Galen Samuel Harrison

Charlottesville, VA, galenharrison.com, me@galenharrison.com

Education

- PhD Computer Science, December 2025 from University of Virginia
  - Dissertation: *Methodological Choices in Artificial Intelligence and Machine Learning*
  - Committee: Madhav Marathe (Advisor), Blase Ur, Mona Sloane, Ferdinando Fioretto, Afsaneh Doryab
- MS Computer Science, March 2020 from University of Chicago
  - Thesis: *Towards Understanding Decisions about Fairness in Machine Learning*
  - Committee: Blase Ur (Advisor), Ben Zhao, Marshini Chetty, Joseph Lorenzo Hall
- Master of Legal Studies, June 2019 from University of Chicago Law School
- BA Mathematics, ΦBK, May 2015 from Reed College.
  - Thesis: *Differentially Private Statistics: Implementation and performance of PTR and objective perturbation methods*
  - Committee: Adam Groce (Advisor), James Fix, Derek Applewhite, Kris Anderson

Awards and Honors

Phi Beta Kappa, Commendation for Excellence in Scholarship 2012 & 2015, UVA Deans Fellowship (2020-2021), University of Virginia Collaborative Excellence in Public Service (Team award)

Publications

JupyterLab in Retrograde: Contextual Notifications That Highlight Fairness and Bias Issues for Data Scientists. **Galen Harrison**, Kevin Bryson, Ahmad Bamba, Luca Dovichi, Aleksander Binion, Arthur Borem, Blase Ur. *Best Paper CHI 2024*

Identifying Complicated Contagion Scenarios from Cascade Data. **Galen Harrison**, Amro Alabsi Aljundi, Jiangzhuo Chen, SS Ravi, Anil Kumar Vullikanti, Madhav Marathe, Abhijin Adiga. *KDD 2023*.

A Network-based Analytics Framework for High-resolution Agent-based Epidemic Simulation Ensembles. Amro Alabsi Aljundi, **Galen Harrison**, Jiangzhuo Chen, Madhav Marathe, Henning Mortveit, Anil Kumar Vullikanti, Abhijin Adiga. *Winter Simulation Conference 2023*.

Synthetic Information and Digital Twins for Pandemic Science: Challenges and Opportunities. **Galen Harrison**, Przemyslaw Porebski, Mandy L. Wilson, Jiangzhuo Chen, Henning Mortveit, Parantapa Bhattacharya, Dawen Xie, Stefan Hoops, Anil Vullikanti, Li Xiong, Madhav Marathe. *TPS-ISA 2023*.

Data-Driven Real-Time Strategic Placement of Mobile Vaccine Distribution Sites. Zakaria Mehrab, Mandy L. Wilson, Serina Chang, **Galen Harrison**, Bryan Lewis, Alex Telionis, Justin Crow, Dennis Kim, Scott Spillman, Kate Peters, Jure Leskovec, Madhav Marathe. *AAAI 2022*

Projected resurgence of COVID-19 in the United States in July—December 2021 resulting from the increased transmissibility of the Delta variant and faltering vaccination.

Shaun Truelove, Claire P Smith, Michelle Qin, Luke C Mullany, ... **Galen Harrison**, ..., Michael C Runge, Cecile Viboud. *eLife 2022*

Files of a Feather Flock Together? Measuring and Modeling How Users Perceive File Similarity in Cloud Storage. Will Brackenbury, **Galen Harrison**, Kyle Chard, Aaron Elmore, Blase Ur *SIGIR 2021.*

An Empirical Study on the Perceived Fairness of Realistic, Imperfect Machine Learning Models. **Galen Harrison**, Julia Hanson, Christine Jacinto, Julio Ramirez, Blase Ur. *FAccT 2020*

Skluma: An extensible metadata extraction pipeline for disorganized data. Tyler Skluzacek, Rohan Kumar, Ryan Chard, **Galen Harrison**, Kyle Chard and Ian Foster. *eScience 2018*

Workshops, Presentations, and other publications

JupyterLab in Retrograde. **Galen Harrison** *Invited Talk at University of the Arts, London*, December 3, 2024

Synthetic Data to Support US-UK Prize Challenge for Developing Privacy Enhancing Methods: Predicting Individual Infection Risk During a Pandemic [Data set]. **Galen Harrison**, Jiangzhuo Chen, Henning Mortveit, Przemyslaw Porebski, Dawen Xie, Mandy L. Wilson, Parantapa Bhattacharya, Anil Vullikanti, Li Xiong, Madhav Marathe. *Tech Report/Dataset*

Predictability of Vaccination Surges and Google Search Trends. **Galen Harrison**, Mark Orr, Srinivasan Venakatramanan, Mandy L. Wilson, Bryan Lewis, Madhav Marathe. *UVA Engineering Research Symposium 2023* (Poster Competition Finalist)

Fake friends: Privacy Implications of Synthetic Population Data. **Galen Harrison**. *4th Annual Symposium on Applications of Contextual Integrity 2022* (Peer reviewed)

Commentary on "Divining Insights: Visual Analytics through Cartomancy" by Mcnutt et al. **Galen Harrison**, *alt.chi* 2020

- Commentary selected by committee for inclusion in published proceedings of *alt.chi* 2020

Towards Supporting and Documenting Algorithmic Fairness in the Data Science Workflow. **Galen Harrison**, Julia Hanson, Blase Ur. *Workshop on Technology and Consumer Protection (ConPro '19), at IEEE S&P* (Peer reviewed)

Where's the Bias? Developing Effective Model Governance. **Galen Harrison**, Natasha Duarte, Joseph Lorenzo Hall. *NeurIPS 2018 Workshop on Challenges and Opportunities for AI in Financial Services: the Impact of Fairness, Explainability, Accuracy, and Privacy* (Peer reviewed)

Various blogs, Center for Democracy and Technology blog, 2018.

