Ilia Popov

Bioinformatics | Data Analyst

+7 999 695-64-52 | iljapopov17@gmail.com | LinkedIn | GitHub | Personal Website

Data analyst and bioinformatician with 3+ years of experience applying Python and R to complex biological and clinical datasets. Developed custom tools and pipelines that streamline biological data processing and visualization, accelerating research and reducing manual workload. Contributed to several international projects. Passionate about transforming complex data into clear, interpretable, and scalable solutions that drive progress.

SKILLS

- **Programming & Data Analysis:** Python (pandas, seaborn, matplotlib, scikit-learn), R (tidyverse, ggplot2), SQL (PostgreSQL), GIS (NaturalEarth, tmap), Snakemake, Bash/Shell, Git
- **Computational & Analytical Skills:** Reproducible research workflows (pytest, CI/CD), advanced statistical analysis (regression, multivariable modeling, hypothesis testing, A/B testing), and applied machine learning (clustering, predictive modeling, feature selection, model evaluation)
- **Bioinformatics:** Metagenomics (16S, shotgun), Genome assembly & annotation, Phylogenetics
- **Languages:** Russian (Native), English (IELTS C1)

EXPERIENCE

Don State Technical University (Sep 2023 – Present) *Rostov-on-Don, Russia*

Research Fellow

- Developed KEGGaNOG, a Python tool for translating eggNOG annotations into KEGG pathway completeness profiles; achieved 21 stars on GitHub and 30,000+ installations | GitHub
- Created KrakenParser, a Python utility to batch-parse Kraken2 reports, reducing multi-sample aggregation time from hours to minutes, accelerating metagenomic insights for 100+ samples per week | GitHub
- Performed multivariable statistical analyses linking bat microbiome composition with viral/pathogen presence using R (tidyverse, MaAsLin2, ggplot2), generating publication-ready visualizations for predictive epidemiology.
- Applied GIS tools (NaturalEarth, tmap) to map pathogen and microbiome sampling sites, delineating geographic ranges of emerging infections for spatial risk modeling.
- Led data analysis, visualization, and manuscript preparation for 8 peer-reviewed publications, integrating bioinformatics with predictive modeling to transform raw metagenomic data into actionable insights.
- Delivered a masterclass “Introduction to Phylogenetic Analysis with Python and R” for 30 students at Mongolian National University of Education | Class materials

PJSC «Tatneft-Biotech» (Dec 2024 – Mar 2025) *Remote, Kazan, Russia*

Bioinformatics Data Analyst

- Built reproducible metagenomic pipelines (Snakemake, Kraken2, QIIME2) that increased weekly sample throughput x10, accelerating identification of corrosion-driving microbes.
- Analyzed taxonomic and functional profiles, producing clear visualizations for 100+ industrial samples per week.

- Assembled and annotated metagenomes to reveal metabolic pathways linked to material degradation, supporting predictive microbiome models.

Bina Nusantara University (Feb 2023 – July 2023) *Jakarta, Indonesia*
Research Intern

- Contributed to clinical trial with 100+ children across 4 cohorts, generating high-quality microbiome and clinical datasets for multivariable analysis.
- Applied R (tidyverse, MaAsLin2, ggplot2) to link microbiome profiles with growth outcomes, producing publication-ready visualizations.
- Annotated *Lpb. plantarum* IS-10506 genome and identified CRISPR-Cas elements, enabling functional insights into probiotic effects.

HACKATHONS

Public Health Hackathon (8-10 Aug 2025) *Almaty, Kazakhstan*
«Gurren Lagman» Team

- Project: «Exploring Patterns in Well-being Perceptions and Healthcare Evaluation in Kazakhstan»
- Led GIS analysis and built an interactive dashboard, deploying results via a live website, enabling real-time exploration of spatial and survey data | Results Website

Bioinformatics bootcamp (27-30 May 2025) *Saint Petersburg, Russia*

- Mentored «BioBytes» and «NGS Go!» teams on metagenomic data analysis, guiding project design, pipeline development, and data visualization

MOST PROUD OF

- **Scholarship** of the President of the Russian Federation for Young Researchers
- **Project Lead** of the Russian Science Foundation project № 25-24-00351 «Epizootological monitoring of white nose syndrome and other fungal infections of bats in the Rostov region»

EDUCATION

Bioinformatics Institute *Saint Petersburg, Russia*
Retraining program, Bioinformatics for Biologists (Sep 2023 – May 2024) GPA: 3.0/4.0

- Project: «The shadow of HIV: searching for indirect signs of HIV infection in cell-free DNA samples»
- Large-scale clinical analysis (n=793 samples: 39 HIV+ vs 754 HIV-), handling extreme class imbalance (~1:20) to extract reproducible pathogen/microbial signals | GitHub

Rostov State Medical University *Rostov-on-Don, Russia*
Medical Doctor Degree (Sep 2017 – June 2023) GPA: 4.85/5.0