



Andrei Illarionov

MSc Student in Biology | Computational Biology & Data Science

Email rival.andrey@yandex.ru

Portfolio [Data Science Portfolio \(Yandex Practicum\)](#)

GPA 5.0/5.0

Language: English B2 (DET 110/160, 2026)

Research interests: molecular dynamics, bioinformatics, genomic machine learning, biomedical data, and explainable AI

EDUCATION

Institution	Degree / Course	Period
Kazan Federal University	MSc student in Biology (Genetics and Genetic Technologies)	2025–2027
University of Tyumen	BSc in Psychology (General and Social Psychology)	2021–2025
Tyumen State Medical University	Incomplete higher education, General Medicine	2019–2021
Yandex Practicum	Professional retraining: Data Science (Advanced)	2025–2027
School 21	Professional training: Backend/Data Science Development	2026–2028
Kazan Federal University	Fundamentals of Bioinformatics and Computational Biology	2025–2026
Novosibirsk State University	Applied Bioinformatics and Molecular Biology	2025–2026

SELECTED RESEARCH PROJECTS

MANIA/WANIA: Python Package for Molecular Dynamics Analysis of NaPi2b/SLC34A2

2026–present | Team research project, Kazan Federal University | Planned MSc thesis

With a research team under the supervision of R. A. Vlasenkova, I am developing MANIA/WANIA as the computational core of my planned MSc thesis. The package provides a reproducible workflow for analysing GROMACS molecular dynamics trajectories of NaPi2b/SLC34A2, including MDAnalysis-based processing, extraction of structural and dynamic descriptors, construction of contact and residue-interaction networks, and feature generation for comparing NORM and TUMOR states. My contributions include package architecture, CLI and configuration design, testing, documentation, and analysis-pipeline logic.

ThermoMPNN + SWARM: Predicting Protein Stability Changes

2026 | Independent computational biology project | [GitHub](#)

Designed and conducted a MegaScale $\Delta\Delta G$ prediction pilot comparing four trainable heads on frozen ThermoMPNN features. Completed 40 training runs with protein-level splits, test evaluation, and paired bootstrap analysis. No convincing advantage of SWARM's repeated information exchange over the static control was found.

Explainable AI for Biomedical Signal Classification

2026 | Independent machine-learning project

Developed a machine-learning pipeline for classifying biomedical signals, including time-series preprocessing, baseline-model training, performance evaluation, and model interpretation using XAI methods. The project focuses on explainable analysis of physiological data (EEG).

Adaptation of DNABERT-2 for Short Genomic Sequences

2026 | Research proposal for AIRI Summer School

Reviewed DNABERT-2 and conducted preliminary fine-tuning experiments on the GUE Core Promoter Detection task. Assessed the effect of model_max_length on classification of short 70-bp promoter sequences; best preliminary MCC: 0.6887, compared with a baseline of 0.6108.

Effects of Rehabilitation after Spinal Cord Injury in Mice

2025 | KFU, Department of Genetics and Genetic Technologies

Participated in the experiment and data analysis under the supervision of T. Ageeva.

ATXN1 RNA Interference and Off-Target Effects in Glioblastoma Cells

2025–2027 | KFU, Department of Genetics and Genetic Technologies

Contributed to laboratory work, experimental procedures, and data analysis; supervisor: A. N. Khabibrakhmanov.

Combined Gene Therapy for an Alzheimer's Disease Model

2025–2026 | KFU, Department of Genetics and Genetic Technologies

Contributed to a grant-funded project developing a combined approach to address mitochondrial dysfunction and endoplasmic-reticulum stress; supervisor: A. N. Khabibrakhmanov.

Relationship between Dream Representations and Meaning-in-Life Orientations

2024–2025 | University of Tyumen, Department of General and Social Psychology

BSc thesis investigating psychological associations; supervisor: I. A. Rusyaeva.

Time Estimation Accuracy and Features of Time Perception in Adolescents

2023–2024 | University of Tyumen, Department of General and Social Psychology

Research project involving data collection and analysis on time estimation and perception; supervisor: V. V. Grishin.

ACADEMIC ACHIEVEMENTS

- “I am a Professional” Olympiad (2025–2026), Biotechnology: Prize-winner (58 points);
- “I am a Professional” Olympiad (2025–2026), Psychology: Winner (73 points);
- ITMO Mega Olympiad (2025–2026), Biotechnology: Winner (81 points);
- Higher League Olympiad (2025–2026), Mathematics: Winner (100 points);
- International Subject Olympiad “Vysshaya Proba Znaniy” (2025), English, Artificial Intelligence, and Psychology: First-degree winner;
- “I am a Professional” Olympiad (2024–2025), Psychology: Winner (90 points);
- Federal Universities Olympiad (2024–2025), Biology: Winner (94 points);
- Federal Universities Olympiad (2024–2025), Psychology: Prize-winner (83 points);
- “Izumrud” Student Olympiad (Ural Federal University, 2024–2025), Humanities: Prize-winner (73 points);
- Lomonosov Universiade (2024–2025), Philosophy: Prize-winner (75 points).

CONFERENCE AND SUMMER SCHOOL PARTICIPATION

Conference / School	Status	Date
XXXII Symposium on Bioinformatics and Computer-Aided Drug Discovery (BCADD-2026) Online	Poster presentation accepted	19-21 Oct 2026
XI International Conference “Mathematical Biology and Bioinformatics” Pushchino, Russia	Presentation accepted for the program	12-15 Oct 2026
Conference abstract, OpenBio	Illarionov, A. A., Akberova, N. I., & Vlasenkova, R. A. MANIA: A reproducible pipeline for molecular dynamics trajectory analysis based on residue-interaction networks.	2026–2027
“Chinese + Intelligent Bioengineering” DUT–BGI International Summer School DUT and BGI College · Dalian & Shenzhen, China	Participant	12 Jul–2 Aug 2026
Law Afterknown: Law Beyond the Edge, University of Tyumen	Oral presentation	2026–2027
III OIC Youth Scientific Congress, Kazan, KFU	Oral presentation	2025–2026
Life Systems Engineering: Achievements and Prospects, Moscow State University	In-person attendance	2025–2026

SKILLS

Programming & data: Python (NumPy, pandas, SciPy, scikit-learn, matplotlib, seaborn), SQL/PostgreSQL, Jupyter; Bash/Linux; Git/GitHub; C and R (basic).

Machine learning: Classification and regression; linear models, SVMs, decision trees and random forests; CatBoost, LightGBM, XGBoost; feature engineering; cross-validation; data-leakage control; model calibration and interpretation (SHAP); Optuna; foundational NLP/LLM workflows.

Statistics & research: Descriptive statistics; correlation and regression analysis; hypothesis testing; experimental design; psychometrics and questionnaire-data analysis.

Bioinformatics: GROMACS and MDAnalysis; molecular-dynamics analysis; sequence analysis and biological-data preprocessing; BLAST, FastQC, Trimmomatic, SPAdes, QUAST, Barnap, RAST, and Mauve.

Laboratory methods: DNA/RNA/protein extraction; PCR, electrophoresis, Western blot; cell culture and transfection; animal models.