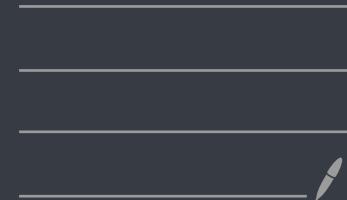


INCLUSIVE DESIGN



SPEAKERS:



WEIWEI PAN
RESEARCH
ASSOCIATE



MATTHEW FINNEY
GRADUATE STUDENT
ADVISORY COMMITTEE



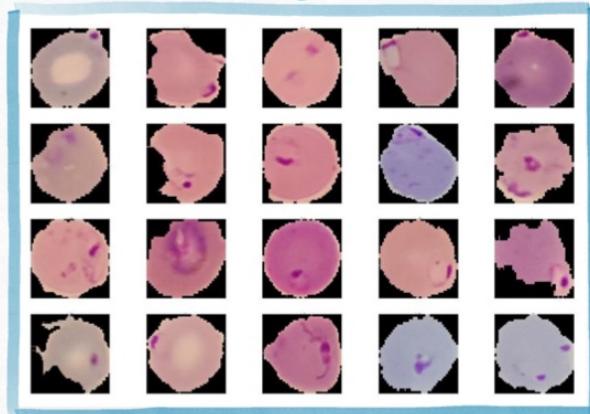
SCARLETT GONG
GRADUATE STUDENT
ADVISORY COMMITTEE



PAUL TEMBO
GRADUATE STUDENT
ADVISORY COMMITTEE

THE POTENTIAL OF MACHINE LEARNING:

Disease Diagnosis



Detecting malaria from blood smears

Drug Discovery



Quickly discovering new drugs for COVID

Agriculture



Precision agriculture for changing Climate

Urban Planning



Predicting and planning for resource needs

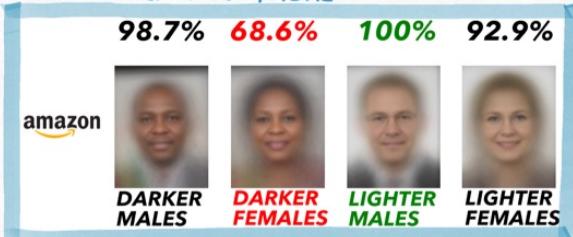
EXAMPLES OF DATA SCIENCE FAILS:

PREDICTIVE POLICING



PREDICTIVE POLICING ALGORITHMS ARE RACIST.
THEY NEED TO BE DISMANTLED, MIT TECHNOLOGY
REVIEW 2020

FACIAL RECOGNITION SYSTEMS



FACIAL RECOGNITION IS ACCURATE, IF YOU'RE A
WHITE GUY, NEW YORK TIMES, 2018

PRE-TRIAL, PAROLE RISK ASSESSMENT

VERNON PRATER

Prior Offenses
2 armed robberies, 1 attempted armed robbery

Subsequent Offenses
1 grand theft

LOW RISK

3

HIGH RISK

8

BRISHA BORDEN

Prior Offenses
4 juvenile misdemeanors

Subsequent Offenses
None

INJUSTICE EX-MACHINA: PREDICTIVE ALGORITHM
IN CRIMINAL JUSTICE, UCLA LAW REVIEW, 2019

PRECISION HEALTH CARE



SEX AND GENDER DIFFERENCES AND BIASES IN
ARTIFICIAL INTELLIGENCE FOR BIOMEDICINE
AND HEALTH CARE, NATURE, 2020

A LONG HISTORY OF DESIGN FAILS:

PROBLEMS WITH DATA SCIENCE IS NOT AN EXCEPTION, NOT A GLITCH, IT'S PART OF A LONG HISTORY OF **SYSTEMATIC** DESIGN & DEPLOYMENT FAILS OF NON-DIGITAL TECH.

RACE AFTER TECHNOLOGY, RUHA BENJAMIN



- PHOTOGRAPHY IS A TECH OF SUBJECTIVE DECISIONS
- CALIBRATION OF COLOR PHOTOS USES A "SHIRLEY" CARD REFERENCE WITH WHITE WOMAN
- RENDER DARKER SKIN FEATURES BLURRY
- 1960'S COMPLAINTS FROM BLACK PARENTS ABOUT POOR QUALITY SCHOOL PHOTOS TO KODAK
- 1960'S, 70'S COMPLAINTS FROM FURNITURE & CHOCOLATE COMPANIES FINALLY CHANGED COLOR CALIBRATION PROCEDURES

