Yaroslav Ozerov

oz.yaroslav@outlook.com | linkedin.com/in/prendsgarde | github.com/RobbyTheFish

Education

National University of Science and Technology "MISIS" Moscow, Russia
Bachelor's Degree, Computational and Applied Mathematics *Expected 2027*

Experience

Technology Support Engineer March 2022 – September 2023
Servicepipe *Moscow, Russia*

- Monitored and investigated DDoS incidents, configured protection rules, and coordinated with clients to resolve security events under SLA constraints
- Automated operational tasks with Python and Bash scripts, reducing manual effort for recurring incident workflows
- Queried and analyzed traffic logs using SQL and the Loki/Grafana stack to support incident triage and post-mortem analysis

Analytics Engineer September 2023 – September 2024
Servicepipe *Moscow, Russia*

- Built analytical dashboards for network traffic analysis in Grafana and Apache Superset, enabling data-driven decisions for security and product teams
- Designed and maintained dbt models and Airflow DAGs for data aggregation, transforming raw ClickHouse events into analytics-ready datasets
- Developed internal tooling and HITL aggregation services for the analyst team, improving data accessibility and reducing ad-hoc query load

Data Engineer September 2024 – Present
Servicepipe *Moscow, Russia*

- Architected and built a data platform from scratch handling **130 TB/month** of web traffic events, supporting analytics, BI, and machine learning workloads
- Reduced end-to-end data latency **from 5 minutes to 15 seconds** by redesigning the ingestion pipeline with Kafka, Apache Spark, and Iceberg on S3
- Implemented unified, access-controlled data layer with role-based permissions, enabling secure self-serve analytics across teams

Projects

Traffic Retrospective Analysis Service | *ClickHouse, dbt, Airflow, Python* Servicepipe (NDA)

- Designed a retrospective analysis system for high-volume web traffic, enabling security analysts to query and replay historical attack patterns
- Built aggregation pipelines transforming raw event streams into structured, queryable datasets optimized for analytical queries in ClickHouse

Data Platform | *Kafka, Spark, Iceberg, ClickHouse, S3, Airflow, dbt* Servicepipe (NDA)

- Built a Lakehouse-based data platform from scratch with unified ingestion, transformation, and serving layers for 130 TB/month of event data
- Established data governance with access control, data contracts, and monitoring, providing a single source of truth for BI and ML teams

Vault-MoreTech | *Python, FastAPI, Docker* | github.com/RobbyTheFish/vault-moretech Hackathon Project

- Developed a HashiCorp Vault alternative for secrets management during a hackathon, implementing token-based authentication and encrypted key-value storage
- Containerized the service with Docker and exposed a REST API via FastAPI for integration with CI/CD workflows

Technical Skills

Languages: Python, SQL, Bash
Data & Streaming: Apache Spark (PySpark), Apache Kafka, Apache Iceberg, dbt Core, DuckDB
Databases: ClickHouse, PostgreSQL
Orchestration & BI: Apache Airflow, Dagster, Apache Superset, Grafana
Infrastructure: Docker, S3/MinIO, Yandex Cloud, Google Cloud, Loki
Backend: FastAPI, Flask, Celery, RabbitMQ