

# Polina Kirichenko

pol.kirichenko@gmail.com ✉  
polkirichenko.github.io 🌐  
Google Scholar, Semantic Scholar 🎓

## Education

---

**Ph.D. degree in Machine Learning, New York University** New York, USA  
Center for Data Science; supervisor: professor [Andrew Gordon Wilson](#) 2019–2024  
Thesis: On the Reliability of Deep Learning Models: Uncertainty and Generalization Under Distribution Shifts

**Graduate student in Operations Research, Cornell University** Ithaca, USA  
Operations Research and Information Engineering department; transferred to NYU 2018–2019

**B.Sc. degree in Computer Science, Higher School of Economics** Moscow, Russia  
Computer Science department; supervisor: professor [Dmitry Vetrov](#) 2014–2018  
Cumulative GPA: 9.1 (10.0 scale), class rank: top 3%

## Work Experience

---

**Meta AI, FAIR** New York, USA  
Research Scientist, supervisor: professor [Kamalika Chaudhuri](#) 2024 – current  
Visiting Researcher (FAIR-NYU AI Mentorship program); mentor: [Mark Ibrahim](#) 2022–2024  
Research Intern at AI Integrity team; mentors: [Hamed Firooz](#), [Randall Balestriero](#) June 2022–Sep 2022

**Princeton University, Visual AI group** Princeton, USA  
Visiting Researcher, supervisor: professor [Olga Russakovsky](#) 2024 – current

**Cold Spring Harbor Laboratory** Cold Spring Harbor, USA  
Research Intern; supervisor: professor [Anthony Zador](#) June 2021–Aug 2021  
Research topics: meta-learning with compressed neural networks

**DeepMind** Mountain View, USA  
Research Scientist Intern; mentors: [Mehrdad Farajtabar](#), [Razvan Pascanu](#), [Balaji Lakshminarayanan](#) June 2020–Oct 2020  
Research topic: continual learning with deep generative models

**École Polytechnique Fédérale de Lausanne (EPFL)** Lausanne, Switzerland  
Machine Learning and Optimization Lab June 2018–Aug 2018  
Research Intern; supervisors: professors [Martin Jaggi](#) and [Dan Alistarh](#)  
Research topic: low precision training of neural networks

**Bayesian Methods Research Group** Moscow, Russia  
Research Assistant; supervisor: professor [Dmitry Vetrov](#) 2016–2018  
Research topic: structured sparsification of deep neural networks

**Google**  
Software Engineering Intern, Google Cloud Platform Team 2017, Seattle, USA  
STEP Software Engineering Intern, Piper Team (Google's version control system) 2016, Munich, Germany

## Publications

---

\* denotes equal contribution

**Pathologies of Out-of-Distribution Detection**  
Yucen Lily Li, Daohan Lu, [Polina Kirichenko](#), Shikai Qiu, Tim Rudner, Bayan Bruss, Andrew Gordon Wilson  
Under Review

**Modeling Caption Diversity in Contrastive Vision-Language Pretraining** [arXiv]  
Samuel Lavoie, [Polina Kirichenko](#)\*, [Mark Ibrahim](#)\*, Mahmoud Assran, Andrew Gordon Wilson,  
Aaron Courville, Nicolas Ballas  
*International Conference on Machine Learning (ICML) 2024*

