Joseph Futoma

Science Center Room 316, School of Engineering & Applied Sciences, Harvard University, Cambridge MA 02138 USA
jfutoma14@gmail.com • +1 (401) 241-0158 • https://jfutoma.github.io/ • US Citizen

EDUCATION

Duke University, Durham, NC

- Ph.D. in Statistical Science — Aug 2013 – May 2018
  - Advisor: Katherine Heller, Ph.D.
  - Dissertation: *Gaussian Process-Based Models for Clinical Time Series in Healthcare*
  - Focus: Bayesian statistics, machine learning, healthcare applications, data analysis of electronic health records.
  - Deep learning algorithm for early detection of sepsis currently implemented and used at Duke University Hospital.
- M.S. in Statistical Science — Aug 2013 – May 2016
  - GPA: 3.93 / 4.00
  - Coursework: Linear Models, Probability & Measure Theory, Bayesian & Modern Statistics, Advanced Machine Learning, Statistical Inference, Statistical Case Studies, Probability & Statistical Models, Spatial Statistics

Dartmouth College, Hanover, NH

- A.B. with High Honors in Mathematics — Sep 2010 – Jun 2013
  - Advisor: Daniel Rockmore, Ph.D.
  - Thesis: *Scalable Inference Algorithms for Clustering Large Networks*
  - *Summa Cum Laude*, GPA: 3.94 / 4.00 (3.96 in major)
  - Coursework: Abstract Algebra, {Real, Complex, Numerical} Analysis, Chaos, Probability, Quantum Mechanics, Networks, Topology, Measure Theory, Machine Learning, Mathematical Statistics, Intro to Applied Math

RESEARCH INTERESTS

Machine learning, reinforcement learning & sequential decision making, Bayesian methods, clinical time series modeling, healthcare applications, implementation & integration of machine learning in medicine, ethical & sociotechnical implications of machine learning in the real world.

ACADEMIC RESEARCH EXPERIENCE

School of Engineering & Applied Sciences, Harvard University

- Postdoctoral Fellow, Center for Research on Computation & Society — Jun 2018 – Present
  - Advisor: Finale Doshi-Velez, Ph.D.
  - Focus: Reinforcement learning & sequential decision-making in healthcare; analysis of intensive care unit (ICU) data.

Dept. of Statistical Science, Duke University

- Research Associate — May 2018 – Present
- NDSEG Research Fellow — Sep 2015 – May 2018
- Statistical & Applied Mathematical Sciences Inst. (SAMSI) Research Fellow — Aug 2014 – May 2015
  - Mentor: David Dunson, Ph.D.
  - Focus: Analysis of high-dimensional biological data from exercise physiology studies.
- First Year Statistical Science Research Fellow — Aug 2013 – May 2014

Dept. of Mathematics, Dartmouth College

- Presidential Scholar Research Assistant, — May 2012 – Aug 2013
  - Focus: Time-varying topic models, scalable network analysis.
  - Advisor: Daniel Rockmore, Ph.D.

INDUSTRY WORK EXPERIENCE

Statistical Consulting — Jan 2017 – Present

- International Farming Corporation.
- Lab of Cynthia Toth, MD at Duke University: Predict disease progression in macular degeneration.

Amazon.com, Inc, Prime Music, Personalization

- Research Intern — May 2015 – Aug 2015
  - Project: Fit large-scale regressions with Apache Spark to solve the source appraisal problem in music recommendation. In online A/B testing, proposed algorithm increased overall music consumption by 1.2%.
  - Supervisor: Charles Thompson

Quintiles Transnational, Predictive Analytics

- Biostatistician Intern — Jun 2014 – Sep 2014
  - Project: Develop statistical models to predict site adherence for risk-based monitoring of clinical trials.
  - Supervisors: Joseph Lucas, Ph.D. and Valerii Fedorov, Ph.D.

