

# Algorithmic bias

Data Science in a Box  
[datasciencebox.org](http://datasciencebox.org)

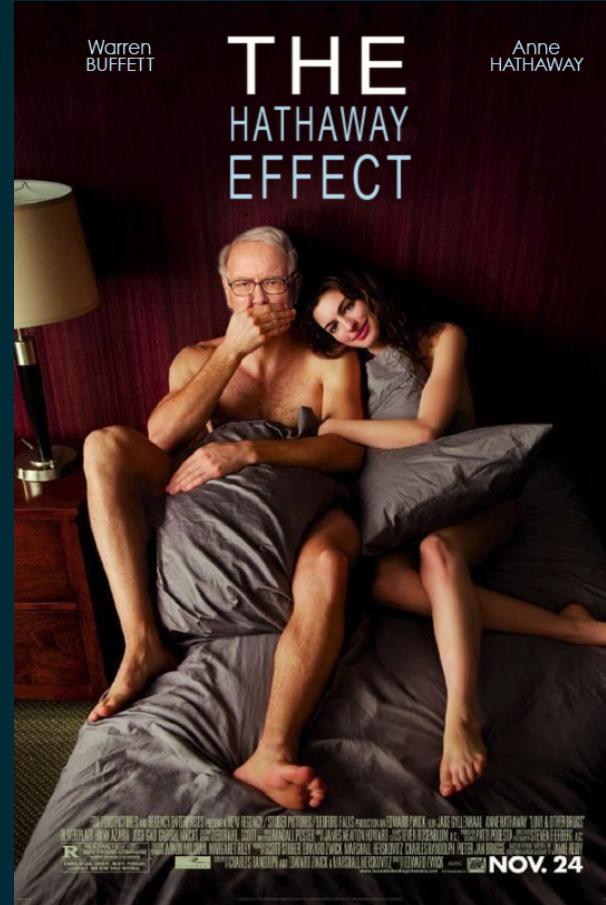


# First a bit of fun...

# The Hathaway Effect



- Oct. 3, 2008: Rachel Getting Married opens, BRK.A up 0.44%
- Jan. 5, 2009: Bride Wars opens, BRK.A up 2.61%
- Feb. 8, 2010: Valentine's Day opens, BRK.A up 1.01%
- March 5, 2010: Alice in Wonderland opens, BRK.A up 0.74%
- Nov. 24, 2010: Love and Other Drugs opens, BRK.A up 1.62%
- Nov. 29, 2010: Anne announced as co-host of the Oscars, BRK.A up 0.25%



Dan Mirvish. The Hathaway Effect: How Anne Gives Warren Buffett a Rise. The Huffington Post. 2 Mar 2011.

# Algorithmic bias and gender



# Google Translate

Turkish Chinese Spanish Detect language ↻ English Chinese (Simplified) Spanish Translate

×

o bir asker	he is a soldier
o bir öğretmen	She's a teacher
O bir doktor	He is a doctor
o bir hemşire	she is a nurse
o bir yazar	he is a writer
o bir köpek	he is a dog
o bir dadi	she is a nanny
o bir kedi	it is a cat
o bir rektör	he is a rector
o bir başkanı	he is a president
o bir girişimci	he is an entrepreneur
o bir Şarkıcı	she is a singer
o bir Öğrenci	he is a student
o bir Tercüman	he is a translator
o çalışan	he is hard working
o tembel	she is lazy
o bir ressam	he is a painter
o bir kuaför	he is a hairdresser
o bir garson	he is a waiter
O bir mühendis	He is an engineer
o bir mimar	he is an architect
o bir sanatçı	he is an artist



# Amazon's experimental hiring algorithm

- Used AI to give job candidates scores ranging from one to five stars -- much like shoppers rate products on Amazon
- Amazon's system was not rating candidates for software developer jobs and other technical posts in a gender-neutral way; it taught itself that male candidates were preferable

Gender bias was not the only issue. Problems with the data that underpinned the models' judgments meant that unqualified candidates were often recommended for all manner of jobs, the people said.

Jeffrey Dastin. Amazon scraps secret AI recruiting tool that showed bias against women.  
Reuters. 10 Oct 2018.



