AnnaElaine (Anna) Rosengart **Updated July 21, 2025**

CONTACT INFORMATION
Department of Statistics and Data Science
Carnegie Mellon University
Pittsburgh, PA USA
website: aerosengart.github.io
email: arosenga@andrew.cmu.edu

SUMMARY

I am a Ph.D. student in the Department of Statistics & Data Science at Carnegie Mellon University. I primarily do work related to analysis of biomedical data. Currently, I am focused on statistical methods for wastewater-based epidemiology.

EDUCATION

Current	**Ph.D. Statistics**, Carnegie Mellon University, Pittsburgh, PA USA *Advised by F. William Townes*
2024	**M.S. Statistics**, Carnegie Mellon University, Pittsburgh, PA, USA
2021	**B.S. Data Science**, University of Michigan, Ann Arbor, MI USA
2021	**B.F.A. Jazz Studies**, University of Michigan, Ann Arbor, MI USA

AWARDS

Phi Beta Kappa. Inducted 2021.

RESEARCH

Preprints

[1] Omelchenko, A. A., Siwek, J. C., Chhibbar, P., Arshad, S., Nazarali, I., Nazarali, K., **Rosengart, A.**, Rahimikollu, J., Tilstra, J., Shlomchik, M. J., Koes, D. R., Joglekar, A. V., & Das, J. (2024). Sliding Window INteraction Grammar (SWING): a generalized interaction language model for peptide and protein interactions. *bioRxiv : the preprint server for biology*, doi:10.1101/2024.05.01.592062.

Peer-Reviewed Journal Articles

[1] **Rosengart AL**, Bidwell AL, Wolfe MK, Boehm AB, Townes FW. Spatiotemporal Variability of the Pepper Mild Mottle Virus Biomarker in Wastewater. *ACS ES&T Water* (2024). doi:10.1038/s41592-024-02175-z.

[2] Rahimikollu J, Xiao H, **Rosengart A**, Rosen ABI, Tabib T, Zdinak PM, He K, Bing X, Bunea F, Wegkamp M, Poholek A, Joglekar A, Lafyatis R, Das J. SLIDE: Significant Latent Factor Interaction Discovery and Exploration across biological domains. *Nature Methods* (2024). doi:10.1021/acsestwater.4c00866.

[3] Wheeler J, **Rosengart A**, Jiang Z, Tan K, Treutle N, Ionides EL. Informing policy via dynamic models: Cholera in Haiti. *PLOS Computational Biology* (2024). doi:10.1371/journal.pcbi.1012032.

[4] Li SR, Moheimani H, Herzig B, Kail M, Krishnamoorthi N, Wu J, Abdelhamid S, Scioscia J, Sung E, **Rosengart A**, Bonaroti J, Johansson PI, Stensballe J, Neal MD, Das J, Kar U, Sperry J, Billiar TR. High-dimensional proteomics identifies organ injury patterns associated with outcomes in human trauma. *Journal of Trauma and Acute Care Surgery* (2023). doi:10.1097/ta.0000000000003880.

[5] Abdelhamid SS, Scioscia J, Vodovotz Y, Wu J, **Rosengart A**, Sung E, Rahman S, Voinchet R, Bonaroti J, Li S, Darby JL, Kar U, Neal MD, Sperry J, Das J, Billiar TR. Multi-Omic Admission-Based Prognostic Biomarkers Identified by Machine Learning Algorithms Predict Patient Recovery and 30-Day Survival in Trauma Patients. *Metabolites* (2022). doi:10.3390/metabo12090774.

Class Projects and Other Unpublished Work

[1] Shen A, Rosengart A. Can ChatGPT Predict the Weather? A Study of the Universality of the Self-Attention Mechanism. Carnegie Mellon 10-716 (Advanced Machine Learning). Spring 2024.

[2] Rosengart A, Scharfstein K, Thivierge G. Variational and Proximal Causal Effect Estimation in the Presence of Unmeasured Confounding. Carnegie Mellon University 10-708 (Probabilistic Graphical Models). Winter 2023.

Talks

Invited

[1] Significant Latent Factor Interaction Discovery & Exploration. CSI Retreat, Center for Systems Immunology, University of Pittsburgh. October 2022.

Formal Presentations

[1] Wastewater-based Epidemiology. CDC Site Visit, Delphi Research Group, Carnegie Mellon University. June 2023.

[2] Characterizing Variability in Pepper Mild Mottle Virus for Wastewater Epidemiology. CDC Center for Forecasting and Analytics Site Visit, Delphi Research Group, Carnegie Mellon University. June 2024.

Work Experience

2025 Summer — **AI Validation Fellow**
Handshake Model Validation Expert Fellowship
Developed prompts for large language model evaluation.

2023 Summer — **Laboratory Researcher**
University of Pittsburgh, Center for Systems Immunology
Applied large language modeling tools to prediction tasks in protein binding. Led code composition and documentation for construction, training, and use of protein binding Transformer.

2022 Spring-Summer — **Laboratory Researcher**
University of Pittsburgh, Center for Systems Immunology
Developed methodology for identification of putative mechanisms of disease from multi-omic biomedical data. Composed and documented code.

2021 Summer — **Trauma ATLAS Intern**
Pittsburgh Trauma Research Center, University of Pittsburgh
Analyzed scale, multi-omics data from experimental trials in trauma patient treatments. Composed sample code scripts for data cleaning, visualization, and gene set enrichment analysis to fellow researchers. Implemented consensus-based feature selection for dimension reduction.

2021 Spring — **Undergraduate Research Trainee**
Department of Statistics, University of Michigan
Trained in performing time series analysis and statistical modeling of stochastic processes for epidemiological study using partially observed Markov processes.

2020 Fall — **Laboratory Researcher**
Center for Biologic Imaging, University of Pittsburgh
Worked on the development of efficient and accurate computational methodologies for analysis of large-scale experimental biomedical imaging data. Adapted preexisting software and the development of a software pipeline for cleaning, manipulating, and mapping terabyte-scale mouse brain imaging slices to the Allen Mouse Brain Atlas.

Teaching Experience

2022 Fall	**36-401 Modern Regression** *Teaching Assistant* Undergraduate course on applied data analysis.
2023 Spring	**36-615 Software for Large-Scale Data** *Teaching Assistant* Graduate course on distributed data systems and analysis.
2023 Spring	**36-616 Computational Methods for Statistics** *Teaching Assistant* Graduate course on computing tools for data analysis.
2023 Fall, 2024 Fall	**36-613 Data Visualization** *Teaching Assistant* Graduate course on methods and practices for data visualization.
2023 Fall, 2024 Fall	**36-614 Data Engineering and Distributed Environments** *Teaching Assistant* Graduate course on computing tools for data science.
2024 Spring, 2025 Spring	**36-315 Statistical Graphics and Visualization** *Teaching Assistant* Undergraduate course on methods and practices for data visualization.

Technical Skills

R, Python, LaTeX, Shell