HONORS & AWARDS

- Center for Research on Computation & Society (CRCS) Postdoctoral Fellowship, Harvard — 2018
- Plenary Presentation (top 3/1540 of abstracts), Society for Hospital Medicine — 2018
- **National Defense Science and Engineering Graduate (NDSEG) Fellowship** — 2015

- **Winner, LinkedIn Economic Graph Challenge** 2015
- 2x Honorable Mention, NSF Graduate Research Fellowship Program 2014, 2015
- First Year Statistical Science Research Fellowship, Duke University 2013 – 2014
- Notable Paper Award (top 2.6% of submissions), AISTATS 2013 2013
- Rufus Choate Scholar (GPA in top 5% of class), Dartmouth College 2010 – 2013

GRANTS

- "Innovative Predictive Model to Anticipate Steroid Induced Hyperglycemia and Guide Insulin Regimens," May 2018 - Apr 2019
  - Co-Investigator (PI: Dr. Ann McGee).
  - Duke Institute for Health Innovation.
  - Total: $25,000.
- "Early Identification of Cardiac Decompensation and Cardiogenic Shock," May 2018 - Apr 2019
  - Co-Investigator (PI: Dr. Ajar Cochar).
  - Duke Institute for Health Innovation.
  - Total: $60,000.
- "Implementation of a Duke-Specific Early Warning System," May 2016 - Apr 2017
  - Collaborator (PI: Dr. Cara O'Brien).
  - Duke Institute for Health Innovation.
  - Total: $50,000.
- "Improving Chronic Disease Management in Duke Primary Care: Building a Virtual Medical Neighborhood for Chronic Kidney Disease," May 2016 - Apr 2017
  - Collaborator (PI: Dr. Blake Cameron).
  - Duke Institute for Health Innovation.
  - Total: $40,000.
- "Chronic Kidney Disease Population Management Tools," Jun 2015 - May 2016
  - Collaborator (PI: Dr. Uptal Patel).
  - Duke Translational Research Institute.
  - Total: $50,000.
- "Duke Connected Care Chronic Kidney Disease Care Improvement Project," May 2015 - Apr 2016
  - Collaborator (PI: Dr. Dev Sangvai).
  - Duke Institute for Health Innovation.
  - Total: $65,000.

PUBLICATIONS

REFEREED JOURNAL PAPERS

[1] M. Sendak, W. Ratliff, D. Sarro, E. Alderton, **J. Futoma**, M. Gao, M. Nichols, M. Revoir, F. Yashar, C. Miller, K. Kester, S. Sandhu, K. Corey, N. Brajer, C. Tan, A. Lin, T. Brown, S. Engelbosch, K. Anstrom, M. Elish, K. Heller, R. Donohoe, J. Theiling, E. Poon, S. Balu, A. Bedoya, C. O'Brien. "Sepsis Watch: A Real-World Integration of Deep Learning into Routine Clinical Care", to appear in *Journal of Medical Internet Research*. doi: https://doi.org/10.2196/15182

[2] N. Brajer, B. Cozzi, M. Gao, M. Revoir, M. Nichols, **J. Futoma**, J. Bae, N. Setji, S. Balu, A. Hernandez, M. Sendak. "Prospective and External Evaluation of a Machine Learning Model to Predict In-Hospital Mortality", to appear in *JAMA Network Open*. doi: https://doi.org/10.1101/19000133.

[3] **J. Futoma**, J. Morris, and J. Lucas. "A Comparison of Models for Predicting Early Hospital Readmissions," *Journal of Biomedical Informatics*, vol. 56, pp. 229–238, Aug 2015.

REFEREED CONFERENCE PAPERS

[1] **J. Futoma**, M. Hughes, F. Doshi-Velez. "POPCORN: Partially Observed Prediction COnstrained Reinforcement LearNing.", to appear at *AISTATS 2020*.

[2] **J. Futoma**, M. A. Masood, F. Doshi-Velez. "Identifying Distinct, Effective Treatments for Acute Hypotension with SODA-RL: Safely Optimized Diverse Accurate Reinforcement Learning.", to appear at *AMIA Informatics Summit 2020*.

[3] M. Sendak, M. Elish, M. Gao, **J. Futoma**, A. Bedoya, M. Nichols, W. Ratliff, S. Balu, C. O'Brien. " "The Human Body is a Black Box": Supporting Clinical Decision-Making with Deep Learning," to appear at *FAT* 2020*.