# Algorithmic bias and race



# Facial recognition



Interview |  
**'A white mask worked better': why algorithms are not colour blind**

Ian Tucker |  
When Joy Buolamwini found that a robot recognised her face better when she wore a white mask, she knew a problem needed fixing

Sun 28 May 2017 13.27 BST

Joy Buolamwini is a graduate researcher at the MIT Media Lab and founder of the Algorithmic Justice League - an organisation that aims to challenge the biases in decision-making software. She grew up in Mississippi, gained a Rhodes scholarship, and she is also a Fulbright fellow, an Astronaut scholar and a Google Anita Borg scholar. Earlier this year she won a \$50,000 scholarship funded by the makers of the film *Hidden Figures* for her work fighting coded discrimination.

Ian Tucker. 'A white mask worked better': why algorithms are not colour blind.  
The Guardian. 28 May 2017.

# Criminal Sentencing

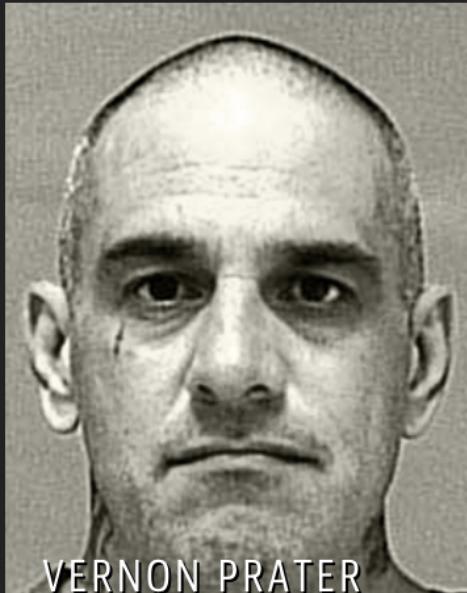
There's software used across the country to predict future criminals.  
And it's biased against blacks.



Julia Angwin, Jeff Larson, Surya Mattu, and Lauren Kirchner. Machine Bias. 23 May 2016. ProPublica.

# A tale of two convicts

## Two Petty Theft Arrests



VERNON PRATER

LOW RISK

**3**



BRISHA BORDEN

HIGH RISK

**8**

*Borden was rated high risk for future crime after she and a friend took a kid's bike and scooter that were sitting outside. She did not reoffend.*



[datasciencebox.org](http://datasciencebox.org)

## Two Petty Theft Arrests

VERNON PRATER

Prior Offenses

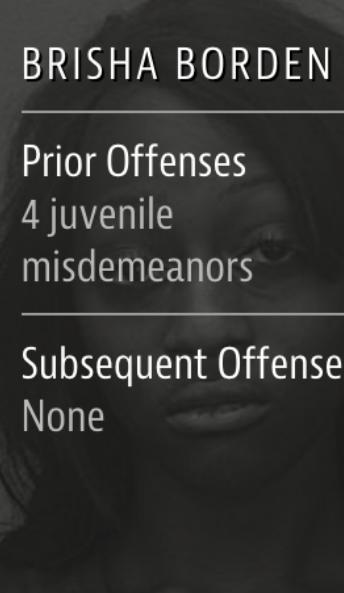
2 armed robberies, 1  
attempted armed  
robbery

Subsequent Offenses

1 grand theft

LOW RISK

**3**



BRISHA BORDEN

Prior Offenses

4 juvenile  
misdemeanors

Subsequent Offenses

None

HIGH RISK

**8**

*Borden was rated high risk for future crime after she and a friend took a kid's bike and scooter that were sitting outside. She did not reoffend.*

“Although these measures were crafted with the best of intentions, I am concerned that they inadvertently undermine our efforts to ensure individualized and equal justice,” he said, adding, “they may exacerbate unwarranted and unjust disparities that are already far too common in our criminal justice system and in our society.”

Then U.S. Attorney General Eric Holder (2014)



# ProPublica analysis

## Data:

Risk scores assigned to more than 7,000 people arrested in Broward County, Florida, in 2013 and 2014 + whether they were charged with new crimes over the next two years



# ProPublica analysis

## Results:

- 20% of those predicted to commit violent crimes actually did
- Algorithm had higher accuracy (61%) when full range of crimes taken into account (e.g. misdemeanors)

	WHITE	AFRICAN AMERICAN
Labeled Higher Risk, But Didn't Re-Offend	23.5%	44.9%
Labeled Lower Risk, Yet Did Re-Offend	47.7%	28.0%

- Algorithm was more likely to falsely flag black defendants as future criminals, at almost twice the rate as white defendants
- White defendants were mislabeled as low risk more often than black defendants

# How to write a racist AI without trying

## How to write a racist AI in R without really trying

2018/09/27

5 min read

Last year, Robyn Speer wrote a really great post [How to make a racist AI without really trying](#). Go read it.

