MARGARITA BOYARSKAYA

PhD Candidate
NYU Stern Technology, Operations, and Statistics dept.

44 West 4th st, KMC 8-181
10012 New York, NY
+1(646)724-4297
mboyarsk@stern.nyu.edu

Education

09.2016 – 05.2023 (expected)
PhD, Information Systems
New York University, New York, United States
Used R and Python to perform data simulations for research projects in causal inference and data science. Cleaned and merged datasets, fit embedding models. Wrote research papers, proved theorems, and developed novel prediction methods. Performed statistical modeling and hypothesis testing. Selected coursework: Modern Statistics & Causal Inference, Causal Inference: Methods for Program Evaluation and Policy Research, Advanced Mathematics of Data Science, Inference & Representation, Natural Language Processing.

06.2009 – 07.2014
MS, **BSc**, Theoretical Mathematics
Lomonosov Moscow State University, Moscow, Russia
Major: Abstract Algebra

Work Experience

09.2022 – present
Senior Research Associate
JPMorgan AI Research, Explainable AI Center of Excellence
New York City, United States

Leading a research project on XAI and fairness at an industry AI research lab. Developing causal models for algorithmic recourse and explainability in consumer credit decisioning.

02.2022 – 06.2022
Cornell University
Course co-designer, assistant instructor, Break Through Tech AI
New York City, United States

Designed Python coding exercises, tests, laboratory assignments as well as written materials that communicated, demonstrated, and tested the understanding of Machine Learning foundations course. Assisted Prof. Brian D'Alessandro (Meta) in designing the course components. Worked closely with the course design team on selecting the topics and modes of instruction for each subject. Prepared datasets, teaching demonstration elements, and hands-on Python coding materials for each subject.

Designed materials for the following topics:
data science life cycle, data cleaning, mathematical computing with Numpy, Neural Networks and Deep Learning, K-Nearest Neighbors, Logistic Regression, Decision Trees, Random Forests, Gradient Boosting, text as data, Word Embeddings, model selection and evaluation metrics, ML debugging: bias, variance, class imbalance, leakage, concept drift; AI Fairness and societal harms, Explainable AI.

03.2020 – 06.2020 Research Intern

Microsoft Research, Fairness, Accountability, Transparency & Ethics group
New York City, United States
Collaborators: Prof. Alexandra Olteanu and Prof. Kate Crawford. Designed and executed a crowdsourced survey of potential harms that autocomplete, image search, chatbot, and search ranking systems may have on users. Wrote code to generate randomized experimental setups. Performed data collection and cleaning on Microsoft's internal crowdsourcing platform and iterated study design refinement. Performed statistical analysis and hypothesis testing. Co-developed a taxonomy and an actionable framework for identifying harms of IR systems to users and data subjects. Work generated attention from Microsoft product development team; a new working group was established to research the identification and mitigation of IR harms at Microsoft.

Co-authored an accepted NeurIPS 2020 NBIAIR workshop paper on the ethics of AI research. Selected presenter at PAI & CIFAR workshop on ethical research review and publishing. Work referenced in the New Yorker magazine. Co-authored an article 'Responsible Computing During COVID-19 and Beyond', published by Communications of the ACM.

05.2019 – 09.2019 Research Intern

Microsoft Research, Fairness, Accountability, Transparency & Ethics group
New York City, United States
Co-authored a paper on one of the central problems of fair ML research – the proxy variables problem (with Prof. Solon Barocas and Prof. Hanna Wallach). Proposed a causal framework for resolving inconsistent treatment of proxy variables. Developed a unifying definition of a proxy variable, identifying differences with the econometric and legal definitions. Surveyed Machine Learning literature as well as U.S. legislature and legal scholarship, and communicated key constraints to implementing statistical optimization solutions to unfairness by proxy variables.
Presented findings at NeurIPS 2019 HCML workshop, ICML 2020 Law & ML workshop; NYU CDS Math & Democracy seminar.

Gave feedback to the tri-lab Fairness, Accountability, Transparency, and Ethics (FATE) group members on ongoing research work. Contributed to the newest literature discussions at the tri-labs FATE reading group. Co-initiated a Causality and RL reading group among interns.

05.2017 – 01.2018 Research Fellow

Jain Family Institute
New York City, United States
Contributed to building the Digital Ethics and Algorithmic Fairness expertise at JFI. Led discussions in the Digital Ethics reading group.
Spearheaded a recommender systems research initiative. Strategized and designed a research program for building an ontology driven cross-domain scientific literature recommender system. Supervised an intern in data scraping. Initiated and led an NLP reading group.

