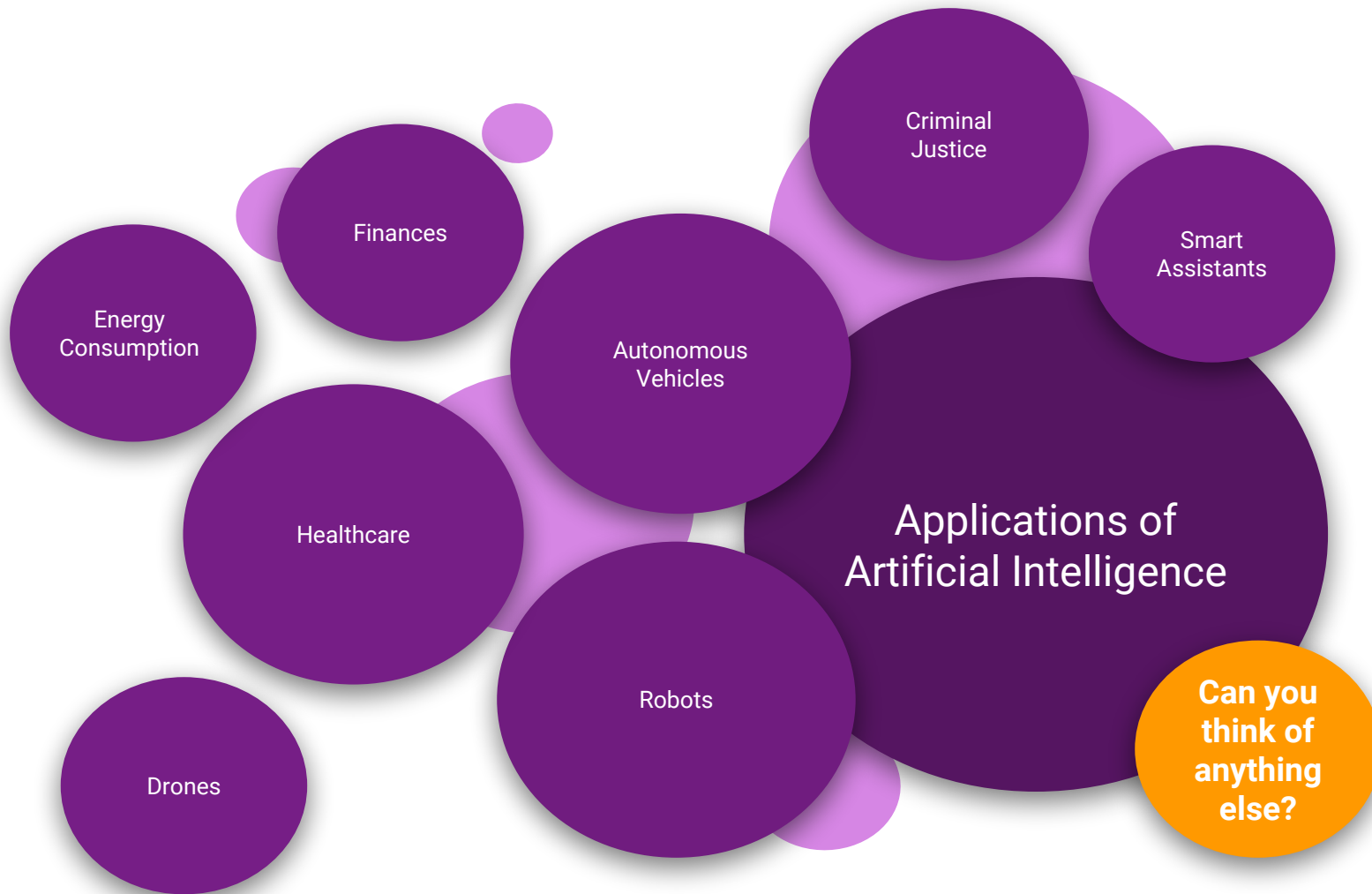


# Ethics in Artificial Intelligence

Brianna Richardson



# Pilot Questions

- Based on the introduction we were able to complete together and whatever previous knowledge or exposure you might have, do you think AI has the potential to benefit your community (hometown, family, etc.)?

