

Race and the Machine

Re-examining Race in Data Analysis

Sophie Chou
9.28.2015

First, some links:

Tweet: @mpetitchou

Paper: bit.ly/racepaper

#d4gx

What is Race?



Essentialist

fixed, immutable, pre-determined.

Constructionist

contextual, behavioral, social.

TL;DR

1. If you are a Data Scientist, be a Social Constructionist*
2. Algorithms can be racist, but we can do better!

Why?

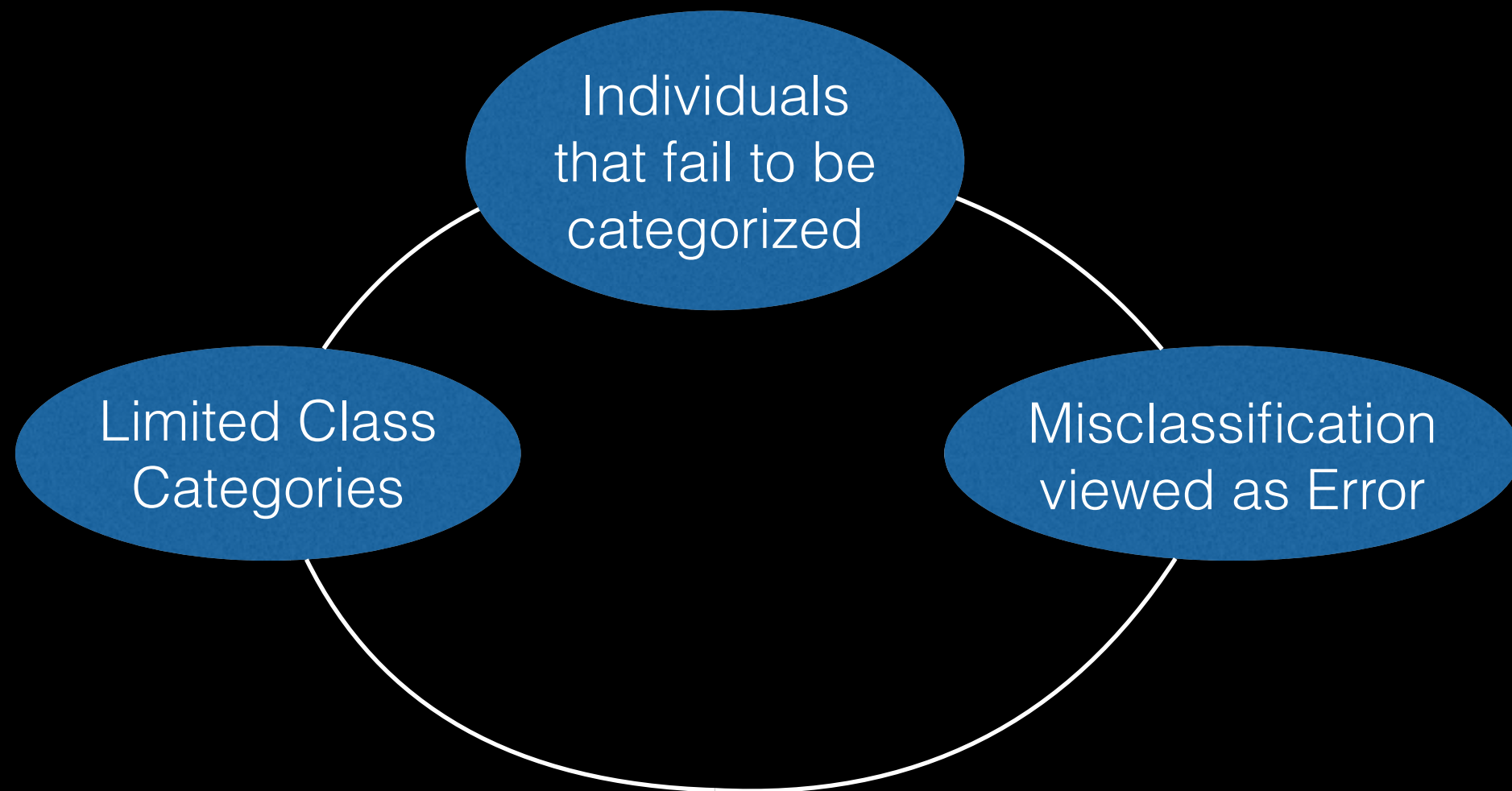
By definition, data mining* is the *mining of observed behaviors within a social context*.

