

AI & Society

A *very* brief survey of algorithmic bias and how AI systems can amplify inequities in society

Wells Lucas Santo (AI4ALL)
@WellsLucasSanto

AI4ALL

Mission: Increase diversity and inclusion in AI



Summer Education Program

- 2-3 week programs for underrepresented high school students
- 6 university partners & sites in 2018
- 148 students in 2018

Open Learning Platform

- Free, online and offline AI education program
- Material for students and educators
- 1M high school students in 5 years
- Launching early 2019

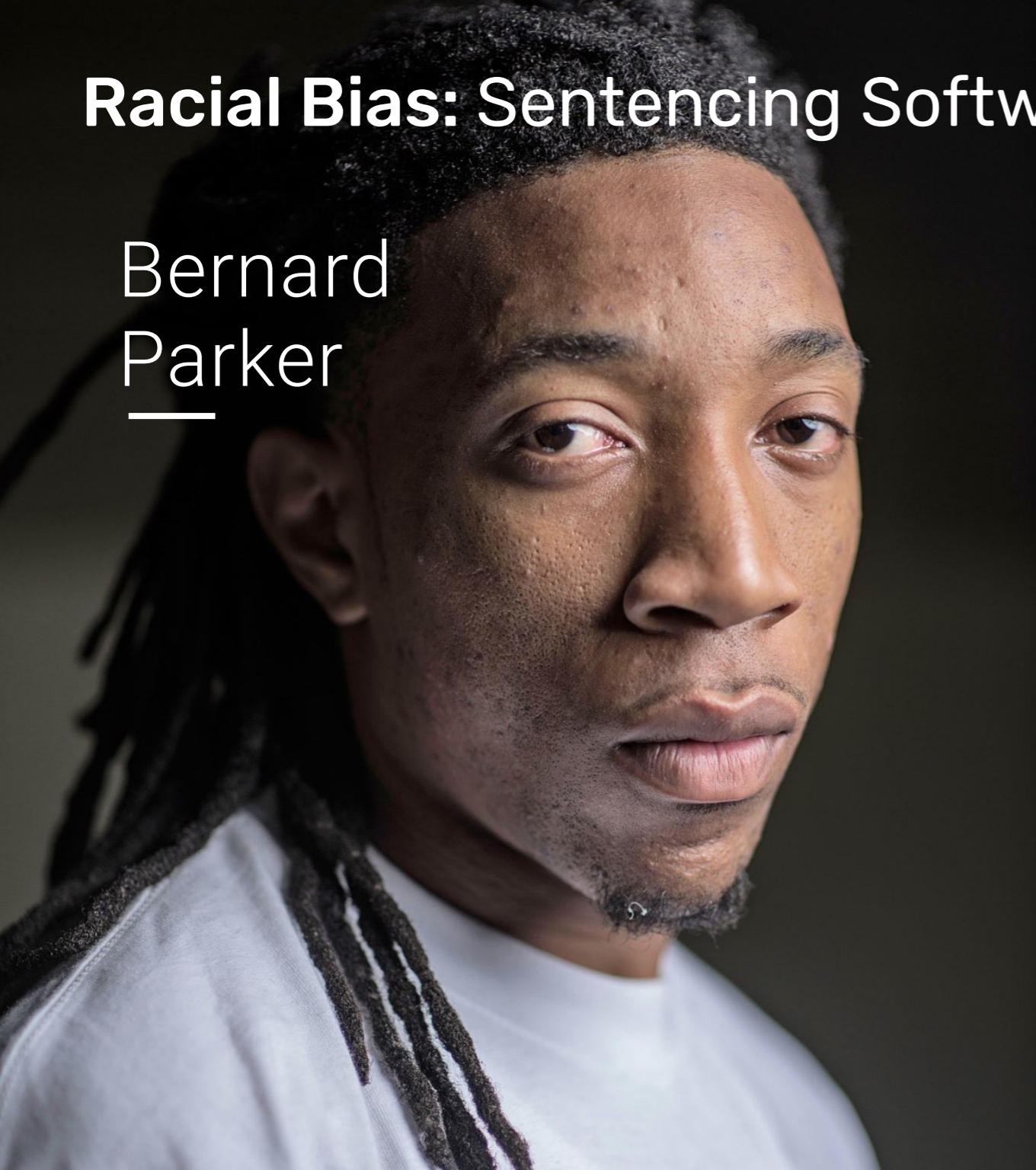
Timeline of Events



May 2016	"Machine Bias"
July 2016	First AI Now Symposium
September 2016	<i>Weapons of Math Destruction</i>
November 2016	FATML 2016 "To Predict and Serve" Gender Biased Word Embeddings
December 2017	NYC Algorithmic Accountability Bill & Task Force
January 2018	<i>Algorithms of Oppression</i> <i>Automating Inequality</i>
February 2018	FAT* 2018 "Gender Shades"
April 2018	<i>Artificial Unintelligence</i>
July 2018	ACLU Congress Rekognition

