engineer

Designed and implemented an ELT modern data stack data warehouse. Technologies used are Snowflake, DBT, Fivetran, S3, and Terraform. Developed data anonymization Spark application using PySpark for masking PII fields. AWS EMR cost optimization, 20% decrease in total EMR cost. Writing documentation such as how-to guides for specific technologies, RFCs for new projects, and post-mortems.

### Letgo – June 2021 – March 2022 - Senior Data Engineer

Developing data pipelines using Terraform in AWS which enables other teams (Business Intelligence, Data Science etc.) to use data with compliance rules applied. Tech stack: Scala, Kafka, Spark, EMR (Hadoop), Cassandra, S3, Redshift, ECS. Building ETLs using Airflow, Python, Scala, Spark, EMR. Manage all data infrastructure using Terraform. Architectural design support for the Data Science team. Cost optimization in AWS.

### Bahcesehir University – Feb 2020 – Feb 2021 - Lecturer

I gave lectures about Big Data and the Apache Hadoop Ecosystem for two semesters at the Master's Degree level. It includes the core concepts of Hadoop and MapReduce, data ingestion tools, Apache Hive, Scala, and Apache Spark. I had 79 students.



## Contact information

Phone: +31 6 25 55 41 48

Email: [murat.cakir@xebia.com](mailto:murat.cakir@xebia.com)