



# Vladyslav Sumatokhin

## Data Engineer

### About

A results-driven Data Engineer with expertise in predictive modeling, ETL processes, and database management. Committed to continuous learning and adapting to the latest industry advancements. Seeking roles that maximize my skills and offer avenues for professional growth.

### CONTACT

+7 993 890 91 02 @dotsum  
me@dotsum.tech @dotsum  
https://dotsum.github.io/

### EDUCATION

Moscow, Russia 2020 - 2022

Master's Degree (graduated with honors)

Bauman Moscow State Technical University

Field of Study:  
Organization and Management of  
High-Tech Industries

Moscow, Russia 2016 - 2020

Bachelor's Degree (graduated with honors)

Moscow Polytechnic University

Field of Study:  
Mechanical Engineering

### LANGUAGES

Russian fluent Ukrainian fluent English C1 German A1

### PROFESSIONAL EXPERIENCE

#### ML/DATA ENGINEER

Moscow, Russia  
06.2022 - 05.2023



##### RTK-Element

The company is absorbed by Rostelecom, Russia's largest provider of digital services and services

- Designed and optimized loaders for raw data from XML files, achieving a more than **2x increase** in loading speed.
  - Developed ETL processes for ML models, creating processes for **over 20 data marts** from scratch.
  - Engineered predictive models to assess the risk of death and disease progression using EHR data, contributing to more effective and proactive patient care.
  - Enhanced model accuracy by reducing Type II error rate to **below 20%**, while ensuring interpretability for healthcare professionals.
  - Collaborated in the migration from PostgreSQL to GreenPlum, optimized procedures and data structures for GreenPlum.
- Administered ML infrastructure: JupyterHub, MLFlow, S3, GreenPlum for a team of four.

#### DATA ANALYST

Moscow, Russia  
05.2021 - 05.2022



##### Element Lab

Government contractor for the digitalisation of healthcare

- Designed and implemented ETL processes for the analytical data warehouse using PostgreSQL and Apache Airflow, successfully developing **over 20 processes** from scratch.
- Administered the Tableau server for **over 20 users**.
- Developed and maintained **over 10 executive dashboards** in Tableau. Actively engaged with end-users to understand their requirements, enhancing the user experience, resulting in the majority of dashboards being accessed daily.
- Prepared scripts for both regular and on-demand reporting. Improved regular reporting refresh frequency from monthly to daily.
- Collaborated in the migration from Oracle DB to PostgreSQL; refactored PL/SQL procedures and ETL processes, successfully migrating **over 10 processes**.

### TECHNICAL SKILLS

Programming Languages	Python, JavaScript/TypeScript, Bash
Infrastructure	Linux, Docker, k8s, JupyterHub
Data Bases	PostgreSQL, GreenPlum, MongoDB, S3, Redis
Data Processing	Pandas/Polars, Dask, Spark, Airflow
Machine Learning	NumPy, SciPy, Sk-learn, CatBoost, PyTorch, TensorFlow, MLFlow, DVC
Visualization	Tableau, matplotlib, seaborn, plotly, bokeh, D3.js
Other Skills	HTML, CSS, React, Redux, GraphQL, FastAPI, git, Jira, prompt engineering

