

Denys Herasymuk

📍 New York, NY, USA ✉ d.herasymuk@nyu.edu 🌐 dherasymuk.com 🏠 Google Scholar 🔗 LinkedIn

RESEARCH INTERESTS

My research interests lie in **AI safety**, **time-series modelling**, and **healthcare applications**. I aim to develop fair, stable, robust, and transparent learning systems that can be deployed responsibly in real-world, high-stakes environments such as clinical decision support.

EDUCATION

New York University, Center for Data Science <i>Advisor: Prof. Julia Stoyanovich</i>	Ph.D. in Data Science <i>Sept. 2026 – Present</i>
Imperial College London <i>Research programme, published a full paper at IEEE AIoT'25 [P1]</i>	Computing Postgraduate (Occasional FT) <i>Oct. 2024 – May 2025</i>
Ukrainian Catholic University <i>GPA 94.3/100, with honors summa cum laude, top-1 CS university in Ukraine</i>	BSc. in Computer Science <i>Sept. 2019 – June 2023</i>
New York University, Tandon School of Engineering <i>GPA 4.0/4.0</i>	Visiting Student <i>May 2022 – Aug. 2022</i>

PUBLICATIONS

* indicates equal contribution/co-first authorship

Peer-Reviewed Papers

- [P1] M. M. Maheri, **D. Herasymuk**, H. Haddadi. *Client Clustering Meets Knowledge Sharing: Enhancing Privacy and Robustness in Personalized Peer-to-Peer Learning*. IEEE Annual Congress on Artificial Intelligence of Things (**IEEE AIoT**), 2025.
- [P2] F. A. Khan*, **D. Herasymuk***, N. Protsiv, J. Stoyanovich. *Still More Shades of Null: An Evaluation Suite for Responsible Missing Value Imputation*. International Conference on Very Large Data Bases (**VLDB**), 2025.
- [P3] **D. Herasymuk**, F. A. Khan, J. Stoyanovich. *Responsible Model Selection with Virny and VirnyView*. International Conference on Management of Data (**SIGMOD**), 2024.

Posters

- [P4] **D. Herasymuk**, D. Lopushansky, D. Eytan, M. Mazwi. *Real-time Deployment of Unsupervised Change-point Detection for Pediatric Cardiac Monitoring*. Symposium on Artificial Intelligence in Learning Health Systems (SAIL), 2026.

Preprints

- [P5] **D. Herasymuk**, A. Mozghova, N. Protsiv, V. Sydorak, J. Stoyanovich. *VirnyFlow: Optimizing ML Pipelines for Accuracy, Fairness, and Stability at Scale* (under review at IEEE BigData'26).
- [P6] M. Kim*, **D. Herasymuk***, D. Lopushansky, M. Mazwi, D. Eytan. *Beyond the AI Translational Gap: Can Foundation Models Finally Bridge Research and Practice in Pediatrics?* (under review at Translational Pediatrics'26).
- [P7] F. A. Khan, **D. Herasymuk**, N. Protsiv, J. Stoyanovich. *An Epistemic and Aleatoric Decomposition of Arbitrariness to Constrain the Set of Good Models*. arXiv preprint arXiv:2302.04525, 2025.

Theses

- [P8] **D. Herasymuk**. *An Experimental Framework for Measuring Model Fairness and Stability in the Machine Learning Lifecycle* (bachelor thesis). Supervisor: Prof. Julia Stoyanovich. Ukrainian Catholic University, Department of Computer Science. Lviv, 2023.

ACADEMIC EXPERIENCE

Center for Responsible AI at NYU

Sept. 2026 – Present

PhD Student

- Doing research on machine learning that encompasses AI safety, time-series modelling, and healthcare applications, supervised by [Prof. Julia Stoyanovich](#)

Seattle Children's Research Institute

June 2025 – Aug. 2026

Research Engineer

- Conducted research at the intersection of machine learning and healthcare under the supervision of [Prof. Mjaye Mazwi](#), focusing on real-world applications built on top of physiological and biomedical signals such as ECGs
- Pioneered responsible AI tools for pediatric cardiology by integrating fairness, stability, and transparency
- Developed and deployed an unsupervised, non-parametric change point detection algorithm for ECG time-series in real clinical environments, adapting the *Optimal Transport*, *Topological Data Analysis*, and *Moving Block Bootstrap* techniques to robustly identify arrhythmia-related transitions [P4]
- Created a multimodal ECG foundation model that combines *contrastive and captioning pretraining* on large-scale ECG-text pairs generated via a *retrieval-augmented generation* (RAG) pipeline, enabling robust and generalizable representations for arrhythmia detection and patient identification
- Collaborated with interdisciplinary teams of clinicians and data scientists to translate research findings into practical healthcare applications
- Contributed to the preparation of academic publications and dissemination of research outcomes in top-tier venues

