

Sergei Volkov

Data analyst/Python developer with an academic background in limnology and hydrophysics, passionate about automation. A long-term fan and user of Linux, Vim, and open-source software. I write efficient code to collect data from APIs, load it into databases, analyze, visualize, and create automated reports with Tableau or other BI tools. I prepare and analyze AB tests, check hypotheses, and describe the research to stakeholders. I value reproducibility (in computing and research), keeping information version-controlled, and writing concurrent, easy-to-read code.

Work experience

- Aug 2023 – **Data Scientist, Crocus Labs GmbH**
Feb 2024 Software development, Software architecture, Product design, Project management, Testing, Team leading
I was hired to build an ETL for collecting data from smart lamps for greenhouses. However, upon my arrival, a complete redesign of the AWS architecture was necessary. I set up CI/CD for the frontend, wrote build instructions for many components, and learned how to set up and deploy various AWS services (API Gateway, DynamoDB, IoT Core, Amplify, etc.). In parallel, I developed a simple Python app to calculate light density from the lamp as a sum of LED chips based on lamp geometry and LED chips' ray files. During the design of the AWS architecture, I also participated in database design. Just before the implementation of the new AWS structure, the company decided to move to a self-hosted solution due to weak network coverage on farms, thus canceling the AWS implementation. I created a prototype of a Linux host with containers communicating via IoT to mimic lamps, but the company decided to cancel the development of smart lamps and focus on creating their own LED dies. Since I was still on probation, my employment was terminated.
- Jun 2021 – **PhD Student, Leibniz-Institut für Gewässerökologie und Binnenfischerei (IGB Berlin)**
Mar 2023 Research design, Data processing, Fieldwork, Linux systems administration, DevOps
I was hired to contribute to joint research on biomass sinking in turbulent stratified environments. I deployed probes, collected and processed data, and participated in biological field and lab work. The main activity was managing and using a stereoscopic video recorder and the MATLAB package for video processing. The package was initially non-functional with obscure code and dependencies on an outdated OpenCV interface. I used Docker to build an environment with MATLAB, compiled all components (including OpenCV with custom build flags), and translated the GUI with an X server on the host machine. Falling behind schedule for a successful thesis defense, I decided to quit.
- Oct 2020 – **Data QA, Playrix**
Feb 2023 Data analysis, Validation, ETL design, Manual and automated QA
I was responsible for the data quality of new tables created by the data engineering team and data migration to the new ETL and DBMS. I verified Tableau dashboards based on game/product data and some based on marketing data. Writing complex SQL queries processing millions of rows was my daily routine. I led a small team of QA engineers, distributed and prioritized tasks, and provided onboarding and mentoring.
- Mar 2020 – **QA Engineer, Playrix**
Oct 2020 Manual QA, Automation, Teaching, Testing documentation writing, MitM traffic sniffing, Statistical analysis
- Nov 2019 – **Data Analyst, Playrix**
Feb 2020 Data analysis, Automation, User tracking events design, Reporting and presenting to stakeholders
SQL, Python, Analytics, DataViz, Big Data

- May 2016 – **Teaching Assistant**, *Petrozavodsk State University*
 Jul 2020 Educational program design, Lecturing, Teaching
Thermodynamics, Heat transfer, and other courses
- Jan 2016 – **Research Assistant**, *Northern Water Problems Institute*
 now Data analysis, Fieldwork, Experiment design, Data processing frameworks development, Academic writing, editing, and typesetting