[4] **J. Futoma**, S. Hariharan, M. Sendak, N. Brajer, M. Clement, A. Bedoya, C. O'Brien, and K. Heller. "An Improved Multi-Output Gaussian Process RNN with Real-Time Validation for Early Sepsis Detection," in *Proceedings of the 2nd Machine Learning for Healthcare Conference (MLHC)*, Boston, MA, Aug 2017.

[5] **J. Futoma**, S. Hariharan, and K. Heller. "Learning to Detect Sepsis with a Multitask Gaussian Process RNN Classifier," in *Proceedings of the 34th International Conference on Machine Learning (ICML)*, Sydney, Australia, Aug 2017.

[6] **J. Futoma**, M. Sendak, C. B. Cameron, and K. Heller. "Predicting Disease Progression with a Model for Multivariate Longitudinal Clinical Data," in *Proceedings of the 1st Machine Learning for Healthcare Conference (MLHC)*, Los Angeles, CA, Aug 2016.

[7] **J. Futoma**, M. Sendak, C. B. Cameron, and K. Heller. "Scalable Joint Modeling of Longitudinal and Point Process Data for Disease Trajectory Prediction and Improving Management of Chronic Kidney Disease," in *Proceedings of the 32nd Conference on Uncertainty in Artificial Intelligence (UAI)*, New York City, NY, Jun 2016.

[8] N. Foti, **J. Futoma**, D. Rockmore, and S. Williamson. "A Unifying Representation for a Class of Dependent Random Measures," in *Proceedings of the 16th Conference on Artificial Intelligence and Statistics (AISTATS)*, Scottsdale, AZ, May 2013.

REFEREED WORKSHOP PAPERS

[1] **J. Futoma**, M. Hughes, F. Doshi-Velez. "Prediction-Constrained POMDPs," in *NeurIPS 2018 Workshop on Reinforcement Learning under Partial Observability*, Montreal, Canada, Dec 2018.

[2] **J. Futoma**, A. Lin, M. Sendak, M. Clement, A. Bedoya, C. O'Brien, and K. Heller. "Learning to Treat Sepsis with Multi-Output Gaussian Process Deep Recurrent Q-Networks," in *NeurIPS 2017 Workshop on Machine Learning for Health*, Long Beach, CA, Dec 2017.

[3] **J. Futoma** and J. Lucas. "Predicting Early Hospital Readmissions using Electronic Health Records," in *NeurIPS 2014 Workshop on Machine Learning for Clinical Data, Healthcare and Genomics*, Montreal, Canada, Dec 2014.

OTHER PAPERS

[1] A. Lin, M. Sendak, A. Bedoya, M. Clement, N. Brajer, **J. Futoma**, H. Bosworth, K. Heller, C. O'Brien. "Evaluating sepsis definitions for clinical decision support against a definition for epidemiological disease surveillance.", doi: https://doi.org/10.1101/648907, May 2019.

[2] **J. Futoma**. "Gaussian Process-Based Models for Clinical Time Series in Healthcare," *Duke University Ph.D. Dissertation*, May 2018.

[3] **J. Futoma**. "Scalable Inference Algorithms for Clustering Large Networks," *Dartmouth College Senior Thesis*, Jun 2013.

PAPERS IN PREPARATION

[1] **J. Futoma**, R. Kamaleswaran, F. Doshi-Velez. "Investigating Generalizability of Machine Learning Models for Prediction of Vasopressor Onset.", in preparation to submit, *Critical Care Medicine*.

[2] A. Bedoya*, **J. Futoma***, M. Clement, K. Corey, N. Brajer, A. Lin, M. Simons, M. Gao, M. Nichols, S. Balu, K. Heller, M. Sendak, C. O'Brien. "Machine Learning for Early Detection of Sepsis: An Internal and Temporal Validation Study," under review at *JAMIA Open*. (*: joint first author)

PRESENTATIONS & TALKS

- Invited Talk, BayesComp 2020. Jan 2020
- Poster Presentation, MLHC 2019. Aug 2019
  - A. Kansal, S. Kashyap, W. Ratliff, K. Sriram, M. Sendak, M. Nichols, M. Gao, **J. Futoma**, M. Revoir, S. Balu, M. Pencina, K. Kester, C. Miller, Z. Wegermann, C. Granger, J. Schroeder, C. Milano, M. Patel, S. Jones, C. Patel, A. Kochar. "Using Predictive Mortality and Cardiogenic Shock Identification Tools to Support Team-Based Treatment Intervention on Adult Cardiology Patients at Duke University Hospital" (Clinical Abstract).

