




# Sergia VOLODIN (Sergei)

sergia94 at protonmail dot com      +41 78 732 01 34

 sergia-ch.github.io / Wehntalerstrasse, 8057 Zürich, Switzerland 

Birth date: 3rd of October 1994 (27 years), Russian 

Sergia (me): artsy recently graduated software engineer and young scientist with interests in open-source democratic and consensual technology and experience in industry, academia and startups research and development

## EDUCATION

**EPFL** Swiss Federal Institute of Technology in Lausanne (EPFL)   
Lausanne, Switzerland

Sep 2017 – Apr 2021

- Master's degree in **Computer Science**, GPA: **5.67/6**
- Minor in Computational **Neurosciences**
- Research Assistant position (2017–2019)
- Thesis  "CauseOccam: Learning Interpretable Abstract Representations in Reinforcement Learning Environments via Model Sparsity"

 **Moscow Institute of Physics and Technology**   
Moscow, Russia

June 2017

Bachelor's degree in **Applied Mathematics**, GPA: **4.84/5**

## SKILLS

**Relevant courses:** **Machine Learning, Software Engineering**, Unsupervised and Reinforcement Learning, Convex Optimization, Distributed Algorithms, Algorithms, Random graph theory, Functional Programming, Set Theory, Random Processes, Functional Analysis, Biological modeling of neural networks, Complexity theory, Learning theory, Neuroscience: behavior and cognition, Neuroprosthetics, Theory and methods for Reinforcement Learning, Optimization for Machine Learning, Computer Vision

**Scientific programming:** **Keras, TensorFlow**, ray/tune/rllib, tf-agents, scikit-learn, PyTorch, Brian 2, MATLAB, Mathematica, R





**Programming languages:** Python, C/C++, Java, Scala, nasm, C#, AVR C++

**Frameworks:** Qt/QML, Django, Android Studio, OpenGL/GLSL, Unity 3D, Blender

**Environment:** Git,  $\text{\LaTeX}$ , Bash, Debian/Ubuntu Linux

**Scientific skills:** **experimental** sections of research papers, working on **theoretical** problems, scientific presentation, data analysis



**Software development:** **agile** software development (Scrum), debugging, design patterns, concurrent and distributed systems, TCP/IP networking, AVR microcontrollers, Arduino platform, team and project management in small startups

**Languages:**  English:  TOEFL iBT 113/120,  French: A1,  Russian: native

## RESEARCH EXPERIENCE

**EPFL, Laboratory for Computational Neuroscience**   
Lausanne, Switzerland


Master's Thesis student  
Oct 2020 – Apr 2021

- Designed  an algorithm with **Python 3, Pytorch and ray** based on the "Consciousness Prior" proposal that finds a simple causal model of an RL environment in the general case. The project is a continuation of my Google Research internship (see below)
- The algorithm works on benchmarks, see my thesis  "CauseOccam: Learning Interpretable Abstract Representations in Reinforcement Learning Environments via Model Sparsity"
- The work includes theoretical results on abstraction learning as well as a code base with tests and documentation
- The thesis defended on the 21st of April 2021 with Adam Gleave (Berkeley/DeepMind) as an external expert






**Center for Human-Compatible AI (CHAI), Berkeley** 

Summer Intern

Berkeley, CA, United States (remote due to COVID-19, from Zurich, Switzerland 

June 2020 – Sep 2020


- Designed  better defenses against adversarial policies in Multi-Agent Reinforcement Learning via alternating training of opponents using **Python 3, Tensorflow, ray, rllib**.
- Ran hyperparameter sweeps on multiple machines with ray and rllib
- Converted legacy code using stable baselines and Tensorflow 1.0 to rllib and Tensorflow 2.0
- Results published as a blog post  "Defending against Adversarial Policies in Reinforcement Learning with Alternating Training" on the Effective Altruism forum

**Google Research** 

Software Engineering Intern



Mountain View, CA, United States


Nov 2019 – Feb 2020

- Designed an algorithm to uncover a linear **Causal Model** of a **Reinforcement Learning** environment using interventions with **Python 3, Tensorflow, tf-agents**, and analyzed the effect of interventions on the quality of exploration
- Used TensorFlow and tf-agents to conduct the experiments with large hyperparameter sweeps
- Results  published as an ICLR CLDM workshop paper



