

# Anton Karazeev

anton.karazeev@gmail.com

+7 (977) 490-21-83



## EDUCATION

**Moscow, Russia**

**Moscow Institute of Physics and  
Technology**

**September 2014 —  
July 2019 (expected)**

- B.Sc. in Computer Science and Physics, Department of Innovation and High Technologies
- Coursework for the state qualification exam in Physics at MIPT: "Molecular Dynamics" [[Presentation](#)] [[Code](#)]
- Intermediate Coursework: "Advanced Parser for Biomedical Texts" [[Presentation](#)] [[Poster at MCCMB'17](#)]
- Undergraduate Coursework: "[Development of a mechanism for anomaly detection](#)" [[Presentation](#)] [[Code](#)]

## EXPERIENCE

**R&D Data Scientist**

**Information Systems  
Development Center**

**July 2019 (expected)**

- Responsible for Natural Language Processing, processing of legal documents.

**Research Fellow**

**Laboratory for Digital Business**

**March 2019 — Present**

- Responsible for research on Anomalies and Outliers Detection.
- Found and fixed a bug concerning model based on Generative Adversarial Active Learning (GAAL) in [PyOD](#) toolkit for outlier detection.

**R&D Data Scientist**

**ChatFirst**

**September 2018 — April 2019**

- Implementing different deep learning models to improve performance of chatbots, reading papers on related topics.
- Responsible for Natural Language Processing.

**Research Fellow**

**Russian Quantum Center**

**January 2018 — March 2019**

- Responsible for Quantum Computing, Machine Learning.
- Developed [a series of seminars](#) concerning Quantum Computing. They include jupyter notebooks with basics of linear algebra, quantum mechanics and also work with QISKit (IBM), pyQuil (Rigetti) and Q# (Microsoft) was demonstrated.
- Supervised research student's bachelor degree diploma "Precision-Guaranteed Single-Qubit Process Tomography".
- Became a member of [IBM Qiskit](#) on GitHub.
- "[Revealing quantum chaos with machine learning](#)" — arXiv preprint.

**Teaching Assistant**

**Laboratory of Neural Networks  
and Deep Learning**

**March — December 2017**

- Responsible for preparing practical and theoretical assignments for the course of [Reinforcement Learning](#) and theoretical assignments for the course of [Natural Language Processing](#) with the number of 100+ enrolled students each.

- Research on protein function analysis.
- Text mining, Natural language processing, Keyword extraction, Machine learning algorithms. As an intermediate result the new method of keywords extraction using Information Theory proposed ([ResearchGate](#)).
- Participated in development of NLP package [SciLK](#) which was designed specifically for text-mining in natural sciences like biology and chemistry.

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

- Responsible for Natural Language Processing projects.
- Participated in preparing the datasets and building baselines for competition [Sberbank Data Science Journey](#) which is based on [SQuAD](#).
- Developed an analogue of Amazon Mechanical Turk to improve experience of colleagues who evaluated the quality of collected datasets (Python, Flask).

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

- Negotiated with IBM engineers and applied some of the IBM Watson's services in tasks of signal processing.
- The system of baby cry recognition was built.

**TECHNICAL EXPERIENCE**

---

**Projects**

- **[Service for Reading](#)** (2019)
- **[Quantum Computing Bot](#)** (2018). Monitoring the load of IBM Q processors from IBM Quantum Experience. Bot is already available inside QISKit workspace in Slack (soon it will be uploaded here: <https://github.com/Qiskit/qiskit-bot>). And it's available on Telegram: <https://t.me/QuantumComputingBot>.
- **[Quantum Keypad](#)** (2018). This keypad allows to easily compose quantum circuits of different kinds. Besides keypad itself, Quantum Keypad consists of a power bank and Raspberry Pi Zero W. As a simulator I used [QISKit](#) package for Python. Inspired by [Model Q](#).
- **[Reverse Engineering in Dispersion Engineering](#)** (2018). With my friend from EPFL we have a project on Dispersion Engineering. Our model predicts parameters of resonator system's simulation.
- **[Frontopolar](#)** (2017). Applied Reinforcement Learning for Stock Trading. State of the art results were reached. Different approaches were tested including Q-learning and Recurrent Reinforcement Learning.

**Contributed to Open source**

- **[PyOD](#)** - PR #108
- **[QISKit](#)** - PR #366
- **[pyQuil](#)** - PR #371
- **[SimulaQron](#)** - PR#90
- **[Gensim](#)** - fixed issue #671
- **[yandexdataschool/Practical\\_RL](#)** - PR #12
- **[My projects on GitHub](#)**

## SKILLS

---

- **Russian:** native, **English:** fluent, **German:** basics (A2)
- **Programming languages:** Python, C/C++, bash, R, experienced with SQL and JavaScript
- **Python libraries:** numpy, sklearn, pandas; **for NLP:** NLTK, Gensim; **for Deep Learning:** TensorFlow, PyTorch; **for Quantum Computing:** QISKit, pyQuil, Q#; **for Web:** Flask
- Experimented with RaspberryPi and Arduino. [Projects](#)
- Started "MIPT SciTech Club" to discuss and share ideas on deep learning topics. Led a few seminars on topics such as "Introduction to bayesian methods"
- Experienced with **3D modeling** (FreeCAD, Blender) and **3D printing** (Ultimaker Cura, Ender 3). [Examples of models](#)

## TEACHING

---

<b>Programming Existing Quantum Computers</b> Cryptography course at <a href="#">Yandex School of Data Analysis</a>	<b>May 8, 2018</b>
<b>Deep Reinforcement Learning</b> course at MIPT, based on <a href="http://rl.berkeley.edu/deeprlcourse/">rl.berkeley.edu/deeprlcourse/</a> <a href="#">Practical assignments</a>	<b>October — December 2017</b>
<b>Deep Learning in Natural Language Processing</b> course at MIPT, based on <a href="http://cs224n.stanford.edu">cs224n.stanford.edu</a> <a href="#">Practical assignments</a>	<b>March — December 2017</b>

