Inioluwa Deborah Raji

(+01) 925-858-4491
rajiinio@berkeley.edu

Education

Electrical Engineering and Computer Science, University of California, Berkeley
Berkeley, U.S.A.

Ph.D. in Computer Science 2021 - 2026 *(expected)*
- Advisor: Prof. Benjamin Recht

Engineering Science, University of Toronto Toronto, Canada
Bachelors of Applied Science in Engineering Science (Robotics Engineering) 2014 - 2019

Notable Awards and Honors

- **Tech for Humanity Prize**, New America 2024
- **Rise25 Award**, Mozilla Foundation 2024
- **Time 100 Most Influential People in AI**, Time magazine 2023
- **AI 100**, Business Insider 2023
- **Forbes 30 Under 30**, Forbes 2021
- **Pioneer/ Barlow Award**, Electronic Frontier Foundation (EFF) 2020
- **35 Under 35 Innovator Award**, MIT Tech Review 2020
- **AI Innovation Award for "AI for Good"**, Venture Beat 2019
- **Best Paper Awards:** AAAI/ACM AI, Ethics & Society (AIES) conference (2019), ML-Retrospectives, Surveys & Meta-Analyses (ML-RSA) workshop at NeurIPS (2020), IEEE Conference on Secure and Trustworthy Machine Learning (2024)

Fellowships and Grants

- **Technology Public Interest X-Grant**, MacArthur Foundation ($500,000) 2022
- **Senior Technology Fellowship**, Mozilla Foundation ($120,000 annually) 2021-2025

Academic Service

Main Organizer for: *Simons Institute workshop on "Bridging Prediction and Intervention Problems in Social Systems" [2026], Banff International Research Station (BIRS) for Mathematical Innovation and Discovery workshop on "Bridging Prediction and Intervention Problems in Social Systems" workshop [2024], SSRC Data Fluencies Research Workshop [2024], ICML Workshop on "Participatory Approaches to Machine Learning"[2020], NeurIPS Workshop on "Navigating the Broader Impacts of AI Research" [2020], REALML workshop [2019, 2021],* **UC Berkeley Initiatives** - *Algorithms, Data & Society (ADS) Reading Group [Weekly; 2022-present], AI & Society Initiative [Bi-weekly; 2024-present]*

Organizing Committee: ***Steering Committee***, *Fairness, Accountability & Transparency (FAccT) Conference [2021-2025];* ***Ethics Review Chair***, *Neural Information Processing Systems (NeurIPS) Conference [2021, 2022];* ***Publicity co-chair***, *Fairness, Accountability & Transparency Conference (FAccT) [2021, 2022]*

Area Chair for: *Fairness, Accountability & Transparency Conference (FAccT) [Data and Algorithm Evaluation; 2021-present]; Neural Information Processing Systems (NeurIPS) Conference [Datasets & Benchmarks track; 2023-present] ,*

Reviewer for: ***Journals*** *- Nature, Cell Patterns, Big Data & Society;* ***Conferences*** *- International Conference of Machine Learning (ICML), Neural Information Processing Systems (NeurIPS), ACM Fairness, Accountability & Transparency (FAccT) conference, AAAI/ACM Artificial Intelligence, Ethics, & Society (AIES) conference;* ***Funders*** *- Mozilla Technology Fund (MTF), Canadian Institute for Advanced Research (CIFAR)*

Lead, Open Source Audit Tooling (OAT) team | Mozilla Foundation   2022 - 2025
- Recruited and led a team of 3 research collaborators and 2 research assistants to lead Mozilla's Open Source Audit Tooling agenda .
- Published two academic peer-reviewed papers, one of which won a Best Paper award. Submitted 3 policy comments, two of which were cited in follow up reports by the US NTIA and the European Commission; gave presentations to US Federal Trade Commission (FTC), UK Ofcom, and OECD.
- OAT members advising selection in two rounds of Mozilla Technology Fund (MTF): 5 teams in inaugural cohort (2022); 8 teams in follow-up cohort (2023) focused on AI audit tooling.

Trustworthy AI Research Fellow | Mozilla Foundation   2021-2023
- Independent researcher working on the topics of algorithmic accountability and AI evaluation.

