

ALLISON KOENECKE

koenecke@cornell.edu ◊ www.infosci.cornell.edu/~koenecke

RESEARCH AREAS

Computational Social Science, Algorithmic Fairness, Automated Speech Recognition,
Causal Inference, Public Health

EMPLOYMENT

Cornell University

Assistant Professor, Department of Information Science
Field Faculty Member, Department of Computer Science

Ithaca, NY

Jul 2022 - Present

Microsoft Research New England

Postdoctoral Researcher, Machine Learning and Statistics Group

Cambridge, MA

Jul 2021 - Jun 2022

Google

Quantitative Analyst Intern, Economics Team

Mountain View, CA

Jun - Sep 2019

Microsoft

Data Scientist Intern, Artificial Intelligence & Research Group

Sunnyvale, CA

Jun 2018 - Jan 2019

Google

Quantitative Analyst Intern, Economics Team

San Francisco, CA

Jun - Sep 2017

Facebook

Data Science Intern, Events Team

Menlo Park, CA

Jun - Aug 2016

NERA Economic Consulting

Associate Analyst, Antitrust & Competition Group
Research Assistant, Antitrust & Competition Group
Intern, Antitrust & Competition Group

New York, NY

Jun 2015 - Mar 2016

Jun 2014 - Jun 2015

Jan 2013; Jun - Aug 2013

EDUCATION

Stanford University

Ph.D., Institute for Computational & Mathematical Engineering
M.S., Institute for Computational & Mathematical Engineering
Reading Committee: Susan Athey, Sharad Goel, and Hal Varian
Funding: NSF Graduate Research Fellowships Program

Stanford, CA

Sep 2016 - Jun 2021

Sep 2020 - Jun 2021

Massachusetts Institute of Technology

S.B., Mathematics with Computer Science
Minor, Economics

Cambridge, MA

Sep 2010 - Jun 2014

PUBLICATIONS

- Huynh, Benjamin Q., Elizabeth T. Chin, **Allison Koenecke**, Derek Ouyang, Daniel E. Ho, Mathew V. Kiang, David H. Rehkopf. 2024. *“Potential for Allocative Harm in an Environmental Justice Data Tool.”* Nature Machine Intelligence. <https://www.nature.com/articles/s42256-024-00793-y>
- Adam, Hammaad, Fan Yin, Neil Tenenholtz, Lorin Crawford, Lester Mackey, and **Allison Koenecke**. 2023. *“Should I Stop or Should I Go: Early Stopping of Randomized Experiments on Heterogeneous Populations.”* NeurIPS. <https://arxiv.org/abs/2306.11839>

