

Emma Willard School

Coded Bias: How Racism Shows Up In Artificial Intelligence Algorithms



Prepared by: Chiara Shah
January 18th, 2021

Sojourner Truth: Ain't I A Woman?



Not only was Sojourner Truth involved with the antislavery movement, she was also involved in the women's rights movement.

In 1851 at the Women's Rights Convention held in Akron, Ohio, Sojourner "delivered what is now recognized as one of the most famous abolitionist and women's rights speeches in American history".



Sojourner Truth: Ain't I A Woman?



Er

AI, Ain't I A Woman?

amazon

Microsoft

IBM

Face++ 旷视

Google



Er

People Assume Algorithms Are Neutral

We think of computers as

Impartial

Random

Non-feeling

Unbiased

Secure




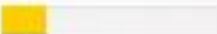




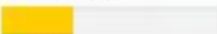





Automated

Joy Buolamwini's Story

Joy's story from the gendershades.org website



Joy Buolamwini's Story

Gender Classifier	Darker Male	Darker Female	Lighter Male	Lighter Female	Largest Gap
 Microsoft	94.0% 	79.2% 	100% 	98.3% 	20.8% 
 FACE++	99.3% 	65.5% 	99.2% 	94.0% 	33.8% 
 IBM	88.0% 	65.3% 	99.7% 	92.9% 	34.4% 



Latanya Sweeney's Story



Dr. Sweeney is a professor of Computer Science at Harvard.

She earned MIT's first Computer Science doctorate awarded to a woman of color.

<https://chicagotribune.search.yahoo.com/search?p=Latanya+Sweeney>



Amazon's Hiring Algorithm



Amazon discontinued a recruiting algorithm after discovering that it led to gender bias in its hiring. (Credit: Brian Snyder/Reuters)

Online retailer Amazon, whose global workforce is 60 percent male and where men hold 74 percent of the company's managerial positions, recently discontinued use of a recruiting algorithm after discovering gender bias.



Bias in criminal justice algorithms

“Our analysis found that black defendants were far more likely than white defendants to be incorrectly judged to be at a higher risk of recidivism, while white defendants were more likely than black defendants to be incorrectly flagged as low risk.”

The COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) algorithm was found to be biased against African-Americans, according to a report from ProPublica.



Bias in word associations



Historically certain words have certain associations. We heard this earlier today Dr. King's "Black is Beautiful" speech.



Bias in word associations

“We have a situation where these artificial-intelligence systems may be perpetuating historical patterns of bias that we might find socially unacceptable and which we might be trying to move away from.”

Biased Bots: Artificial-intelligence systems echo human prejudices

Princeton University-based researchers set an AI loose on a trove of documents on the Internet containing 840 billion words.



More examples of bias in decisions made by AI

AI in **Self-Driving Cars**: difficulty spotting pedestrians of color

AI in **Medical Diagnoses**: difficulty diagnosing melanoma on skin of color

AI in **Law Enforcement**: scanning crowds of people, mistakes happen more often than a correct identification (no matter the skin color of the person)

AI in **Healthcare Predictions**: in trying to predict which patients would need extra medical care, the AI favored white patients over black patients (meaning, it approved procedures and tests more often for whites than blacks). The reason this happened was because it looked at prior healthcare spending (and white people spent more than black people).

AI in the **College Admissions Process**



Potential Harm

POTENTIAL HARMS FROM ALGORITHMIC DECISION-MAKING

INDIVIDUAL HARMS			COLLECTIVE SOCIAL HARMS
ILLEGAL DISCRIMINATION	UNFAIR PRACTICES		
HIRING		LOSS OF OPPORTUNITY	
EMPLOYMENT			
INSURANCE & SOCIAL BENEFITS			
HOUSING			
EDUCATION			
CREDIT		ECONOMIC LOSS	
DIFFERENTIAL PRICES OF GOODS			
LOSS OF LIBERTY		SOCIAL STIGMATIZATION	
INCREASED SURVEILLANCE			
STEREOTYPE REINFORCEMENT			
DIGNATORY HARMS			

Where is the bias coming from?

1. Prejudiced assumptions made during the algorithm development process
2. Prejudices in the training data
3. Lack of complete training data
4. Too much training data that is skewed in a particular direction

Where is the bias coming from?

“...algorithms with too much data, or an over-representation, can skew the decision toward a particular result. Researchers at Georgetown Law School found that an estimated 117 million American adults are in facial recognition networks used by law enforcement, and that African-Americans were more likely to be singled out primarily because of their over-representation in mug-shot databases.”

Will AI ever be completely unbiased?

In theory...yes.

But practically, probably not.

What we can try to do is *minimize* the bias.

Solution?

Get the word out. Call out injustices.

- Joy is doing that with her spoken word piece, Algorithmic Justice League, etc.
- Movie “Coded Bias”

Revise Computer Science curriculum.

- new teaching methods and sense of responsibilities
- maybe have computer scientists take some sort of “oath” that’s similar to the hipocratic oath.



Solution?

Public Interest Technologist.

- technology practitioners who focus on social justice, the common good, and/or the public interest.
- Should have a background in law, technology, and policy, and not necessarily a computer science degree.