- Decomposed Evaluations of Geographic Disparities in Text-to-image Models** [arXiv]  
Abhishek Suresddy, Dishant Padalia, Nandhinee Periyakaruppa, Oindrila Saha, Adina Williams, Adriana Romero-Soriano, Megan Richards\*, **Polina Kirichenko\***, Melissa Hall\*  
*Trustworthy Multi-modal Foundation Models Workshop at ICML 2024*; **Outstanding paper award**
- Does Progress On Object Recognition Benchmarks Improve Real-World Generalization?** [arXiv]  
Megan Richards, **Polina Kirichenko**, Diane Bouchacourt, Mark Ibrahim  
*ICML Data-centric Machine Learning Research Workshop 2023*  
**International Conference on Learning Representations (ICLR) 2024**
- Understanding the Detrimental Class-level Effects of Data Augmentation** [arXiv]  
**Polina Kirichenko**, Mark Ibrahim, Randall Balestrieri, Diane Bouchacourt, Rama Vedantam, Hamed Firooz, Andrew Gordon Wilson  
*ICML Workshop on Spurious Correlations, Invariance, and Stability 2023*  
**Neural Information Processing Systems (NeurIPS) 2023**
- Last Layer Re-Training is Sufficient for Robustness to Spurious Correlations** [arXiv, code]  
**Polina Kirichenko\***, Pavel Izmailov\*, Andrew Gordon Wilson  
*ICML Workshop on Spurious Correlations, Invariance, and Stability 2022*; **oral presentation**  
**International Conference on Learning Representations (ICLR) 2023**; **spotlight (notable-top-25%)**
- On Feature Learning in the Presence of Spurious Correlations** [arXiv, code]  
Pavel Izmailov\*, **Polina Kirichenko\***, Nate Gruver\*, Andrew Gordon Wilson  
First presented at *ICML Workshop on Principles of Distribution Shift 2022*  
**Neural Information Processing Systems (NeurIPS) 2022**
- Chroma-VAE: Mitigating Shortcut Learning with Generative Classifiers** [arXiv]  
Wanqian Yang, **Polina Kirichenko**, Micah Goldblum, Andrew Gordon Wilson  
**Neural Information Processing Systems (NeurIPS) 2022**
- Does Knowledge Distillation Really Work?** [arXiv, code]  
Samuel Stanton, Pavel Izmailov, **Polina Kirichenko**, Alexander A. Alemi, Andrew G. Wilson  
**Neural Information Processing Systems (NeurIPS) 2021**
- Task-agnostic Continual Learning with Hybrid Probabilistic Models** [arXiv, poster]  
**Polina Kirichenko**, Mehrdad Farajtabar, Dushyant Rao, Balaji Lakshminarayanan, Nir Levine, Ang Li, Huiyi Hu, Andrew Gordon Wilson, Razvan Pascanu  
*ICML Workshop on Invertible Neural Networks and Normalizing Flows 2021*; **spotlight talk**
- Why Normalizing Flows Fail to Detect Out-of-Distribution Data** [arXiv, code]  
**Polina Kirichenko\***, Pavel Izmailov\*, Andrew G. Wilson  
First presented at *Workshop on Invertible Neural Networks and Normalizing Flows at ICML 2020*  
**Neural Information Processing Systems (NeurIPS) 2020**
- Semi-Supervised Learning with Normalizing Flows** [arXiv, poster, code]  
Pavel Izmailov\*, **Polina Kirichenko\***, Marc Finzi\*, Andrew G. Wilson  
First presented at *Workshop on Invertible Neural Networks and Normalizing Flows at ICML 2019*  
**International Conference on Machine Learning (ICML) 2020**
- Subspace Inference for Bayesian Deep Learning** [arXiv, poster, slides, code]  
Pavel Izmailov\*, Wesley Maddox\*, **Polina Kirichenko\***, Timur Garipov\*, Dmitry Vetrov, Andrew G. Wilson  
First presented at *Workshop on Uncertainty & Robustness in Deep Learning at ICML 2019*; **contributed talk**  
**Uncertainty in Artificial Intelligence (UAI) 2019**
- SWALP: Stochastic Weight Averaging in Low Precision Training** [PMLR, code]  
Guandao Yang, Tianyi Zhang, **Polina Kirichenko**, Junwen Bai, Andrew G. Wilson, Christopher De Sa  
**International Conference on Machine Learning (ICML) 2019**

## Service

---

Conference reviewing:	NeurIPS 2019 (top 400 highest-scoring reviewers), ICLR 2020, ICML 2020 (top 33% reviewer), UAI 2020, NeurIPS 2020, AISTATS 2021, AISTATS 2022, NeurIPS 2022, ICML 2023, NeurIPS 2023, ICML 2024
Workshop reviewing:	NeurIPS 2019 WiML, NeurIPS 2019 BDL, ICML 2020 UDL, NeurIPS 2020 HAMLETS, ICML 2021 INN+ , ICML 2021 UDL, NeurIPS 2021 BDL, NeurIPS 2022 DistShift, ICML SCIS 2023, NeurIPS ATTRIB 2023
Workshop organizing:	HAMLETS at NeurIPS 2020, HAMLETS at NeurIPS 2021