Data isn't static, and neither are people's behavior.

Making the assumption that race *is not* performative limits analysis.

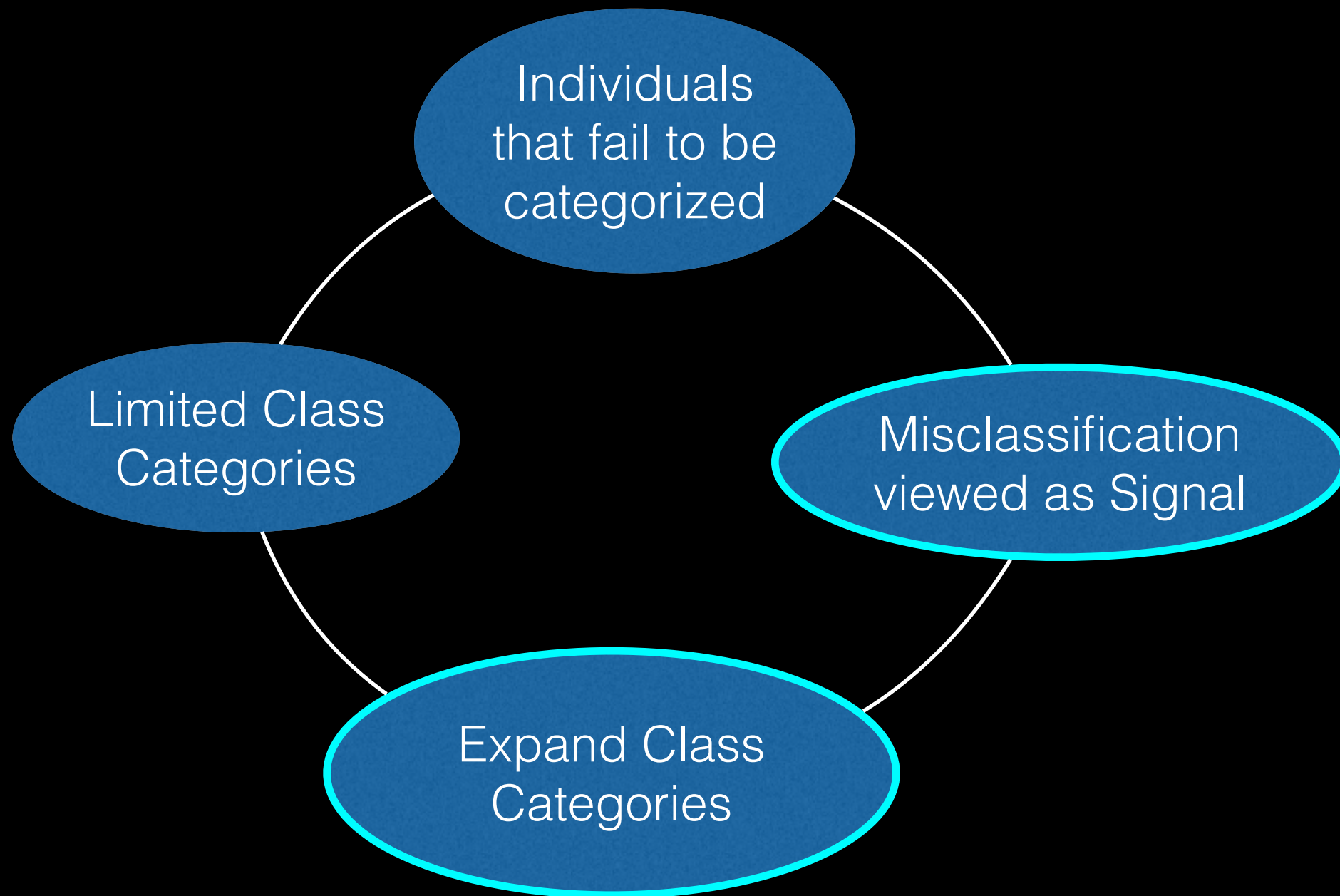
Social constructionism takes that into account.