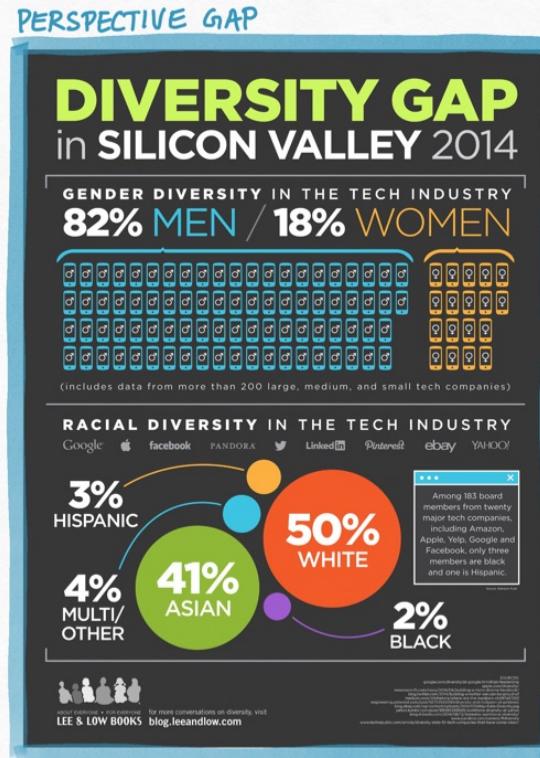
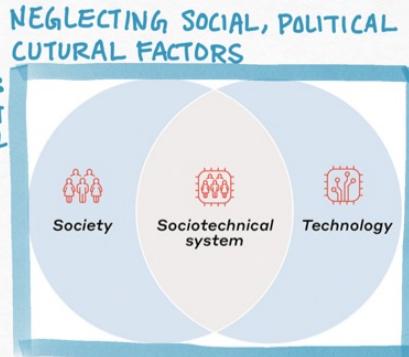
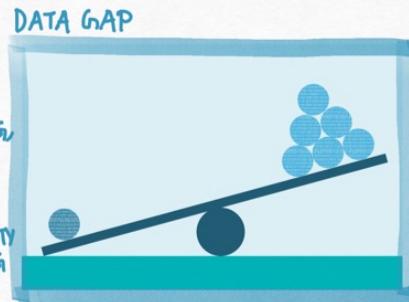
INVISIBLE WOMEN, CAROLINE CRIADO PEREZ



- OVER 80% OF GLOBAL POPULATION STILL USE TRADITIONAL STOVES & BIOMASS FUEL.
- TRADITIONAL STOVES GIVE OFF TOXIC FUMES, EQUIVALENT TO >100 CIGARETTES A DAY IN UNVENTILATED ROOM.
- TRADITIONAL STOVES ARE INEFFICIENT, WOMEN COOK FOR 3-7 HOURS A DAY.
- EFFORTS TO DESIGN & POPULARIZE EFFICIENT, CLEAN STOVES STARTED IN 1950'S, LARGELY RESULTED IN FAILURES

WHY IS THIS HAPPENING?

BIAS IN TECH IS NOT AN ALGORITHMIC, MATH PROBLEM; IT IS A PEOPLE PROBLEM, IT IS ABOUT NEGLECTING THE HUMAN DIMENSION AND INEXTRICABLY LINKED WITH DIVERSITY & REPRESENTATION.



- DESIGN CHOICES ARE NOT NEUTRAL NOR OBJECTIVE BUT ENCODE VALUES AND PERSPECTIVES
- PRIORITIZE TECHNICAL PARAMETERS OVER USER NEEDS
- COMMUNITIES THAT ARE IMPACTED BY TECH ARE NOT REPRESENTED AS DESIGNERS, POLICY MAKERS, ENGINEERS

RACE AFTER TECHNOLOGY,
RUMA BENJAMINE
INVISIBLE WOMEN,
CAROLINE CRIADO PEREZ

INVISIBLE WOMEN READING:

- PERSPECTIVE GAP
- DATA GAP
- NEGLECT OF SOCIO-CULTURAL FACTORS

DOES FAIRNESS RESEARCH FIX IT?

FAIRNESS DEFINED MATHEMATICALLY ABSENT REAL-LIFE CONTEXT AND ENGAGEMENT OF STAKEHOLDERS CAN REINFORCE EXISTING BIASES AND INJUSTICES: FAIRNESS WASHING

BEYOND BIAS: REIMAGINING THE TERMS OF ETHICAL AI IN CRIMINAL LAW



- FAIRNESS RESEARCH HAPPENING IN ACADEMIA, MAJOR TECH CORPS, FAR FROM Affected COMMUNITIES.

