Derek M. Walkama

Boston, MA
(508)-273-3496
Derek.M.Walkama@gmail.com

Data Scientist
web: sites.tufts.edu/derekwalkama
LinkedIn: linkedin.com/in/derekwalkama

Interests and Programming Languages

Data scientist currently working with causal ML in the target discovery space who is deeply interested in ML/AI, bioinformatics, causal and statistical modeling, and image analysis/processing. Have deep experience processing and modelling multi-omic data including: genomics (WGS), transcriptomics (NGS and single-cell), and EHR data.

Languages:

R iGraph, xts, gmlnet, caret, foreach, doParralel, tidyverse, seurat

Python numpy, scipy, pandas, tensorflow

Matlab image processing toolbox, curve fitting toolbox, parallel processing toolbox

unix StarCluster, ParallelCluster, AWS, BASH

Education

Ph.D. Physics. Tufts University, 2021
Thesis: *Elastic instability of flows and structures at low Reynolds number*
Advisor: Professor Jeffrey S. Guasto

Sc.M. Physics. Tufts University, 2017
Graduated Magna Cum Laude

B.A. Physics & Mathematics Clark University, 2015
Graduated Cum Laude with Departmental Honors

Professional Experience

Senior Data Scientist, Early Development Feb 2022 – present
Aitia, Somerville, MA

- Oversaw the development of large scale, multi-modal causal models in multiple disease areas such as cancer, neurodegenerative, and cardiac. This includes the development of disease-specific multi-omic data pipelines.
- Led the design and development of a flagship model building pipeline that integrates an open-source data science tool, DVC (data version control), and proprietary causal AI software, REFS.
- Provided HPC support for platform migration from StarCluster to AWS ParallelCluster.

Data Scientist, Early Development Feb 2021 – Feb 2022
Aitia, Somerville, MA

- Constructed Bayesian network models using proprietary Causal AI technology (REFS) to exploit the recent explosive growth of multi-omics and clinical data in order to create "virtual" (in silico) patients in oncology, neurodegeneration, and immunology.
- Disease specific data-sets with tens of thousands of features are curated and pre-processed for model building.
- Processed and modelled novel data such as "single-cell data" which is $\sim$10,000$\times$ larger than standard transcriptomic data.

- Developed and supported tools for, and improvements to, the platform which are currently used by various departments allowing for future scalability.
- Presented important findings to both internal audiences and clients.

Graduate Research Assistant 2016 – 2021
Guasto Lab, Department of Mechanical Engineering, Tufts University, Medford, MA

- Currently studying localized stretching structures of viscoelastic fluids in porous media via DNA visualization techniques.
- Using particle tracking and Lagrangian statistical tools, discovered that dispersion is regulated by flow geometry in viscoelastic flows.
- Designed and conducted Monte-Carlo Langevin simulations of swimming cells in a viscosity gradient to support experimental discoveries.
- Discovered that disordering flow geometry affects the local flow type experienced by viscoelastic fluids and hinders a critical flow instability responsible for chaotic velocity fluctuations.
- Utilizing high speed microscopy, invertebrate sperm flagellum buckling was studied in a microfluidic extensional flow. An in-house flagella tracking algorithm was developed to investigate flagellum curvature.

Data Science Consultant Jan 2020 – Sept 2020
Gene Network Sciences Inc, Cambridge, MA

- Curated real-world financial and weather data was processed and vetted to build causal and predictive models using a proprietary causal machine learning platform (REFS).
- Identified quality issues with customer-provided data by developing a bespoke outlier detection algorithm, leading to customer modification of their internal data pipelines.
- Using repeated & stratified cross-validation, demonstrated value of integrating customer data with multi-modal financial data for building predictive models and assisting with go/no-go decisions for modeling.
- Findings were reported to clients and executives via intuitive graph visualizations (iGraph) and presentations. Described technical methods to lay audiences.

Certifications

2023 *Key Topics in Causal Inference*, Harvard T.H. Chan School of Public Health
Workshop given by leaders in the field of Causal Inference: Jamie Robbins, Miguel Hernán, Judith Lok, Eric Tchetgen Tchetgen, and Tyler Vanderweele.

Honors and Awards

2020 *Burlingame Fellowship in Physics*, Tufts University
Recognizes outstanding achievement by senior graduate student in Physics.