Graduate Fellow | Stanford RegLab   June 2022- Present
- Working with legal collaborators on regulatory proposals for algorithm auditing. Supervised by Prof. Daniel Ho.

Harms Research Fellow | Algorithmic Justice League   Oct 2020 - Sept 2021
- Developing projects on designing reporting systems for harms discovery in algorithmic auditing.

Technology Fellow | AI Now Institute, New York University   Oct 2019 - Oct 2020
- Research and policy work in facial recognition and AI accountability practice.

Research Fellow & Affiliate | Partnership on AI   June 2019 - Jan 2020
- Setting transparency standards, benchmarking norms and legal conditions for AI deployment.

Student Researcher | Google AI   June 2018 - Nov 2019
- Selected for 12 month Google AI Research Mentorship Program, paired with the Ethical AI team.

Advisory Roles

AI Advisory Working Group
Federation of American Scientists (FAS) — 2025 - Present

Safe and Secure AI Advisory Group
Canadian Artificial Intelligence Safety Institute(CAISI) — 2025 - Present

SAIGE Council
Partnership on AI (PAI) — 2025 - Present

OECD AI Advisory Expert Network
OECD AI Policy Observatory — 2024 - Present

Network of Experts for UNSG's AI Advisory Body
United Nations High-Level Advisory Body on Artificial Intelligence — 2024 - Present

AI Policy and Governance Working Group
Science, Technology, and Social Values (ST&SV) Lab, Institute of Advanced Study — 2023 - Present

Center for Civil Rights and Technology Advisory Council
Leadership Conference on Civil and Human Rights — 2023 - Present

TeachAI Advisory Committee
Practitioner network led by Code.org, ETS, ISTE, and Khan Academy. — 2023 - Present

CDT AI Governance Lab Advisory Committee
Center for Democracy and Technology AI Governance Lab — 2023 - Present

RealML Advisory Committee
Practitioner network of AI accountability advocates — 2023 - Present

Health AI Partnership (HAIP) Advisory Council
Practitioner network led by Duke Health and Mayo Clinic — 2023 - Present

Health AI Partnership (HAIP) Leadership Council
Practitioner network led by Duke Health and Mayo Clinic — 2022 - 2023

Select Policy Engagements

Public Interest AI Working Group, Paris AI Action Summit
French Government, Envoy to the AI Action Summit — 2024 - 2025

Principal Participant, AI Academic Roundtable
Organized by United States House of Representative Ro Khanna — 2024

Speaker, Conference on Frontier AI Safety Frameworks
Organized by U.K. AI Safety Institute — 2024

Participant, Inaugural Convening of the International Network of AI Safety Institutes
Organized by International Network of AI Safety Institutes, chaired by U.S. AISI — 2024

Panelist, Artificial Intelligence Hearing
Office of the National Coordinator for Health IT (ONC) — 2024

Delegate, AI Safety Summit
U.K. Government, Department for Science, Innovation and Technology — 2023

Principal Participant, Civil Rights Roundtable
Organized by United States Senator Cory Booker — 2023

Speaker, Congressional Briefing on Artificial Intelligence and Civil Rights
Organized by Lawyer's Committee For Civil Rights — 2023

Principal Participant, Inaugural AI Insight Forum
Organized by United States Senate Leader Chuck Shumer — 2023

TEACHING & COURSE DEVELOPMENT

"Case Studies in Prediction, AI, and Public Policy" (PP290), | Fall 2025 (Upcoming)
- UC Berkeley, Goldman School of public policy course, open both to policy (e.g., MPP and MPA) students interested in data science as well as data science, EECS and statistics PhD students and advanced undergraduate students interested in policy.
- Developed full syllabus and teaching materials (teaching slides, reading assignments, course evaluation, class pre-read) ; will manage grading, and office hours.
- Invited and will manage roster of 7 guest speakers throughout the course.
- Will lecture regularly, with co-Instructor Prof. Avi Feller, about once a week.

"Machine Learning Evaluation" (CS281B/Stat241B) | Spring 2025
- UC Berkeley, Electrical Engineering and Computer Science (EECS) graduate course for CS/Stats/Engineering Masters and PhD students.
- Developed full syllabus and teaching materials (teaching slides, reading assignments, course evaluation); managed grading, and office hours.
- Lectured regularly, with co-Instructor Prof. Benjamin Recht, about once a week.

