



Wonkwon Lee

📍 111 Sylvan Ave, Englewood Cliffs, NJ 07632 📞 646-469-7805 @ wonkwon.lee@nyu.edu 🔗 <https://wonkwonlee.github.io>

Profiles	<div> wonkwonlee</div> <div> wonkwon-lee</div>	
Research Interests	<p><<<<<TEMP>>>>></p> <p>My research interests focus on building robust, safe, and interpretable AI systems, with a particular emphasis on their applications in healthcare. I am passionate about exploring out-of-distribution robustness in vision models, ensuring that AI systems can maintain performance in unpredictable environments. My past work includes differentially private synthetic data generation and benchmarking AI systems for privacy, utility, and reproducibility. Additionally, I am interested in applying causal inference and reasoning to AI systems to address alignment challenges, ensuring AI can understand human intentions and provide reliable, human-aligned decision-making.</p>	
Education	<div><div>New York University, Courant Institute</div><div>Computer Science</div><div>3.48</div></div> <div>Advisor: Professor Julia Stoyanovich</div> <div>Affiliate: Center for Responsible AI</div> <div>Publication: <i>Epistemic Parity: Reproducibility as an Evaluation Metric for Differential Privacy</i></div> <div>Coursework: Computer Vision, Natural Language Processing, Responsible Data Science, Data Science for Healthcare</div>	09/2021 - 05/2023 MS
	<div><div>University of Manchester</div><div>Computer Science and Mathematics</div><div>Upper Second-Class Honours</div></div> <div>Advisor: Eva M. Navarro-Lopez</div> <div>Publication: <i>Models of Neurons and Neuronal Networks</i></div> <div>Coursework: Convex Optimization, Linear Algebra, Partial Differential Equations, Real and Complex Analysis, Image Processing, Cryptography</div>	09/2015 - 06/2018 BSc
Research Experience	<div><div>Center for Responsible AI, New York University</div><div>Graduate Research Assistant</div><div>🔗 https://github.com/DataResponsibly/SynRD</div><div><ul style="list-style-type: none">Conducted research under Professor Julia Stoyanovich on evaluating differentially private (DP) synthetic data generation methods.Developed “Epistemic Parity,” an evaluation metric based on the likelihood of reproducibility of quantitative claims in social science research.Created SynRD, an open-source DP synthetic data benchmarking Python package that organizes the Epistemic Parity workflow, existing papers, and datasets.</div></div>	September 2022 - May 2023 New York, NY
	<div><div>McDevitt Lab, NYU</div><div>Graduate Research Assistant</div><div><ul style="list-style-type: none">Performed predictive diagnostic research under Professor John T. McDevitt and conducted statistical and meta-analysis.Preprocessed and visualized biomarker data from microfluidic chip sensors using R, Matplotlib, and Pandas.</div></div>	October 2021 - February 2022 New York, NY
Publications	<div><div>Epistemic Parity: Reproducibility as an Evaluation Metric for Differential Privacy</div><div>ACM SIGMOD Record, 53(1), 65-74</div><div>🔗 https://dl.acm.org/doi/10.14778/3611479.3611517</div><div>Lucas Rosenblatt, Bernease Herman, Anastasia Holovenko, Wonkwon Lee, Joshua Loftus, Elizabeth McKinnie, Taras Rumezhak, Andrii Stadnik, Bill Howe, Julia Stoyanovich</div></div>	06/2024
	<div><div>Epistemic parity: reproducibility as an evaluation metric for differential privacy</div><div>Proceedings of the VLDB Endowment, 16(11), 3178-3191</div><div>🔗 https://arxiv.org/abs/2301.10750</div><div>Lucas Rosenblatt, Bernease Herman, Anastasia Holovenko, Wonkwon Lee, Joshua Loftus, Elizabeth McKinnie, Taras Rumezhak, Andrii Stadnik, Bill Howe, Julia Stoyanovich</div></div>	08/2023
	<div><div>Out of distribution performance of state of art vision model</div><div>arXiv preprint arXiv:2301.10750</div><div>🔗 https://arxiv.org/abs/2301.10750</div><div>Salman Rahman, Wonkwon Lee</div></div>	01/2023

