Christine R. **Herlihy**

Researcher · Data Scientist · Developer

christine.herlihy@gmail.com | crherlihy | christineherlihy | @crherlihy | U.S. Citizen

Objective

Computer science researcher with a focus on natural language processing, machine learning, and graphical modeling for the purpose of: (1) multi-modal information extraction and reconciliation; and (2) the development of explainable, hierarchical predictive models, in which micro-level behaviors influence, and are influenced by, macro-level trends. I am particularly interested in how computational methods and distributed systems can be used to analyze and solve coordination problems and market failures, and/or identify optimal corrective interventions.

Education

Georgia Institute of Technology — *Atlanta, GA*

M.S. in Computer Science | Concentration: Machine Learning | GPA: 3.6/4.0 — *Jan. 2017 - Dec. 2018*

- Courses: Machine Learning | Artificial Intelligence | Reinforcement Learning | Machine Learning for Trading | High-Performance Computing Big Data Analytics for Healthcare | Video Game Design | Data Analytics using Deep Learning | Information Security | Computability and Algorithms.

Johns Hopkins University — *Baltimore, MD*

Graduate Coursework in Computer Science | GPA: 3.67/4.0 — *May 2015 - Jan. 2017*

- Completed pre-requisite and grad-level coursework in CS prior to transferring to GT; TA for Data Structures course.
- Courses: Programming Using Java | Data Structures | Computer Organization | Software Engineering | Algorithms | Computer Architecture.

University of Maryland, College Park — *College Park, MD*

Graduate Coursework in International Political Economy | GPA: 3.9/4.0 — *Aug. 2014 - May 2015*

- Completed graduate-level courses in quantitative methods and political economy before switching fields to CS.

Georgetown University — *Washington, DC*

B.A. in Government; Minor in Spanish | GPA: 3.76/4.0 — *Aug. 2007 - May 2011*

- Received academic honors during every semester of attendance; studied abroad in Santiago, Chile; graduated *magna cum laude*.

Work Experience

Georgia Tech Research Institute — *Atlanta, GA*

Research Scientist, High Performance Computing and Data Analytics Branch — *Jan. 2017 - PRESENT*

- Technical task lead and/or core contributor for a range of applied 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 fake news.

Econometrica, Inc. — *Bethesda, MD*

Research Associate II & Project Manager, Data Analytics — *Aug. 2015 - Jan. 2017*

- Econometric analysis, data visualization, healthcare policy evaluation, and research process automation using Python, R, Stata, and LaTeX.
- Primary project was an impact evaluation contract that involved using various econometric methodologies, a genetic matching algorithm, and survey design schemes to detect and determine changes in quality of care, medical outcomes, Medicare costs, and unintended consequences associated with emerging bundled payment and gainsharing mechanisms in health care systems.

University of Maryland, College Park — *College Park, MD*

Graduate Research & Teaching Assistant — *Aug. 2014 - Aug. 2015*

- RA for Dr. Joel Simmons; developed a data extraction pipeline; assisted with feature engineering and econometric analysis in Python and R.

U.S. Department of the Treasury, Office of Financial Research — *Washington, DC*

Graduate Research Intern; assisted Dr. Benjamin Kay and Dr. Ayeh Bandeh-Ahmadi — *Jan. 2015 - Jul. 2015*

- Used SQL and Python to conduct quantitative analysis related to executive compensation and financial risk.
- Explored methodological challenges associated with parsing Twitter data to predict/detect systemic risk across a range of asset classes.

Frontier Strategy Group — *Washington, DC*

Management Consultant & Research Analyst, Latin America — *Aug. 2012 - Aug. 2014*

- Conducted econometric research to advise C-suite executives at over 200 multinationals on resource allocation and risk management in LATAM.
- Built forecast models for key macroeconomic indicators; helped drive research process automation and data analysis efforts in R and Python.

Technical Skills

Programming	**Proficient:** Python, R, Java, SQL, LaTeX | **Familiar:** Julia, C/C++, Scala, Spark, bash, C#, Unity, Ethereum, MPI, Assembly (MIPS)
ML | DL | NLP | etc.	scikit-learn, pandas, numpy, OpenAI Gym, PyTorch, spaCy, gensim, textacy, Neo4j, git, Docker, Postgres, FHIR, OMOP, Tableau
Domain Interests	Healthcare, Cybersecurity, Finance/Econometrics, Behavioral Economics, Smart Contracts, Quantitative Social Science
Languages	English (Native), Spanish (Fluent), Portuguese (Intermediate), Farsi (Beginner)

Select Projects

ClarityNLP: An NLP framework for clinical phenotyping
Georgia Tech Research Institute

Programmer | Integration of Machine Learning Algorithms to Platform || Sponsor: Celgene — *Jul. 2018 - PRESENT*

- Integrated supervised and unsupervised learning techniques into the ClarityNLP platform to aid clinical researchers in tasks associated with computational phenotyping, including mining of unstructured clinical notes for the purpose of feature engineering and/or cohort identification.