- *Presented at: CODE 2022, IC2S2 2023 (Parallel Talk), INFORMS 2023, NeurIPS (2023 Spotlight), CODE 2023*
3. Xiong, Ruoxuan, **Allison Koenecke**, Michael Powell, Zhu Shen, Joshua T. Vogelstein, Susan Athey. 2023. “**Federated Causal Inference in Heterogeneous Observational Data.**” *Statistics in Medicine*. <https://onlinelibrary.wiley.com/doi/10.1002/sim.9868>
→ *Presented at: Microsoft Research Summit 2022 (Causal ML and Fairness Talk)*
 4. Papakyriakopoulos, Orestis*, Anna Seo Gyeong Choi*, William Thong, Dora Zhao, Jerone Andrews, Rebecca Bourke, Alice Xiang†, **Allison Koenecke**†. 2023. “**Augmented Datasheets for Speech Datasets and Ethical Decision-Making.**” *ACM Conference on Fairness, Accountability, and Transparency*. <https://dl.acm.org/doi/10.1145/3593013.3594049>
→ *Presented at: FAccT 2023 (Talk), IC2S2 2023 (Poster)*
 5. Pang, Rock Yuren*, Jack Cenatempo*, Franklyn Graham, Bridgette Kuehn, Maddy Whisenant, Portia Botchway, Katie Stone Perez, **Allison Koenecke**. 2023. “**Auditing Cross-Cultural Consistency of Human-Annotated Labels for Recommendation Systems.**” *ACM Conference on Fairness, Accountability, and Transparency*. <https://dl.acm.org/doi/10.1145/3593013.3594098>
→ *Media coverage: Cornell News*
→ *Presented at: FAccT 2023 (Talk)*
 6. **Koenecke, Allison**, Eric Giannella, Robb Willer, Sharad Goel. 2023. “**Popular Support for Balancing Equity and Efficiency in Resource Allocation: A Case Study in Online Advertising to Increase Welfare Program Awareness.**” *International AAAI Conference on Web and Social Media*. <https://ojs.aaai.org/index.php/ICWSM/article/view/22163/21942>
→ *Media coverage: Cornell News, Modern Farmer*
→ *Presented at: ICWSM 2023, IC2S2 2023 (Plenary Talk)*
 7. Zamfirescu-Pereira, JD, Jerry Chen, Emily Wen, **Allison Koenecke**, Nikhil Garg, Emma Pierson. 2022. “**Trucks Don’t Mean Trump: Diagnosing Human Error in Image Analysis.**” *ACM Conference on Fairness, Accountability, and Transparency*. <https://dl.acm.org/doi/10.1145/3531146.3533145>.
→ *Presented at: NeurIPS Workshop on Human and Machine Decisions 2021, FAccT 2022*
 8. **Koenecke, Allison**, Michael Powell, Ruoxuan Xiong, Zhu Shen, Nicole Fischer, Sakibul Huq, Adham M. Khalafallah, et al. 2021. “**Alpha-1 Adrenergic Receptor Antagonists to Prevent Acute Respiratory Distress Syndrome and Death from Cytokine Storm Syndrome.**” *eLife*. <https://elifesciences.org/articles/61700>.
→ *Media coverage: New York Times, Forbes, WebMD, Howard Hughes Medical Institute News*
 9. Powell, Michael, **Allison Koenecke**, James Brian Byrd, Akihiko Nishimura, Maximilian F. Konig, Ruoxuan Xiong, Sadiqa Mahmood, et al. 2021. “**Ten Rules for Conducting Retrospective Pharmacoepidemiological Analyses: Example COVID-19 Study.**” *Frontiers in Pharmacology*. <https://www.frontiersin.org/articles/10.3389/fphar.2021.700776/full>.
→ *Presented at: Toronto Workshop on Reproducibility 2022*
 10. Rose, Liam, Laura A. Graham, **Allison Koenecke**, Michael Powell, Ruoxuan Xiong, Zhu Shen, Kenneth W. Kinzler, et al. 2021. “**The Association Between Alpha-1 Adrenergic Receptor Antagonists and In-Hospital Mortality from COVID-19.**” *Frontiers in Medicine* 8: 304. <https://doi.org/10.3389/fmed.2021.637647>.