2018 *Tufts Polymer Art Competition Winner*

Journal Publications

1. D.M. Walkama, N. Waisbord, M. Kumar, A. Ardehkani, and J.S. Guasto. Geometry regulates dispersion in viscoelastic porous media flows. *In preparation for submission* (2023).
2. M. Stehnach, N. Waisbord, D.M. Walkama, and J.S. Guasto. Viscotaxis Aggregation of swimming cells in non-uniform viscous environments. *Nat. Phys.* **17**, 926–930 (2021).
3. D.M. Walkama, N. Waisbord, and J.S. Guasto. Disorder suppresses chaos in viscoelastic flows. *Physical Review Letters* **124**, 164501 (2020).

4. M. Kumar, D.M. Walkama, J.S. Guasto, and A.M. Ardekani. The buckling of sperm flagellum in an extensional flow. *Physical Review E* **100**, 063107 (2019).
5. N. Waisbord, N. Stoop, D.M. Walkama, J. Dunkel, and J.S. Guasto. Anomalous percolation flow transition of yield stress fluids in porous media. *Physical Review Fluids* **4**, 063303 (2019).

Contributed Presentations

1. B. Nathasingh, D.M. Walkama, L.M. Mayhew, K. Loh, J. Latourelle, B.W. Church, Y.E. Wang. AACR 2023. Orlando, Florida
   "Infer cancer cell gene dependency in multiple myeloma using causal AI in-silico patient model"
2. D.D. Vagie, D.W. Walkama, L.M. Mayhew, T.E. Oakland, B.W Church. ASH Meeting 2022. New Orleans, Louisiana
   "Causal ai in silico patient model identifies minichromosome (MCM) Family Genes as novel predictors for overall survival in multiple myeloma"
3. D.W. Walkama, N. Waisbord, J.S. Guasto. APS-DFD Meeting 2019. Seattle, WA
   "Geometric disorder regulates dispersion in viscoelastic porous media flows"
4. D.W. Walkama, N. Waisbord, J.S. Guasto. APS March Meeting 2019. Boston, MA
   "Disorder suppresses chaotic viscoelastic flow"
5. D.W. Walkama, N. Waisbord, J.S. Guasto. APS-DFD Meeting 2018. Atlanta, GA
   "Disorder suppresses viscoelastic instability"
6. D.W. Walkama, N. Waisbord, J.S. Guasto. Tufts Polymer Research Symposium 2018. Medford, MA.
   "Disorder suppresses viscoelastic instability" (Poster)
   ∗ Polymer art competition winner
7. D.W. Walkama, N. Waisbord, J.S. Guasto. New England Complex Fluids Workshop 2017. Boston, MA
   "geometric disorder suppresses elastic turbulence"
8. D.W. Walkama, N. Waisbord, J.S. Guasto. New England Complex Fluids Workshop 2017. Medford, MA
   "pattern formation in viscoelastic flow through porous media"

Teaching & Mentoring

Assistant Teaching

2016 fall	PHY-0042, Electricity & Magnetism I, Course Assistant
2016 summer	PHY-0042, Electricity & Magnetism I, Recitation Instructor
2016 spring	PHY-0012, General Physics II, Lab Instructor
2015 fall	PHY-0011, General Physics I, Lab Instructor

Research Mentoring and Training

Mentored one undergraduate and one graduate students on independent research topics and provided training on microfabrication, microfluidics, microscopy, and image analysis.

Professional Service and Society Membership

Journal Peer Review

Conducted peer review for submissions to: *Soft Matter, Science Advances,* and *Proceedings of the National Academy of Sciences.*

Society Membership

1. Graduate Physics and Astronomy Student Society (President, 2017)
2. American Physical Society - Division of Fluid Dynamics (APS-DFD)
3. New England Complex Fluids Workgroup (NECFW)

References

Professor Jeffrey S. Guasto
Tufts University
Medford, Massachusetts, USA
Associate Professor of Mechanical Engineering
Ph.D. Advisor
email: Jeffrey.Guasto@tufts.edu

Dr. Laurel Mayhew
Aitia
Somerville, MA
Senior Director, Early Development
Supervisor
email: LMayhew.Career@gmail.com

Dr. Jignesh Parikh
J Square Labs
Greater Boston Area, Massachusetts, USA
Founder
Supervisor and lead Computational Scientist
email: jiggy@jsquarelabs.com

Dr. Daniel Vagie
Aitia
Somerville, MA
Data Scientist
R&D colleague
email: vagiedd@outlook.com