

**Context & Refinement**

**Extreme Digital Vetting of Visitors to the U.S. Moves Forward Under a New Name**

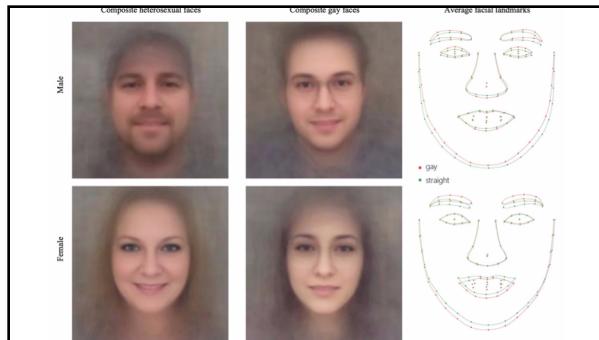


ICE officials have invited tech companies, including Microsoft, to develop algorithms that will track visa holders' social media activity.

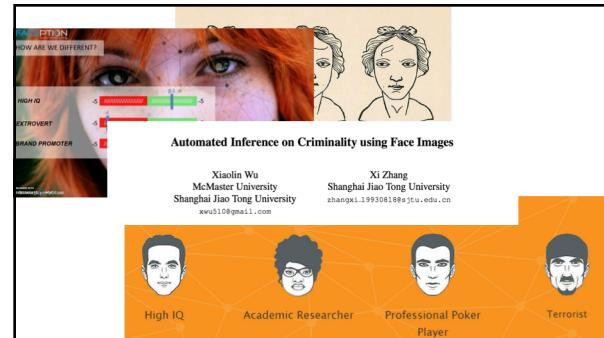
by George Joseph, Nov. 22, 8 a.m. EST

**Context & Refinement**

**A Stanford scientist says he built a gaydar using "the lamest" AI to prove a point**



## Proxies

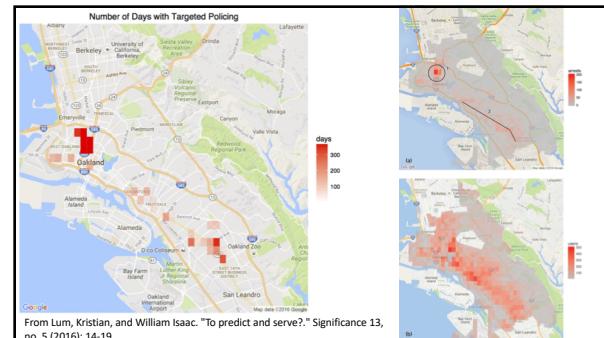


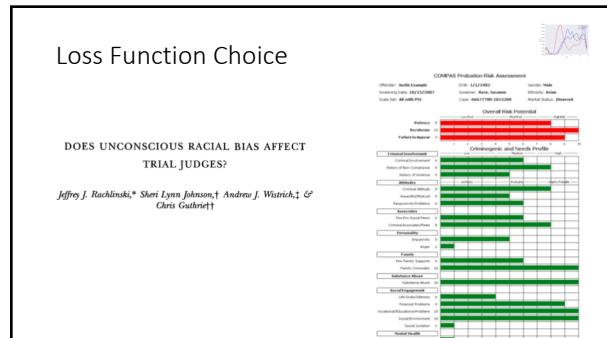
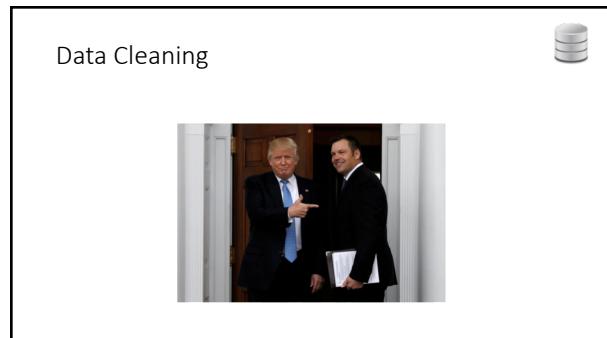
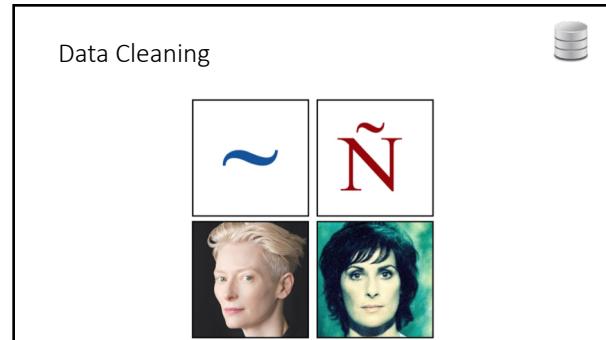
## Data Collection

