Ethics in Computer Science

COMP4920

Week 5: Bias and Fairness

Flora Salim

Acknowledgment of Country

I would like to acknowledge the Bedegal people that are the Traditional Custodians of this land. I would also like to pay my respects to the Elders both past and present and extend that respect to other Aboriginal and Torres Strait Islanders who are present here today.

Agenda

- Fairness and Transparency (a quick glance)
- Bias
- Fairness

Transparency

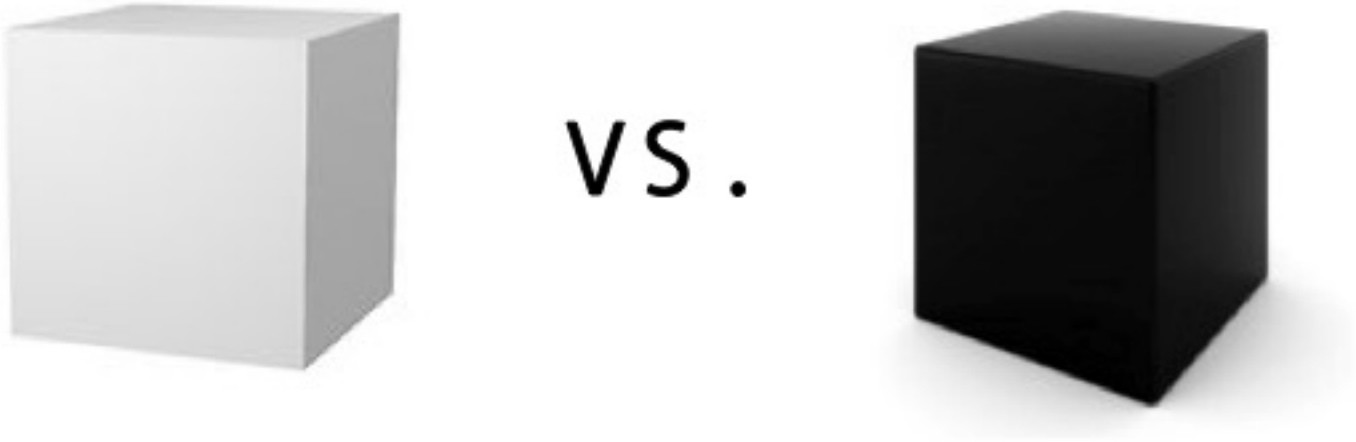
Transparency in AI

Transparency is one of seven key requirements for the realisation of 'trustworthy AI' (EU Commission's High-Level Expert Group on AI (AI HLEG) in April 2019)

"Transparency" is the single most common, and one of the key five principles emphasised in the vast number – a recent study counted 84 – of ethical guidelines addressing AI on a global level (Jobin et al., 2019).

Larsson, S. & Heintz, F. (2020). Transparency in artificial intelligence. Internet Policy Review, 9(2). https://doi.org/10.14763/2020.2.1469; https://policyreview.info/concepts/transparency-artificial-intelligence

Black box vs white box algorithms



Norbert Wiener, 1948, Cybernetics: or Control and Communication in the Animal and the Machine

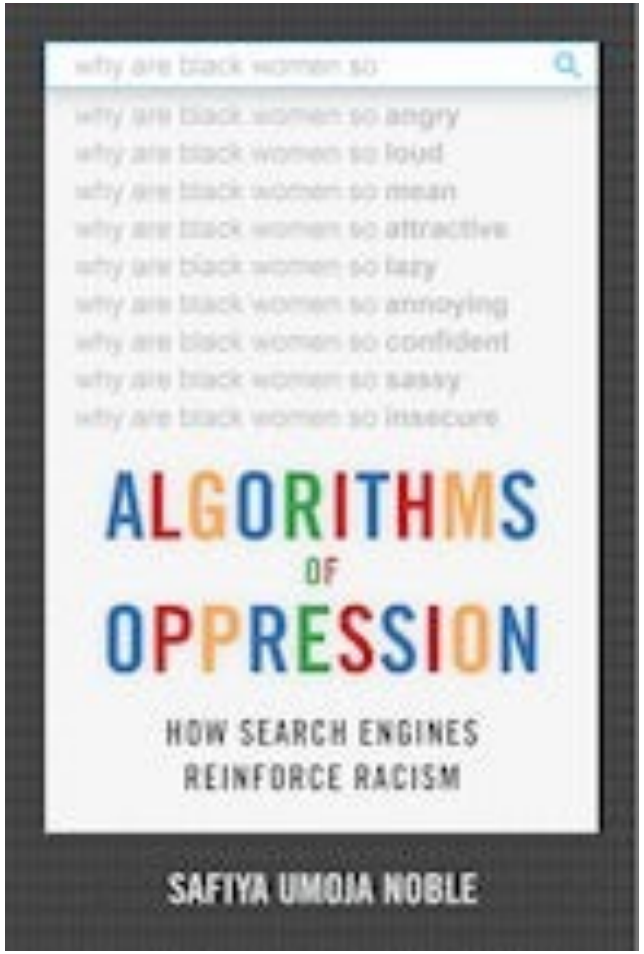
The need for Transparency and Explainable AI (XAI)