## PUBLICATIONS

---

<b>"Revealing Quantum Chaos with Machine Learning"</b> <a href="#">arXiv preprint</a>	<b>February 2019</b>
<b>"Neural Network Quantum State Tomography"</b> <a href="#">Superconducting Quantum Technologies (SQT)</a> Poster (based on <a href="https://github.com/RQC-QApp/NNQST">https://github.com/RQC-QApp/NNQST</a> )	<b>July 30 — August 3, 2018</b>
<b>"Precision-guaranteed quantum process tomography: Application to IBM Quantum Experience"</b> <a href="#">Central European Workshop on Quantum Optics (CEWQO)</a> Poster	<b>May 21 — 25, 2018</b>
<b>"Generative Adversarial Networks (GANs): Engine and Applications"</b> <a href="#">Medium Story</a>	<b>August 2017</b>
<b>"Advanced Parser for Biomedical Texts"</b> <a href="#">Moscow Conference on Computational Molecular Biology (MCCMB)</a> <a href="#">Poster</a> , <a href="#">Thesis</a>	<b>July 27 — 30, 2017</b>

## ADDITIONAL EDUCATION

---

<b>"Quantum Computing" course at Skoltech</b>	<b><a href="#">Quantum Computing</a></b>	<b>February 1 — March 16, 2018</b>
• Final Project - <a href="#">Quantum walks and Variational algorithm</a> for 3- and 4-level systems.		
<b>"Summer school on Bayesian Methods in Deep Learning"</b>	<b><a href="#">DeepBayes Summer School</a></b>	<b>August 26 — 30, 2017</b>

- Participated in a hackathon during the school. [Project](#).

**"Natural Language Processing"  
course (based on  
[cs224d.stanford.edu](https://cs224d.stanford.edu))****DeepHack Lab****September — December 2016**

- Accepted a proposal to become a Teaching Assistant after the end of the course.

**"Supercomputer technologies  
for atomistic modelling" course****Igor Morozov ([IHED RAS](#))****September — December 2015**

- Final Project - [Molecular Dynamics](#) is a program written in C using OpenMP framework for parallel computing. Used [VMD](#) for visualisation.

**HACKATHONS**

---

**Kasárne/Kulturpark, Košice****Hack Kosice****March 30 — 31, 2019**

- [Efficient and Faster Care](#) challenge
- Implemented [healthcare system](#) using Zebra wristband printer and QR code scanner to identify patients easily
- Reduced the amount of time needed to register a new patient
- [Presentation](#)

**Aalto University, Helsinki****Junction****November 23 — 25, 2018**

- [Applications with Bluetooth Mesh](#) challenge
- Worked with [Zephyr RTOS](#) and [reel board](#) that has built-in Bluetooth transceiver
- Developed simple Industrial Internet of Things (IIoT) project which demonstrates the advantage of Bluetooth Mesh network

**Tochka Kipeniya, Moscow****Space Apps Challenge****October 20 — 21, 2018**

- [Firefighter Bot for Telegram](#)
- Implemented during [Space Apps 2018 Challenge](#) using data from NASA including [Active Fire Data](#) by NASA
- Python, python-telegram-bot

**Volkshotel, Amsterdam****Quantum Internet Hackathon****October 13 — 14, 2018**

- Worked with framework for Quantum Internet called [SimulaQron](#)
- Contributed to the [SimulaQron project](#) on GitHub
- [Implemented](#) quantum leader-election algorithm

**Skoltech, Moscow****Quantum Hackathon****May 18, 2018**

- 1<sup>st</sup> [place](#)
- There were problems on (1) quantum process tomography and (2) solving 3-SAT problem with QAOA
- Python, Quantum Information Toolkit (QIT)

**Aalto University, Helsinki****Junction****November 24 — 26, 2017**

- [LegalEngine](#) - [website/telegram chat-bot/email notification system](#), "qqmbr" team member, challenge by [Castrén & Snellman](#)
- Our solution makes the client-attorney interaction easier with the use of telegram chat-bot and email notifications, the attorney's work and billing more transparent to the client.
- Python, Flask library, html, css

- 1<sup>st</sup> place in challenge by [SGS](#), "NN:Nerds" team member, [Presentation](#)
- Solution allows quick access to the main concepts found in documents.
- Responsible for development of telegram-bot and processing documents using IBM Watson service for Natural Language Understanding. [Devpost](#).
- Python, IBM Watson API, Telegram API

- 4<sup>th</sup> place, "App in the Restaurant" iOS application, [Demo](#), [Presentation](#)
- 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.
- Python, Flask library

- 2<sup>nd</sup> place, [@SenseOfSpeech\\_bot](#) telegram-bot, [Presentation](#)
- Solution allows to extract emotions from user's recorded speech. Also it helps to train selected emotion with samples from TED talks.
- Speech Emotion Recognition (SER) module by [Vokaturi](#) was used as a core for telegram-bot based system to help users improve speech during performances.
- Python, Telegram API

- "Prediction of Experimental Metadata from Gene Expression"
- Used Machine learning algorithms to predict phenotype by gene expression. Distinguish with high accuracy samples of male and female tissues of [Mus musculus](#) organism. Datasets from Gene Expression Omnibus were used. [Project](#).

- Text Mining, parsing the records from [PubMed](#) and [UMLS](#).
- Analysis of research trends of chemical compounds and diseases during period of 1990-2015 using parsed information from PubMed database. [Project](#).
- Python

- Used a python wrapper around the Twitter API and Topic Modeling of tweets (gensim).