Sofya Dymchenko, Ph.D.

Grenoble, France | sofya.dymchenko@gmail.com | svdcvt.github.io

Summary

Postdoctoral researcher specialising in AI for Science and HPC, with a focus on data-driven surrogate modelling for scientific computing. Passionate about developing AI methodologies from a data-centric and compute efficiency perspective and applying them to impactful real-world problems, particularly, in CFD engineering, earth sciences, and molecular dynamics. I believe in the importance of interdisciplinary collaboration, open and inclusive science, and reproducible research.

Research Experience

Doctoral Researcher, Grenoble, France — Mar 2022 – Apr 2026
Datamove team, INRIA and University of Grenoble-Alpes (UGA)

- Developed novel active learning methodology for training deep PDE surrogates: created method for PINNs outperforming existing approaches (offline); pioneered extension to online setting for data-driven surrogate where training occurs concurrently with running simulations in HPC environments.
- Extended open-source Melissa framework; supervised engineer and 2 interns contributing to the project.
- Publications: 2 first-author workshop papers (NeurIPS'23, SC'24), 1 under review.

Research Intern, Moscow, Russia — Jun 2021 – Dec 2021
Computational Intelligence Lab, Skolkovo Institute of Science and Technology (Skoltech)

- Continued research on optimal transport algorithms from master's thesis; explored extensions to computational geometry applications and computer vision tasks.

Research Intern, Moscow, Russia — Feb 2020 – Aug 2020
Media Algorithms Laboratory, Huawei R&D Center

- Developed and trained generative models with attention mechanisms for audio processing and personalized noise reduction; contributed to R&D pipeline development in a product team environment.

Teaching

Teaching Assistant & Seminar Tutor — Sep 2019 – Feb 2024

- Led Python programming course for 3 consecutive years (100+ students): designed curriculum, managed teaching assistants, created assignments, and delivered lectures; awarded "Best Course" (2020).
- Taught seminars in Data Science, Machine Learning, and Python for bachelor, master, and professional development students at HSE (2019-2021) and UGA (2022-2024).

Education

PhD in Computer Science, Grenoble, France — Mar 2022 – Apr 2026
Laboratory of Informatics of Grenoble, UGA

- Thesis: High Performance Online Deep Neural Network Training from Synthetic Data with Active Learning for Scientific Computing. [link].

MSc in Applied Mathematics and Computer Science, Moscow, Russia — Sep 2019 – Jun 2021
Statistical Learning Theory program, Skoltech jointly with HSE (with honors, GPA: 4.89/5)

- Led hackathon-winning university DS projects for industrial partners (Sberbank, Tinkoff bank, McKinsey).
- Thesis: Developed modified optimal transport algorithm reducing complexity from quadratic to log-linear, applied to computational mesh generation [link].

BSc in Mathematics, Moscow, Russia — Sep 2015 – Jun 2019
Faculty of Mathematics, HSE (GPA: 7.87/10); Minor in Data Science, Faculty of Computer Science, HSE

- Thesis: Collected and preprocessed 10-hour noisy YouTube dataset from scratch for rare language classification; trained multimodal CNN-RNN models achieving 77.6% accuracy [link].

Publications

- **S. Dymchenko**, A. Purandare, and B. Raffin. MelissaDL x Breed: Towards Data-Efficient On-line Supervised Training of Multi-parametric Surrogates with Active Learning. In *SC-W 2024 - Workshops of The International Conference on High Performance Computing, Network, Storage, and Analysis*, pages 1–9, Atlanta (Georgia), United States, Nov. 2024. IEEE [paper] [presentation video]
- **S. Dymchenko** and B. Raffin. Loss-driven sampling within hard-to-learn areas for simulation-based neural network training. In *MLPS 2023 - Machine Learning and the Physical Sciences Workshop at NeurIPS 2023 - 37th Conference on Neural Information Processing Systems*, pages 1–5, New Orleans, United States, Dec. 2023 [paper]
- P. Cesar, **S. Dymchenko**, A. Purandare, and B. Raffin. Learning where to simulate: Generative active sampling for online pde surrogate training. In *under review*, 2026

Additional Experience & Skills

Leadership & Service:

- Supervised 2 master students on research internships (Feb – Sep 2025).
- Reviewed for Workshop on Advancing Neural Network Training at NeurIPS23 and ICML24.
- Contributed to Grenoble AI for Physical Sciences 2024 workshop [link] organization (event support, outreach).
- Active in doctoral school student council and laboratory synergy initiatives at UGA.

Technical skills: Python (PyTorch, JAX), Git, Bash, SLURM, Linux/HPC environments, C++ (basic).

Languages: English (fluent), French (intermediate), Russian (native).

Personal interests: bouldering, table tennis, chess, arts and crafts, volunteering at animal shelter.