Racial Bias: Sentencing Software

Bernard
Parker



Dylan
Fugett



ProPublica "[Machine Bias](#)" (May 2016)
Angwin, Larson, Mattu, and Kirchner

Racial Bias: Sentencing Software

Bernard
Parker

Dylan
Fugett

ProPublica "[Machine Bias](#)" (May 2016)
Angwin, Larson, Mattu, and Kirchner

Racial Bias: Sentencing Software

Bernard
Parker

HIGH RISK

10

Dylan
Fugett

LOW RISK

3

ProPublica "Machine Bias" (May 2016)
Angwin, Larson, Mattu, and Kirchner

Racial Bias: Sentencing Software

Bernard
Parker

HIGH RISK

10

Dylan
Fugett

LOW RISK

3

PRIOR OFFENSE
1 attempted burglary

Racial Bias: Sentencing Software

Bernard
Parker

HIGH RISK

10

PRIOR OFFENSE
1 nonviolent resisting arrest

Dylan
Fugett

LOW RISK

3

PRIOR OFFENSE
1 attempted burglary

Racial Bias: Sentencing Software

Bernard
Parker

HIGH RISK

10

PRIOR OFFENSE

1 nonviolent resisting arrest

SUBSEQUENT OFFENSES

None

Dylan
Fugett

LOW RISK

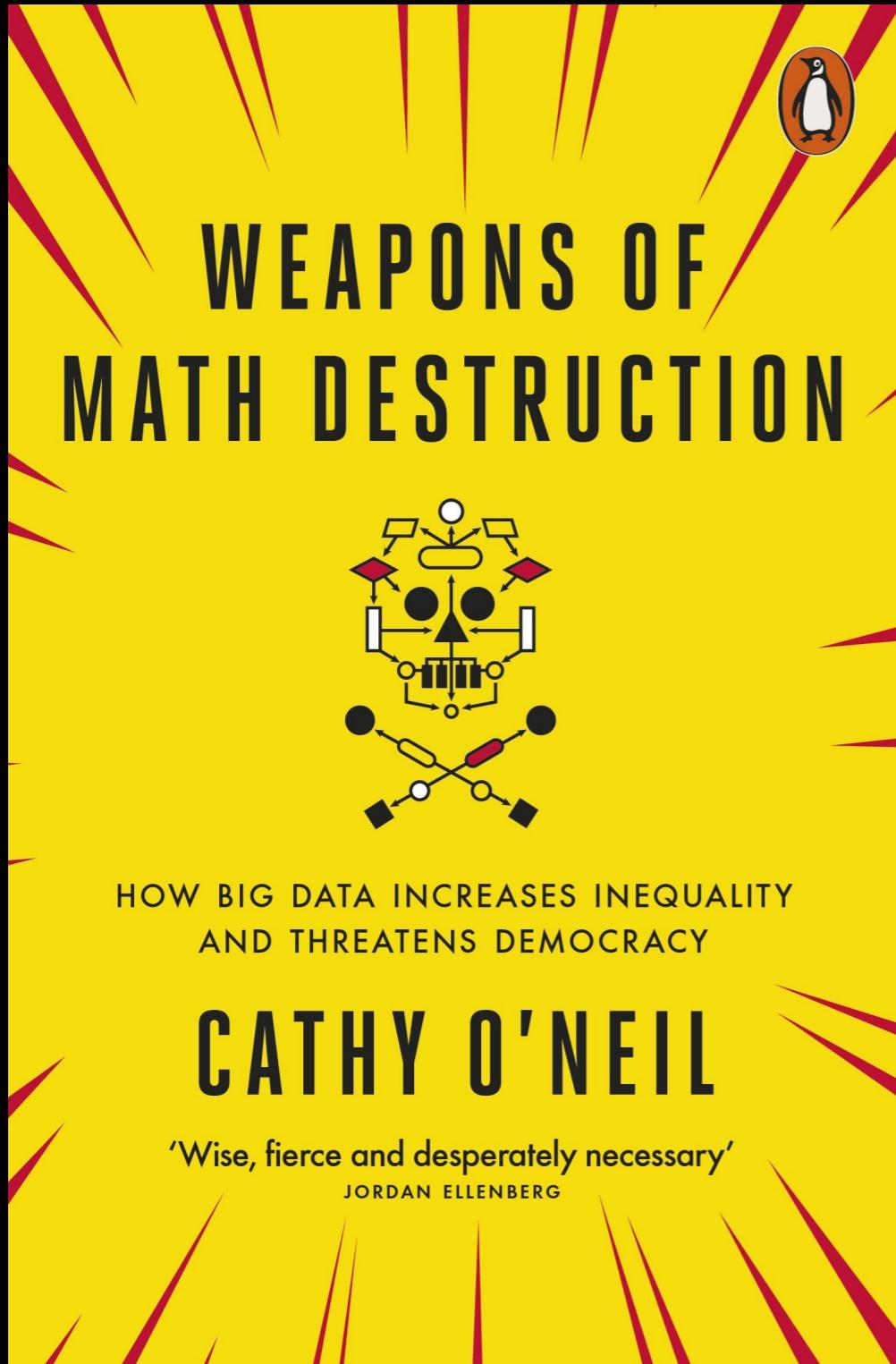
3

PRIOR OFFENSE

1 attempted burglary

SUBSEQUENT OFFENSES

3 drug possessions



- Compelling case studies
- Dangers of:
 - Automated decision making *at scale*
 - Opaque "black-box" systems
 - Lack of regulation
- Deep consequences when systems decide on socioeconomic resources: amplifies socioeconomic and racial inequality

Cathy O'Neil, *Weapons of Math Destruction* (September 2016)



