Christine Herlihy

: crherlihy.github.io
: christine.herlihy@gmail.com
: christineherlihy
: github.com/crherlihy
: crherlihy

Education

University of Maryland, College Park — College Park, MD
Ph.D. candidate in Computer Science — 2019–present

- *Coursework*: Algorithms in ML: Guarantees and Analyses, Applied Mechanism Design, Computational Linguistics I & II, Common-sense Reasoning and Natural Language Understanding, Program Analysis.
- *Advised by*: John Dickerson | *Research areas*: sequential and combinatorial decision-making under uncertainty (multi-armed bandits, RL); algorithmic fairness; graph-based knowledge representation/reasoning; healthcare.

Georgia Institute of Technology — Atlanta, GA
M.S. in Computer Science | Concentration: Machine Learning — 2017–2018

- *Coursework*: Machine Learning, Artificial Intelligence, Reinforcement Learning, ML for Trading, HPC, Big Data Analytics for Healthcare, Data Analytics using Deep Learning, Info. Security, Computability and Algorithms

Georgetown University — Washington, DC
B.A. in Government; minor in Spanish — 2007–2011

- *Coursework*: Political economy, institution/mechanism design, quantitative and qualitative methods, linguistics.
- Received academic honors during every semester of attendance; studied abroad in Santiago, Chile.

Experience

Microsoft Research — Remote
Summer Researcher, Augmented Learning and Reasoning group — 06/2023—present

- *Advised by*: Adith Swaminathan and Jennifer Neville | Algorithm development for complex recommender ecosystems.

Google Research — Remote
Student Researcher, Responsible AI team — 09/2022–02/2023

- *Advised by*: Emmanuel Klu | causal modeling ∪ algorithmic fairness in dynamic environments.

Microsoft Research — Remote
Summer Researcher, FATE group — 06/2022–09/2022

- *Advised by*: Miro Dudík and Alexandra Chouldechova | sample-efficient disaggregated model evaluation.

Amazon Robotics — Remote
Summer Research Scientist — 05/2021 - 08/2021

- *Advised by:* Victor Amelkin and Alex Barbosa | graph-based deep reinforcement learning algorithm development.

Georgia Tech Research Institute — Atlanta, GA
Research Scientist II, High-Performance Computing & Data Analytics Branch — 01/2017 - 08/2019

- Technical task lead and core contributor for a range of research projects and proposals, including: patient-level predictive modeling; computational phenotyping; application of unsupervised learning and NLP techniques on unstructured text to develop machine phenotypes and detect spatially/temporally co-occurring machine failures; development of models to predict geopolitical conflict and detect misinformation.

Econometrica, Inc. — Bethesda, MD
Research Associate II, Health Data Analytics — 08/2015 - 01/2017

- Patient and population-level predictive modeling; healthcare policy evaluation using Python, R, and Stata.
- Primary project was an impact evaluation contract that involved using various econometric methodologies, a genetic matching algorithm, and survey design schemes to detect changes in medical outcomes and unintended consequences associated with bundled payment and gainsharing mechanisms in health care systems.

Frontier Strategy Group Washington, DC
Macroeconomic Research Analyst, Latin America 08/2012 - 08/2014

– Conducted econometric research and built forecast models using Python and R to advise C-suite executives at over 200 multinationals on resource allocation and risk management in Latin America.

Technical Skills

- **Proficient:** Python, R, Julia, SQL, LaTeX | **Familiar:** Java, C++, Scala, Spark, bash, Coq
- **Libraries & tools:** scikit-learn, pandas, NumPy, SciPy, OpenAI Gym, PyTorch, TensorFlow, spaCy/scispacy, nltk, gensim, textacy, AllenNLP, NetworkX, Deep Graph Library (DGL), Neo4j, ray, git, Gurobi, Postgres, Tableau
- **Languages:** English (Native), Spanish (Fluent), Portuguese (Intermediate), Farsi (Beginner)

Conference Papers

[1] **C. Herlihy** and J. Dickerson, "Networked Restless Bandits with Positive Externalities", in *AAAI Conference on Artificial Intelligence*, 2023. [Online]. Available: https://arxiv.org/abs/2212.05144.

[2] **C. Herlihy**, A. Prins, A. Srinivasan, and J. Dickerson, "Planning to Fairly Allocate: Probabilistic Fairness in the Restless Bandit Setting", in *ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, (to appear), 2023. [Online]. Available: https://arxiv.org/abs/2106.07677.

[3] **C. Herlihy** and R. Rudinger, "MedNLI Is Not Immune: Natural Language Inference Artifacts in the Clinical Domain", in *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, Online: Association for Computational Linguistics, Aug. 2021, pp. 1020–1027. [Online]. Available: https://aclanthology.org/2021.acl-short.129.

[4] M. Halter, **C. Herlihy**, and J. P. Fairbanks, "A Compositional Framework for Scientific Model Augmentation", in *Proceedings Applied Category Theory 2019, ACT 2019, University of Oxford, UK, 15-19 July 2019*, J. Baez and B. Coecke, Eds., ser. EPTCS, vol. 323, 2019, pp. 172–182. [Online]. Available: https://doi.org/10.4204/EPTCS.323.12.