- FOCUS ON TECHNICAL DEFINITION OF PROBLEM, E.G. DATA IMBALANCE, SUBGROUP DISPARITY, AND TECHNICAL SOLUTIONS, NOT ON SOCIO-TECH SYSTEM:

- ↳ DEFINITION OF CRIME: STREET V.S. WHITE COLLAR
 - ↳ CONFLATION OF ARREST & DANGER
 - ↳ RACIAL DISPARITIES IN POLICING, ARREST, CHARGING, LEGAL REPRESENTATION SENTENCING PRACTICES

ETHICAL LIMITS OF ALGORITHMIC FAIRNESS SOLUTIONS IN HEALTH CARE MACHINE LEARNING



- FAIRNESS METRICS DO NOT TAKE INTO ACCOUNT COMPLEX CAUSAL RELATIONSHIPS BTW BIOLOGICAL ENVIRONMENTAL & SOCIAL FACTORS OF MEDICAL CONDITIONS.
- DIFFERENCE DOES NOT ENTAIL INEQUITY IN MEDICAL TREATMENT. BLINDING DECISIONS TO PROTECTED ATTRIBUTES NOT ALWAYS GOOD
- FAIRNESS METRICS CAN FORCE ALIGNMENT BTW HETEROGENEOUS GROUPS
- FAIR DOES NOT MEAN HARMLESS, WHICH IS THE FIRST GOAL OF MEDICINE

BETTER DESIGN THROUGH PARTICIPATION:

MACHINE LEARNING IS MAKING A MOVEMENT AWAY FROM TRADITIONAL DESIGN FRAMEWORKS.

QUESTIONS FOR ETHICAL ML

I. IDENTIFY THE STAKEHOLDERS

- ↳ WHO ARE THE USERS?
- ↳ WHO ARE THE AFFECTED COMMUNITIES?

II. WHAT TYPES OF HARM CAN YOUR TECH DO?

- ↳ WHAT KIND OF HARM CAN TECH FAILURES CAUSE
- ↳ WHAT KIND OF HARM CAN THE SOCIO-TECHNICAL SYSTEM CAUSE

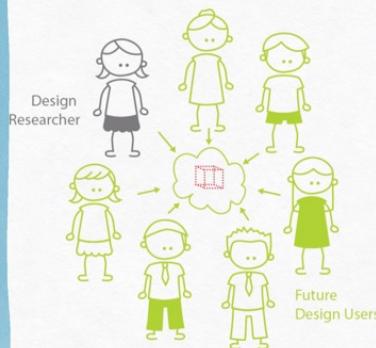
III. WHAT TYPES OF GOOD CAN YOUR TECH DO?

- ↳ WHAT KINDS OF NEEDS DO YOUR USERS HAVE
- ↳ WHAT KINDS OF CONSTRAINTS DO YOUR USERS HAVE

IV. WHAT ARE YOUR OWN ETHICAL PROFESSIONAL RESPONSIBILITIES?

- ↳ WHAT IS ENGINEERING ETHICS?
- ↳ DO ENGINEERS NEED ETHICS?

PARTICIPATORY FRAMEWORKS FOR MACHINE LEARNING



- PARTICIPATORY DESIGN IS DESIGN THAT ACTIVELY INCORPORATE FEED BACK FROM STAKE-HOLDERS
- CO-DESIGN IS A COLLABORATIVE DESIGN PROCESS BTW DESIGNERS & STAKE-HOLDERS
- PARTICIPATORY ACTION RESEARCH IS COMMUNITY EMBEDDED RESEARCH THAT PARTNER WITH AFFECTED COMMUNITIES TO ACHIEVE SOCIAL CHANGE

LESSONS FROM PARTICIPATORY MACHINE LEARNING

1. DON'T MAKE SYSTEMS MORE FAIR AT THE EXPENSE OF MAKING THEM MORE JUST. GO BEYOND TECHNICAL FRAMINGS OF HARM, E.G. ALGORITHMIC BIAS, DATA IMBALANCE.
2. COMMUNITIES DERIVE MOST VALUES FROM LOCALIZED SOLUTIONS RATHER THAN SCALABLE, GENERALIZABLE SOLUTIONS.
3. MANY MEANINGFUL INTERVENTIONS TOWARDS MORE EQUITABLE SYSTEMS ARE NON-TECHNICAL.

EXERCISE: USE THE PARTICIPATORY DESIGN
FRAMEWORK TO DESIGN PROGRAMMING
FOR COMMUNITAS.