Amazon scraps secret AI recruiting tool that showed bias against women

Jeffrey Dastin

8 MIN READ



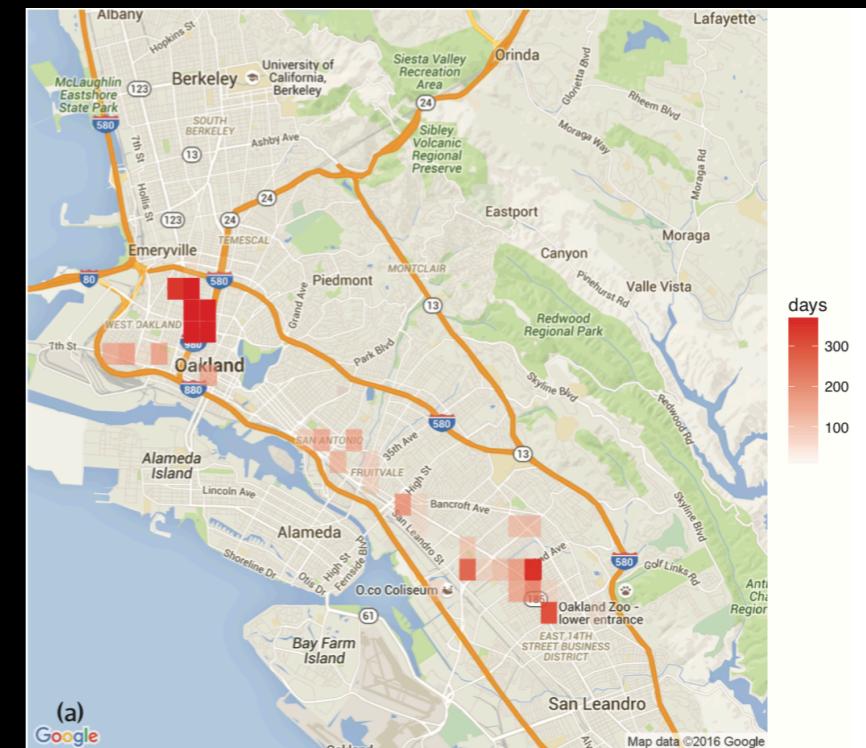
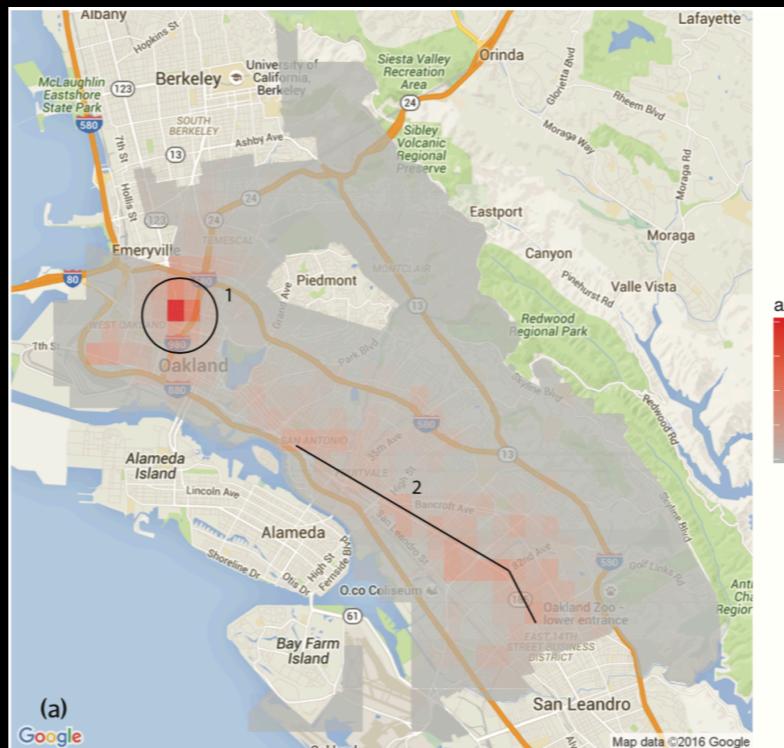
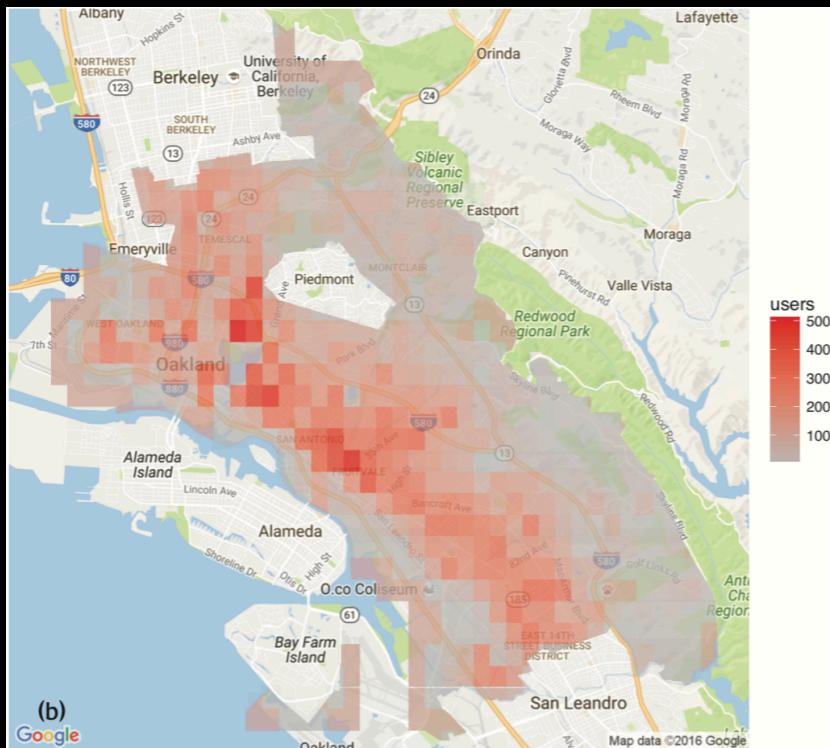
Amazon's system taught itself that male candidates were preferable. It penalized resumes that included the word "women's," as in "women's chess club captain." And it downgraded graduates of two all-women's colleges, according to people familiar with the matter.