NLP Pipeline to Detect Cascading and/or Co-occurring Machine Failures
Georgia Tech Research Institute

NLP Task Lead | Programmer | Report Co-Author || Sponsor: NCR — *Apr. 2018 - PRESENT*

- Developed a Python-based NLP pipeline to ingest, process, and analyze a large corpus of unstructured work-order notes | Implemented topic modeling and analyzed variation in corpus composition over time in an effort to detect machine failures that co-occur within space and/or time.
- Constructed a semantic graph to allow stakeholders to reason over and query for problem-solution tuples that are similar in the vector space.

Blockchain Platform to Facilitate Patient-Directed Provider-to-Provider Record Sharing
Georgia Tech Research Institute

Principal Investigator | Programmer | Report Author | HIVE $25K Seed Grant Winner || Sponsor: GTRI — *July 2017 - June 2018*

- Developed a Python and Ethereum-based prototype to illustrate how synthetically generated patients, healthcare providers, and patient-designated stakeholders might interact with an escrow smart contract to conduct patient-directed record sharing transactions.

CSE 6250: Chest X-ray Disease Diagnosis with Deep Convolutional Neural Networks
Georgia Institute of Technology

Programmer | Predictive Model Development | Report Co-Author || Graduate Coursework — *Jan. 2018 - May 2018*

- Developed a PyTorch-based pipeline to use CNNs to detect and localize the 14 thoracic pathologies present in the NIH Chest X-ray dataset.

Patient Level Prediction Modeling
Georgia Tech Research Institute

Programmer | Predictive Model Development | Report Co-Author || Sponsor: GTRI — *Jan. 2017 - Oct. 2017*

- Worked with a team of clinical and computer science researchers to develop a Python-based patient-level prediction model for sepsis using Lasso regression and random forest | Contributed to a web-based clinical decision support tool to help hospitals determine which ICU patients are most at-risk | Presented results during poster session of the 2017 OHDSI Symposium.

Predicting Geopolitical Conflict and Cooperation Using GDELT Data
Georgia Tech Research Institute

Programmer | Predictive Model Development | Report Co-Author || Sponsor: GTRI — *Jan. 2017 - Apr. 2017*

- Developed a set of models (GLM; k-NN) to predict dyad-level geopolitical conflict and cooperation at the month-dyad level using the Global Database of Events, Language, and Tone (GDELT) dataset and CAMEO codes | Presented results at UT Austin's Conflict Conference.

Development of a Mobile App to Gamify Museum Visits Using Location-Aware Beacons
UMD, College Park | The Newseum

Team Lead | UMD FIA-Deutsch Seed Grant Fellow || Sponsor: UMD Future of Information Alliance — *Aug. 2014 - May 2015*

- Delivered successful pitch presentation resulting in a $25K seed grant | Managed a team of engineers to develop a prototype iOS app with a Parse back-end that communicated with Estimote beacons to deliver scavenger-hunt style, location-specific content to Newseum visitors.

Teaching & Mentorship

2018	**Co-Instructor: Machine Learning Short Course**, Georgia Tech Professional Education	*Atlanta, GA*
2016	**Teaching Assistant: Data Structures**, Johns Hopkins University	*Baltimore, MD*
2014-2015	**Teaching Assistant: Political Theory; International Relations**, University of Maryland, College Park	*College Park, MD*
2012-2014	**Volunteer English as a Second Language (ESL) Teacher**, Latino Student Fund	*Washington, DC*

Presentations & Posters

Observational Health Data Sciences and Informatics (OHDSI) Symposium
Bethesda, MD

- Presented poster titled *Implementing Real-Time Patient Level Predictions Using PLP Models* — *October 18-20, 2017*

The Conflict Conference
Austin, TX

- Presented work titled *Using Big Data to Predict and Analyze Cooperation and Conflict* — *April 7-9 2017*

Honors & Awards

2017-2018	**GTRI HIVE $25K Research Grant Winner**, Georgia Tech Research Institute Python, public-key cryptography, & Ethereum-based platform to facilitate patient-directed record sharing	*Atlanta, GA*
2014-2015	**UMD FIA-Deutsch $25K Seed Grant Fellow**, FIA-Deutsch Seed Grant Competition Flip the Museum: A Platform to Extend the Audience Engagement Life Cycle Through Gamification of Content	*College Park, MD*
2014-2015	**Dean's Fellowship; academic honors**, University of Maryland, College Park	*College Park, MD*
2007-2011	**Merit-based scholarship recipient; academic honors**, Georgetown University	*Washington, DC*
2010	**Gilman International Scholarship Recipient**, United States Department of State	*Washington, DC*