

# Pratyusha Ria Kalluri

PhD Candidate in the Stanford AI Lab

riakalluri.com | pkalluri@stanford.edu

## Education

Ph.D. Computer Science, Stanford, 2017-present  
Visiting Researcher, Complutense University of Madrid, 2016-2017  
B.S. Computer Science, Massachusetts Institute of Technology, 2012-2016

## Fellowships and awards

Open Philanthropy AI Fellowship, 2019  
National Science Foundation Graduate Research Fellowship\*, 2019  
\*I was selected and respectfully declined this fellowship to take a sabbatical focused on social & ethical dimensions of AI  
Stanford Enhancing Diversity in Graduate Education Fellowship, 2017  
Stanford School of Engineering Fellowship, 2017  
Paul and Daisy Soros Fellowship for New Americans, 2017  
Stanford Community Impact Award, 2021  
Best paper, ACM Conference on Fairness Accountability and Transparency (FAcCT), 2022

## Publications

The Surveillance AI Pipeline  
**P. Kalluri\***, A. Birhane\*, W. Agnew\*, M. Cheng\*, K. Owens\*  
Submitted to *Nature*. In review.

Dialect prejudice predicts AI decisions about people's character, employability, and criminality  
V. Hoffman, **P. Kalluri**, D. Jurafsky, S. King  
In preparation.

Algorithms and humans censor self-reports of discrimination  
C. Lee\*, K. Gligorić\*, **P. Kalluri\***, E. Durmus\*, M. Perry\*, N. San, D. Tse, D. Jurafsky, J. Eberhardt  
In preparation.

Modern machine learning models reify essentialist racial and gender categories  
**P. Kalluri\***  
In preparation.

Vision-language models merge perceptual and semantic knowledge  
**P. Kalluri\***, M. Yuksekgonul\*, J. Zhou, D. Jurafsky  
In preparation.

Easily Accessible Text-to-Image Generation Amplifies Demographic Stereotypes at Large Scale  
**P. Kalluri\***, F. Bianchi\*, E. Durmus\*, F. Ladhak\*, M. Cheng\*, D. Nozza, T. Hashimoto, D. Jurafsky, J. Zou, A. Caliskan  
ACM Conference on Fairness Accountability and Transparency (FAcCT), 2023.

The Values Encoded in Machine Learning Research  
**P. Kalluri\***, A. Birhane\*, D. Card\*, W. Agnew\*, R. Dotan\*, M. Bao\*  
**Best paper.** ACM Conference on Fairness Accountability and Transparency (FAcCT), 2022.

When and Why Vision-Language Models Behave like Bag-of-words, and what to do about it?  
M. Yuksekgonul, F. Bianchi, **P. Kalluri**, D. Jurafsky, J. Zou

International Conference on Learning Representations (ICLR), 2023.

The Underlying Values of Machine Learning Research

**P. Kalluri\***, A. Birhane\*, D. Card\*, W. Agnew\*, R. Dotan\*, M. Bao\*

Conference on Neural Info. Processing Systems (NeurIPS), Resistance AI Workshop, 2020.

Don't ask if artificial intelligence is good or fair, ask how it shifts power

**P. Kalluri**

Nature, 2020.

If We're Not Careful, Tech Could Hurt the Fight against COVID-19

**P. Kalluri**, L. Gillespie, A. Foryciarz, W. Elhai, S. Srivastava, A. Panezi, L. Einstein

Scientific American, 2020.

Learning Controllable Fair Representations

J. Song\*, **P. Kalluri\***, A. Grover, S. Zhao, S. Ermon

International Conference on Artificial Intelligence and Statistics (AISTATS), 2019.

\*Co-authors. In many cases, co-authors have made incredible and incomparable contributions, especially incomparable given the interdisciplinarity of these significant contributions, and in these cases we collectively and enthusiastically follow the convention of each author listing themselves first in their own curriculum vitae.

## Teaching

Instructor • Stanford "*Ethical and Social Issues in Natural Language Processing*", 2023

Guest lecturer • Elisava Barcelona School of Design & Eng., *Master in Design for Responsible AI*, 2023

Guest lecturer • Johns Hopkins University "*AI Ethics and Social Impact*", 2023

Guest lecturer • Stanford "*Design for AI*", 2023

Guest lecturer • Stanford "*Human-Centered Natural Language Processing*", 2023

Guest lecturer • U'Mista Cultural Centre, 2023

Guest lecturer • Stanford "*Design for AI*", 2022

Guest lecturer • City College of San Francisco "*Build Abolition 101*", 2021

Mentor • Stanford research mentor, 2021-2022

Guest lecturer • Princeton "*Black Mirror: Race, Technology, and Justice*", 2020

Facilitator • FAccT CtrlZ.AI "*Our Relationship with AI*" workshop, 2020

Facilitator • Stanford Women in AI series, 2019

Facilitator • Stanford Humans of AI series, 2019

Mentor • Stanford STEM Fellows mentor, 2019

Guest lecturer • MIT "*Introduction to Artificial Intelligence*", 2016

Teaching Assistant • MIT "*Mathematics for Computer Science*", 2014

Teaching Assistant • MIT "*Introduction to Electrical Engineering and Computer Science*", 2013