Industry Experience	LG CNS America System Engineer	04/2024 - Present Englewood Cliffs, NJ
	<ul style="list-style-type: none"> Designed and implemented network infrastructure to enhance system performance and security, collaborating with cross-functional teams to troubleshoot and resolve complex networking issues. Monitored and maintained network systems to ensure optimal uptime and reliability 	
	Stealth Project (EPLIA) Co-founder / Software Engineer	01/2023 – 12/2023 San Francisco, CA
	<ul style="list-style-type: none"> Co-founded a startup developing a telemedicine app that connects patients with doctors fluent in their native language, enhancing accessibility and patient care. Designed and implemented core application features using Next.js for the front-end, AWS for cloud services, and WebRTC for real-time communication. 	
	Pricewaterhouse Coopers NLP Data Scientist	06/2022 - 08/2022 New York, NY
	<ul style="list-style-type: none"> Implemented and fine-tuned a BERT-based relation extraction model using PyTorch to classify semantic relationships between entities. Designed data annotation protocols and ML pipelines; deployed the models to AWS for scalable production use. 	
Projects	Out-of-Distribution Robustness Evaluation Of Vision Models	09/2022 - 01/2023
	🔗 https://github.com/wonkwonlee/out-of-distribution-vision-model <ul style="list-style-type: none"> Conducted in-depth out-of-distribution robustness comparison research among 58 computer vision models, including ViT, convolution, attention-convolution hybrid, sequence-based, network-based, and multi-layer perception. 	
	Collaborative-Filter Based Recommender System	02/2022 – 05/2022
	🔗 https://github.com/wonkwonlee/als-recommender-system <ul style="list-style-type: none"> Implemented a collaborative-filter-based movie recommender system using PySpark's alternating least square method and achieved mean average precision of 0.066 	
	Landslide Prediction Modeling	04/2021 – 08/2021
	🔗 https://github.com/wonkwonlee/landslide-prediction-modeling <ul style="list-style-type: none"> Preprocessed GIS and time-series climate data and implemented XGBoost and LightGBM models using TensorFlow Won 6th out of 150 teams in a national data science competition. (top 4%) 	
Awards	SIGMOD Research Highlight Awards ACM SIGMOD 2024 🔗 sigmod-award	06/2024
	Best Experiment, Analysis & Benchmark Paper Runner-up VLDB 2023 🔗 https://vldb.org/2023/?conference-awards	08/2023
	Wasserman Center Career Grant New York University	11/2021
	KMA Landslide Prediction Modeling Contest Korea Meteorological Administration	09/2021
	Computer Science Final-Year Project Award University of Manchester	07/2018
	International Mathematical Excellence Scholarship University of Manchester	09/2015 - 09/2017

Skills	Programming Languages		
	Python, Java, C/C++, R, MATLAB, JavaScript, SQL		
	Frameworks and Libraries		
	PyTorch, TensorFlow, scikit-learn, Pandas, Numpy, SciPy, Flask, Django, Apache Spark		
	Research Tools and Methodologies		
	Jupyter Notebooks, Git/GitHub, Docker, AWS, IBM Cloud, LaTeX		
Languages	Korean Native	English Fluent	Japanese Fluent
References	Julia Stoyanovich		
	Associate Professor, Computer Science and Engineering		
	https://stoyanovich.org		
	New York University 370 Jay Street Brooklyn, NY 11201 stoyanovich@nyu.edu		
	Martin Lotz		
	Associate Professor, Mathematics		
	https://homepages.warwick.ac.uk/staff/Martin.Lotz		
	Zeeman Building Coventry CV4 7AL United Kingdom martin.lotz@warwick.ac.uk		
	Rizos Sakellariou		
	Professor, Department of Computer Science		
	http://www.cs.man.ac.uk/~rizos/		
	University of Manchester Oxford Road Manchester M13 9PL U.K. rizos@manchester.ac.uk		