Stereotyping



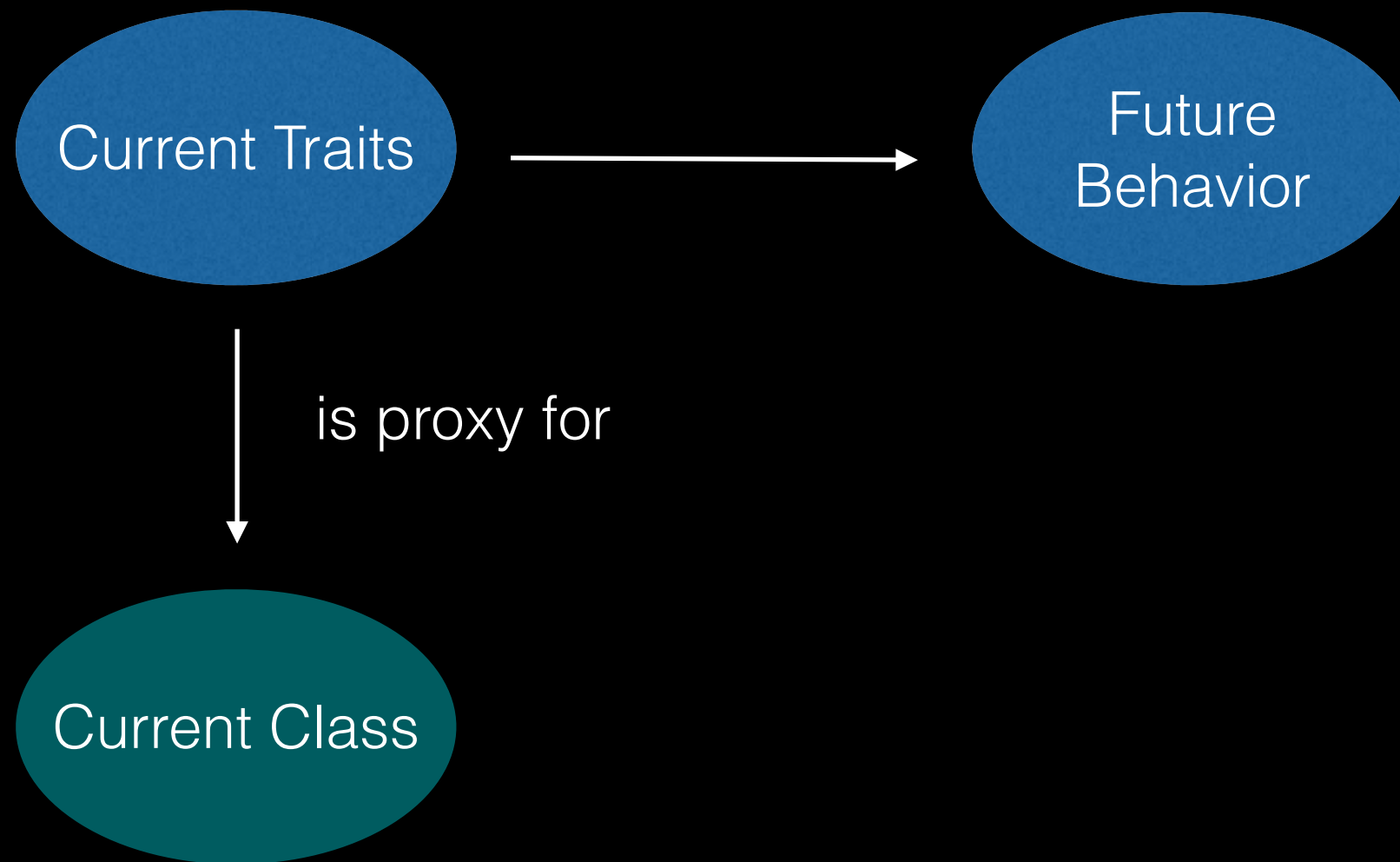
Assumption that behavior explains race, rather than race **is behavioral**.

Breaking the Cycle



Behavior is a **snapshot in time**.
People might belong to several classes.

Indirect Profiling



Latanya Sweeney Truth

www.instantcheckmate.com/

Looking for Latanya Sweeney? Check Latanya Sweeney

Ads by Google

Latanya Sweeney, Arrested?