Workshop Papers

[1] **C. Herlihy**, P. Goel, and J. Dickerson, *Networked Restless Bandits with Positive Externalities*, Disinformation Countermeasures and Machine Learning Workshop, ICML, 2022.

[2] **C. Herlihy**, A. Prins, A. Srinivasan, and J. Dickerson, *Planning to Fairly Allocate: Probabilistic Fairness in the Restless Bandit Setting*, Responsible Decision Making in Dynamic Environments Workshop, ICML, 2022.

[3] A. Prins, **C. Herlihy**, and J. Dickerson, *What Should I Grow Today so I Make Money Tomorrow? Supporting Small Farmers' Crop Planning with Social, Environmental, and Market Data*, Practical ML for Developing Countries Workshop, ICLR, 2022.

[4] **C. Herlihy**, S. Huang, M. Diep, N. Johnson, N. Sehgal, J. Dickerson, D. Jackson, and C. Baur, *An mHealth Intervention for African American and Hispanic Adults: Preliminary Field Test Results about User-reported Article Relevance*, Workshop on Machine Learning in Public Health, NeurIPS, 2021.

Additional Publications

[1] A. Moreland, **C. Herlihy**, M. Tynan, G. Sunshine, R. McCord, C. Hilton, J. Poovey, A. Werner, C. Jones, E. Fulmer, A. Gundlapalli, H. Strosnider, A. Potvien, M. García, S. Honeycutt, and G. Baldwin, "Timing of State and Territorial COVID-19 Stay-at-Home Orders and Changes in Population Movement - United States, March 1-May 31, 2020", *Morbidity and Mortality Weekly Report*, vol. 69, no. 35, Sep. 2020.

[2] D. Ruiz, D. Stout, and **C. Herlihy**, "Use of Genetic Matching in Program Evaluation: The Case of RAD", *Cityscape*, vol. 19, no. 2, pp. 337–350, 2017, ISSN: 1936007X. [Online]. Available: http://www.jstor.org/stable/26328344.

Select Projects

- Constrained Resource Allocation in the Restless Bandit Setting | UMD (Python, 2020-present)
  Developing algorithms to provide fairness guarantees and exploit spillover effects/structural relationships among arms.
- A Compositional Approach to Representing and Manipulating Scientific Models | GTRI (Julia, 2018-2019)
  Developed program analysis and category-theoretic techniques for epi model extraction and composition from text and code.
- ClarityNLP: An open-source NLP framework for clinical phenotyping | GTRI & Celgene (Python, 2018-2019)
  Developed supervised and unsupervised methods for clinical text mining, feature engineering, and cohort identification.
- NLP Pipeline to Detect Cascading and Co-occurring Machine Failures | GTRI (Python & Neo4j, 2018-2019)
  Statistical analysis and topic modeling of work order notes; constructed KG to identify spatial/temporal failure patterns.
- Chest X-ray Disease Diagnosis with Deep Convolutional Neural Networks | Georgia Tech (PyTorch, 2018)
  Used CNNs to detect and localize the 14 thoracic pathologies present in the NIH Chest X-ray dataset.

Teaching

- **Co-Instructor** at Georgia Tech Professional Education — 2018
  Machine Learning Short Course
- **Teaching Assistant** at Johns Hopkins University — 2017
  Algorithms and Data Structures
- **Teaching Assistant** at University of Maryland, College Park — 2014–2015
  Political Theory; International Relations
- **Volunteer Teacher** at Latino Student Fund — 2012–2014
  English as a Second Language (ESL) for Adult Learners

Service

- **Program Committee Member,** EAAMO 2023 — 2023
- **Program Committee Member,** AI for Social Good workshop (held at AAAI 2023) — 2023
- **Graduate Student Peer Mentor,** University of Maryland CS Department — 2022–2023
- **Co-organizer**, Pasteur's Quadrant AI for Social Good Seminar Series — 2021–2022
- **Reviewer,** AMIA (2020); NeurIPS (2022–pres.) — 2020–pres.
- **Student Advisory Board Member,** Iribe Initiative for Inclusion & Diversity in Computing — 2019–2021

Personal Interests

- **Hiking & running**: I'm partial to half-marathons, but currently working on improving speed.
- **Learning languages**: Working on Farsi, American Sign Language (ASL), and Haskell.
- **Training my puppy**: Reward shaping for fun and ~~*profit*~~ *slightly* less chaos— every day is a new (mis)adventure.

Scholarships and Awards

- **Grace Hopper Conference Scholarship Recipient** | University of Maryland, College Park — 2021
- **Professional Research Experience Program (PREP) Researcher** | NIST — 2019–2021
- **Dean's Fellowship (CS)** | University of Maryland, College Park — 2019–2020
- **HIVE $25K Research Grant Winner** | Georgia Tech Research Institute — 2017
- **FIA-Deutsch $25K Seed Grant Fellow** | University of Maryland, College Park — 2014–2015
- **Dean's Fellowship (GVPT)** | University of Maryland, College Park — 2014–2015
- **Merit-based scholarship recipient; academic honors** | Georgetown University — 2007–2011
- **Gilman International Scholarship Recipient** | United States Department of State — 2010