SELECT GUEST LECTURES

My work has been featured in a variety of syllabi across a wide range of courses, and I am often invited to give guest lectures as a result. Here is a sample of recent guest lectures:

"Decision-making beyond Prediction" course
Cornell University   Fall 2025 (Upcoming)

AI Policy Summer School
Brown University   Summer 2025

Computational Social Science Summer School
Stanford University   Summer 2025

"Ethical Machine Learning in Human Deployments" course
MIT EECS   Spring 2024

"Quantitative Evaluation of Language Technologies" course
Carnegie Mellon University   Spring 2024

"Data Science for Social Change" course
University of California, Berkeley   Fall 2024

"Ethics of NLP" seminar
UCLA   Fall 2024

"Responsible AI Innovation & Management" course
UC Berkeley   Spring 2023

"Practical Principles for Designing Fair Algorithms" course
Cornell University   Fall 2023

"Free Speech, Democracy and the Internet" course
Stanford Law School   Fall 2022

"Internet Governance and Human Rights" course
Columbia University School of International and Public Affairs   Spring 2022

"Special Topics in Computer Science: Algorithm Audits" course
Northeastern University   Spring 2021

"SPRING POLICY LAB: AI and Implicit Bias" course
University of Pennsylvania   Spring 2021

SELECT POLICY
REPORTS &
COMMENTS

1. **Co-Author**: International AI Safety Report, 2025 (with UK Government, Department for Science, Innovation and Technology); The California Report on Frontier AI Policy, 2025 (commissioned by Governor Newsom as part of the Joint California Policy Working Group on AI Frontier Models); A Different Approach to AI Safety: Proceedings from the Columbia Convening on AI Openness and Safety, 2025 (Working group member); Beyond open vs. closed: Emerging consensus and key questions for foundation AI model governance, 2024 (with Carnegie Endowment for International Peace), and more.

2. **Led policy comment submissions with Mozilla Open Source Audit Tooling (OAT) team**: US National Telecommunication and Information Administration (NTIA) Comment (cited several times in their following "Artificial Intelligence Accountability Policy" Report!); US National Institute on Standards & Technology (NIST) Comment; EU Digital Service Act (DSA), Article 40 Comment (cited several times in report for Delegated Act on data access provided for in Digital Services Act).

3. Participated in policy comment submissions on open source, AI accountability and third party auditor access, with the IAS ST&SV Lab AI Policy and Governance Working Group, Stanford HAI, Carnegie Endowment for International Peace (CEIP), and Mozilla.

Work cited in the following select government reports, letters and official communications: *Economic Report of the President (2024); Letter from House Committee on House Committee on Science, Space, and Technology (2023), House Task Force on Artificial Intelligence Final Report (2024) The Bipartisan Senate AI Working Group report (2024); NIST AI Risk Management Framework Playbook (2023); NIST's Trustworthy and Responsible AI Resource Center (2023); National Institute of Standards and Technology (NIST) Face Recognition Vendor Test (FRVT) Part 3: Demographic Effects (2019) and more.*

SELECT
PUBLICATIONS

Full list of publications is available on my Google Scholar page.

Key research themes: Internal AI auditing & documentation; External AI auditing & evaluation practices under limited access; Re-thinking machine learning evaluation methods; Designing a policy ecosystem for effective AI auditing; Understanding AI functionality in deployment settings; Research ethics in machine learning.

Inioluwa Deborah Raji and Joy Buolamwini. Actionable auditing: Investigating the impact of publicly naming biased performance results of commercial ai products. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, pages 429–435, 2019

Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timnit Gebru. Model cards for model reporting. In *Proceedings of the conference on fairness, accountability, and transparency*, pages 220–229, 2019

Alice Xiang and Inioluwa Deborah Raji. On the legal compatibility of fairness definitions. In *Workshop on Human-Centric Machine Learning at the 33rd Conference on Neural Information Processing Systems*, 2019

Inioluwa Deborah Raji and Jingying Yang. About ml: Annotation and benchmarking on understanding and transparency of machine learning lifecycles. In *Human-Centric Machine Learning workshop at Neural Information Processing Systems conference 2019*, 2019

