Sergei Volkov

Data analyst/Python developer with academic background (limnology and hydrophysics), passion for automation, fan of Linux, Vim and open source. I am able to write efficient code to collect the data from APIs, load it into DB, analyze, visualize, create automated reports with Tableau or other BI, prepare and analyze AB test, check hypotheses and describe the research to stakeholders. I love reproducibility (in computing and research), keeping information version controlled, writing concurrent and easy-to-read code.

Work experience

Aug 2023 – Feb 2024 **Data scientist**, *Crocus Labs GmbH*

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 (for further analysis) but upon my arrival the complete redesign of the AWS architecture was necessary. I started by setting up CI/CD for the frontend and writing building instructions for many components as well and learning how to set up and deploy different AWS services (API gateway, DynamoDB, IoT core, Amplify etc). In parallel I developed the simple Python app to calculate light density from the lamp as a sum of LED chips based on lamp geometry and LED chips' rayfiles. During design of AWS acrhitecture I also participated in databases design. Just before the start of implementation of the new AWS structure the company desided to go to self-hosted solution (due to weak network coverage on farms) thus the implementation was canceled. I created a simple prototype of the Linux host and bunch of containers communicating via IoT to immitate lamps but the company suddenly desided to cancel development of the smart lamps and focus on creation of their own LED dies. And as soon as I was still on a probation period, my employment was terminated.

Jun 2021 – Mar 2023 **PhD Student**, *Leibniz-Institut für Gewässerökologie und Binnenfischerei (IGB Berlin)*

Research design, Data processing, Field work, Linux systems administration, DevOps

I was hired carrying a physical part of the joined research on biomass sinking in the turbulent stratified environment. I had to deploy the probes, collect and process their data and participate in biological field and lab work. The main activity was to manage and use a stereoscopic video recorder and the MATLAB package for video processing. Though when I started the package didn't work, the code was obscure (manyfolded classes without documentation) and relied on the third party interface to OpenCV lib which development was stopped. I had to use Docker to build an environment with MATLAB inside to compile all the components (including OpenCV with custom build flags) and translate the GUI with X server on the host machine. I was falling behind the schedule for a successful thesis defense and I decided to quit the job.

Oct 2020 – Feb 2023 **Data QA**, *Playrix*

Data analysis, Validation, ETL design, Manual and automated QA

I was responsible for data quality of all new tables created by data engineering team, all the data duting migration to the new ETL and new DBMS. Also I checked all the dashboards in Tableau that were based on the game/product data and some based on the marketing data. Writing complex SQL queries processing millions of rows was my daily routine. Also I leaded a small team of QA engineers, distributed and prioritized tasks, made on-boarding mentoring.

Mar 2020 – Oct 2020 **QA engineer**, *Playrix*

Manual QA, Automation, Teaching, Testing documentation writing, MitM traffic sniffing, Statistical analysis

Nov 2019 – Feb 2020 **Data analyst**, *Playrix*
Data analysis, Automation, User tracking events design, Reporting and presenting to stakeholders
SQL, Python, Analytics, DataViz, BigData

May 2016 – Jul 2020 **Teaching assistant**, *Petrozavodsk state university*
Educational program design, Lecturing, Teaching
Thermodynamics, heat transfer and other courses

Jan 2016 – now **Research assistant**, *Northern water problems institute*
Data analysis, Field work, 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
Field work: 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 team of developers
- Task management and distribution
- Project management

Projects

2015 – 2017 **Lake Onego: life under the ice**, *NWPI*
Research and field work
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*
Joined interdisciplinary project researching sinking and degradation of biomass in the turbulent environment (lake water column)

Awards - Certifications - Licenses

2017 **Machine Learning**, *Stanford University at 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 direct analytical calculation procedure for heat exchanger design, avoiding usage 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. Zdorovennova, S. Volkov, R. Zdorovennov, 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. Zdorovennova, N. Palshin, T. Efremova, R. Zdorovennov, 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.