## Service

Stanford Community Impact Award, 2021

Resistance AI Workshop, organizer and speaker, 2020

FAccT CtrlZ.AI - Our Relationship with AI workshop, organizer, 2020

NeurIPS, top 5% of reviewers, 2020

NeurIPS Queer in AI, speaker, 2019

Stanford Inclusion in AI, creator and organizer, 2019

Stanford Humans of AI, creator and organizer, 2019

Stanford Queer in AI, organizer, 2019

Stanford Women in AI, organizer, 2019  
Stanford STEM Fellows, mentor, 2019  
ICML Queer in AI, speaker, 2019  
Radical AI Network, creator and organizer, 2019  
Stanford NLP Retreat, organizer, 2018  
MIT LGBTQ+ Rainbow Lounge, contributor, 2016  
MIT LGBTQ+ Affiliated, organizer and speaker, 2016  
MIT Class Vice President, 2013  
MIT Think Inspire Motivate Talk, “Women Who Have Inspired Me”, speaker, 2013  
MIT CityDays institute-wide service day, organizer, 2012  
Bezos “Be the Change” Local Ideas Festival, organizer, 2012

## Talks

AI and power  
Elisava Barcelona School of Design & Eng., Master in Design for Responsible AI, 2023

AI and power  
Johns Hopkins University, AI Ethics and Social Impact, 2023.

Artificial intelligence: A story about values, politics, and power  
Stanford, Design for AI, 2023.

AI and colonialism  
U'Mista Cultural Centre, 2023.

Machine learning and power  
Stanford, Human-centered Natural Language Processing, 2023.

Easily Accessible Text-to-Image Generation Amplifies Demographic Stereotypes at Large Scale  
ACM Conference on Fairness Accountability and Transparency (FAcCT), 2023.

Biased Perspective, Unequal Effects: AI & Healthcare  
National Science Ethics and Policy Symposium, 2022.

Biased Intelligence  
NPR WHYY, AI Nation, 2021.

Beyond Fairness in Machine Learning  
Conference on Neural Information Processing Systems (NeurIPS), 2021.

From fairness to justice: centering marginalized communities  
IEEE / CVF Computer Vision and Pattern Recognition Conference (CVPR), 2021.

Abolition and the Tech Industry  
City College of San Francisco (CCSF), Build Abolition 101, 2021

Demystifying AI: How AI Works and How AI Does Not Work

Princeton, Black Mirror: Race, Technology, and Justice, 2020.

Designing Machine Learning Models That Shift Power  
Columbia University, The American Assembly, We Be Imagining, 2020

Future of Artificial Intelligence: Bias, Policy, and More  
UCSF, Hard Questions, 2020.

Our Relationship with AI: Exploring the Present and Dreaming up Radical Futures  
ACM Conference on Fairness Accountability & Transparency (FAccT), CtrlZ.AI, 2020.

The Hierarchy of Meaning in GPT-2  
OpenAI, 2019.

How Machine Learning Research Shifts Power  
Conference on Neural Information Processing Systems (NeurIPS), 2019.

Fairness, Accountability, Transparency, and Ethics and How it Affects the Queer Community  
International Conference on Machine Learning (ICML), 2019.

Inclusion in AI  
Stanford, Inclusion in AI, 2019.

Fairness and Ethics in AI  
Y Combinator AI Conference, 2018.

The Paradox of Privacy  
PD Soros Fellowship Conference, 2018.

Science Intuition: The Higgs Boson  
Best presenter. MIT, Science Intuition Conference, 2015

## In the news

*A sample of news coverage.*

AI image generators show gender and racial bias  
CBS News Prime Time, 2023

This is how AI image generators see the world  
The Washington Post, 2023

How the “Surveillance AI Pipeline” Literally Objectifies Human Beings  
404 Media, 2023

Towards a Standard for Identifying and Managing Bias in Artificial Intelligence  
National Institute for Standards and Technology, 2022

Mozilla Internet Health Report: AI, in Real Life

Mozilla, 2022

Study finds that few major AI research papers consider negative impacts  
VentureBeat, 2021

The term 'ethical AI' is finally starting to mean something  
VentureBeat, 2020

The battle for ethical AI at the world's biggest machine-learning conference  
Nature, 2020

An AI conference once known for blowout parties is finally growing up  
MIT Tech Review, 2019

Three MIT alumni win 2017 Paul and Daisy Soros Fellowships for New Americans  
MIT News, 2017

Priya's Freedom to Give Back  
The Huffington Post, 2017