## Data Collection



<http://bit.ly/ds100-sp18-eth>



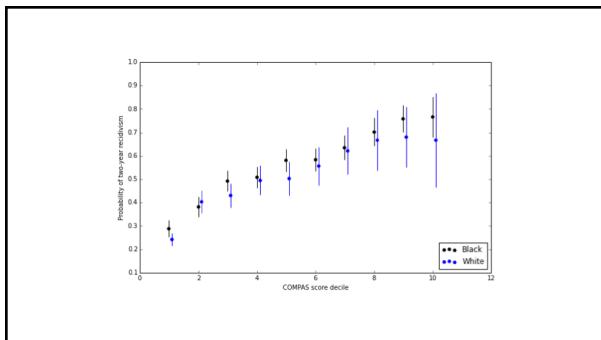
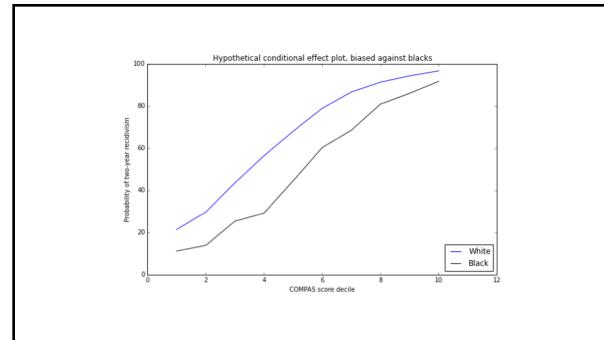
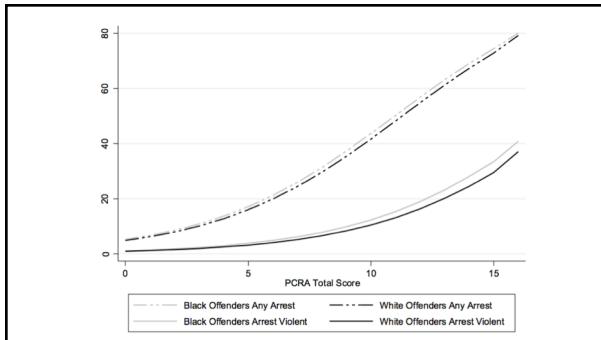


## Machine Bias

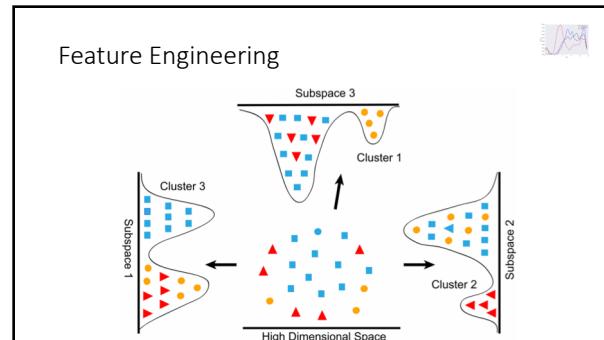
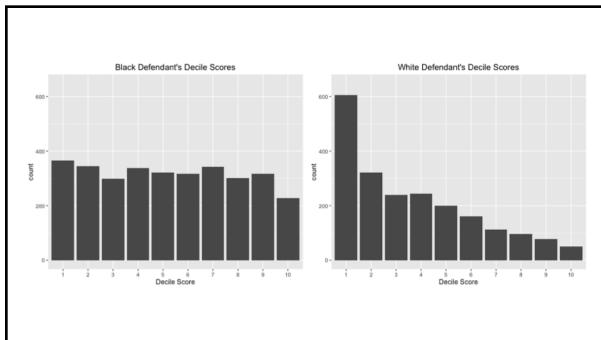
There's software used across the country to predict future criminals. And it's biased against blacks.