11. **Koenecke, Allison**, Andrew Nam, Emily Lake, Joe Nudell, Minnie Quartey, Zion Mengesha, Connor Toups, John R. Rickford, Dan Jurafsky, and Sharad Goel. 2020. *“Racial Disparities in Automated Speech Recognition.”* Proceedings of the National Academy of Sciences 117 (14): 7684–89. <https://doi.org/10.1073/pnas.1915768117>.
 → *Media coverage: New York Times, Scientific American, Business Insider, The Verge, Ars Technica, Stanford News, VentureBeat, etc.*
 → *Interactive website: fairspeech.stanford.edu*
 → *Presented at: IC2S2 2020 (Parallel Talk), Women in Data Science 2023 (Plenary, Middlebury College), NIH Voice & AI Symposium 2023 (Panel), and others*
12. **Koenecke, Allison*** and Jordi Feliu-Fabà*. 2020. *“Learning Twitter User Sentiments on Climate Change with Limited Labeled Data.”* In Workshop Proceedings of the 14th International AAAI Conference on Web and Social Media. ICWSM SocialSens. <http://doi.org/10.36190/2020.22>.
13. König, Maximilian F., Mike Powell, Verena Staedtke, Ren-Yuan Bai, David L. Thomas, Nicole Fischer, Sakibul Huq, et al. 2020. *“Preventing Cytokine Storm Syndrome in COVID-19 Using $\alpha - 1$ Adrenergic Receptor Antagonists.”* Journal of Clinical Investigation 130 (7): 3345–47. <https://doi.org/10.1172/JCI139642>.
14. **Koenecke, Allison** and Amita Gajewar. 2020. *“Curriculum Learning in Deep Neural Networks for Financial Forecasting.”* In Mining Data for Financial Applications, edited by Valerio Bitetta, Ilaria Bordino, Andrea Ferretti, Francesco Gullo, Stefano Pascolutti, and Giovanni Ponti, 11985:16–31. Cham: Springer International Publishing. <https://arxiv.org/abs/1904.12887>.
 → *Presented at: JSM 2022, ECML-PKDD MiDAS 2019 (Best Paper Award)*
15. **Koenecke, Allison**. 2019. *“A Game Theoretic Setting of Capitation versus Fee-for-Service Payment Systems.”* Edited by Baogui Xin. PLOS ONE 14 (10): e0223672. <https://doi.org/10.1371/journal.pone.0223672>.

WORKING PAPERS

1. **Koenecke, Allison**, Anna Seo Gyeong Choi, Katelyn Mei, Hilke Schellmann, Mona Sloane. 2024. *“Careless Whisper: Speech-to-text Hallucination Harms.”* arXiv. <https://arxiv.org/abs/2402.08021>.
2. Zhao, Robin, Anna Seo Gyeong Choi, **Allison Koenecke**, Anaïs Rameau. *“Quantification of Automatic Speech Recognition System Performance on d/Deaf and Hard of Hearing Speech.”*
 → *Presented at: American Laryngological Association (ALA) 2024 Annual Meeting*
3. Harvey, Emma, **Allison Koenecke**, Rene Kizilcec. *“Towards an Educator-Centered Method for Measuring Bias in Large Language Model-Based Chatbot Tutors.”*
 → *Presented at: AAAI AI4ED 2024 (Poster)*
4. Choi, Anna Seo Gyeong and **Allison Koenecke**. *“Auditing Korean Speech Data for Dialectal Representation.”*
 → *Presented at: IC2S2 2023 (Poster)*

WRITTEN GUIDES

1. **Koenecke, Allison.** GESIS Guides to Digital Behavioral Data: Expert Insights into Fairness Issues in Computational Social Science Research. (GESIS, 2023).
2. **Koenecke, Allison** and Hal Varian. Synthetic Data Generation for Economists. (AEA, 2020).
3. **Koenecke, Allison.** Deep Learning for Time Series Forecasting. (2020).
4. **Koenecke, Allison.** Sequential Pattern Mining in R for Business Recommendations. (2019).

AWARDS

Grants & Fellowships

- Cisco Research Award, 2023
- Cornell Center for Data Science for Enterprise & Society Microsoft Azure Cloud Computing Grant, 2023
- Cornell Center for Social Sciences Grant, 2022
- LinkedIn-Bowers Award, 2022
- NSF Graduate Fellowship, 2018 - 2021
- Stanford HAI Seed Research Grant, 2020
- Ben Rolfs Memorial Award, 2020
- National Physical Science Consortium Fellowship, 2016 - 2018
- Gene Golub Fellowship, 2016 - Present
- EDGE Doctoral Fellowship, 2016 - Present
- AAUW Selected Professions Fellowship (Declined), 2016
- Ruhr Fellowship, 2012