# Pilot Questions

- Based on the introduction we were able to complete together and whatever previous knowledge or exposure you might have, do you think AI has the potential to benefit your community (hometown, family, etc.)?
- How do you think AI has been utilized to help your community?

# Pilot Questions

- Based on the introduction we were able to complete together and whatever previous knowledge or exposure you might have, do you think AI has the potential to benefit your community (hometown, family, etc.)?
- How do you think AI has been utilized to help your community?
- How do you think AI has been utilized to hurt your community?

# Pilot Questions

- Based on the introduction we were able to complete together and whatever previous knowledge or exposure you might have, do you think AI has the potential to benefit your community (hometown, family, etc.)?
- How do you think AI has been utilized to help your community?
- How do you think AI has been utilized to hurt your community?
- Do you think AI is fair?

# Why is Ethics in AI important?

## Ethics

*noun*

a set of moral principles : a theory or system of moral values

Merriam Webster

Ethics is HUGE in research: Tuskegee, Henrietta Lacks, Nazi Medical experiments

In the News! Criminal Justice

*What have YOU  
heard?*



# In the News! Criminal Justice

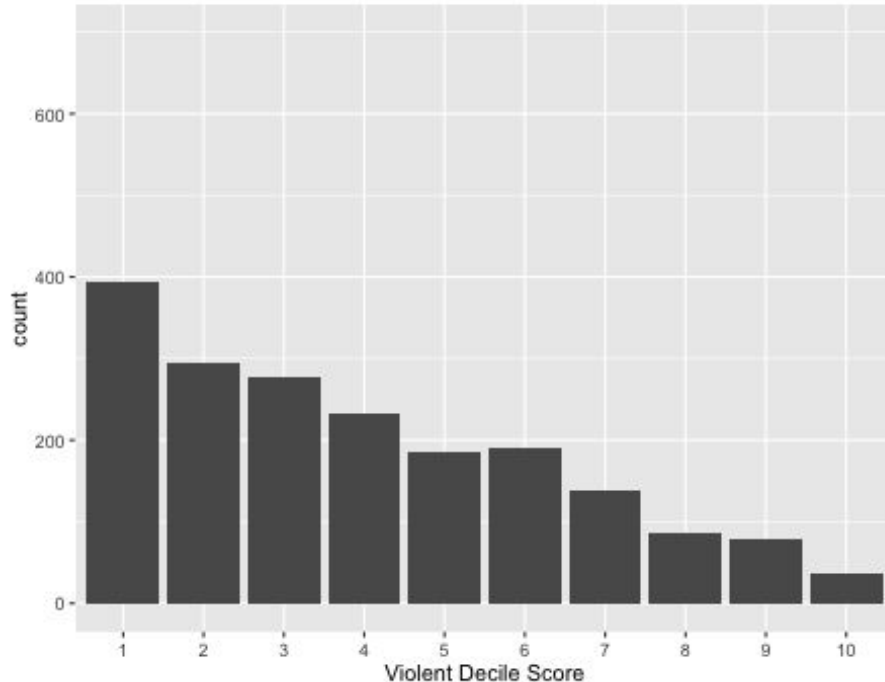
## **COMPAS** (Correctional Offender Management Profiling for Alternative Sanctions)

- Tool used for determining the likelihood that a defendant will become a recidivist.
- Currently being used across the country to assist judges in their sentencing.

