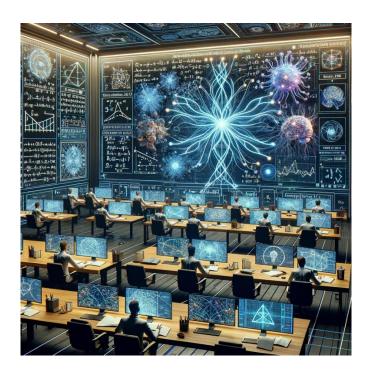# Statistics in the Age of AI

May 9-11, 2024
The George Washington University

Sponsored By:

Thursday, May 9, 2024

Rome Hall 351
801 22nd St NW
Washington, DC 20052

8:00 am-9:15 am	**Short Course Registration**
9:15 am-12:00 pm	**Short Course I:** *Conformal Prediction*
	Jing Lei, Carnegie Mellon University
12:00 pm-1:30 pm	**Break**
1:30 pm-4:15 pm	**Short Course II:** *Causal Inference*
	Peng Ding, University of California, Berkeley

Friday, May 10, 2024

Jack Morton Auditorium
Media and Public Affairs Building
805 21st St NW
Washington, DC 20052

8:00 am-9:00 am	**Conference Registration**
9:00 am-9:10 am	**Welcome and Opening Remarks**

- *Xiaoke Zhang, George Washington University*
- *Pamela Norris, George Washington University*

9:10 am-10:25 am   **Session I.** Chair: *Paromita Dubey, University of Southern California*

- *David Donoho, Stanford University, "Widespread Panic over Model Collapse."*
- *Jianqing Fan, Princeton University, "Causal Learning from Heterogeneous Environments."*
- *Richard Samworth, University of Cambridge, "Optimal Convex M-Estimation via Score Matching."*

10:25 am-10:45 am   **Coffee Break, Art Gallery Lobby (2nd floor)**

10:45 am-12:00 pm   **Session II.** Chair: *Raymond Wong, Texas A&M University*

- *Steve Marron, University of North Carolina, Chapel Hill, "Object Oriented Data Analysis."*
- *Hongyu Zhao, Yale University, "Interpretable and Spatially-Aware Integration of Spatial Transcriptomics Datasets."*
- *Naisyin Wang, University of Michigan, "Utilizing Synthetic Components to Balance Privacy Protection and Data Utility."*

12:00 pm-1:30 pm   **Lunch: District House B114 (2121 H St NW, B1 level)**

1:30 pm-2:45 pm   **Session III.** Chair: *Joan Jian-Jian Ren, University of Maryland, College Park*

- *Qiwei Yao, London School of Economics and Political Science, "Autoregressive Networks with Dependent Edges."*
- *Xiaotong Shen, University of Minnesota, "Boosting Data Analytics with Synthetic Volume Expansion."*
- *Grace Yi, University of Western Ontario, "Label Correction of Crowdsourced Noisy Annotations with an Instance-Dependent Noise Transition Model."*

2:45 pm-3:05 pm   **Coffee Break, Art Gallery Lobby (2nd floor)**

3:05 pm-4:20 pm   **Session IV.** Chair: *Feifang Hu, George Washington University*

- *Lan Wang, University of Miami, "Doubly Robust Uncertainty Quantification for Quantile Treatment Effects in Sequential Decision Making."*
- *Mei-Ling Ting Lee, University of Maryland, College Park, "First-Hitting-Time Threshold Regression and Neural Network."*
- *Byeong U. Park, Seoul National University, "Nonparametric Causal Additive Models with Smooth Backfitting."*

4:20 pm-5:20 pm   **Panel I: Statistical Research in the Age of AI.**
                Moderator: *Damla Senturk, University of California, Los Angeles*

- *Haoda Fu, Eli Lily and Company*
- *Jonas Mueller, Cleanlab*
- *Annie Qu, University of California, Irvine*
- *Huixia Judy Wang, George Washington University*
- *Xiao Wang, Purdue University*

Saturday, May 11, 2024

Jack Morton Auditorium
Media and Public Affairs Building
805 21st St NW
Washington, DC 20052

8:00 am-9:00 am   **Conference Registration**

9:00 am-10:15 am   **Session V.** Chair: *Yaqing Chen, Rutgers University*

- *Iain Johnstone, Stanford University, "Expectation Propagation and Maximum Likelihood in Generalized Linear Mixed Models."*
- *Xihong Lin, Harvard University, "Build an End-to-End Scalable Data Science Ecosystem Using Statistics, ML, and AI for Whole Genome Sequencing Analysis."*
- *Runze Li, Penn State University, University Park, "How Can Statistics Help Students Prepare for Better-Paid Jobs? "*

10:15 am-10:35 am   **Coffee Break, Art Gallery Lobby (2nd floor)**

10:35 am-11:50 am   **Session VI.** Chair: *Hua Liang, George Washington University*

- *Tailen Hsing, University of Michigan, "An RKHS Approach for Variable Selection in High-Dimensional Functional Linear Models."*
- *Victor Panaretos, EPFL, "Graphical Models in Infinite Dimensions."*
- *Yehua Li, University of California, Riverside, "Machine Learning with Functional Predictors and Applications to Crop Yield Prediction."*

12:00 pm-1:00 pm   **Lunch: Continental Ballroom (University Student Center, 800 21st St NW, 3rd Floor)**

1:00 pm- 2:00 pm   **Poster Session: Student Center 309 ((University Student Center, 800 21st St NW, 3rd Floor)**