- On Managing Risk in Machine Learning Projects

Publicly Fair Machine Learning. **Galen Harrison**. Trustworthy Algorithms Workshop 2017. (Peer reviewed)

Poking at the Cloud: Identifying Factors Behind Selective Cloud Uploading. Amanda Aizuss, Max Chen, **Galen Harrison**, Sotiri Komissopoulos, Blase Ur. *Poster, SOUPS 2017.* (Peer reviewed)

Patterns of Criminal Justice: The Old Bailey from 1674-1913. **Galen Harrison**. *Presentation at Law and the Justice System in the Era of Big Data 2014.*

Reviewer

CSCW 2020, DIS 2020, TOCHI 2021, HCAI@NeurIPS 2022, FAccT 2023, RegML 2023-2024, KDD 2023 & 2025 (Excellent Reviewer Commendation), SDM 2023, Minds & Machines 2023, 2024, AIES 2025, EAAMO 2025

Professional Service

University of Chicago Independent Review Committee 2017-2018, Fulbright Interview Committee 2017, 2018, and 2019, UVA Graduate School Mentor 2022-2023, 2023-2024, 2024-2025

Research Experience

- Data Scientist Intern (PhD), Integra FEC (Summer 2024)
  - Perform quantitative research to identify and trace suspicious cryptocurrency flows
  - Built reproducible, documented data pipelines in BigQuery to identify and extract relevant patterns from large datasets
  - Used Python (Pandas, Matplotlib, Selenium) to refine analyses, and communicate findings to stakeholders
- Research Assistant, University of Virginia Biocomplexity Institute (September 2020 - Present)
  - Led data integration and cross-source inference for COVID-19 vaccination analytics, contributing to a public-facing dashboard that received formal commendation for public service from University of Virginia
  - Designed reference solutions for White House-sponsored US-UK Prize Challenge on Privacy-Enhancing Technologies (PETS), adapting graph neural network approaches to infection prediction on massive networks
  - Developed and implemented large-scale graph-learning frameworks handling graphs with up to 61M nodes, employing parallel computing techniques and HPC environment
  - Published first-author paper in KDD demonstrating novel methods to extract cascade parameters from aggregate structural properties, advancing state-of-the-art in network influence modeling
  - Led evaluation framework development for LLM-generated technological dependency networks in manufacturing, creating robust validation methodologies for assessing graph quality and utility
  - Developed and implemented data synthesis algorithms using PyTorch and JAX, advancing understanding of data similarity metrics
- Research Assistant, University of Chicago, Department of Computer Science (August 2016 - August 2020)
  - Conceptualized and developed a novel JupyterLab extension for data scientists that tracked workflow actions and provided real-time fairness notifications, leading to a Best Paper award at CHI 2024
  - Built full-stack implementation using TypeScript and custom code analysis tools to log, analyze, and provide contextual feedback during the data science process
  - Designed and executed user studies evaluating human decision-making in fairness-related AI contexts, resulting in first-author publication at FAccT 2020

  - Contributed significantly to successful NSF Fairness in Artificial Intelligence grant, drafting key sections of the proposal
  - Founded and facilitated a design methods reading group, fostering interdisciplinary collaboration among researchers
  - Mentored undergraduate and high school students in research methodologies, resulting in three students achieving co-author status on published papers
  - Applied quantitative and qualitative research methods to analyze user study data, including statistical analysis and visualization development
- Internet Architecture Intern, Center for Democracy and Technology (Summer 2018)
  - Researched and wrote a policy paper on machine learning model governance in FinTech (Accepted at NeurIPS Workshop)
  - Wrote blog posts translating hardware side-channel vulnerabilities for policy makers
  - Compiled and communicated research about other various technical topics for policy makers
- Data Programmer/Pre-doctoral associate, Yale School of Management (August 2015-August 2016)
  - Project: Extract the p-values from large numbers of papers, in order to estimate the p-value distribution
  - Project: Calculate sub networks and network statistics for a network of physicians with shared patients
  - Other research support as appropriate, including web-scraping, data organization, and providing advice on approaches and feasibility

Teaching Experience

- Teaching Assistant, University of Virginia, Department of Computer Science (Fall 2022 - Spring 2024)
  - Teaching assistant for graduate and undergraduate versions of "Privacy in the Internet Age"
  - Developed automated grading infrastructure
  - Lectured on differential privacy
- Mentor, Center for Data and Computation, University of Chicago (Summer 2019, Summer 2020)
  - Mentored undergraduate and high school students engaged in summer research experiences
  - Provided advice and instruction about the conduct of human subjects research
  - Taught mentees about survey design, qualitative coding, academic writing
  - High School Mentee contributions lead to co-authorship credit
- Teaching Assistant, University of Chicago, Department of Computer Science (Fall 2017, Spring 2020)
  - Courses: Ethics, Fairness, Responsibility and Privacy in Data Science, Introduction to Computer Security

- Created hands-on assignments in fair machine learning and differential privacy
- Aided students in clarifying their understanding of core concepts in computer security
- Graded students' submitted work which consisted of problem sets and cryptographic implementations

Other Experience

- Research Assistant to Prof. Adam Groce, Reed College (June 2015-August 2015)
- Student Body Senator, Reed College (2014-2015)
- Judicial Board, Reed College. (2012-2014), Chair (2013-2014)
- Research Experience for Undergraduates with Simon DeDeo, Santa Fe Institute (Summer 2013)
- Research Assistant, Linac Coherent Light Source, Stanford Linear Accelerator (Summer, 2012)