1) Enter Name and State. 2) Access Full Background
Checks Instantly.

www.instantcheckmate.com/

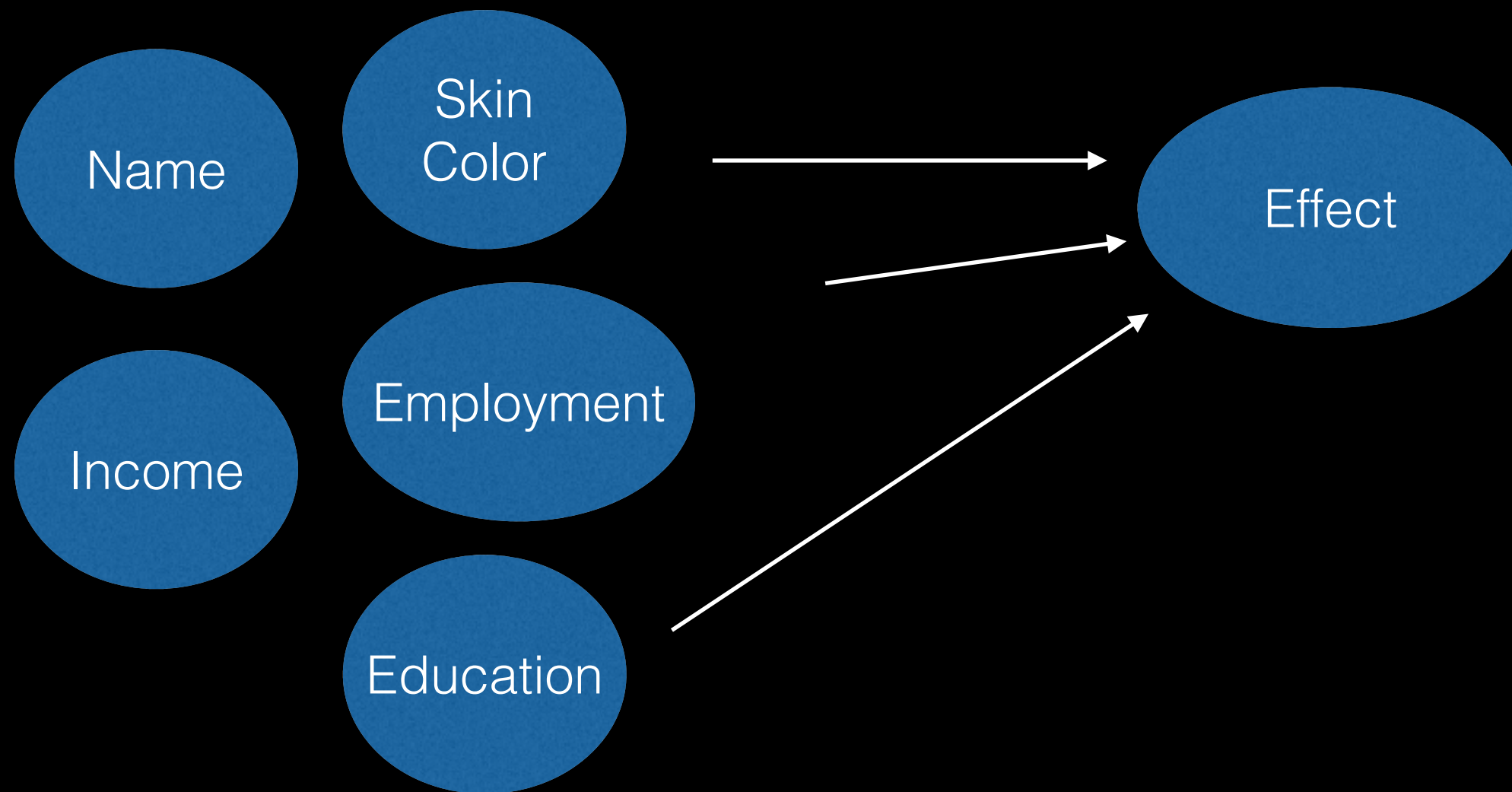
Latanya Sweeney

Public Records Found For: Latanya Sweeney

Unbundling Race



Unbundling Race



How can we do better?

- View outliers as signal
- Include contextual, temporal considerations
- Unbundle race
- Monitor Algorithms
- Your idea here!

How we think about race vs.
what historical data tells us:

for more, see
bit.ly/census-race