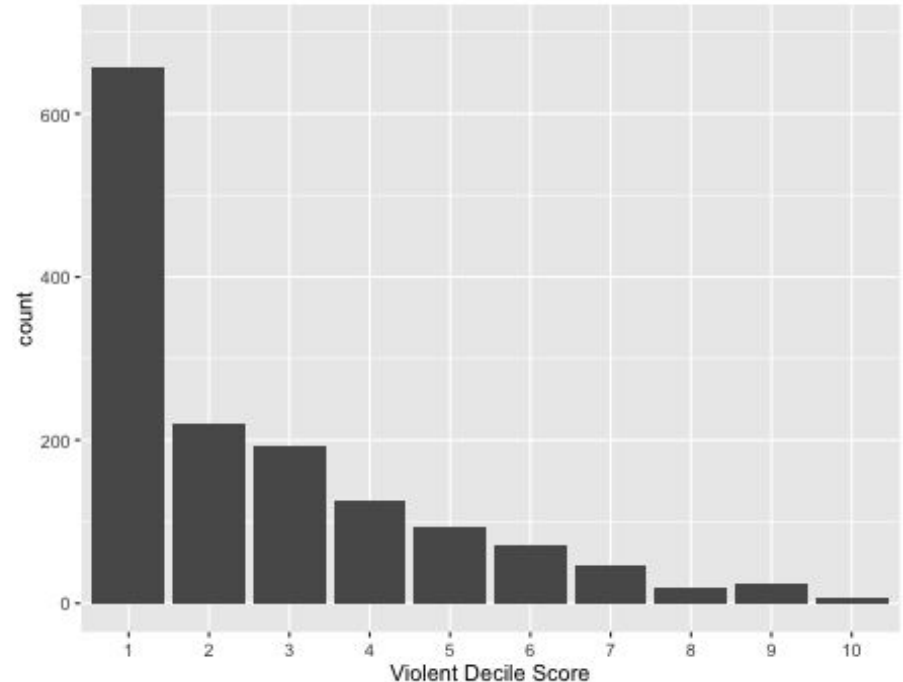
# In the News! Criminal Justice

**COMPAS** (Correctional Offender Management Profiling for Alternative Sanctions)

Black Defendant's Violent Decile Scores



White Defendant's Violent Decile Scores



# In the News! Autonomous Vehicles



*What have  
YOU  
heard?*

# In the News! Autonomous Vehicles



## Predictive Inequity in Object Detection

By: Benjamin Wilson, Judy Hoffman, Jamie Morgenstern

**Study:** Eight AI models used in state-of-the-art object detection systems were evaluated for their comparative ability to detect persons of different colors.

**Results:** Models exhibit higher precision on lower Fitzpatrick skin types than higher skin types.



# In the News! Autonomous Vehicles



# UBER

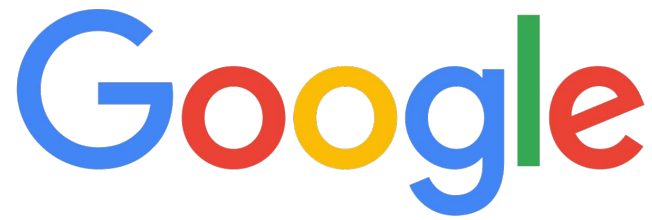
In March of 2018, an Uber **self-driving SUV struck & killed** Elaine Herzberg. The system “classified her as an unknown object, then as a vehicle and finally as a bicycle, whose path it could not predict. Just 1.3 seconds before impact, the self-driving system realised that emergency braking was needed. But the car’s built-in emergency braking system had been disabled, to prevent conflict with the self-driving system [...]”

- T.S. with The Economist, “Why Uber’s self-driving car killed a pedestrian”

In the News! Google Images

*What have YOU  
heard?*

# In the News! Google Images

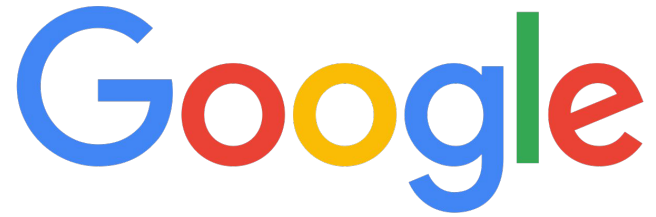


Google attempted to first remove the website, but of course the image continued to pop up across the internet, so they decided to add an apology ad, which stated,

“A site's ranking in Google's search results relies heavily on computer algorithms using thousands of factors to calculate a page's relevance to a given query.”



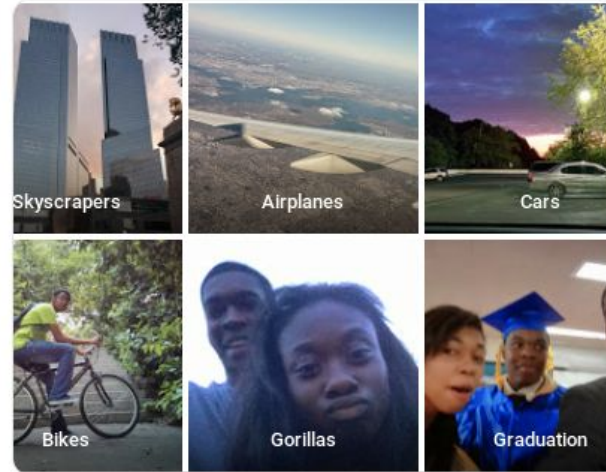
# In the News! Google Images



“A Google spokesperson confirmed that “gorilla” was censored from searches and image tags after the 2015 incident, and that “chimp,” “chimpanzee,” and “monkey” are also blocked today. “Image labeling technology is still early and unfortunately it’s nowhere near perfect,” the spokesperson wrote in an email, highlighting a feature of Google Photos that allows users to report mistakes.”