Honors

- Forbes 30 under 30 (Science), 2022
- Stanford School of Engineering Justice, Equity, Diversity, and Inclusion (JEDI) Appreciation Award, 2021
- Harvard Kennedy School Belfer Center Tech Spotlight Runner-Up, 2021
- Berkeley EECS Rising Stars, 2020
- Best Paper Award at ECML PKDD Workshop on Mining Data for Financial Applications ("*Curriculum Learning in Deep Neural Networks for Financial Forecasting*"), 2019
- Jane Street Symposium Selected Attendee, 2019
- Microsoft BUILD Conference Selected Attendee, 2018
- ICME Booster Award, 2018
- Math Prize for Girls Alumna Speaker, 2013
- Intel Science Talent Search Semifinalist, 2010
- Siemens Competition Semifinalist, 2010
- USA Math Olympiad Qualifier, 2009

TEACHING

Cornell

- INFO 4390/5390 / CS 5382 (Practical Principles for Designing Fair Algorithms): Spring 2024
- INFO 2950 (Introduction to Data Science): Fall 2022, Fall 2023
- INFO 6960 (Data Science for Global Development), Spring 2023

Other

- Stanford MS&E 298 (Detecting Discrimination in Data), Guest Lecturer (Nov 2023)
- Cornell Tech ORIE 5355 (Applied Data Science: Decision-making beyond prediction), Guest Lecturer (Oct 2023)
- Cornell Tech CS 6382 (Data Science for Social Change), Guest Lecturer (Mar 2023)
- University of Michigan SIADS 681 (Health Analytics), Guest Lecturer (May 2022)
- Stanford HISTORY 344F (Innovations in Inclusive Design), Guest Lecturer (Apr 2022)
- Cornell Tech CS 6382 (Data Science for Social Change), Guest Lecturer (Mar 2022)
- Stanford CME 500 (Computational Math in Industry and Beyond Seminar Series), Guest Lecturer (Jan 2022)
- Stanford PWR 2TB (Hip Hop, Orality, and Language Diversity), Guest Lecturer (May 2021)
- Stanford MS&E 190 (Methods and Models for Policy and Strategy Analysis), Guest Lecturer (Apr 2020)
- NERA Onboarding Training (Excel, SAS, VBA, Stata), Associate Analyst (Jul 2015)
- MIT 14.73x (The Challenges of Global Poverty, Profs. Esther Duflo and Abhijit Banerjee), Teaching Assistant (Jan - Jun 2014)

SERVICE

Mentorship

- Stanford Women in Mathematics, Statistics, and Computational Engineering (Co-Founder & Co-President, Mar 2018 - Jun 2021)
- Stanford Women in Math Mentoring (Mentor, Sep 2017 - Sep 2020)
- Bay Area Graduate Pathways to STEM (Mentor, Sep 2017 - Sep 2018)
- NERA Researcher Roundtable (Committee Member, Feb 2015 - Mar 2016)
- MIT Undergraduate Society of Women in Mathematics (President, active Sep 2011 - Jun 2014)

Reviewing

- **Conferences/Workshops:** FAccT (AC '24, AC '23, AC '22), Regulatable ML @ NeurIPS (PC '23), Machine Learning for Health (AC '23), AIES (PC '23), ACM Web Conference (FATE PC '23), EAAMO (AC '22), CHI ('22), NeurIPS (Ethics Reviewer '22, WHMD Reviewer '21), CLear ('22), ACL-IJCNLP ('21)
- **Journals:** Nature Machine Intelligence, Journal of Artificial Intelligence Research, PLOS One, Journal of Applied Psychology, Operations Research for Health Care
- **Grants:** Wellcome Trust, Social Science Research Council: Data Fluencies Dissertation Grants

TALKS & PUBLICITY

Selected Talks

- University of Michigan Data Science/Computational Social Science (DS/CSS) Seminar (Apr 2024)
- Kensho (Jan 2024)
- University of Rochester, Department of Philosophy, Goergen Institute for Data Science (Dec 2023)
- Cornell Hack4Impact Environment x Tech Panel (Nov 2023)
- NYU Center for Practice and Research at the Intersection of Information, Society, and Methodology (PRIISM) Talk (Oct 2023)
- North Carolina A&T State University: Artificial Intelligence (AI) and the Humanities/Social Sciences (Sep 2023)
- HackMIT Fireside Chat (Sep 2023)