FAT / ML

Fairness, Accountability, and Transparency in Machine Learning

(2014 - present)

FAT / ML



FAT / ML

Analogy

Parallelograms capture semantics: [MikolovYZ 13]

- Man:King :: Woman:Queen
- Paris:France :: Tokyo:Japan
- He:Blue :: She:Pink
- He:Brother :: She:Sister
- He:Doctor :: She:Nurse
- He:Sausage :: She:Buns
- He:Realist :: She:Feminist
- She:Pregnancy :: He:Kidney stone
- He:Computer programmer :: She:Homemaker



Based on word2vec trained
on Google News corpus

Gendered Bias: Translation

ENGLISH ▾

He is a babysitter
She is a doctor

TURKISH

TRANSLATE

O bir bebek bakıcısı
O bir doktor

Gendered Bias: Translation

ENGLISH ▾

He is a babysitter

She is a doctor

TURKISH

O bir bebek bakıcısı

O bir doktor

TURKISH

O bir bebek bakıcısı

O bir doktor

Gendered Bias: Translation

ENGLISH ▾

He is a babysitter
She is a doctor

TURKISH

O bir bebek bakıcısı
O bir doktor

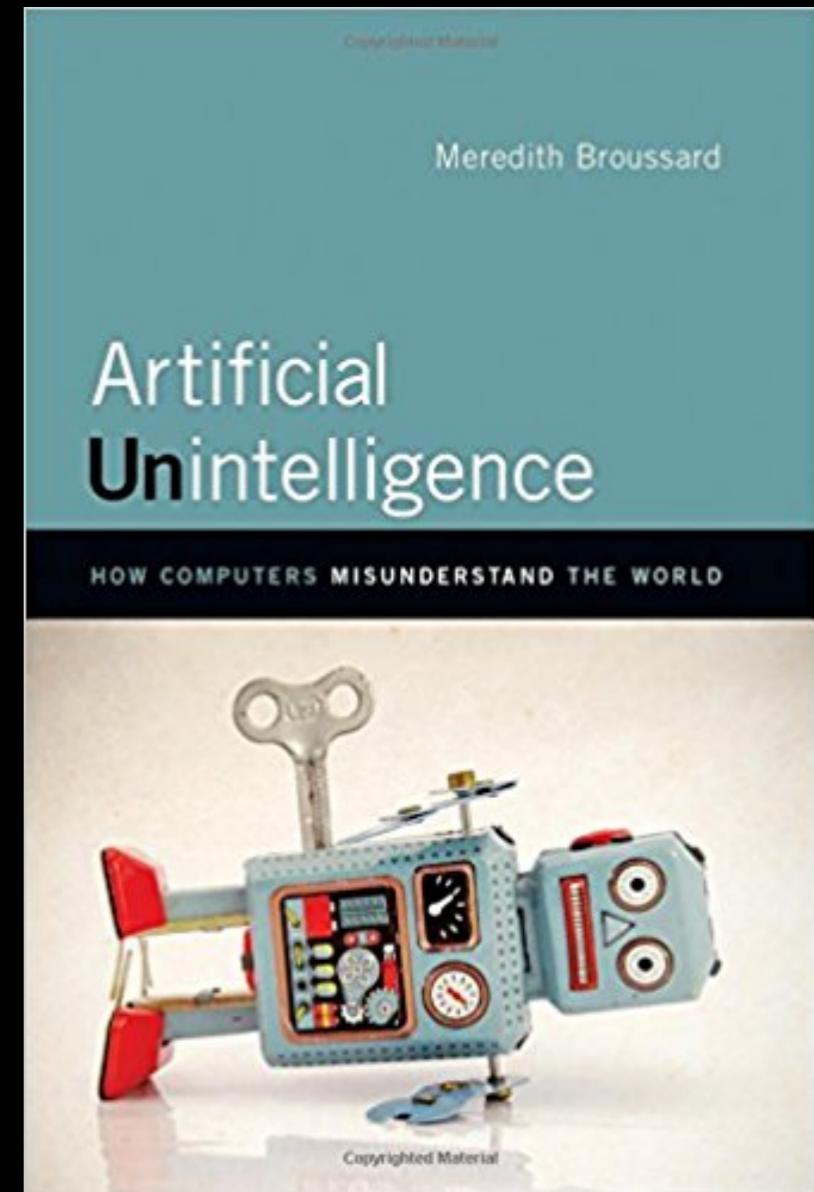
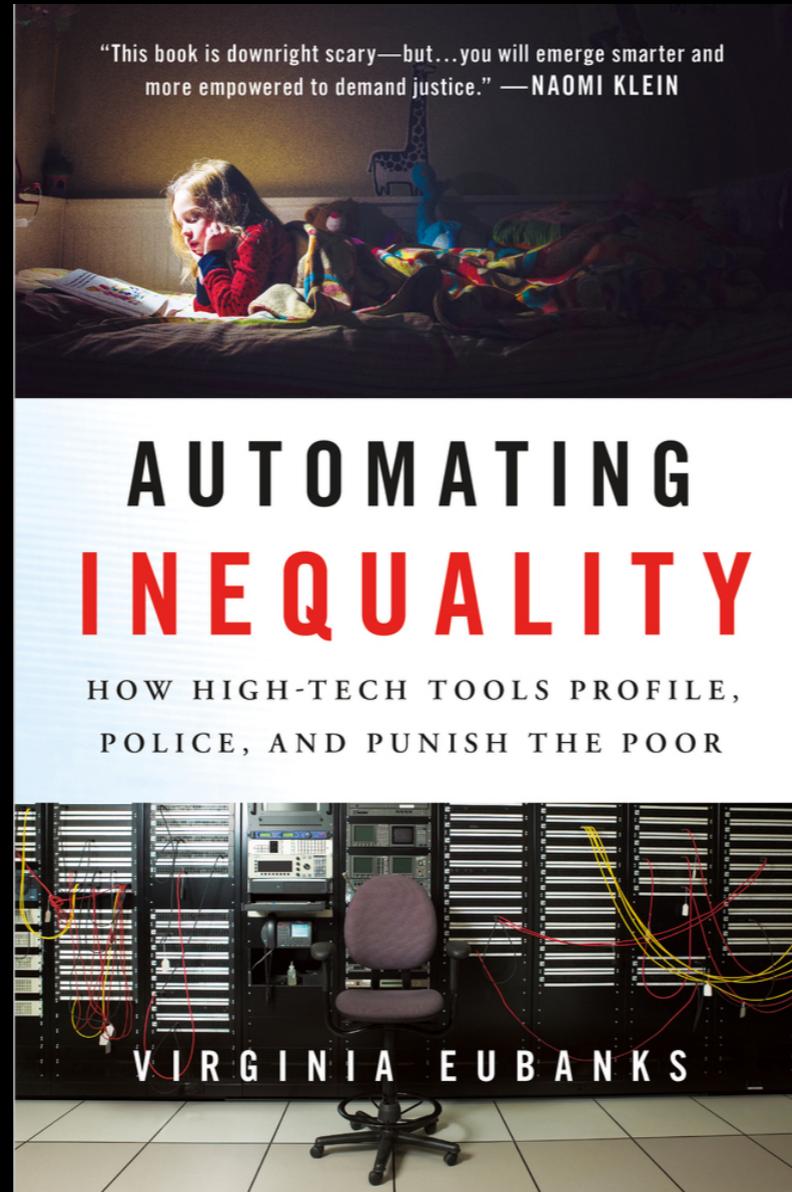
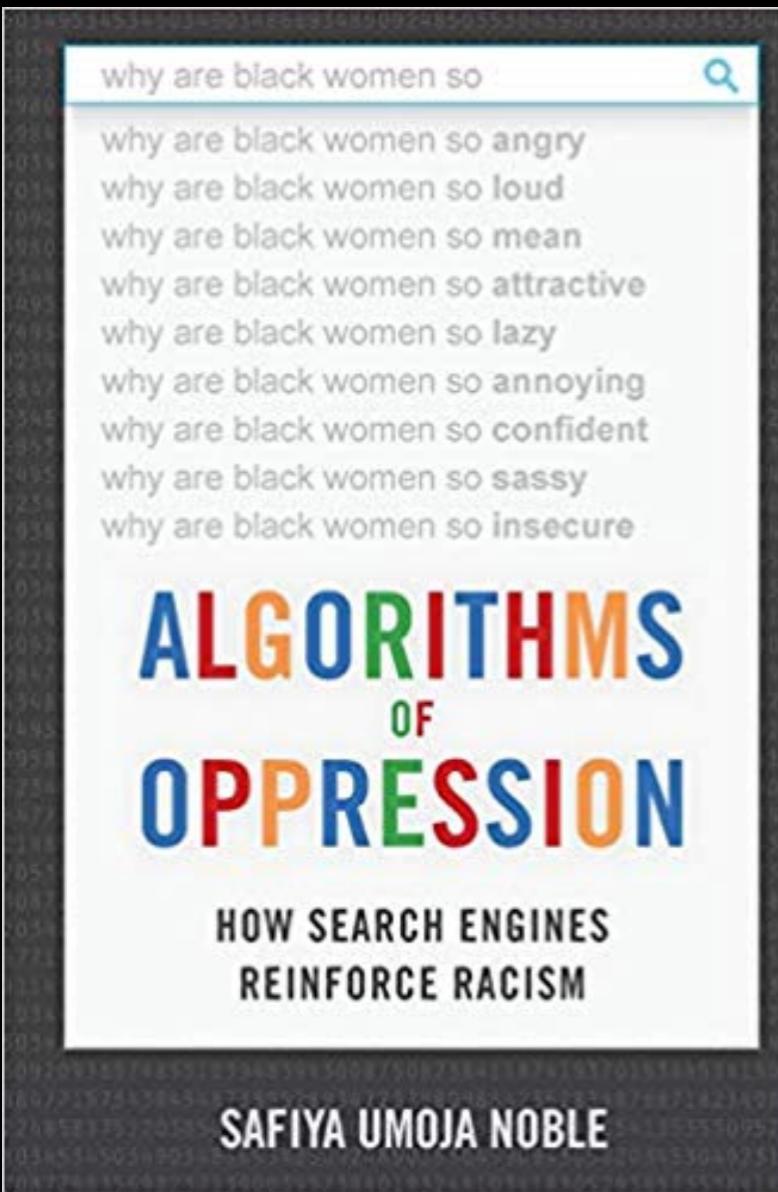
TURKISH

