

# Nikita Kazeev

*Research scientist with 7+ years of experience at the intersection of AI and physics. Specializing in the development of novel machine learning models to solve bespoke challenges.*

## Work Experience

National University of Singapore: postdoc under [Kostya Novoselov](#) 2022 – present

### Leadership

- Speaker search and selection for the 500-person [AI4X 2025 conference](#)
- The main organizer of the [ICLR 2025 workshop](#) on multiscale machine learning

### Transformer for generating symmetric crystals; [ICML 2025](#); [code](#)

- Created a generative model for material design based on symmetry inductive bias;
- Coded it in pytorch
- The best generation quality and diversity among space group – conditioned models;
- Lead a team of 6

### ML for predicting properties of defects in 2D crystals; papers ([1](#), [2](#)); [patent](#); [conference](#); [code](#)

- Created sparse representation of crystals with defects, achieving 4x energy prediction error reduction
- Developed code for reproducible parallel running of machine learning experiments
- Lead a team of 3

CERN [Yandex → HSE University] 2014 – 2022

### Received 2025 Breakthrough Prize in Fundamental Physics [as a member of LHCb](#)

### Generative models uncertainty estimation; [paper](#); [conference](#); [code](#) ([1](#), [2](#))

- Developed first ever methods for estimating uncertainty of conditional GANs
- Lead a team of 3 students

### Generative models for fast simulation; papers ([1](#), [2](#), [3](#)); [code](#); news ([1](#), [2](#), [3](#), [4](#), [5](#))

- Developed a machine learning model (GAN) trained on tabular data for high-fidelity simulation of a Cherenkov detector achieving  $\sim 10^5$  speed-up compared to the ab-initio simulator;

### Machine learning on noisy data; papers ([1](#), [2](#)); [code](#); [presentation video](#)

- Developed a way to rigorously train ML algorithms on data with label noise for the noise model common in high-energy physics

### Muon identification at the LHCb experiment at CERN; [paper](#); [poster](#)

A classification problem over tabular data with noisy labels under timing constraints

- Developed a gradient boosting model that provides 30% false positive rate reduction in the critical low track momentum region
- Integrated the model into LHCb production (mostly C++)
- Packaged it as a problem for a data science competition ([IDAO-2019](#))

### CatBoost aka fighting biases with dynamic boosting; [paper](#)

- Set up a distributed system for running experiments for the team with bash and Python
- Studied gradient boosting enchantments with experiments on toy data

## Education

**Higher School of Economics (HSE University)** 2016 – 2020

PhD in Computer Science, supervisor [Andrey Ustyuzhanin](#);  
[Machine Learning for particle identification in the LHCb detector](#)

**Sapienza – Università di Roma** 2016 – 2020

PhD in Physics (double degree with the HSE), supervisor [Barbara Sciascia](#)

**Product Management course at Yandex** spring 2018

**Yandex School of Data Analysis** 2013 – 2015

Master's level CS course: Algorithms, Machine Learning, Deep Learning, Distributed Systems

**Moscow Institute of Physics and Technology**

MS in Physics 2014 – 2016

Optimisation of data processing of the LHCb experiment

BS in Physics 2010 – 2014

Study of the quantum states of the electrons in nonideal plasma and selected molecules using wave packet molecular dynamics with packet splitting

## Mentorship

- Mentored 9 students ([1](#), [2](#), [3](#), [4](#), [5](#), [6](#), [7](#), [8](#), [9](#)), 3 interns, student workshop projects ([1](#), [2](#))
- Student council member 2013 – 2016, led several dormitory-scale IT projects
- Co-PI of a \$3.4m AI Singapore grant multiscale machine learning

## Teaching and outreach

- Pitch my work to the general public: major Russian newspaper ([1](#), [2](#)), blog ([1](#), [2](#)), [science slam](#), [interview](#), [interview](#), [talk](#)
- Taught at Machine Learning in High Energy Physics Summer School in [2015](#), [2016](#), [2017](#), [2018](#), [2019](#), [2020](#), [2021](#), [2023](#)
- Lecturer at the [award-winning](#) Coursera Advanced Machine Learning [specialisation](#)
- Introductory data analysis [course](#) at HSE - teaching & course development
- [Science battle](#): a pitch of my research to the general public, in Russian
- >20 [conference talks](#)

## Service

- Reviewer for RSC Advance, Machine Learning: Science and Technology, AI for Accelerated Materials Design Workshop @ NeurIPS & ICLR 2023 – 2025;
- Peer Staff Supporter at NUS, serving as a first line support for mental wellbeing

## Other

- An aspiring stand-up comedian; bombed more than 20 times at open mics
- International Junior Science Olympiad – 2008, Korea, silver medal

## Skills

- **Machine Learning & Artificial Intelligence**, with a focus on models for structured non-text domain-specific data
- **Programming**: research coding in Python and pytorch; production in C++; Linux administration
- **Research writing & speaking**
- **Physics**: deep understanding of natural science
- **Collaboration, leadership & mentorship**
- **Languages**: Russian (native), English (fluent, IELTS 8)

## Academic publications

All: [Google Scholar](#)

### Selected:

1. **Kazeev, Nikita**, et al. "[WyckoffTransformer: Generation of Symmetric Crystals.](#)" ICML 2025
2. **Kazeev, N.**, Al-Maeeni, A.R., Romanov, I. et al. [Sparse representation for machine learning the properties of defects in 2D materials.](#) *npj Comput Mater* 9, 113 (2023).
3. [Anderlini, L., Chimpoesh, C., Kazeev, N., Shishigina, A., & LHCb collaboration. \(2023, February\). Generative models uncertainty estimation. In Journal of Physics: Conference Series \(Vol. 2438, No. 1, p. 012088\). IOP Publishing.](#) **I'm the corresponding author**
4. M. Borisyak and N. Kazeev. "Machine Learning on data with sPlot background subtraction". In: *Journal of Instrumentation* 14.08 (Aug. 2019), P08020–P08020. [doi: 10.1088/1748-0221/14/08/p08020](#). **I'm the corresponding author**
5. Derkach, D., Kazeev, N., Ratnikov, F., Ustyuzhanin, A., & Volokhova, A. (2019). Cherenkov detectors fast simulation using neural networks. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*. <https://doi.org/10.1016/j.nima.2019.01.031> **I'm the corresponding author**