- *Satarupa Bhattacharjee, Penn State University, University Park, "Causal Inference for Non-Euclidean Object Data with Continuous Treatments."*
- *Na Bo, University of Pittsburgh, "Pseudo-Outcome-Based Meta-Learners for Analyzing Treatment Heterogeneity in Survival Data: Application to Pediatric Asthma Care Under COVID-19 Disruption."*
- *Yasin Khadem Charvadeh, Memorial Sloan Kettering Cancer Center, "Bridging Data Gaps in Healthcare: Effective Healthcare Data Analytics with Variational Autoencoder Imputation."*
- *Yan-Yu Chen, University of California, Davis, "Testing the Causal Effects of Continuous Treatments Learned by Deep Neural Networks."*
- *Yiqun Chen, Stanford University, "GenePT: A Simple but Effective Model for Genes and Cells Built from ChatGPT."*
- *Hyo Young Choi, University of Tennessee Health Science Center, "ArtistR: A Novel Framework for Systematically Detecting Alternative Transcript Initiation by Integrating ATAC-seq and RNA-seq."*
- *Yihong Gu, Princeton University, "Environment Invariant Linear Least Squares."*
- *Kentaro Hoffman, University of Washington, "Some Models Are Useful, but for How Long?"*
- *Ju-Sheng Hong, University of California, Davis, "SAND: Smooth Imputation of Sparse and Noisy Functional Data with Transformer Networks."*

- *Su I Iao, University of California, Davis, "Deep Frechet Regression."*
- *Wei Jin, Johns Hopkins University, "A Bayesian Decision Framework for Optimizing Sequential Combination Antiretroviral Therapy in People with HIV."*
- *Gauri Kamat, Brown University, "Bayesian Record Linkage with Variables in One File."*
- *Annela Kelly, Regis College, "Statistical Analysis of Trends of Student Mindsets Amidst the Pandemic."*
- *Salil Koner, Duke University, "A Conformal Prediction Framework for Functional Data Suitable for Both Sparse and Dense Designs."*
- *Poorbita Kundu, University of California, Davis, "Decomposition-Based Intrinsic Modeling of Shape-Constrained Functional Data."*
- *Tsz-Kit Lau, University of Chicago, "AdAdaGrad: Adaptive Batch Size Schemes for Adaptive Gradient Methods."*
- *Yonghoon Lee, University of Pennsylvania, "Simultaneous Conformal Prediction of Missing Outcomes with Propensity Score Discretization."*
- *Jiayi Li, University of California, Los Angeles, "Geometry of Polynomial Neural Networks."*
- *Yicheng Li, Tsinghua University, "Generalization Error Curves for Analytic Spectral Algorithms under Power-law Decay."*
- *Anusha Natarajan, Columbia University, "Utilizing Supervised Machine Learning to Comprehending Popularity of Digital Content over Public Discourse on National and Global Events."*
- *Seongoh Park, Sungshin Women's University, South Korea, "High-dimensional Missing Data Imputation Via Undirected Graphical Model."*
- *Grace Ringlein, Johns Hopkins University, "Impacts of Predictive Uncertainty in Healthcare Prior Authorizationpr."*
- *Wookyeong Song, University of California, Davis, "Inference for Dispersion of Random Objects and Curvature of Underlying Metric Spaces."*
- *Bohan Wu, Columbia University, "Extending Mean-Field Variational Inference via Entropic Regularization: Theory and Computation."*
- *Tiantian Yang, University of Idaho, "An Interpretable Graph Neural Network for Disease Classification in Omics Data Analysis."*
- *Dapeng Yao, Johns Hopkins University, "A Deep Reinforcement Learning Algorithm for Optimizing Antiretroviral Therapy in People with HIV."*
- *Xianli Zeng, University of Pennsylvania, "Bayes-Optimal Fair Classification with Linear Disparity Constraints via Pre-, In-, and Post-processing."*
- *Yi Zhang, George Washington University, "Independence-Encouraging Subsampling for Nonparametric Additive Models."*
- *Qixian Zhong, Xiamen University, "Dynamic Modeling for Multivariate Functional and Longitudinal Data."*
- *Yidong Zhou, University of California, Davis, "Dynamic Modeling of Sparse Longitudinal Data and Functional Snippets with Stochastic Differential Equations."*
- *Hang Zhou, University of California, Davis, "Statistical Inference for Object Data."*
- *Tian-Yi Zhou, Georgia Institute of Technology, "Classification of Unbounded Data Generated by Gaussian Mixture Models via Deep ReLU Networks."*
- *Yingchao Zhou, Iowa State University, "Locally Adaptive Conformal Prediction for Regression."*
- *Changbo Zhu, University of Notre Dame, "Testing Independence for Sparse Longitudinal Data."*

2:10 pm-3:25 pm   **Session VII.** Chair: *Xiucai Ding, University of California, Davis*

- *Haiyan Huang, University of California, Berkeley, "Integrative Deep Multi-Learning for Biclustering and Predicting Cancer Drug Responses: Leveraging Omics and Drug Molecular Data."*
- *Lily Wang, George Mason University, "Functional Regression through Distributed Learning: An Application to Brain Imaging Studies."*
- *Johannes Schmidt-Hieber, University of Twente, "A Statistical Analysis of an Image Classification Problem."*

3:25 pm-3:45 pm   **Coffee Break, Art Gallery Lobby (2nd floor)**

3:45 pm-4:45 pm   **Panel II: Statistical Education and Practice in the Age of AI.**
                        Moderator: *Fang Yao, Peking University*

- *John Aston, University of Cambridge*
- *Mina Karzand, University of California, Davis*
- *Xuming He, Washington University at St Louis*
- *Tian Zheng, Columbia University*
- *Hongtu Zhu, University of North Carolina, Chapel Hill*

4:45 pm-5:00 pm   **Closing Remarks**
                        Huixia Judy Wang, George Washington University