## Awards

---

<b>Outstanding paper award at ICML workshop on Trustworthy Foundation Models</b>	2024
<b>Spotlight at ICLR</b>	2023
<b>DeepMind Fellowship</b>	2019
New York University Center for Data Science Graduate Fellowship	2019
Golden HSE Award	<a href="#">[link]</a> , 2019
HSE Alumni Academic Fellowship	<a href="#">[link]</a> , 2019
NeurIPS Travel Award	2019
ICML Travel Award	2019
Cornell Operations Research and Information Engineering Graduate Fellowship	2018
Travel Grant for Women in Data Science Conference	<a href="#">[link]</a> , 2018, 2019
Ilya Segalovich Scholarship (Yandex)	<a href="#">[link]</a> , 2016, 2017
<b>Google Generation Scholarship EMEA (Google Anita Borg Memorial Scholarship)</b>	<a href="#">[link]</a> , 2015
Google Travel Grant for the Grace Hopper Celebration of Women in Computing	2015

## Talks

---

“Addressing robustness to biases in vision foundation models”	
- <b>Invited talk at the ECCV Workshop on Uncertainty Quantification for Computer Vision</b>	2024
“Towards robust and reliable deep learning”	
- Princeton, Visual AI Lab seminar	2023
- FAIR Labs, Meta AI	2023
- Microsoft Research, AI Frontiers labs	2023
“Distribution shifts in machine learning”	
- Guest lecture at the “Introduction to Data Science” course at NYU	2023
“Leveraging Large Scale Models for Identifying and Fixing Deep Neural Networks Biases”	
- WiML Un-Workshop at ICML	2023
“Last Layer Re-Training is Sufficient for Robustness to Spurious Correlations”	
- <b>Spotlight talk at ICLR</b>	<a href="#">[link]</a> , 2023
- Oral Presentation at the ICML Workshop on Spurious Correlations, Invariance, and Stability	<a href="#">[link]</a> , 2022
- Genentech, AI seminar	2022
“Robustness of Deep Learning Models to Distribution Shift”	
- WiML Un-Workshop at ICML	2022
“Why Normalizing Flows Fail to Detect Out-of-Distribution Data”	
- ML Collective, Deep Learning: Classics and Trends (with Robin T. Schirmer)	<a href="#">[slides]</a> , 2021
- Facebook AI Research, Uncertainty team	2021
- CogSys Talks, Technical University of Denmark	<a href="#">[video]</a> , 2020
- Capital One, Machine Learning seminar	2020
- NeurIPS 2020	<a href="#">[video]</a> , 2020
- INN+ : Invertible Neural Networks and Normalizing Flows workshop at ICML	<a href="#">[video]</a> , 2020
“Applications of normalizing flows: semi-supervised learning, anomaly detection, and continual learning”	
- <b>Keynote talk at ICML Workshop on Representation Learning for Finance</b>	<a href="#">[video]</a> , 2021
“Does your model know what it doesn’t know?”	
- WiML Un-Workshop at ICML	2021

- “Task-agnostic Continual Learning with Hybrid Probabilistic Models”  
 - ICML INN+ workshop [video], 2021
- “Continual Learning in Neural Networks”  
 - Bayesian Methods Research Group seminar [video], 2021
- “Anomaly detection via Generative Models”  
 - ODS DafaFest 2020, Uncertainty Estimation in ML Workshop [video], 2020
- “Uncertainty Estimation in Bayesian Deep Learning”  
 - WiML Un-Workshop at ICML 2020
- “Subspace Inference for Bayesian Deep Learning”  
 - University of Paris-Saclay, UQSay seminar [link], 2021  
 - Uncertainty and Robustness in Deep Learning workshop at ICML [video], 2019  
 - Higher School of Economics [video], 2019
- “How do we build neural networks we can trust?”  
 - **Broad Institute of MIT and Harvard** [video], 2019

## Technical Skills

---

Programming languages: Python, C++  
 Deep learning frameworks: PyTorch, TensorFlow

## Teaching & Mentoring

---

- AI4ALL [ai-4-all.org](http://ai-4-all.org)  
 Mentor for undergraduate students working on their first ML research projects 2024  
 Cornell University Ithaca, USA  
 Teaching Assistant for “Data Science for All” course Jan 2019–May 2019
- Bayesian Methods for Machine Learning on Coursera [course link], Moscow, Russia  
 Teaching Assistant; helped prepare assignments and quizzes Sep 2017–Aug 2018  
 The specialization of the course received Coursera [Outstanding Educator Award](#)
- Higher School of Economics Moscow, Russia  
 Teaching Assistant for “Probability Theory and Statistics” (Sep 2016–June 2017),  
 “Introduction to Data Analysis” (Jan 2016–June 2016), “Introduction to Programming” (Sep 2015–Dec 2015)