- Poster Presentation, MLHC 2019. Aug 2019
  - N. Brajer, B. Cozzi, M. Gao, M. Nichols, M. Revoir, S. Balu, K. Whalen, **J. Futoma**, C. O'Brien, C. Patel, P. Setji, A. Hernandez, M. Sendak. "Leveraging Machine Learning to Decrease In-Hospital Mortality Rates" (Clinical Abstract).
- Poster Presentation, MLHC 2019. Aug 2019
  - M. Simons, **J. Futoma**, K. Corey, M. Gao, M. Nichols, K. Whalen, M. Sendak, F. Doshi-Velez, A. McGee, T. Setji. "Development of a Clinical Decision Tool and Protocol for Identification and Treatment of Corticosteroid Induced Hyperglycemia" (Clinical Abstract).
- Poster Presentation, American Diabetes Association Scientific Sessions. Jun 2019
  - M. Simons, **J. Futoma**, M. Gao, K. Corey, M. Sendak, K. Whalen, F. Doshi-Velez, A. McGee, T. Setji. "Predictive Model for Hyperglycemic Events after High Dose Corticosteroid Administration".
- Invited Talk, NeurIPS 2018 Workshop: All of Bayesian Nonparametrics. Dec 2018
- Spotlight & Poster Presentation, NeurIPS 2018 Workshop on RL under Partial Observability. Dec 2018
- Poster Presentation, MLHC 2018. Aug 2018
  - A. Lin, **J. Futoma**, A. Bedoya, N. Brajer, M. Sendak, F. Yashar, M. Nichols, M. Gao, M. Clement, K. Heller, C. O'Brien. "Leveraging Deep Learning and Rapid Response Team Nurses to Improve Sepsis Management" (Clinical Abstract).
- Poster Presentation, ISBA 2018 World Meeting. Jun 2018
- Poster Presentation, American Thoracic Society International Conference 2018. May 2018
  - A. Lin, M. Sendak, A. Bedoya, M. Clement, **J. Futoma**, M. Nichols, M. Gao, K. Heller, C. O'Brien. "What Is Sepsis: Investigating the Heterogeneity of Patient Populations Captured by Different Sepsis Definitions".
- Plenary Presentation, Society for Hospital Medicine 2018. Aug 2018
  - **Top 3 / 1540 submitted abstracts.**
  - A. Lin, **J. Futoma**, M. Sendak, A. Bedoya, M. Clement, M. Gao, M. Nichols, K. Heller, C. O'Brien. "Deeply-Personalized Medicine: Bringing Deep Learning to Sepsis Care".
- Poster Presentation, NIPS 2017 Workshop on ML for Health. Dec 2017
- Invited Talk, Epic Data Science Forum (Madison, WI). Oct 2017
- Oral Presentation, Duke Dept. of Statistical Science, Seminar Series. Sep 2017
- Spotlight Presentation, MLHC 2017. Aug 2017
- Oral Presentation, ICML 2017. Aug 2017
- Oral Presentation, INFORMS Healthcare 2017. Jul 2017
- Spotlight Presentation, MLHC 2016. Aug 2016
- Invited Talk, UAI 2016 Workshop on Bayesian Applications. Jun 2016
- Poster Presentation, UAI 2016. Jun 2016
- Oral Presentation, Bayesian Young Statisticians Meeting. Jun 2016
- Poster Presentation, ISBA 2016 World Meeting. Jun 2016
- Poster Presentation, Society of General Internal Medicine. May 2016
  - M. Sendak, C. B. Cameron, E. Komives, **J. Futoma**, E. Huang, K. Heller, D. Sangvai L. E. Boulware, and U.D.Patel. "Developing a Data-Driven Workflow for Population Health Rounding".
- Oral Presentation & Contributed 1 page paper, NIH-IEEE 2015 Strategic Conference on Healthcare Innovations and Point-of-Care Technologies for Precision Medicine. Oct 2015
  - Z. Sun, **J. Futoma**, M. Sendak, E. Lorenzi, S. Brown, O. Huang, K. Heller, J. Thacker, C. Mantyh, and E. Huang. "Precision Medicine in Point-Of-Care Management of Surgical Complications."
- Poster Presentation, NIPS 2015 Workshop on ML for Clinical Data, Healthcare & Genomics. Dec 2014
- Invited Talk, Applied and Computational Mathematics Seminar Series, Dartmouth College. Apr 2013