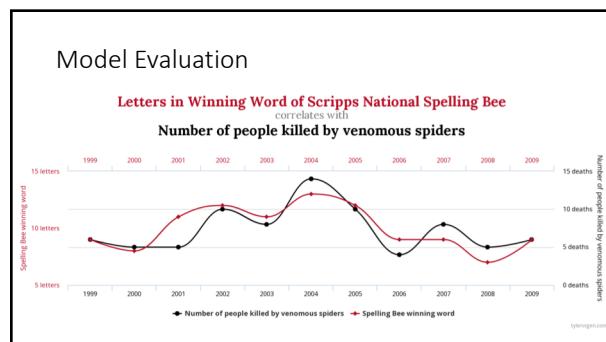
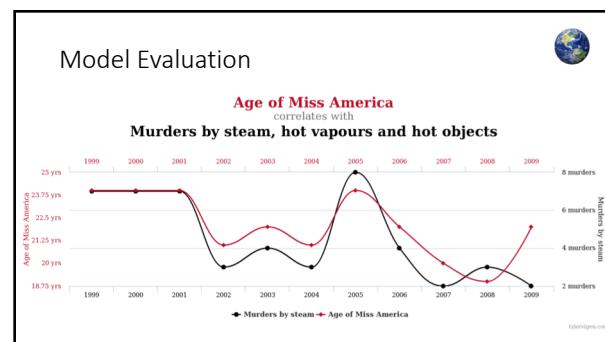
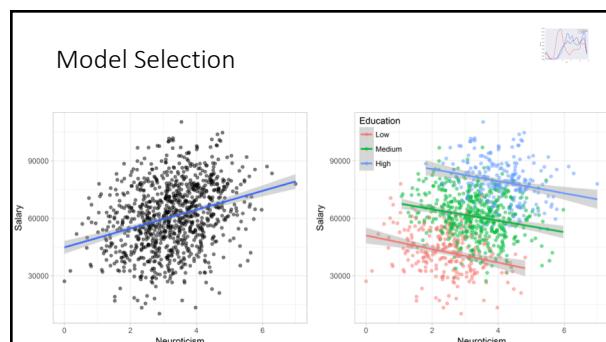
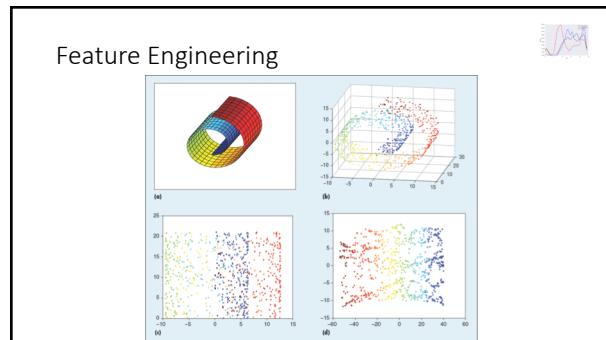
by Julia Angwin, Jeff Larson, Saroya Matto and Lauren Kirchner; ProPublica  
May 23, 2016

**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.



So why does ProPublica think  
COMPAS is biased?





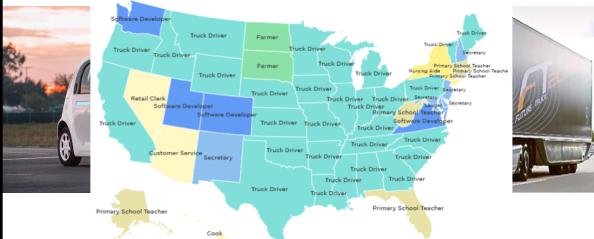
So what do we do about this?

## Ethical Lessons

1. Know what the data are telling you
  - Understand your model, and validate it
  - Involve domain experts; “raw data” is an oxymoron
  - Communicate not just model *capabilities*, but also *limits* and *assumptions*
2. Data about people are about people
  - The data do not exist on their own, someone made them
3. *Always* consider why your model might be wrong in systematic ways
  - Don’t just **trust** the technology, make it **trustworthy**
  - **Close** the data science lifecycle: the world changes, and so should your models

## Ethics beyond the Data Science Lifecycle

### Should we deploy technology at all?



### When is it ethical to collect data?

RYAN SINGEL SECURITY 02.17.12 03:02 PM  
**GOOGLE BUSTED WITH HAND IN SAFARI-BROWSER COOKIE JAR**



### When is it ethical to collect *sensitive* data?

- Without knowing sensitive attributes, it’s not possible to know how to make good decisions.
  - E.g., in a credit-granting context, there’s no reason to believe that qualified minority applicants will be similar to qualified members of the majority
- But having sensitive data means it can be repurposed for other uses

### Inclusion and Representation



## Resources

- Take a course at Berkeley on people and technology
- Visit <https://fatml.org> and <https://fatconference.org>
  - Reading list
  - Principles for Accountable Algorithms
- Books
  - *Weapons of Math Destruction*, by Cathy O'Neil
  - *How to Lie With Statistics*, by Darrell Huff
- Courses around the world: [tinyurl.com/ethics-classes](http://tinyurl.com/ethics-classes)

*"Remember that all models are wrong;  
the practical question is how wrong do  
they have to be to not be useful."*

-George E. P. Box

