

NTA TEAM

Tore Ziken

Торе Зикен · Руководитель направления данных

Head of Data

Astana

> *A data engineer with 8+ years of experience: warehouses and batch and streaming pipelines for retail, mining, finance and research, including two years in Ireland. At NTA he owns data: the places catalogue, search, analytics for businesses and reporting for the city.*

ROLE AT NTA

- Catalogue data pipelines: importing and refreshing places, photos and transport
- Data quality and catalogue search
- Pilot dashboards, reports for the city, analytics for businesses
- Anti-fraud: NTC rewards, reviews, bookings

EXPERIENCE → NTA TASKS

WHAT THEY DO	WHERE NTA NEEDS IT
Airflow / Airbyte, 10+ external APIs, web scraping	→ **Regular imports of places, photos, transport and stations**
Automated quality checks: 1,000+ anomalies a month	→ **A clean catalogue: duplicates, addresses, opening hours, coordinates**
PostgreSQL: migrations, query tuning	→ **Our PostgreSQL + pgvector database, catalogue search**
CDC on Debezium + Kafka, real-time orders	→ **Availability sync with OTAs, an event stream for analytics**
Data Vault, dbt, BI: Superset, Power BI, QuickSight	→ **Analytics in the business dashboard, reports for the Bureau and the akimat**
A graph model for anti-fraud (Amazon Neptune)	→ **Protection against reward farming, fake reviews and bookings**

EXPERIENCE

05.2026 — present

Data Engineer
Andersen Lab, Astana

- A graph layer on Amazon Neptune for fraud analysis
- A bronze/silver/gold lakehouse on Azure Data Lake
- A QuickSight BI layer with anomaly detection

04.2025 — 05.2026

Data Engineer
Mechta.kz, Astana

- A 10+ TB Data Vault 2.0 warehouse
- Airflow and Airbyte pipelines: −40% manual work
- Inventory CDC on Debezium and Kafka, real-time online orders

07.2023 — 04.2025

Data Engineer
Eurasian Resources Group, Astana

- PostgreSQL migrations with 100% accuracy
- Airflow: pipeline reliability +30%
- Monitoring and alerting: −25% failures

09.2021 — 07.2023

Data Engineer
Samruk Business Academy, Astana

- ETL in Python and SQL, Superset dashboards
- Ingestion from APIs, files and web scraping

11.2019 — 08.2021

Data Engineer
AIE Ireland, Dublin

- ETL on GCP and BigQuery: throughput +50%
- Spark over 10–100M records per run

08.2018 — 08.2019

Data Engineer
Kazakhstan Institute for Strategic Studies

- 10 interactive dashboards in Power BI and Superset

EDUCATION

2019 — 2022

M.A. Sustainable Development Management
Rhine-Waal University, Germany

2014 — 2018

B.A. International Relations
L. N. Gumilyov Eurasian National University, Astana

SKILLS

Python · SQL · Airflow · Airbyte · dbt · Spark · Databricks · Kafka · Debezium · PostgreSQL · BigQuery · AWS · Azure · GCP · Superset · Power BI

LANGUAGES

Kazakh · Russian · English · German — intermediate