Technical Skills

- Development *Python*: NumPy, Pandas, SQLAlchemy, Matplotlib, concurrency, Plotly
SQL: Analytic/window functions, query optimization
Data: Spark, Snowflake, Tableau, Looker, Plotly Dash, Spreadsheets
NoSQL: DynamoDB, MongoDB
Linux: Bash, system management, SSH, Nix
Containers: Docker, systemd-nspawn, LXC, Kubernetes, chroot
General: Git, Vim, Emacs, Jupyter
CI/CD: GitHub Actions, TeamCity, GitLab
Cloud: AWS, GCP
Typesetting: Markdown, LaTeX, Pandoc
Other: Julia, MATLAB, minor knowledge of JS, C/C++
- Analytical AB test design and analysis
 Cohort analysis
 Time-series analysis
 Descriptive and inferential statistics
 Machine learning basics
 Data wrangling
- Research *Fieldwork*: Organizing, sampling, working with probes, logistics, hand and power tools
Writing: Academic writing, editing, publishing, proofreading
Data processing: Writing ad-hoc frameworks, automation, visualization

Communication skills

- Presenting
- Communicating with stakeholders
- Public speaking
- Teaching/Mentoring
- Feedback sharing
- Active listening
- Remote team building

Organisational / managerial skills

- Agile, Kanban, Scrum
- Managing small teams of developers
- Task management and distribution
- Project management

Projects

- 2015 – 2017 **Lake Onego: Life Under the Ice, NWPI**
 Research and fieldwork
Joint interdisciplinary project to study the under-ice life
- 2021 – 2023 **Lake Pycnoclines Trap Organic Particles Forming Hot Spots of Accelerated Carbon Cycling in the Water Column (PycnoTrap), IGB Berlin**
Joint interdisciplinary project researching sinking and degradation of biomass in the turbulent environment (lake water column)

Awards - Certifications - Licenses

- 2017 **Machine Learning, Stanford University on Coursera**
 Certificate available at <https://www.coursera.org/account/accomplishments/verify/U4VGJWXFGMS9>

Voluntary

- 2013 – 2015 **Garbage Collection, On multiple sites near Onego Lake, Petrozavodsk, Russia**
- 2011 – now **OpenStreetMap, Data surveys and contribution, Republic of Karelia, Russia**

Education

- 2010 – 2015 **Specialist Degree in Energy Supplies, Petrozavodsk State University**

Thesis

- 2015 **Noniterative Heat Exchanger Calculation**
 Supervisor: Professor Sergei Bogdanov
 The thesis proposes a direct analytical calculation procedure for heat exchanger design, avoiding the use of optimization algorithms

Personal

- Citizenship **Russian**
- Languages **Mother tongue: Russian**
English: C1
German: B1
- Driving license **European Driving License: B, C**

Publications

- 2021 **Full Reynolds Stress Tensor of Convective Turbulence Estimated with Paired Acoustic Doppler Current Profilers, S. Bogdanov, G. Kirillin, S. Volkov, G. Zdorovennova**
- 2019 **Fine scale structure of convective mixed layer in ice-covered lake, S. Volkov, S. Bogdanov, R. Zdorovennov, G. Zdorovennova, A. Terzhevik, N. Palshin, D. Bouffard, G. Kirillin**
- 2019 **Under-ice convection dynamics in a boreal lake, D. Bouffard, G. Zdorovennova, S. Bogdanov, T. Efremova, S. Lavanchy, N. Palshin, A. Terzhevik, L. Vinnå, S. Volkov, A. Wüest, R. Zdorovennov, H. Ulloa**

- 2019 **Structure and dynamics of convective mixing in Lake Onego under ice-covered conditions**, *S. Bogdanov, G. Zdorovenova, S. Volkov, R. Zdorovenov, N. Palshin, T. Efremova, A. Terzhevik, D. Bouffard*
- 2018 **Albedo of a Small Ice-Covered Boreal Lake: Daily, Meso-Scale and Interannual Variability on the Background of Regional Climate**, *G. Zdorovenova, N. Palshin, T. Efremova, R. Zdorovenov, G. Gavrilenko, S. Volkov, S. Bogdanov, A. Terzhevik*

In compliance with the art. 13 GDPR 679/16, I hereby authorize the recipient of this document to use and process my personal details for the purpose of recruiting and selecting staff.