EPFL, Distributed Computing Laboratory   
Lausanne, Switzerland


Research Assistant  
Sep 2018 – Oct 2019

- Investigated **fault tolerance** of a neural network using **Taylor approximation**
- Introduced the *continuous limit* to  bound the error, and compared to the Neural Tangent Kernel limit case
- Conducted  experiments to test the theory using **Keras** including the **implementation** of custom layers and regularizers


EPFL, Computer-Human Interaction in Learning and Instruction laboratory   
Lausanne, Switzerland

Research Assistant  
Sep 2017 – Aug 2018


- Created  a **library** QML-AR for seamless **augmented reality** using **OpenCV**, **Qt/C++** and **Qt/QML** with competitive performance on Android and small visual negative impact
- Designed an  activity for kids for learning math using AR, tested the application in a classroom setting, analyzed the obtained data

Skolkovo Institute of Science and Technology,  
Center for Energy Systems   
Moscow, Russia

Research Intern  
Sep 2016 – Jul 2017


- Characterized using **numerical optimization** and **theoretically** the structure of the set of boundary non-convexities of an image of a quadratic map in case the number of non-convexities is infinite
- Designed and implemented  the Convexity Analysis of Quadratic Maps **library** using **MATLAB** which gives approximate solutions to a number of problems involving quadratic maps

## PUBLICATIONS

Lê-Nguyen Hoang, Louis Faucon, Aidan Jungo, **Sergei Volodin**, Dalia Papuc, Orfeas Liossatos, Ben Crulis, Mariame Tighanimine, Isabela Constantin, Anastasiia Kucherenko, Alexandre Maurer, Felix Grimberg, Vlad Nitu, Chris Vossen, Sébastien Rouault, El-Mahdi El-Mhamdi.  Tournesol: A quest for a large, secure and trustworthy database of reliable human judgments, 2021. Code for the platform (backend, ML, frontend), experiments, part of data analysis, writing

 **ICLR** **Sergei Volodin**, Nevan Wichers, Jeremy Nixon.  Resolving Spurious Correlations in Causal Models of Environments via Interventions, 2020. Topic choice, experiments, theory, writing.  ICLR CLDM workshop 2020.

El-M. El-Mhamdi, R. Guerraoui, A. Kucharyav, **S. Volodin**.  The Probabilistic Fault Tolerance of Neural Networks in the Continuous Limit, 2019. Experiments, theory, writing.


A. Dymarsky, E. Gryazina, B. Polyak, **S. Volodin**.  Geometry of quadratic maps via convex relaxation, 2018. Exp-s, theory, writing.

A. Petrov, **S. Volodin**  Janibekov's effect and the laws of mechanics. Doklady Akademii Nauk, 2013. Graphics for the article, experiments, **first year** of my BSc

## WORK EXPERIENCE

 Fave For Fans 

Sep 2021 – May 2022


Platform dedicated to passionate fans, Zürich, Switzerland 



Software Engineer

- **Responsible** for the backend development with microservices on Cloudflare Workers with TypeScript, an ArangoDB-based database, automatic data schema validation, and CI/CD with integration tests. First full-time engineer at the company.
- **Proposing** and discussing ways to create more ethical and democratic social media sharing mechanisms
- **Research** into ways of obtaining data from third-party services with privacy guarantees
- **Conducted** analysis of the database to create better ranking results

 Tournesol 


May 2020 – May 2021


Startup designing better recommender systems, Lausanne, Switzerland 

- **Co-founded a startup** working on expert-driven recommender systems
- **Responsible** for  back-end engineering using **Django**, and Machine Learning engineering with **TensorFlow**, the API server
- **Responsible** as well for system administration (Debian), (partially) front-end development with **React.js**
- **Co-authored**  the paper with our results

 EscapeControl 

Jul 2015 – Feb 2016

Own b2b startup for escape rooms, Moscow, Russia 

- **Created a startup** selling software and hardware for  real-world escape room games which allows to speed up the construction and reduce maintenance costs
- **Responsible** for back-end software engineering (**C++/Python**), servers administration, sales and customer support
- **Managed** a team of two web developers until a successful launch of the web interface
- Sold more than forty solutions which are currently running in different countries across the globe and provided remote support