Fellowships & Awards

2021 Doctoral Fellow
NYU Fubon Center for Technology, Business, and Innovation

2019, 2020, 2021 Student Fellow
NYU Law School, Information Law Institute, **Privacy Research Group**

Research

[In preparation] *Causal Models for Counterfactually Fair Classification Under Selection Bias.*
Joint work with Prof. Joshua Loftus (LSE).
Work-in-progress presented at: **AAAI AIES 2020** (one of 15 selected student posters), **MLSS Tuebingen 2020**. One of 8 selected oral presentations at **WiML 2019**.

[R&R in progress for a top-ranked business journal] *Theoretical Characteristics of Delta-Margin Voting in Crowdsourcing.*
Joint work Prof. Panos Ipeirotis (NYU Stern).
Work-in-progress presented at: **HCML 2019**, **CI 2019** (selected oral presentation).

[In progress] *What Is a Proxy and Why Is It a Problem?*
Joint work with Prof. Solon Barocas and Prof. Hanna Wallach (Microsoft Research).
Work-in-progress presented at: **ICML 2020 Law & Machine Learning** workshop, NYU CDS Math & Democracy seminar 2019, **NeurIPS 2019 HCML** workshop. A version of this work supplemented by the independent research of Michael Carl Tschantz was presented at **FAccT 2022**.

[Accepted article] *Overcoming failures of imagination in AI system development and deployment.*
Joint work with Prof. Alexandra Olteanu and Prof. Kate Crawford (Microsoft Research).
Presented at the **NeurIPS 2020 NBIAIR** workshop. Selected presentation at **Partnership on AI & CIFAR workshop** on ethical scientific reviewing.

[Published *perspectives* article] *Responsible Computing During COVID-19 and Beyond.*
Barocas S., A. J. Biega, **M. Boyarskaya**, K. Crawford, H. Daumé III, M. Dudík, B. Fish, M. L. Gray, B. Hecht, A. Olteanu, F. Poursabzi-Sangdeh, L. Stark, J. Wortman Vaughan, H. Wallach, M. Zepf
Communications of the ACM, July 2021, Vol. 64 No. 7.

Service & Initiative

2021, 2023 — ACM Fairness, Accountability, and Transparency conference (**FAccT**)
Program Committee member, **reviewer**.

2020 – present — **Causality & Fairness Reading Group**
Founded a monthly, cross-continental faculty and student-led reading group on the intersection of data science and social good. Co-organized, curated literature selection, and managed membership and outreach with Prof. Jennifer Hill and Ravi Sojitra.

2019 — Women in Machine Learning (**WiML**) workshop at NeurIPS 2019
Reviewer

Teaching

03.2022 – 06.2022 — **Co-designed an online Break Through Tech AI course** offered by **Cornell Tech** to underrepresented learners. Created all Machine Learning and AI exercises, wrote reading materials and quizzes, and advised MOOC content programming.

06.2019 – 09.2019 — Taught an undergraduate course **Introduction to Programming & Data Science** at NYU Stern.
Course evaluation: 5.0/5.0, instructor evaluation: 5.0/5.0

03.2019 – 04.2019 — Assisted Prof. Solon Barocas on the **Data Privacy & Ethics** class taught to the NYU Stern MSBA cohort. Graded all assignments and essays, authored long-form feedback to students.

02.2019 – 03.2019 — Prepared teaching materials on AI Fairness for Prof. Natalia Levina's Digital Innovation course taught to NYU Stern MBA cohort.

Additional Education

OxML 2021 — Selected participant (15% acceptance rate)
Oxford Machine Learning Summer School
Co-organized by AI for Global Goals & CIFAR
University of Oxford, UK

MLSS Tübingen 2020	Selected participant (120 out of 1600 applicants) **Machine Learning Summer School** **Max Planck Institute,** Tuebingen, Germany
SICSS 2020	Selected participant, invited as a speaker **Summer Institute in Computational Social Science** **UCLA**, California
M^2LSS 2020	Selected participant **Mediterranean Machine Learning Summer School** Milan, Italy

SKILLS

Natural languages	English (native-equivalent), Russian (native), Persian (intermediate), French (basic), East Circassian (basic/heritage speaker)
Programming & tools	Python, R, Matlab, SQL. NumPy, SciKit-Learn, PyTorch. Past experience in C, C++, JavaScript, HTML, SageMath, High Performance Computing (HPC)