Center for Responsible AI at NYU

Sept. 2022 – Jan. 2026

Research Fellow

- Researched model stability, uncertainty, and algorithmic fairness under the supervision of [Prof. Julia Stoyanovich](#)
- Built a comprehensive benchmark called [Shades-of-Null](#) for responsible missing value imputation. The paper was recently accepted and published at [VLDB'25](#) [P2], a top-tier data management venue
- Engineered an open-source library called [Virny](#) for in-depth profiling of model performance across overall and disparity dimensions. The paper was published at [SIGMOD'24](#) [P3], a top-tier data management systems venue
- Designed and implemented [VirnyFlow](#), a scalable design space for responsible model development that optimizes ML pipelines for accuracy, fairness, stability, and efficiency (under review at KDD'26) [P5]
- Conducted a large-scale study of stability and stability-disparity along the machine learning pipeline [P7]
- Wrote a bachelor thesis about a framework for measuring model fairness and stability in the ML lifecycle ([thesis](#))

I-X & Networks and Systems Lab, Imperial College London

Oct. 2024 – May 2025

Research Assistant

- Conducted research on trustworthy machine learning that encompasses privacy, robustness, personalization, and federated learning, supervised by [Prof. Hamed Haddadi](#)
- Created a decentralized FL framework that provides a personalized model for each client under differential privacy guarantees and security against malicious clients. The paper was recently accepted to [IEEE AIoT'25](#) [P1]

INDUSTRY EXPERIENCE

Harmix

Dec. 2025 – Jul. 2026

Senior Research Engineer

- Built core components of the LLM-based architecture behind [Pam](#), a persistent memory layer that self-onboards into company tools, reconciles conflicting sources, and serves organizational context to AI agents over MCP
- Created [Pam Benchmark](#), an open industrial benchmark for AI memory that evaluates competing systems under a uniform harness, on which Pam consistently outperformed other memory solutions — e.g., **+32.4 pt** over Markdown-file agent memory on *LoCoMo* with 1,986 questions, scored by using LLM-as-a-Judge ([article](#))

EPAM Systems

Jul. 2021 – Dec. 2025

Data Engineer → Senior Data Engineer

- Designed large-scale data infrastructure and machine learning pipelines on AWS for 100+ heterogeneous agricultural data sources (10+ TB), enabling predictive analytics and experimentation at scale
- Developed a graph-based data product using graph databases and graph neural networks to model complex relationships in crop lifecycle quality and yield forecasting ([talk](#))
- Built and maintained end-to-end data workflows to support decision-making in agricultural R&D

Harmix

Dec. 2019 – Sept. 2020

Lead Python Developer

- Developed an application layer for a complex deep learning system at an innovative AI-based startup ([website](#))
- Led and trained a team of three developers to build a Python-based Telegram chatbot and a React web application
- In 8 weeks, we received 30.1K+ users in 90+ countries and 130+ monthly subscriptions

LEADERSHIP & ACHIEVEMENTS

- *Mentored* a team of two students in the [RAI for Ukraine](#) program during the Fall 2025 semester
- Organized and fundraised for an annual *UCU CS Alumni Scholarship* ([post](#)); served as interviewer and jury member
- Worked as a *Teaching Assistant* for the Software Architecture course during the Spring semester of 2022–2023
- Awarded “*Motivational*” *scholarships* as top-1 CS student (2021–2022) and ranked in the top 5 (2020–2021)
- Earlier achievements: *National Olympiad awards* in Mathematics, Physics, and Astronomy in Ukraine (2019); gold medal for *educational excellence* in high school; 1st place at the *European Karate Championships* (2017–2018)

SKILLS

- **Technical:** *Languages:* Python, Java, C/C++, Go, SQL, R, Bash, JS; *ML:* PyTorch, TensorFlow, Pandas, PySpark; *Databases:* PostgreSQL, MongoDB, Neo4j, Redshift; *Cloud/DevOps:* AWS, GCP, Docker, Kubernetes, Linux, Git
- **Certificates:** [AWS Data Analytics](#), [Databricks DE Associate](#), [UCU Data Engineering](#), [Microsoft AI Challenge](#)
- **Languages:** English (Advanced, [TOEFL: 106](#)), Ukrainian (Native)

SELECTED INVITED TALKS

Smarter Together, Safer Apart: Achieving Privacy and Robustness in Personalized ML

- ELLIS Workshop on Safe and Secure Artificial Intelligence May 2026

Responsible Approaches to Missing Value Imputation and Model Selection in Machine Learning

- Ukrainian Catholic University, *UCU & NYU RAI Meetup*, “*RAI for Ukraine*” Program Oct. 2024

Responsible Model Selection with Virny and VirnyView

- Ukrainian Catholic University, *ELLIS Lightning Talk*, *ICTERI Conference* Sept. 2024

Unlocking the Power of Graphs: Graph Data Analytics and Graph Data Science

- EPAM Systems, *EPAM Ukraine Career Webinar* Feb. 2025
- Ukrainian Catholic University, *UCU Data Engineering Program Webinar* Oct. 2024

Reliability and Safety in Responsible AI

- Ukrainian Catholic University, *Lecture for master’s CS students* Dec. 2024

Asynchronous Communication

- National University of Kyiv-Mohyla Academy, *Lecture for master’s CS students* Apr. 2024

Databases: NoSQL in System Design

- National University of Kyiv-Mohyla Academy, *Lecture for master’s CS students* Mar. 2024

Last updated: August 15, 2026