TEACHING & MENTORING EXPERIENCE

Team Member Jun 2019 – Present

- Embedded EthiCS Program at Harvard: bringing ethical reasoning into the computer science curriculum.

Mentored Students Jan 2015 – Present

- Henry Wang (Masters Student, Harvard Computer Science)
- Jianzhun Du (Masters Student, Harvard Computer Science)
- Sanjana Narayanan (Undergraduate, Harvard Computer Science)
- Kristine Zhang (Undergraduate, Harvard Computer Science)
- Yash Nair (Undergraduate, Harvard Computer Science)
- Faraz Yashar (Undergraduate, Duke Computer Science)
- Ouwen Huang (Undergraduate, Duke Computer Science)

- Sanjay Hariharan (Masters Student, Duke Statistical Science)
- Brian Cozzi (Masters Student, Duke Statistical Science)
- Morgan Simons (Medical Student, Duke University School of Medicine)
- Anthony Lin (Medical Student, Duke University School of Medicine)
- Nathan Brajer (Medical Student, Duke University School of Medicine)

Guest Lecturer Sep 2014 – Sep 2014

- Taught 2 weeks of a freshman-level data science course as part of Duke University's Data Expeditions program. Curated a baseball-themed dataset and guided students through simple exploratory analysis.

PROFESSIONAL SERVICE

Area Chair 2019 – Present

- MLHC 2020

Peer Reviewer 2015 – Present

- AISTATS (2017-2020), ICML (2017-2020), MLHC (2016-2019), AMIA (2015-2019), NeurIPS Workshops (2015-2019), NeurIPS (2017-2019), IEEE TPAMI, JMLR, JMIR

Red Judge 2019

- IBM Watson AI XPrize

SKILLS

Python (TensorFlow, PyTorch), MATLAB, R, BUGS/JAGS, Apache Spark, Scala, Julia, UNIX/Linux shell scripting, HTML/CSS, LaTeX, Mathematica, Microsoft Office.

REFERENCES

- **Finale Doshi-Velez, Ph.D.**
  John L. Loeb Associate Professor in Computer Science
  Harvard University John A. Paulson School of Engineering & Applied Sciences
  finale@seas.harvard.edu

- **Katherine Heller, Ph.D.**
  Assistant Professor
  Duke University: Dept. of Statistical Science & Center for Cognitive Neuroscience
  Research Scientist: Google Brain
  kheller@gmail.com

- **Suresh Balu, M.S., M.B.A.**
  Associate Dean for Innovation & Partnership
  Duke University School of Medicine
  Program Director: Duke Institute for Health Innovation
  suresh.balu@duke.edu

- **Cara O'Brien, M.D.**
  Associate Vice Chair of Inpatient Operations
  Department of Medicine
  Duke University School of Medicine
  cara.obrien@duke.edu

- **Barbara Grosz, Ph.D. (joint reference with Alison Simmons)**
  Higgins Professor of Natural Sciences
  Harvard University: John A. Paulson School of Engineering & Applied Sciences
  Faculty Founder: Embedded EthiCS at Harvard
  grosz@eecs.harvard.edu

- **Alison Simmons, Ph.D. (joint reference with Barbara Grosz)**
  Samuel H. Wolcott Professor of Philosophy
  Harvard University: Dept. of Philosophy
  Faculty Founder: Embedded EthiCS ast Harvard
  asimmons@fas.harvard.edu

[*CV compiled on 2020-01-06*]