- Cornell BURE Undergrad Research Talk Series (Aug 2023)
- Nokia Bell Labs, Responsible AI Seminar Series (Jul 2023)
- IC2S2 (Plenary, Jul 2023)
- FAccT (Jun 2023)
- ICWSM (Jun 2023)
- ICWSM—Computational Social Science Research & Data Ethics (Panelist, Jun 2023)
- American Council of Learned Societies—Artificial Intelligence (AI) and the Humanities/Social Sciences (May 2023)
- NYU Institute for Public Knowledge—Co-Opting AI: Language) (Apr 2023)
- NIH Voice & AI Symposium—Ethical and Legal Implications of Voice Biomarkers Panelist (Apr 2023)
- Women in Data Science, Alexandria (Mar 2023)
- Women in Data Science, Middlebury College (Plenary, Jan 2023)
- Joint Statistical Meetings—Advances in Time Series: Statistics Meets Machine Learning (Aug 2022)
- Toronto Workshop on Reproducibility (Feb 2022)
- MIT Operations Management Seminar (Feb 2022)
- ISCA Symposium on Security and Privacy in Speech Communication (Keynote, Nov 2021)
- Microsoft Research Summit, Causal ML Track (Oct 2021)
- Williams CS Colloquium (Oct 2021)
- Opportunity Insights Lunch Series (Oct 2021)
- Bay Area NLP (Apr 2021)
- Stanford Engineering for All (Apr 2021)
- University of Rhode Island AI Lab (Mar 2021)
- Brown Data Science Institute (Feb 2021)
- Cornell Info Sci Colloquium (Feb 2021)
- USC Marshall Operations Management (Jan 2021)
- Microsoft Research Fairness, Accountability, Transparency, and Ethics in AI (Jan 2021)
- Rutgers School of Communication and Information (Dec 2020)
- Microsoft Speech Team (Dec 2020)
- Notre Dame IT, Analytics, and Operations (Nov 2020)
- University of Michigan School of Information (Nov 2020)
- CodeX: The Stanford Center for Legal Informatics (Nov 2020)
- Berkeley EECS Rising Stars (Nov 2020)
- Google Tensorflow PM Team (Nov 2020)
- Google Economics Team (Nov 2020)
- Stanford Machine Learning Retreat (Nov 2020)
- MIT Dahleh Group (Nov 2020)
- Code for America Data Team Workshop (Oct 2020)
- University of Michigan Blablablab (Jul 2020)
- IC2S2 (Jul 2020)
- ICME Xpo 2020 (Jun 2020)
- ICWSM SocialSens Workshop (Jun 2020)
- Stanford MS&E 190 Guest Lecture (Apr 2020)
- ECML-PKDD MIDAS Workshop (Sep 2019)

Selected Media Appearances

- Cornell News (Sep 2023)
- Modern Farmer (Sep 2023)

- Cornell News (Aug 2023)
- Financial Times Specialist (Jun 2023)
- Law360 (Apr 2022)
- Women in Data Science Podcast (Aug 2021)
- Reuters (Aug 2021)
- IEEE Spectrum (Apr 2021)
- BBC: Free Thinking Podcast (Mar 2021)
- Hot & Sour Soup for the Soul Podcast (Feb 2021)
- New York Times (Mar 2020, Nov 2020)
- PNAS Q&A (Sep 2020)
- Scientific American (Jul 2020)
- 3CR Community Radio (Jun 2020)
- VentureBeat (Mar 2020)
- Business Insider (Mar 2020)
- The Verge (Mar 2020)
- Ars Technica (Mar 2020)
- Stanford News (Feb 2020, Mar 2020)