ITBrat 

Jul 2015 – Feb 2016

Algorithmic trading startup, Moscow, Russia 

- **Developed** algorithmic trading application from initial discussion with the team to deployment and supporting in **C++**
- Added low-level user-space networking to the project which allowed to decrease latency and increase profit
- **Responsible** for the performance of the code

## RESEARCH INTERESTS

Artificial Intelligence, Machine Learning, Artificial Intelligence Safety/Ethics, Causal Reasoning, Adversarial policies, Mathematical Optimization, Robotics, Consciousness research

## SCHOLARSHIPS

- 🔗 Research Scholars, a paid **Research Assistant** position, Swiss Federal Institute of Technology in Lausanne (EPFL), 2017 – 2019
- 🔗 Abramov Fund's scholarship for excellent **grades**, 2014

## PROJECTS

- Safe Proximal Policy Optimization** 2019  
*EPFL EE-618 course project, advised by Dr. Kamalaruban Parameswaran and Prof. Volkan Cevher* 🇨🇭
  - 🔗 Added a projection step to the **Proximal Policy Optimization** algorithm to comply with requirements of **Constrained Markov Decision Processes**
  - 🔗 Implemented code in Tensorflow and tested it in simple environments
  - Presented the project at the RLSS 2019 summer school (Lille, France)
- Quadcopter drone from scratch project** 🇨🇭 2012 – 2014
  - Developed 🔗 an algorithm in C++ for stabilization of a quadcopter drone from scratch using AVR microcontrollers, IMU sensors and PID regulators
  - Managed** the project consisting of 2-5 developers
  - Conducted the analysis of launches to improve 🔗 flying quality
  - Results were **published** as a 🔗 popular science article (*in Russian*)

## CONFERENCES AND SUMMER SCHOOLS

- 🇨🇭 🔗 Machine Learning Summer School, 2020 (virtual due to COVID-19), **poster presenter**
- 🇫🇷 🔗 Reinforcement Learning Summer School, 2019 (Lille, France), **poster presenter**, *selected to receive financial help*
- 🇫🇷 🔗 Data science summer school, 2019 (Paris, France), **poster presenter**
- 🇮🇹 🔗 QtDay 2019 (Firenze, Italy), **speaker**, *one hour session on qml-ar*
- 🇫🇷 🔗 P.A.I.S.S. AI Summer School, INRIA Grenoble, 2018, *participant in tutorials given by top experts*; 🔗 *selected to receive financial aid*
- 🇷🇺 🔗 Information Technologies and Systems (Saint-Petersburg, Repino, 2016), **speaker**, *poster presenter*

## COMPETITIONS

- 🇨🇭 🔗 Google HashCode Qualification round coding contest, **top 6%** (team EPFL\_Noobs), managed the team, developed algorithms and did the coding, 2019
- 🇷🇺 🔗 DeepHack.RL hackathon on Deep **Reinforcement** Learning for Atari games, managed the team and developed an 🔗 evolutionary algorithm with an autoencoder, MIPT, Moscow, Russia, 2017

## INTERESTS

**Effective Altruism**, Philosophy, Running (1/2 marathon 2018), Snowboarding, Swimming, Dancing Rock'n'Roll

## VOLUNTEERING

- Effective Altruism Lausanne** 2019  
*Local* 🔗 *EA community* Lausanne, Switzerland 🇨🇭
  - Co-founding the group, 🔗 introduction workshop speaker, running a 🔗 discussion group on AI safety and theory, newsletter management and writing, Facebook events announcements, managing open discussions
- Artificial Intelligence Governance Forum** 2019, 2020  
🔗 *AI governance conference* Geneva, Switzerland (2019), virtual due to COVID-19 (2020) 🇨🇭
  - Time-keeping, technical support, small tutorial on neural networks
- Applied Machine Learning Days** 2019  
*Machine learning* 🔗 *conference* Lausanne, Switzerland 🇨🇭
  - Technical help for presenters, badge check
- Anti-corruption foundation (A. Navalny)** 2017  
*A* 🔗 *non-profit aimed at investigating corruption* Moscow, Russia 🇷🇺
  - Conveyed the results of the investigations by talking to people on the streets as a volunteer