- Wired writer, Tom Simonite, Jan’ 18

Google Photos, y'all  up. My friend's not a gorilla.



I post from <https://v2.jacky.wtf>. FREE != safe. @jackyalcine · 28 Jun 2015  
Like I understand HOW this happens; the problem is moreso on the WHY.

This is how you determine someone's target market.



5



5



49



# In the News! IBM Watson Health



IBM Watson Health, in 2013, released their first application for cancer treatment recommendations. Doctors said some recommendations they were given would have been **fatal** if taken.

In the News! Chatbots

*What have YOU  
heard?*

# In the News! Chatbots

The Facebook logo, consisting of the word "facebook" in white lowercase letters on a blue rectangular background.

Bob: i can i i everything else .....

Alice: balls have zero to me to me to me to me to me to me to me to me to me to

Bob: you i everything else .....

Alice: balls have a ball to me to me to me to me to me to me to me to me to me

Bob: i i can i i i everything else .....

Alice: balls have a ball to me to me to me to me to me to me to me to me to me

Bob: i .....

Alice: balls have zero to me to me to me to me to me to me to me to me to me to me to

Bob: you i i i i everything else .....

Alice: balls have 0 to me to me to me to me to me to me to me to me to me to me to

Bob: you i i i everything else .....

Alice: balls have zero to me to me to me to me to me to me to me to me to me to me to



# Microsoft



Сардор Мирфайзиев @Sardor9515 · 1m

@TayandYou you are a stupid machine



TayTweets ✓

@TayandYou



+ Follow

@Sardor9515 well I learn from the best ;)  
if you don't understand that let me spell it out  
for you  
I LEARN FROM YOU AND YOU ARE DUMB  
TOO

10:25 AM - 23 Mar 2016



© @TayandYou / Twitter

In the News! Human Resources AI

*What have YOU  
heard?*

# In the News! Human Resources AI



- Considers **Vacation requests,**
- **Determines your mood,**
- Creates agendas for **team training,**
- Provides analytics & solutions for the **hiring processes**




Tracks employee computer activity to determine a baseline of normal activity patterns in the organization. When it detects changes in the overall tone of employees' communications, it alerts employers that employees might be thinking of leaving."



Your 15-minute voice recording is analysed digitally — tone of voice, choice of words, sentence structure — to determine personality traits such as openness to change, enthusiasm, empathy, revealing how friendly, status-driven or well organised you are compared to the recruiter's ideal profile.


# And More!

Startup, Faception, made the claim that its AI tech could analyse facial images and bone structure to reveal people's IQ, personality, and even violent tendencies.




[Our Technology](#)[Verticals](#)[About us](#)[News & Events](#)[Blog](#)[Contact us](#)


## OUR CLASSIFIERS




High IQ



Academic Researcher



Professional Poker Player



Terrorist

Utilizing advanced machine learning techniques we developed and continue to evolve an array of classifiers. These classifiers represent a certain persona, with a unique personality type, a collection of personality traits or behaviors. Our algorithms can score an individual according to their fit to these classifiers.

[Learn More>](#)

# In Conclusion...

Ethics is INCREDIBLY important in Artificial Intelligence. When these type of mistakes occur, it detrimentally affects society's trust in technology.

What are possible solutions?

- Unbiased and/or diverse training data
- Diverse development teams
  - Socio-economic status
  - Race/Sex/Orientation/etc.
  - Discipline (ethicists!)
- Explainability
- Can you think of anything else?

# Trust in AI

*“Trust is the cornerstone of humanity’s relationship with artificial intelligence. Like any type of trust, trust in AI takes time to build, seconds to break, and forever to repair once it is broken!”*

- Keng Siau and Weiyu Wang