The idea is to do sentiment analysis with obvious, off-the-shelf tools. As the post says

So that's what we're going to do here, following the path of least resistance at every step, obtaining a classifier that should look

Thomas Lumley. How to write a racist AI in R without really trying.  
Biased and Inefficient. 27 September 2018.



# Further reading



# Machine Bias

## Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica

May 23, 2016

**O**N A SPRING AFTERNOON IN 2014, Brisha Borden was running late to pick up her god-sister from school when she spotted an unlocked kid's blue Huffy bicycle and a silver Razor scooter. Borden and a friend grabbed the bike and scooter and tried to ride them down the street in the Fort Lauderdale suburb of Coral Springs.

Just as the 18-year-old girls were realizing they were too big for the tiny conveyances — which belonged to a 6-year-old boy — a woman came running after them saying, "That's my kid's stuff." Borden and her friend immediately dropped the bike and scooter and walked away.

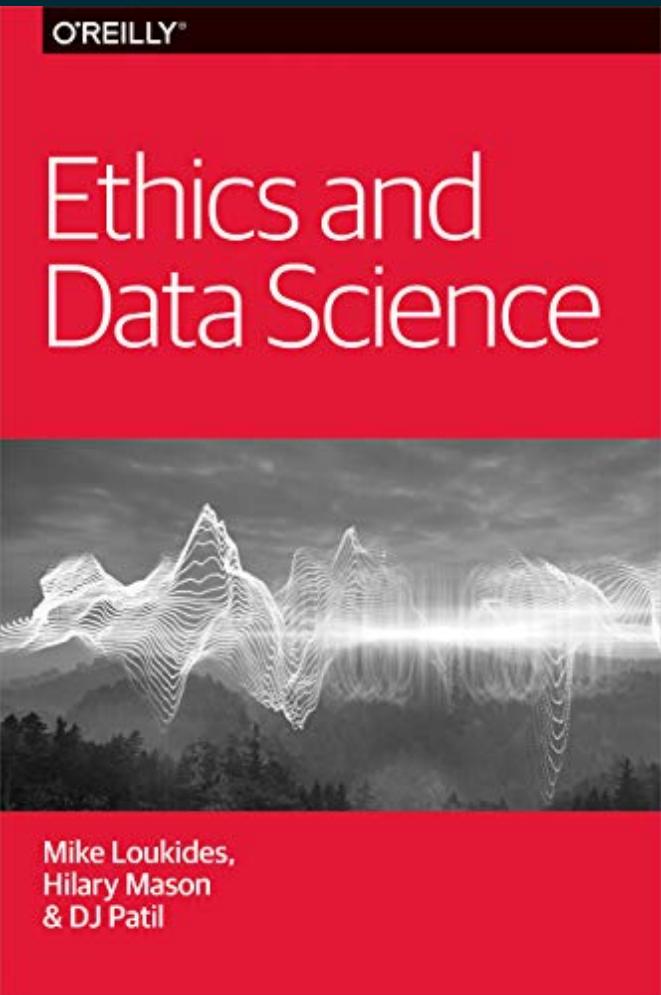
But it was too late — a neighbor who witnessed the heist had already called the police. Borden and her friend were arrested and charged with burglary and petty theft for the items, which were valued at a total of \$80.

