

Isabel Silva Corpus

Brooklyn, New York | isc36@cornell.edu | <https://isabelsilvacorpus.github.io>

EDUCATION

Cornell University Ann S. Bowers College of Computing and Information Science

Ph.D. in Information Science

Committee: Mor Naaman (chair), Allison Koenecke (co-chair), Angelina Wang

August 2024 - Present

New York, NY

Yale University

B.S. in Statistics and Data Science with Distinction

Honors: Distinction in Statistics and Data Science (2021)

August 2017 - May 2021

New Haven, CT

RESEARCH AREAS

Human-AI Interaction, Socio-technical Audits, Algorithmic Fairness, Computational Social Science

PUBLICATIONS

- **Isabel Corpus**, Eric Gilbert, Allison Koenecke, and Mor Naaman. "Introducing AI to an Online Petition Platform Changed Outputs but not Outcomes." [Accepted in Principle at Nature Human Behavior] <https://arxiv.org/pdf/2511.13949>
- **Isabel Corpus** and Allison Koenecke. "Into the Unknown: Accounting for Missing Demographic Data when Mitigating Ad Delivery Skew." [Accepted at FAccT 2026]
- **Isabel Corpus**, Eric Giannella, Allison Koenecke, and Don Moynihan. 2025. As Government Outsources More IT, Highly Skilled In-House Technologists Are More Essential. *Communications of the ACM* 68, 7 (July 2025), 37–40. <https://doi.org/10.1145/3727635>
- Lex Beattie, **Isabel Corpus**, Lucy Lin, and Praveen Ravichandran. 2025. Understanding Sensitive Attribute Association Bias in Recommendation Embedding Algorithms. *ACM Transactions on Recommender Systems* <https://doi.org/10.1145/3777548>
- Navin Kumar, **Isabel Corpus**, Meher Hans, et al. "COVID-19 vaccine perceptions in the initial phases of US vaccine roll-out: an observational study on reddit." *BMC Public Health* 22, no. 1 (2022): 1-14. <http://dx.doi.org/10.1186/s12889-022-12824-7>

PROFESSIONAL EXPERIENCE

Spotify, Data Scientist II New York, NY

July 2021 – December 2023

Algorithmic Responsibility

- Developed and applied methodology for disentangling social factors in high dimensional embeddings, producing forthcoming research and productionization of new metrics within Spotify to assess fairness in recommendation systems.
- Evaluated Spotify recommendation systems for representation and allocation bias. Audited interpretability tooling and methodology (e.g. SHAP, TFMA fairness indicators) using Spotify ML use cases, presented recommendations for best practices.
- Led a project to understand partisan podcasts. Applied SVM's and PCA to linearly separate high-dimensional representations of partisan podcasts, constructed a graph and created metrics to quantify diversity in listening patterns.

Machine Learning Platform

- Productionized and maintained data models using Google BigQuery and dbt for reproducible analyses.
- Prototyped ML models in coordination with teams across Spotify. Designed and led workshops attended by 300+ Spotify employees, teaching principles of machine learning and driving development of ML for novel use cases across Spotify.
- Developed strategy for ML governance at Spotify through interviewing engineers, liaising across legal and technical teams, and developing tracking for machine learning models.

INTERNSHIPS

Action Research Partners: Research Intern, Brooklyn, NY

June 2019 – August 2019

- Evaluated national child welfare systems by writing SPSS and R scripts to identify trends in welfare negligence using datasets with records of tens of millions of individuals. Presented findings to members of the Administration for Children's Services.

InsideSchools: Data Analyst Intern, New York, NY

June 2018 – August 2018

- Combined qualitative interviews and tours of schools with quantitative survey data to report on school metrics and write school reviews for InsideSchools.