Inioluwa Deborah Raji, Andrew Smart, Rebecca N White, Margaret Mitchell, Timnit Gebru, Ben Hutchinson, Jamila Smith-Loud, Daniel Theron, and Parker Barnes. Closing the ai accountability gap: Defining an end-to-end framework for internal algorithmic auditing. In *Proceedings of the 2020 Conference on Fairness, Accountability and Transparency*, pages 33–44, 2020

Inioluwa Deborah Raji, Timnit Gebru, Margaret Mitchell, Joy Buolamwini, Joonseok Lee, and Emily Denton. Saving face: Investigating the ethical concerns of facial recognition auditing. In *Proceedings of the 2020 AAAI/ACM Conference on AI, Ethics, and Society*, pages 145–151, 2020

Inioluwa Deborah Raji. The discomfort of death counts: mourning through the distorted lens of reported covid-19 death data. *Patterns*, 1(4), 2020

Inioluwa Deborah Raji. Handle with care: lessons for data science from black female scholars. *Patterns*, 1(8), 2020

Amandalynne Paullada, Inioluwa Deborah Raji, Emily M Bender, Emily Denton, and Alex Hanna. Data and its (dis) contents: A survey of dataset development and use in machine learning research. *Patterns*, 2(11), 2021

Inioluwa Deborah Raji and Genevieve Fried. About face: A survey of facial recognition evaluation. In *AAAI 2020 Workshop on AI Evaluation*, 2021

Inioluwa Deborah Raji, Morgan Klaus Scheuerman, and Razvan Amironesei. You can't sit with us: Exclusionary pedagogy in ai ethics education. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*, pages 515–525, 2021

Thomas Liao, Rohan Taori, Inioluwa Deborah Raji, and Ludwig Schmidt. Are we learning yet? a meta review of evaluation failures across machine learning. In *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2)*, 2021

Inioluwa Deborah Raji, Emily M Bender, Amandalynne Paullada, Emily Denton, and Alex Hanna. Ai and the everything in the whole wide world benchmark. In *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track*, 2021

BOOK CHAPTERS	1. **Inioluwa Deborah Raji.** "The Anatomy of AI Audits: Form, Process, and Consequences." *The Oxford handbook of AI governance*, edited by Justin B. Bullock, Yu-Che Chen, Johannes Himmelreich, Valerie M. Hudson, Anton Korinek, Matthew M. Young, Baobao Zhang, Oxford University Press (2024).

1. **Inioluwa Deborah Raji.** "The Anatomy of AI Audits: Form, Process, and Consequences." *The Oxford handbook of AI governance*, edited by Justin B. Bullock, Yu-Che Chen, Johannes Himmelreich, Valerie M. Hudson, Anton Korinek, Matthew M. Young, Baobao Zhang, Oxford University Press (2024).

2. **Inioluwa Deborah Raji**, Danaë Metaxa. "A Tale of Two Audits." *AI and Assembly: Coming Together and Apart in a Datafied World*, edited by Toussaint Nothias and Lucy Bernholz, Stanford University Press (2025).

3. **Inioluwa Deborah Raji**, Sasha Costanza Chock, Joy Buolamwini. "Change from the outside: towards credible third-party audits of AI systems." *Missing links in AI governance*, edited by Edited by Benjamin Prud'homme, Catherine Régis, Golnoosh Farnadi, Vanessa Dreier, Sasha Rubel, Charline d'Oultremont, United Nations Educational, Scientific and Cultural Organization (UNESCO) (2023).

4. **Inioluwa Deborah Raji.** "Facing the Tech Giants." *The Black Agenda: Bold Solutions for a Broken System*, edited by Anna Gifty Opoku-Agyeman, St. Martin's Publishing Group (2022).

5. **Inioluwa Deborah Raji.** "The bodies underneath the rubble." *Fake AI*, edited by Frederike Kaltheuner, Meatspace Press (2021).

