

# Anton Karazeev

---

Moscow, Russia  
+7-977-490-21-83 • anton.karazeev@phystech.edu  
<https://www.linkedin.com/in/akarazeev>  
<https://github.com/akarazeev/>

## POSITIONS

**Laboratory of Neural Networks and Deep Learning** October 2017 — Present  
*Researcher*

**Laboratory of Functional analysis of the Genome** June 2016 — Present  
*Research Assistant*

Research on protein function analysis.

Text mining, Natural language processing, Keyword extraction, Machine learning algorithms (the **Laboratory** is located in **BioPharmCluster** at MIPT, <http://www.generesearch.ru/Members.html>).

**Sberbank-Technology** August — October 2017  
*Data Scientist*

Responsible for Natural Language Processing projects.

**Laboratory of Neural Networks and Deep Learning** March — June 2017  
*Teaching Assistant*

"Deep Learning in Natural Language Processing" course. Seminar quizzes preparing and home assignments checking (the **Laboratory** is located in **BioPharmCluster** at MIPT).

**HiQE Group** March — June 2017  
*R&D Data Scientist*

Responsible for: negotiations with IBM engineers to determine which IBM tools including Watson services are useful to HiQE Group's projects, audio signal processing using machine learning methods.

## TEACHING

**Deep Reinforcement Learning** October — Present 2017  
*Course at MIPT, based on [rll.berkeley.edu/deeprlcourse/](http://rll.berkeley.edu/deeprlcourse/)*

Practical assignments: <https://github.com/deepmip/ipt/deep-rl-seminars>

**Deep Learning in Natural Language Processing** March — Present 2017  
*Course at MIPT, based on [cs224n.stanford.edu](http://cs224n.stanford.edu)*

Practical assignments: <https://github.com/deepmip/deep-nlp-seminars>

## EDUCATION

**Moscow Institute of Physics and Technology** 09.2014 — 08.2018 (expected)  
*Department of Innovation and High Technologies,  
Undergraduate student (B.Sc.)*

- Discrete mathematics: mathematical logic, discrete analysis, probability theory, mathematical statistics (with practical assignments in Jupyter notebooks), random processes
- Mathematics: mathematical analysis, linear algebra, differential equations, computational mathematics
- Computer science: Python, SQL, C/C++ with OpenMP and MPI frameworks, algorithms and data structures, OOP and design patterns, multithreading and concurrency, distributed computing, MapReduce, Hadoop and Hive
- General physics (lectures, seminars, lab work), theoretical physics
- Other: business communications, chemistry

ADDITIONAL EDUCATION	<b>DeepBayes Summer School</b> , Moscow "Summer school on Bayesian Methods in Deep Learning" <a href="http://deepbayes.ru">http://deepbayes.ru</a>	August 26 - 30, 2017
	<b>Bioinformatics Summer School</b> , Moscow "Big Data in Bioinformatics" <a href="http://bioinformaticsinstitute.ru/summer2017">http://bioinformaticsinstitute.ru/summer2017</a>	July 31 - August 5, 2017
	<b>Natural Language Processing</b> (based on cs224d.stanford.edu) by <i>DeepHack Lab</i>	2016
	<b>Supercomputer technologies for atomistic modelling</b> by Igor Morozov ( <i>IHED RAS</i> ) Molecular Dynamics - program written in C using OpenMP framework for parallel computing. Used VMD for visualization. Code on GitHub - <a href="https://github.com/akarazeev/MolecularDynamics-3sem-MIPT-2015">https://github.com/akarazeev/MolecularDynamics-3sem-MIPT-2015</a> .	2015
CONFERENCES	<b>Moscow Conference on Computational Molecular Biology</b> Moscow, Russia "Advanced Parser for Biomedical Texts" Poster: <a href="https://akarazeev.github.io/data/poster_mccmb2017.pdf">https://akarazeev.github.io/data/poster_mccmb2017.pdf</a> Thesis: <a href="https://akarazeev.github.io/data/thesis_mccmb2017.pdf">https://akarazeev.github.io/data/thesis_mccmb2017.pdf</a>	July 27 - 30, 2017
PUBLICATIONS	<b>Medium Story</b> "Generative Adversarial Networks (GANs): Engine and Applications"	August, 2017
HACKATHONS	<b>LauzHack</b> , EPFL, "NN:Nerds" team member 1st place in challenge by <a href="#">SGS</a> , <a href="#">presentation</a> Responsible for development of telegram-bot (Python) and processing documents using IBM Watson service for Natural Language Understanding.	November 11 - 12, 2017
	<b>mABBYlity</b> , Phystechpark "App in the Restaurant" iOS application, <a href="#">demo</a> , <a href="#">presentation</a> App allows to recognise entities from restaurant menus using smartphone's camera and translates them. ABBYY Real-Time Recognition SDK, ABBYY Lingvo API and Spoonacular API were used. Responsible for backend (Python).	October 7 - 8, 2017
	<b>Neurocampus</b> , Skolkovo Moscow School of Management "S.o.S. - Sense of Speech" telegram-bot, <a href="#">@SenseOfSpeech_bot</a> , <a href="#">2nd place</a> Speech Emotion Recognition (SER) module was used as a core for telegram-bot based system to help users improve speech during performances. Responsible for development (Python).	September 22 - 24, 2017
	<b>Bioinformatics Summer School</b> , Moscow "Prediction of Experimental Metadata from Gene Expression" Used Machine learning algorithms to predict phenotype by gene expression. Datasets from Gene Expression Omnibus were used. Link to project on GitHub - <a href="https://github.com/BioinfoGroup/MetaDataPredictionFromGE">https://github.com/BioinfoGroup/MetaDataPredictionFromGE</a>	August 3 - 4, 2017
	<b>BioHack 2017</b> , Saint Petersburg Text Mining, parsing the records from PubMed and UMLS. Link to project on GitHub - <a href="https://github.com/akarazeev/BioHack2017">https://github.com/akarazeev/BioHack2017</a>	March 3 - 5, 2017
	<b>Junction 2016</b> , Helsinki, "Dreamteam" team member Used a python wrapper around the Twitter API and Topic Modeling of tweets (gensim).	November 25 - 27, 2016
PROJECTS	<b>Frontopolar</b> , Moscow Applied Reinforcement Learning for Stock Trading.	February - June, 2017

Different approaches were tested including Q-learning and Recurrent Reinforcement Learning. Link to project's wiki on GitHub - <https://github.com/FRTP/Algorithms/wiki>

**Contributed to Open source:**

- <https://github.com/akarazeevprojects>
- **Gensim** - fixed issue #671
- [yandexdataschool/Practical\\_RL](https://github.com/yandexdataschool/Practical_RL) - PR #12

**MOOC**

**Machine Learning**

2016

*by Yandex & MIPT on coursera.org*

**Neural Networks**

*by Bioinformatics Institute on stepik.org*

**Molecular Biology and Genetics**

*by Bioinformatics Institute on stepik.org*

**Discrete Structures**

2015

*by Alex Dainiak (MIPT) on stepik.org*

**SKILLS**

- **Russian:** native, **English:** fluent, **German:** basics (A2)
- Started "MIPT Deep Learning Club" (<https://www.facebook.com/dlmipt>, <https://vk.com/dlmipt>) to discuss and share ideas on deep learning topics. Led a few seminars on topics such as "Introduction to bayesian methods". Recordings can be found here: [youtube.com](https://www.youtube.com) and [pscp.tv](https://www.pscp.tv)
- Experimented with RaspberryPi and Arduino. Projects are located on GitHub (<https://github.com/akarazeev> and <https://github.com/akarazeevprojects>)