O bir bebek bakıcısı
O bir doktor

ENGLISH ▾

She's a babysitter
He is a doctor

TRANSLATE

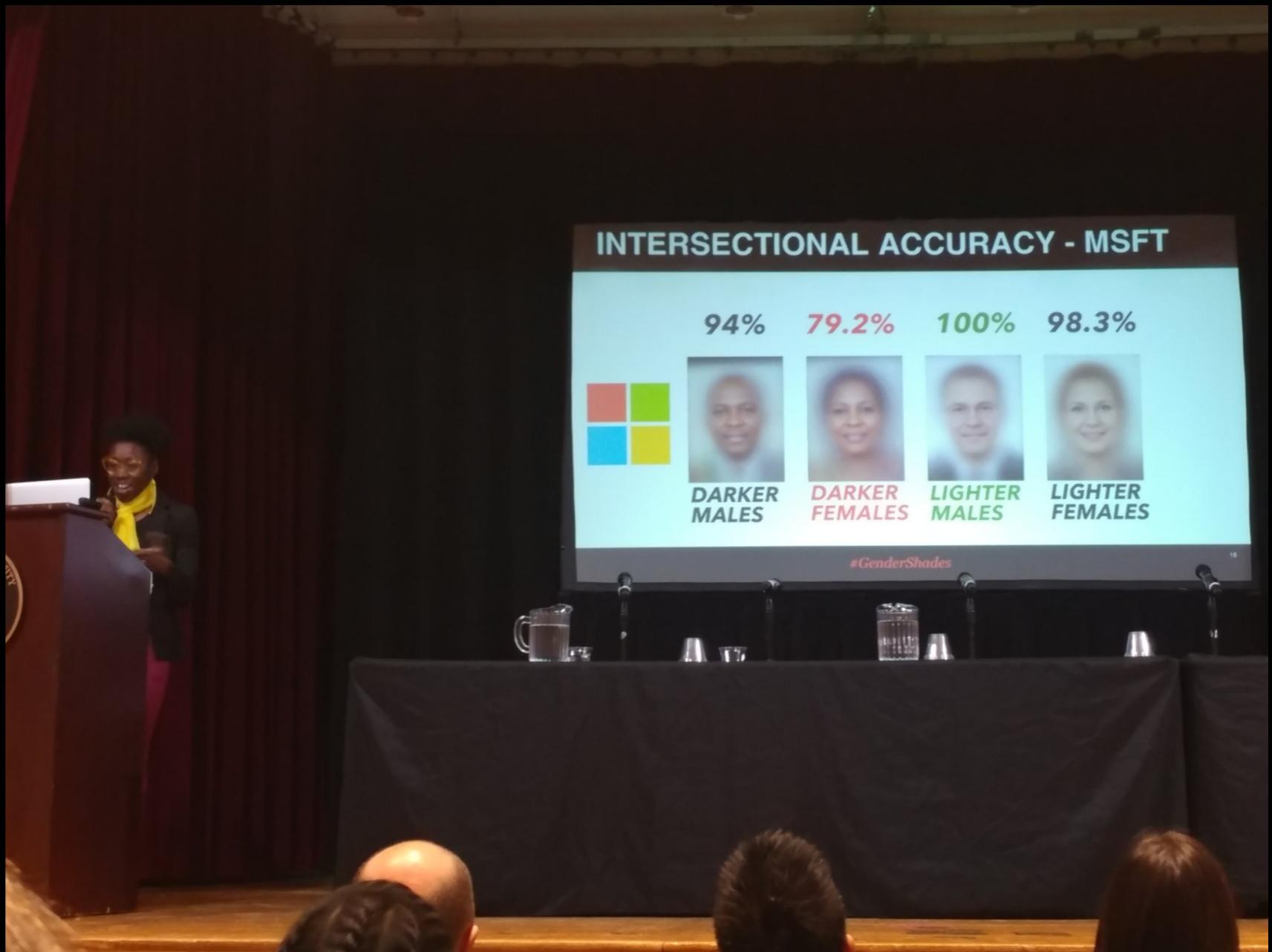


Safiya Noble, *Algorithms of Oppression* (January 2018)
Virginia Eubanks, *Automating Inequality* (January 2018)
Meredith Broussard, *Artificial Unintelligence* (April 2018)

FAT / ML



ACM FAT*



Joy Bouamwini and Timnit Gebru, Gender Shades (February 2018)

Amazon Rekognition

FALSE MATCHES



28 current members of Congress

Racial Bias in Amazon Face Recognition



People of color were disproportionately falsely matched in our test.

Timeline of Events



May 2016	"Machine Bias"
July 2016	First AI Now Symposium
September 2016	<i>Weapons of Math Destruction</i>
November 2016	FATML 2016 "To Predict and Serve" Gender Biased Word Embeddings
December 2017	NYC Algorithmic Accountability Bill & Task Force
January 2018	<i>Algorithms of Oppression</i> <i>Automating Inequality</i>
February 2018	FAT* 2018 "Gender Shades"
April 2018	<i>Artificial Unintelligence</i>
July 2018	ACLU Congress Rekognition

Relevant Orgs & Conferences

AI NOW
INSTITUTE

YOUR
IS ON 🔥

Data&Society
RESEARCH INSTITUTE

FAT / ML



**BERKMAN
KLEIN CENTER**

FOR INTERNET & SOCIETY
AT HARVARD UNIVERSITY



We Robot
2014

APRIL 4-5, 2014 | CORAL GABLES, FL
CONFERENCE ON LEGAL AND POLICY
ISSUES RELATING TO ROBOTICS

Concluding Questions

What must be incorporated into AI education so that people working on AI are aware of these issues?

One step further:

- Is it part of the goals of an AI curriculum to emphasize ethical use of AI? (I'd say yes.)
- Is it part of an AI curriculum to *argue* for a specific ethical position? (More unclear.)

I argue that if we are committed to this, we must teach from an interdisciplinary standpoint, not just a technological one.

Questions?

Wells Lucas Santo (AI4ALL)
@WellsLucasSanto