## Machine Bias

by Julia Angwin, Jeff Larson, Surya Mattu, and Lauren Kirchner



# Ethics and Data Science



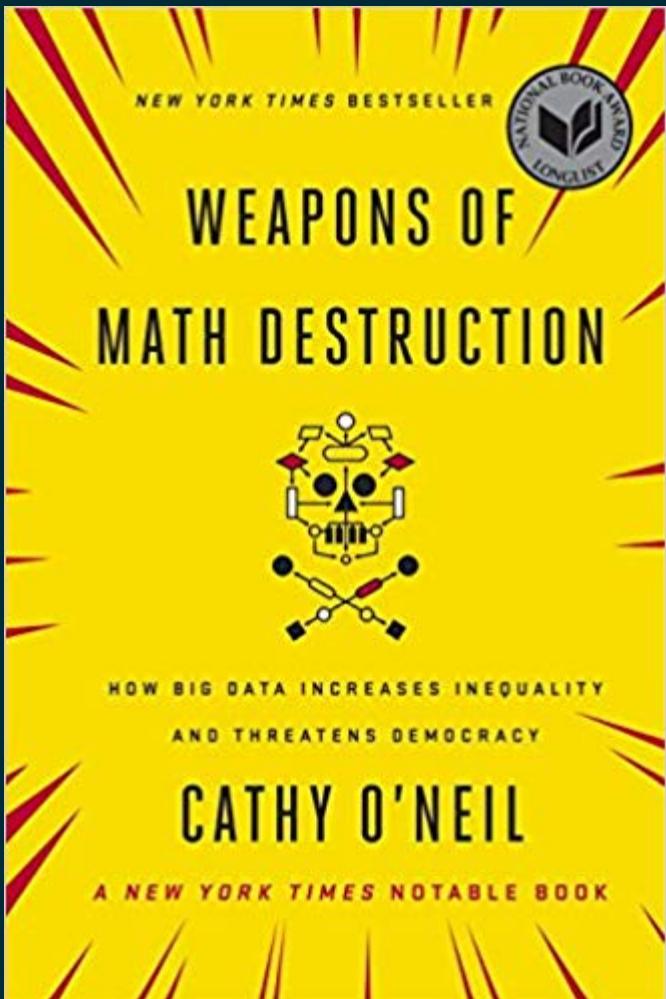
Ethics and Data Science

by Mike Loukides, Hilary Mason, DJ Patil  
(Free Kindle download)



[datasciencebox.org](http://datasciencebox.org)

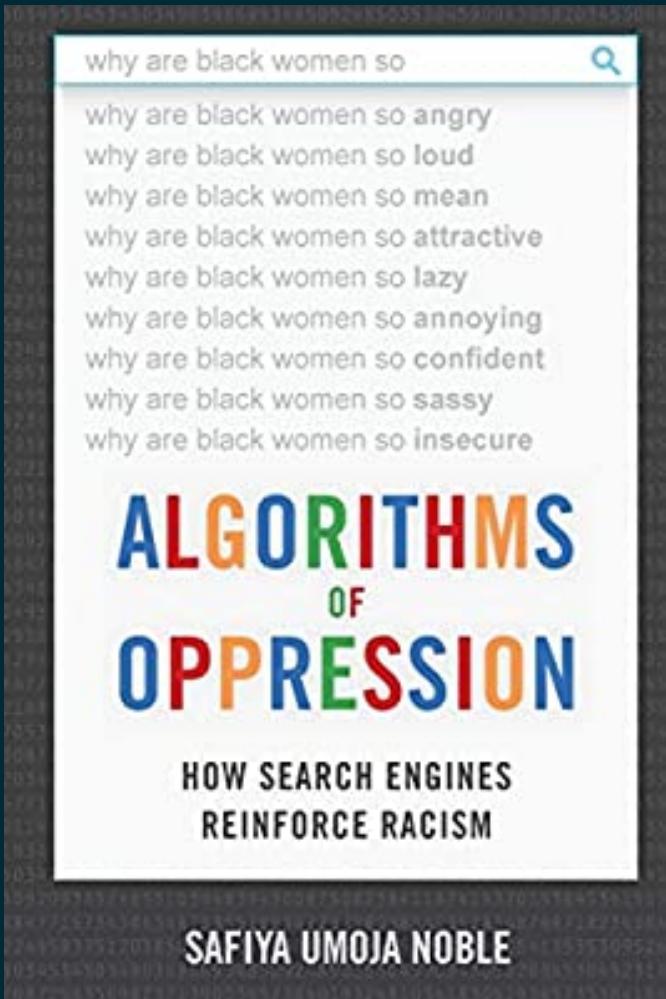
# Weapons of Math Destruction



Weapons of Math Destruction  
How Big Data Increases Inequality and  
Threatens Democracy

by Cathy O'Neil

# Algorithms of Oppression



Algorithms of Oppression  
How Search Engines Reinforce Racism

by Safiya Umoja Noble

# Parting thoughts

- At some point during your data science learning journey you will learn tools that can be used unethically
- You might also be tempted to use your knowledge in a way that is ethically questionable either because of business goals or for the pursuit of further knowledge (or because your boss told you to do so)

How do you train yourself to make the right decisions (or reduce the likelihood of accidentally making the wrong decisions) at those points?