“It’s important for algorithm operators and developers to always be asking themselves: Will we leave some groups of people worse off as a result of the algorithm’s design or its unintended consequences?”

“Algorithmic Bias Detection and Mitigation: Best Practices and Policies to Reduce Consumer Harms”



MITIGATION PROPOSALS

- Nondiscrimination and other civil rights laws should be updated to interpret and redress online disparate impacts
- Operators of algorithms must develop a bias impact statement
- Diversity-in-Design
 - Within work teams
 - In the training data
 - Culturally sensitive decision making



MITIGATION PROPOSALS (continued)

- Operators of algorithms should regularly audit for bias
 - Facebook recently did this and as a result they banned white nationalist content
- Operators of algorithms must rely upon cross-functional work teams and expertise
- Increase human involvement in the design and monitoring of algorithms
- Consumers need better algorithmic literacy



Krita Sharma TED Talk “How to keep human bias out of AI”



For More Information...



For More Information...



Coded Bias movie. Not on Netflix yet...but when it is available, I highly recommend it.



Questions?



Sources

"AI, Ain't I Woman?" *YouTube*, uploaded by Joy Buolamwini, Google, 28 June 2018, www.youtube.com/watch?v=QxuyfWoVV98. Accessed 18 Jan. 2021.

Buolamwini, Joy, and Timnit Gebru. "Gender Shades." *Gender Shades*, MIT Media Lab, 2018, gendershades.org/. Accessed 18 Jan. 2021.

"Dr. King Said It: I'm Black and I'm Proud." *YouTube*, www.youtube.com/watch?v=Suw_CQ3zfTY&feature=youtu.be. Accessed 18 Jan. 2021.

Hadhazy, Adam. "Biased Bots: Artificial-intelligence Systems Echo Human Prejudices." *Princeton University*, 18 Apr. 2017, www.princeton.edu/news/2017/04/18/biased-bots-artificial-intelligence-systems-echo-human-prejudices. Accessed 18 Jan. 2021.

"How to Keep Human Bias Out of AI." *YouTube*, uploaded by TED, Google, 12 Apr. 2019, www.youtube.com/watch?v=BRRNeBKwvNM&feature=youtu.be. Accessed 18 Jan. 2021.

Kantarci, Atakan. "Bias in AI: What It Is, Types & Examples, How & Tools to Fix It." *AIMultiple*, 9 Jan. 2021, research.aimultiple.com/ai-bias/. Accessed 18 Jan. 2021.



Larson, Jeff, et al. "How We Analyzed the COMPAS Recidivism Algorithm." *ProPublica*, 23 May 2016,
www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm. Accessed 18 Jan. 2021.

Lee, Nicol Turner, et al. *Algorithmic Bias Detection and Mitigation: Best Practices and Policies to Reduce Consumer Harms*. Brookings
Institute, 22 May 2019. *Brookings*,
www.brookings.edu/research/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/.
Accessed 18 Jan. 2021.

Pangburn, DJ. "Schools Are Using Software to Help Pick Who Gets In. What Could Go Wrong?" *Fast Company*, 17 May 2019,
www.fastcompany.com/90342596/schools-are-quietly-turning-to-ai-to-help-pick-who-gets-in-what-could-go-wrong. Accessed 18 Jan.
2021.

Perry, Andre M., and Nicol Turner Lee. "AI Is Coming to Schools, and If We're Not Careful, so Will Its Biases." *The Avenue*, Brookings
Institute, 26 Sept. 2019,
www.brookings.edu/blog/the-avenue/2019/09/26/ai-is-coming-to-schools-and-if-were-not-careful-so-will-its-biases/. Accessed 18 Jan.
2021.



"Race, Technology, and Algorithmic Bias." *YouTube*, uploaded by Harvard University, Radcliffe Institute, Google, 7 May 2019, www.youtube.com/watch?v=Y6fUc5_whX8. Accessed 18 Jan. 2021.

Schneier, Bruce, editor. "Public-Interest Technology Resources." *Public Interest Technology*, WordPress, 16 Jan. 2021, public-interest-tech.com/. Accessed 18 Jan. 2021.

"Sojourner Truth: Ain't I a Woman." *National Park Service*, 17 Nov. 2017, www.nps.gov/articles/sojourner-truth.htm. Accessed 18 Jan. 2021.

"Sojourner Truth's 'Ain't I a Woman': Nkechi at TEDxFiDiWomen." *YouTube*, uploaded by TEDx Talks, Google, 7 Feb. 2013, www.youtube.com/watch?v=eUdxsQ0Qsrc&feature=youtu.be. Accessed 18 Jan. 2021.

Sweeney, Latanya. "Discrimination in Online Ad Delivery." *Data Privacy Lab*, President and Fellows Harvard University, 2013, dataprivacylab.org/projects/onlineads/index.html. Accessed 18 Jan. 2021.

Truth, Sojourner. "Ain't I a Woman?" May 1851. *Andrew Jackson's Hermitage*, Andrew Jackson Foundation, thehermitage.com/wp-content/uploads/2016/02/Sojourner-Truth_Aint-I-a-Woman_1851.pdf. Accessed 18 Jan. 2021. Speech.

