

Nikita Yurasov

Senior Software Engineer

[LinkedIn](#) • [GitHub](#) • Berlin, Germany

Senior Software Engineer with 6+ years of experience across **data platform engineering, MLOps, and scalable backend systems**. Built a full data intelligence platform from scratch at a VC firm; architected ML inference services and ETL frameworks handling 9k daily operations and 1.6 PB of data. Experienced in modern data stacks (dlt, SQLMesh, BigQuery), infrastructure-as-code (Terraform, GKE), and end-to-end system delivery.

- Designed and shipped production data platforms spanning ingestion, transformation, orchestration, and serving layers.
- Engineered distributed data systems at petabyte scale with fault-tolerant migration and multi-cloud architectures.
- Open-source contributor ([dbxio](#), [dmp-af](#)).

Experience

Earlybird VC, early-stage venture capital

Berlin, Germany

Senior Software Engineer

Jun 2025 – present

Built EagleEye — a proprietary data intelligence platform from the ground up aggregating 7+ external data sources into a unified analytics layer for the investment team.

- **Designed and built the entire data platform end-to-end:** 9 ingestion pipelines (Crunchbase, Dealroom, Harmonic, Affinity CRM, and others) with dlt, ~190 SQLMesh transformation models across a 4-layer medallion architecture (raw → bronze → silver → gold), and a serving layer on BigQuery.
- **Provisioned and configured a production GKE cluster** with Helm-based deployments, enabling containerized orchestration workloads with isolated failure domains, audit-gated production runs, and zero-downtime rollouts.
- Adopted **Terragrunt across 18 Terraform projects** on GCP (Cloud Run, GKE, BigQuery, Cloud Build, Secret Manager), eliminating configuration drift and enabling DRY infrastructure management across environments.
- Integrated **AI capabilities into existing platform services:** built an MCP server (Model Context Protocol) with OAuth 2.1 for secure AI agent access to the data warehouse, added webhook-driven event processing, and embedded LLM-powered enrichment into data pipelines.

Toloka.AI, empower companies with high quality data to develop AI

Berlin, Germany

Senior Software Engineer

Aug 2024 – May 2025

Architected **ML inference services**, reducing costs by 20% and supporting 8 customer-facing projects.

- Reduced LLM inference costs by **20%** by architecting a **fault-tolerant, distributed delayed inference service** and providing a unified Python client interface for all registered models across diverse environments.
- Enabled **8 customer-facing projects** for cross-functional teams by integrating 8 LLM producers into a single ML inference service, improving security and system compatibility.
- Saved **3 hours per data type creation** for developers by engineering a **code generation tool** for **protobuf, SQL, and Python**, ensuring strong typing integration and CI/CD registration across the company's data platform.

Toloka.AI

Software Engineer

Belgrade, Serbia
Jan 2023 – Aug 2024

Developed scalable **ETL** and migration tools, handling 9k operations daily and 1.6 PB of data.

- Engineered the [dbt-af](#) package to orchestrate thousands of dbt models via atomic operations integrated with Airflow DAGs. Enabled handling of **8,900 operations daily** across **4,700 tables** and **130 DAGs**, resulting in a robust and scalable ETL framework.
- Designed a **distributed data migration tool** for fault-tolerant data transfer across multiple clouds, handling **1.6 PB** of data during migrations.
- Accelerated delivery for anti-fraud, analyst, and R&D teams by improving **base infrastructure** and implementing CI/CD processes, unblocking teams from managing infrastructure.

Yandex, Russia's #1 employer

ML Developer

Moscow, Russia
Aug 2021 – Jan 2023

- Designed a **new approach to detect fraud** on the platform, achieving a **46% increase** in accurate fraud identification and preventing **78%** of potential financial thefts for this fraud type.
- Improved ETL pipeline performance by resolving data preparation inefficiencies, enhancing **data versioning** and consistency, and optimizing resource usage through clustering algorithms.

Projects

dmp-cloud · [platform.dmp-af.cloud](#)

Multi-tenant SaaS platform for dbt teams — domain observability with lineage visualization, branch-based file editing, GitHub App integration, and an MCP server with 14 AI tools. Full-stack: FastAPI, Next.js, PostgreSQL, Dagger CI/CD.

Education

BSc in Computer Security
HSE University

Moscow, Russia
2017 – 2023

Skills

Python, TypeScript, SQL, Bash, HCL

dlt, SQLMesh, dbt, BigQuery, Airflow, Spark, PostgreSQL, DuckDB

GCP, Terraform, Terragrunt, Docker, Kubernetes (GKE), Helm, Dagger

FastAPI, Next.js, Hono, MCP (Model Context Protocol)

Languages

English, working proficiency

German, B1 (intermediate)

Fun fact: I always remember the number e up to nine decimal places because it contains Lev Tolstoy's birth year twice.