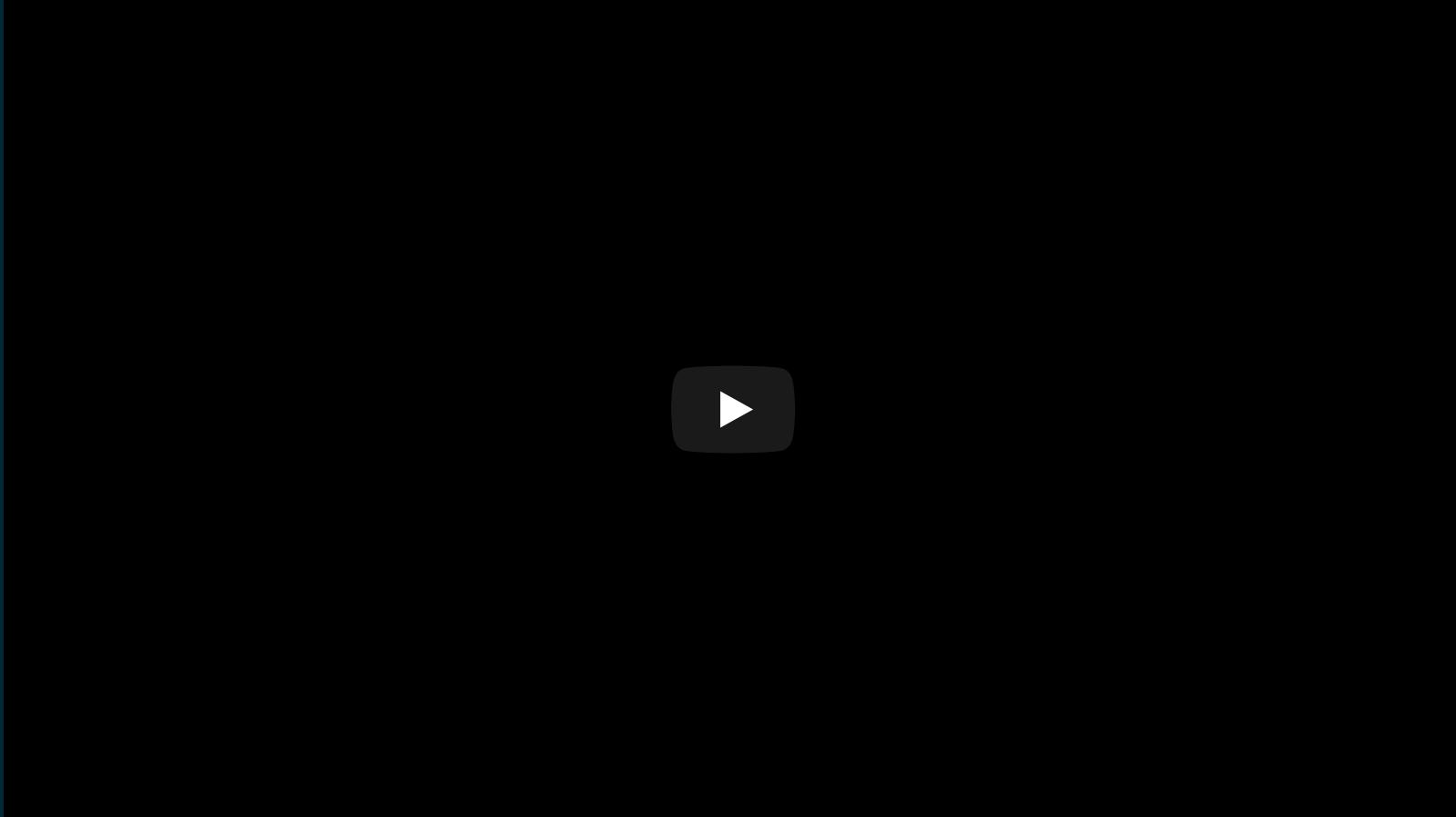


# Do good with data

- Data Science for Social Good:
  - The Alan Turing Institute
  - University of Chicago
- DataKind: DataKind brings high-impact organizations together with leading data scientists to use data science in the service of humanity.
- Sign the Manifesto for Data Practices: [datapractices.org/manifesto](http://datapractices.org/manifesto)



# Further watching



Julien Cornebise. AI for Good in the R and Python ecosystems. useR 2019.



[datasciencebox.org](https://datasciencebox.org)