THINK ABOUT:

- WHAT DOES THE GRAD STUDENT COMMUNITY NEED?
- HOW CAN COMMUNITAS SUPPORT THESE NEEDS?

DOES PARTICIPATORY DESIGN WORK?

- WHAT PERSPECTIVES WERE REPRESENTED AT YOUR TABLE?
- WHICH WERE MARGINALIZED?
- WHO SPOKE THE MOST?
- WHO SPOKE THE LEAST?
- WHOSE PERSPECTIVE WAS MOST REPRESENTED BY YOUR SUMMARY?
- WHOSE PERSPECTIVE WAS LEAST REPRESENTED?
- WHAT WERE POINTS OF DISAGREEMENTS? HOW DID YOU RESOLVE THEM?
- DID YOU FEEL COMFORTABLE TO SPEAK AND DISAGREE?
- DO YOU THINK OTHER PEOPLE FELT COMFORTABLE SPEAKING & DISAGREEING? DID YOU CHECK?
- DID YOU UNDERSTAND OTHER PEOPLE'S PERSPECTIVES? DID YOU CHECK?

READING LIST:

1. Angwin, Julia, et al., Machine Bias: There's software used across the country to predict future criminals. And it's biased against blacks, *ProPublica* (2016)
2. Barabas, Chelsea. "Beyond Bias: Re-Imagining the Terms of 'Ethical AI' in Criminal Law." *Chelsea Barabas, Beyond Bias: Re-imagining the Terms of" Ethical AI" in Criminal Law* 12 (2019).
3. Benjamin, Ruha. "Race after technology: Abolitionist tools for the new jim code." *Social Forces* (2019).
4. Black, Emily, and Michael A. Madaio. "A Call for Universities to Develop Requirements for Community Engagement in AI Research."
5. Brown, Anna, et al. "Toward algorithmic accountability in public services: A qualitative study of affected community perspectives on algorithmic decision-making in child welfare services." *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 2019.
6. Criado-Perez, Caroline. *Invisible Women: Data Bias in a World Designed for Men*. New York: Abrams Press, 2019.
7. Hill, Kashmir. Wrongfully Accused by an Algorithm, *The New York Times* (2020)
8. Katell, Michael, et al. "Toward situated interventions for algorithmic equity: lessons from the field." *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. 2020.
9. Lohr, Steve. Facial Recognition Is Accurate, if You're a White Guy, *The New York Times* (2018).
10. McCradden, Melissa D., et al. "Ethical limitations of algorithmic fairness solutions in health care machine learning." *The Lancet Digital Health* 2.5 (2020): e221-e223.
11. Obermeyer, Ziad, et al. "Dissecting racial bias in an algorithm used to manage the health of populations." *Science* 366.6464 (2019): 447-453.
12. Sloane, Mona, et al. "Participation is not a Design Fix for Machine Learning." *arXiv preprint arXiv:2007.02423* (2020).
13. "Predictive Policing Algorithms Are Racist. They Need to Be Dismantled." MIT Technology Review. 2020.
14. Green, Ben. "Data science as political action: grounding data science in a politics of justice." Available at SSRN 3658431 (2020).
15. Chen, Irene, Fredrik D. Johansson, and David Sontag. "Why is my classifier discriminatory?." Advances in Neural Information Processing Systems. 2018.

PROBLEM: WE HAVE DESIGN BLINDSPOTS

- PERSPECTIVE GAP
- DATA GAP
- NEGLECT OF THE SOCIAL SYSTEM

SOLUTION: PARTICIPATORY DESIGN

- DIVERSIFY THE DECISION MAKING SPACE
- BUILT ON LIVE-EXPERIENCE

PROBLEM: THERE ARE BARRIERS TO PARTICIPATION!

- YOU DON'T EVEN KNOW WHO TO TALK TO!

"VISIBLE VS INVISIBLE AXES OF MARGINALIZATION"

↳ "GENDER PRESENTATION U.S. GENDER IDENTITY"

- PEOPLE DON'T FEEL COMFORTABLE PARTICIPATING
 - "MICRO-AGGRESSION"
 - ↳ "GATE-KEEPING"
 - ↳ "STEREOTYPE THREAT/LIFT"
- IT TAKES A LOT OF WORK TO PARTICIPATE!
 - "INVISIBLE LABOUR"
 - ↳ EMOTIONAL, TIME, UNPAID, UNRECOGNIZED
 - . BECAUSE YOU'RE HARD TO TALK TO!
 - "PRIVILEGE HAZARD"
 - ↳ "SUCCESS HAZARD"
 - "SURVIVOR HAZARD"