Human Nature Lab: Technical Lead & Research Assistant, New Haven, CT

September 2020 – January 2021

- Created a novel Reddit dataset about COVID-19 vaccine hesitancy and modeled Reddit comments with LDA topic modeling. Organized and instrumented web scraping, data cleaning, topic modeling, data visualization in Python. Culminated in publication.
- Led technical project management through task delegation, timeline creation and organization, facilitating communication between technical and non-technical researchers.

STEM PERL Lab: Research Assistant, New Haven, CT

January 2020 – May 2020

- Contributed to research on educational best practices through data cleaning and analysis of large data sets using R. Analyzed survey data and streamlined the data cleaning process for the lab by introducing data exploration techniques.

TEACHING EXPERIENCE**Yale Statistics & Data Science: Undergraduate Learning Assistant & Course Coordinator**

January 2019 – January 2021

- Managed 20 teaching assistants and course affairs for an introductory statistics class of 400+ students. Improved the efficiency of the course by delegating tasks, creating timeline for course affairs, and designing a course roadmap.
- Spent 10 hours per week grading problem sets and holding weekly office hours.
- Courses taught: S&DS 10X: Introduction to Statistics, S&DS 230: Data Exploration and Analysis, S&DS 365: Data Mining and Machine Learning

SERVICE

Conference Organizing

- ACM FAccT 2026 Scheduling Co-Chair
- ACM FAccT 2026 Volunteer (June 2026)
- Cornell Thought Summit: LLMs and Society Volunteer (May 2025)

Broadening Participation

- Speaker: Hunter College High School AI Day (April 2026)
- Panelist: Data Science in Industry panel for Cornell graduate course “CS 5304: Data Science in the Wild” (April 2026)
- Committee member: Cornell Broadening Participation in Information Science (Spring 2025)
- Organizer and panelist: Breaking into Research Panel with Cornell URM x NSBE (Spring 2025)
- Volunteer reviewer for Cornell Student-Applicant Reading Program (SARP) (Fall 2024)
- Co-founder and organizer of the Spotify for Women in Tech Mentorship Program (80+ participants in 2023; 800 cumulative participants as of 2025). Organized career development events and matched mentors to mentees.
- Mentor for Latinas in Tech (2023-2024)

Reviewing

- Reviewer for IC2S2 2026
- Reviewer for ICWSM 2026
- Reviewer (and emergency reviewer) for FAccT 2026
- Reviewer for ACM CHI 2026
- Co-reviewer for European chapter of ACL 2023 SRW

RESEARCH TALKS & POSTERS

- Parallel talk at IC2S2 2026 (Summer 2026)
- Poster at NYU Responsible AI Day (Spring 2026)
- Parallel talk and poster at CODE@MIT (Fall 2025)
- Poster at Yale New Directions in Social Algorithms Research (Fall 2025)
- Cornell Artificial Intelligence, Policy, and Practice (AIPP) Initiative (Spring 2025)

AWARDS & FELLOWSHIPS**2026 Digital Life Initiative Doctoral Fellowship:** \$5,000 (2026)

Cornell Digital Life Initiative

2026 Frederic and Susan Rubinstein PiTech PhD Impact Fellowship: \$19,676 (2026)

Siegel Family Endowment PiTech PhD Impact Fellowship at Cornell University

Dean's Excellence and Hopper-Dean Fellowship: \$146,664 (2024)
Cornell University Ann S. Bowers College of Computing and Information Science

2026 DLI Conscientious Tech Design Workshop: \$600 (2026)
Cornell Digital Life Initiative

ACM FAccT 2026 Financial Support Award: \$640 (2026)
ACM FAccT 2026

Cornell Graduate School Conference Grant: \$300 (2025)
Cornell Graduate School

SKILLS

- Python, R, dbt, Github, SQL
- Technical Coursework: Discrete Math, Probability Theory, Theory of Statistics, Data Mining and Machine Learning, Linear Algebra, Neural Networks and Language
- Socio-technical Coursework: The Structure of Information Networks, Engineering Societal Systems, Non-Ideal Algorithmic Fairness