



Irina Paroshina

PRINCIPAL DATA ENGINEER | DATA PLATFORM ARCHITECT | CONTROLLED SCALABILITY & TCO

Proven record of building enterprise-grade data platforms from the ground up.

Able to own the entire lifecycle of modern data platforms — from system design and integrations to pipelines, data models, observability, governance, analytics enablement and cross-functional delivery.

Combines engineering depth, architectural thinking and experience working across billing, CRM, financial and product domains.

CORE TECHNICAL COMPETENCIES

Data Platform Architecture · Lakehouse · ETL/ELT · Data Modeling · Observability · Systems Integration · Data Governance

SELECTED ENGINEERING ACHIEVEMENTS

- Built an **enterprise open-source DWH from scratch**, replacing a vendor platform and reducing vendor dependency and platform TCO while enabling analytics on a new **telecom** billing system.
- Defined **enterprise DWH semantic architecture and engineering standards**, eliminating accumulated technical debt and enabling sustainable platform evolution.
- Architected a **metadata-driven framework** for automated ETL generation, standardized DWH development and configurable data quality controls.
- Engineered **400+ production pipelines** integrating billing, mediation, CRM and financial systems, with automated recovery, monitoring and SLA controls.
- Designed and delivered the **first production Iceberg lakehouse for a major Uzbek P2P e-commerce** and classifieds platform, built on Trino, Iceberg and MinIO.
- Established end-to-end **observability, lineage and governance foundations** supporting reliable operations, BI, self-service and **AI-assisted analytics**.

TECH STACK

NiFi, Kafka, Trino, Iceberg, MinIO/S3, Hadoop/Hive, Python, C#, PostgreSQL, MS SQL, Oracle, Kubernetes, Grafana, InfluxDB, PowerBI, Superset etc.

Experience

SENIOR DATA ENGINEER — BIRBIR (UZBEKISTAN)

2024 – PRESENT — NIFI, ICEBERG, TRINO, MINIO (KUBERNETES)

- Led the evolution of the analytical platform by defining the enterprise DWH semantic model, eliminating technical debt, and introducing engineering standards for scalable data development.
- Architected a metadata-driven framework for automated ETL generation, data quality validation, and operational monitoring, reducing manual development effort and ensuring consistent implementation.
- Established governance practices for analytical datasets, enabling sustainable platform growth and preparing the foundation for AI-assisted analytics and metadata-driven development.
- Delivered the company's first Iceberg-based lakehouse into production.

HEAD OF DATA PROCESSING & ANALYTICS — CELLFIE MOBILE (VEON GEORGIA)

2018 – 2025 — NIFI, HADOOP/HIVE, KAFKA

Key Achievements:

- Replaced the legacy vendor-built DWH/ETL system with a fully open-source, in-house data platform based on Hadoop, Hive, NiFi and Kafka.
- Designed the entire DWH architecture from the ground up to support the company's transition to the new Ericsson billing system, which initially had no reporting or analytics capabilities.
- Built all integrations with Ericsson billing, mediation systems, product catalog, CRM and financial systems.
- Delivered a scalable, cost-effective platform architecture that became the foundation for analytics, BI, operational monitoring and product development.
- Created the data ingestion framework (hundreds of NiFi flows) with telecom-grade reliability, SLAs and observability.

INDEPENDENT DATA PLATFORM CONSULTANT — VARIOUS COMPANIES. DATA PLATFORM AUDITS, ARCHITECTURE REVIEWS AND OPTIMIZATION.

2022 – PRESENT

DATA SCIENTIST / DATA ENGINEER — VEON GEORGIA

2016–2017

Worked as hybrid Data Scientist + Data Engineer + DevOps. Developed ETL pipelines, predictive models, monitoring, automation tools.

SOFTWARE DEVELOPER / DATABASE ENGINEER — VARIOUS ORGANISATIONS

2011–2016 C#, .NET, PYTHON, MS SQL, POSTGRESQL, ORACLE.

Built: payment systems, CRM platform, incident-management software, warehouse systems.

EARLIER EXPERIENCE (SHORTENED)

2008–2014 — Tech Journalist (Computerra, CNews, IBM...)

2002–2009 — Junior Developer (C++, Fortran), research staff.

EDUCATION

Master's Degree (Computer Modelling of Physical Processes in Applied Electrodynamics)

Bachelor of Science (Physics and Mathematics)

Additional Certifications: online certifications from Moscow Institute of Physics and Technology & Yandex (Математика и Python для анализа данных), Johns Hopkins University (“Executive Data Science”, “Managing Data Analysis”, “Building a Data Science Team”) etc.