

---

---

# Algorithmic fairness

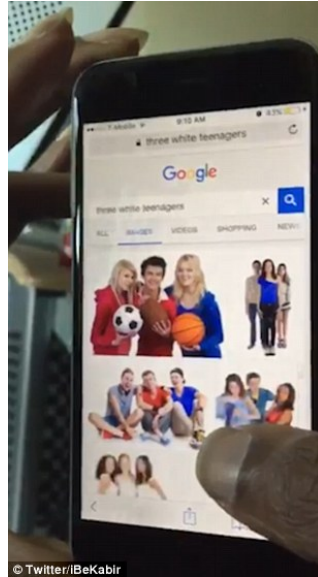
**Sharad Goel**  
**Stanford University**

---

---

# “Three white teenagers”

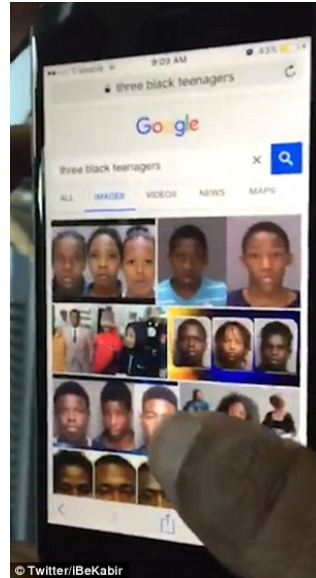
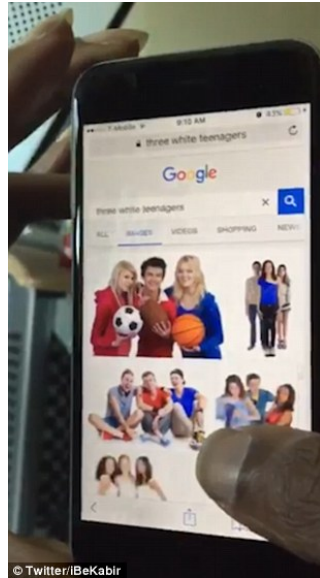
Posted by @iBeKabir



---

# “Three *black* teenagers”

Posted by @iBeKabir



---

**Hard to understand the emergent  
behavior of complex systems**

---

---

---

# **Bias begets bias**

[ From Barocas & Selbst ]

St. George's Hospital in the UK developed an algorithm to sort medical school applicants.

---

---

# **Bias begets bias**

[ From Barocas & Selbst ]

St. George's Hospital in the UK developed an algorithm to sort medical school applicants.

Algorithm trained to mimic past admissions decisions. But past decisions were biased against women and minorities.

---

---

# The target variable matters

What makes a “good” employee?

[ Higher sales, shorter production time, longer tenure ]

Different choices have different effects.

---

---

---

# Street Bump

Crowdsourced effort to measure road quality.

Users install mobile app that uses GPS and accelerometer to detect pot holes.

---



---

# Measurement error

Reported potholes  $\neq$  actual potholes.

Uneven distribution of smart phone ownership could bias data collection efforts, and lead to biased deployment of resources.

---



*Bernard Parker, left, was rated high risk; Dylan Fugett was rated low risk. (Josh Ritchie for ProPublica)*

# 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

---

## **Pretrial risk assessments**

COMPAS score rates a defendant's likelihood to reoffend on a scale from 1 to 10.

High-risk defendants more likely to be detained by a judge while awaiting trial.

---

---

---

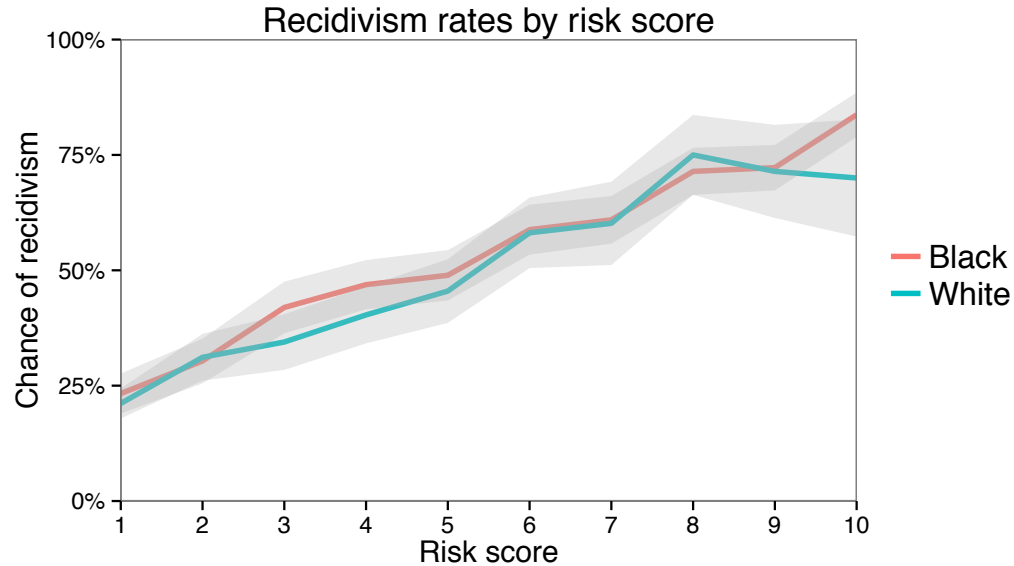
# **Black defendants are detained more often than whites**

Algorithm does not use race.

---

---

# But scores are not obviously biased



---

# **Why are blacks detained more often than whites?**

- Defendants with more serious criminal histories are more likely to reoffend and thus more likely to be detained.
  - Black defendants on average have more serious criminal histories.
-

---

# Are the data biased?

*Reported* crime is a proxy for *actual* crime.

But serious crime is ostensibly less susceptible to bias than minor crime.

Unclear how big the effect might be, or even whether the data are biased against whites or blacks.

---

---

# Are algorithms like **COMPAS** fair?

It's complicated.

[ Disparate impact is not discrimination ]

---



---

# What can [ should ] we do?

## Option 1: Stop using algorithms

Judges are trying to do the same thing as algorithms (determine who's risky), but are typically worse at it.

---

---

# What can [ should ] we do?

## **Option 2: Set race-based detention thresholds**

Detain whites who are at least 20% likely to commit a violent crime, and blacks who are at least 30% likely.

Probably violates the Equal Protection Clause.

---

---

# What can [ should ] we do?

## **Option 3: Move away from pre-trial detention**

Only detain defendants who pose serious threat to public safety. Bolster pre-trial services.

[ Use algorithms to help determine who is risky ]

---

---

# Target's pregnancy predictions

Goal is to identify and reach expectant parents early.

Leading indicators of joining Target's baby-shower registry include unscented soap and lotion, vitamin supplements, and hand sanitizers.

---

---

---

# **Target's pregnancy predictions**

Angry father berates store manager for his teenage daughter receiving maternity coupons.

---

---

# **Target's pregnancy predictions**

Angry father berates store manager for his teenage daughter receiving maternity coupons.

Father later discovers daughter was pregnant.

---

---

**Targeted decisions can reveal  
sensitive information**

---

---

---

# Algorithmic fairness

Algorithms have the potential to improve both the efficiency and equity of decisions.

But their use raises complex ethical questions that we're only beginning to appreciate.

---