Work featured in the following books: *Genius Makers: The Mavericks Who Brought AI to Google, Facebook, and the World (Cade Metz, 2022); Your Face Belongs to Us: A Tale of AI, a Secretive Startup, and the End of Privacy (Kashmir Hill, 2024); Unmasking AI: My Mission to Protect What Is Human in a World of Machines (Joy Buolamwini, 2024); The AI Con: How to Fight Big Tech's Hype and Create the Future We Want (Dr. Emily M. Bender, Dr. Alex Hanna, 2025); Empire of AI: Dreams and Nightmares in Sam Altman's OpenAI (Karen Hao, 2025); More Everything Forever: AI Overlords, Space Empires, and Silicon Valley's Crusade to Control the Fate of Humanity (Adam Becker, 2025), and more.*

SELECT NAMED LECTURES & SEMINARS

Distinguished Speaker series	
University of Colorado Boulder	Fall 2025 (Upcoming)
William Pierson Field Lecture	
Princeton University	Spring 2024
Distinguished Speaker series	
Harvard Kennedy School Program in Criminal Justice Policy and Management	Spring 2024
Fred Kan Distinguished Lecture	
University of Toronto	Fall 2023
Speaker, AI.Humanity seminar	
Emory University	Fall 2023
Speaker, VCAI colloquium	
University of Maryland	Fall 2023
Speaker, AI Cyber Lunch	
Harvard Kennedy School	Spring 2023
Speaker, Yale ISP Ideas Lunch	
Yale Law School	Spring 2023
Distinguished Speaker Seminar Series	
Yale Center for Biomedical Data Science (CBDS) at Yale School of Medicine	Spring 2021

Select Conference Keynotes, Panels & Invited Talks

Keynote
Canadian AI Safety Institute (CAISI) Research Program Annual Meeting 2025 (Upcoming)

Panelist, "The States Shaping AI's Fate: How Local Leadership Is Shaping Global Governance"
Berkeley Public Policy Conference 2025 (Upcoming)

Panelist, "AI and geopolitical power"
Financial Times Future of AI Summit 2025 (Upcoming)

Keynote, "Understanding and Addressing Gender Bias in Artificial Intelligence Systems"
United Nations Women AI-Gender Academic Conference 2025

Panelist
Conference on Health, Inference, and Learning (CHIL) 2025

Panelist, "How To Put The Missing Human Values Back Into AI"
American Medical Informatics Association 2025

Speaker, "A Primer on Algorithm Bias"
GovAI Coalition Summit 2025

Keynote, "Safety, By Any Other Name"
Next Generation of AI Safety workshop, ICML 2024

Keynote
Conference on Health, Inference, and Learning (CHIL) 2024

Panelist
National Artificial Intelligence Advisory Committee (NAIAC) Panel on AI Safety 2024

Keynote
State Legislative Leaders Foundation Conference 2024

Keynote
2nd IEEE Conference on Secure and Trustworthy Machine Learning 2024

Panelist
Workshop on Regulating Machine Learning, NeurIPS 2023

Keynote, "Grounded Evaluations for Assessing Real-World Harms"
Socially Responsible Language Modelling Research (SoLaR) Workshop, NeurIPS 2023

Keynote
ACM/AAAI AI, Ethics and Society (AIES) Conference 2022

Keynote, "Ethical Challenges of Data Collection & Use in Machine Learning Research"
DataPerf: Benchmarking Data for Data-Centric AI, ICML 2022

Panel Moderator, "How Should a Machine Learning Researcher Think About AI Ethics?"
Neural Information Processing Systems (NeurIPS) 2021

Keynote, "Radical Proposal: Third-Party Auditor Access for AI Accountability"
2021 HAI Fall Conference on Policy & AI 2021

Keynote, "The Need for Ethical Oversight in Machine Learning Research"
Journal of Opportunities, Unexpected limitations, Retrospectives, Negative results, and Experiences, MLSys 2021

Keynote, "When AI & Advocacy Collide"
Neural Information Processing Systems (NeurIPS) 2020

Keynote, "AI Resistance and the Five Stages of Corporate Grief"
Resistance AI Workshop, NeurIPS 2020

Keynote, "Classic examples of engineering responsibility"
MLRetrospectives: A Venue for Self-Reflection in ML Research workshop, ICML 2020

Keynote, ""Debugging" Discriminatory ML Systems"
Debugging Machine Learning Models workshop, ICLR 2019