



CEYHUN AKAR

DATA ENGINEER

Data Engineer with experience building and managing robust, end-to-end data systems designed for scalability and real-time performance. Proficient in the full data lifecycle, from automated data ingestion and complex stream/batch processing to API development and customer-facing analytical applications. Specializes in cloud-native, microservice architectures leveraging Kafka, Spark, Flink, and Kubernetes to solve complex data challenges.

CONTACT

- 📞 +90 531 746 75 94
- 📍 Istanbul - Turkey
- ✉ akarceyh@gmail.com
- 🐙 github.com/akarce
- 🌐 akarce.github.io
- 📖 medium.com/@akarce
- 🔗 linkedin.com/in/akarce
- 📺 youtube.com/@ceysdatahub

KEY SKILLS

PROGRAMMING LANGUAGES

- Python
- Bash Scripting
- SQL

TOOLS & TECHNOLOGIES

- Docker
- FastAPI
- Apache Kafka
- Confluent
- Databricks
- AWS
- Grafana
- Kubernetes
- Rancher
- Apache Airflow
- Apache Spark
- Apache Flink
- NiFi
- Hive

DATABASES

- PostgreSQL
- MySQL
- Cassandra
- MongoDB
- Trino
- ClickHouse
- CouchDB
- Elasticsearch

BI & VISUALIZATION TOOLS

- Apache Superset
- Tableau
- Power BI
- Kibana
- IBM Cognos
- AWS Redshift

LANGUAGES

- Turkish: Native
- English: C1 - 7.0 IELTS Score

EDUCATION

Istanbul Technical University
Bachelor of Science, Materials Science

PROFESSIONAL EXPERIENCE

Mediastream AG | Data Engineer

Oct 2024 - Present

Engineered real-time data pipelines on Kubernetes using Kafka and Flink to process player transactions for a centralized anti-fraud and AML monitoring system.

Architected a large-scale web filtering platform using Spark and Airflow to process DNS data, classify URLs, and dynamically block illegal gambling sites.

Optimized batch ETL workflows into ClickHouse and Cassandra, enabling high-performance analytics and regulatory reporting for internal and external stakeholders.

PROJECTS

Data Pipeline for Real-Time Analysis of Yelp Reviews

Sep 2024

[Github Repo](#) / [Youtube Tutorial](#)

Engineered a real-time streaming pipeline using Kafka, MongoDB, and Elasticsearch for Yelp review sentiment analysis with DistilBERT, visualized in Kibana.

Reddit Data Engineering Pipeline

Aug 2024

[Github Repo](#) / [Medium Article](#)

Developed a full ETL pipeline to process Reddit data using NiFi, MinIO, Hive, and Trino, with visualizations in Tableau and Superset.

End to End Data Engineering OTP Project

July 2024

[Github Repo](#) / [Medium Article](#)

Built an end-to-end OTP generation pipeline orchestrated by Airflow, using Kafka for messaging and Cassandra/MongoDB for storage, with integrated alerting.

CERTIFICATIONS & COURSES

Elasticsearch v8 - BTK Academy

Sep 2024

- Elasticsearch, Kibana, Logstash

NoSQL - Document Databases - BTK Academy

Jul 2024

- MongoDB, Couchbase

Apache Kafka - BTK Academy

Jun 2024

- Kafka Streams, Kafka Connect, ksqlDB

Data Engineering Nanodegree - Udacity

Mar 2024

- AWS, AWS S3, PySpark, NoSQL, Data Lakes

Learn SQL Basics for Data Science - Coursera

Jan 2024

- Exploratory Data Analysis, Pyspark, A/B Testing

IBM Data Engineering Specialization - Coursera

July 2023

- Data Warehousing, Linux, NoSQL, Shell Scripting