- The problems of accountability as computing technologies becoming more complex and less intelligible (Helen Nissenbaum).
- The opacity in Machine Learning Systems (Jean Burrell, 2016) due to :
  - Trade secrets
  - Limited people with the knowledge of programming languages and ML
  - The complexity and high dimensionality of data for decision making no longer match human-scale reasoning
- Institutional transparency, public values, regulations
  - Customer's rights for explanations (GDPR Article 15(1))
  - Requirement for human in the loop (GDPR Article 22)
  - Requirement for algorithmic auditing (US Algorithmic Accountability act)

Source: Jake Goldenfein, 'Algorithmic Transparency and Decision-Making Accountability: Thoughts for buying machine learning algorithms' in Office of the Victorian Information Commissioner (ed), Closer to the Machine: Technical, Social, and Legal aspects of AI (2019), Available at SSRN: https://ssrn.com/abstract=3445873, https://ovic.vic.gov.au/wp-content/uploads/2019/08/closer-to-the-machine-web.pdf p.45-65

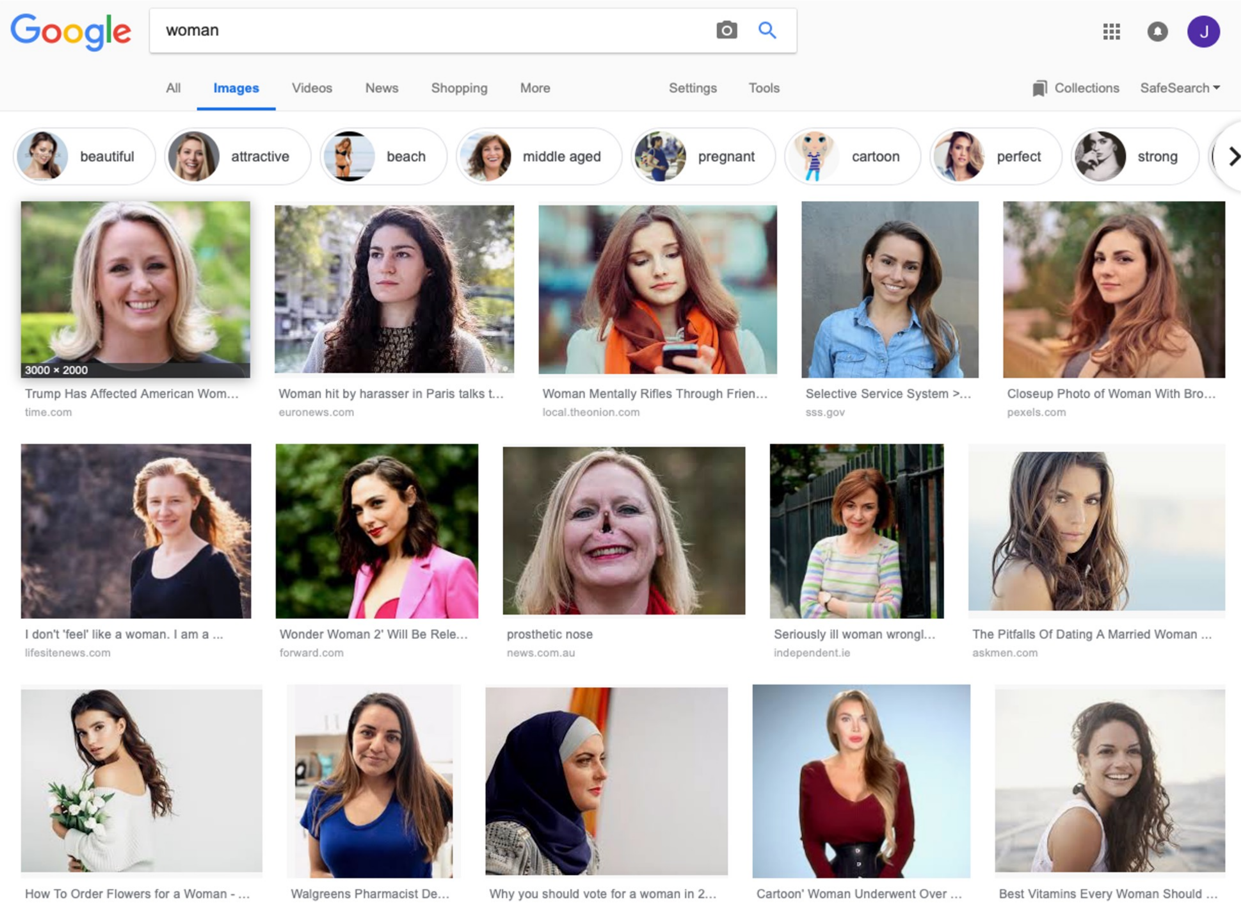
Fairness

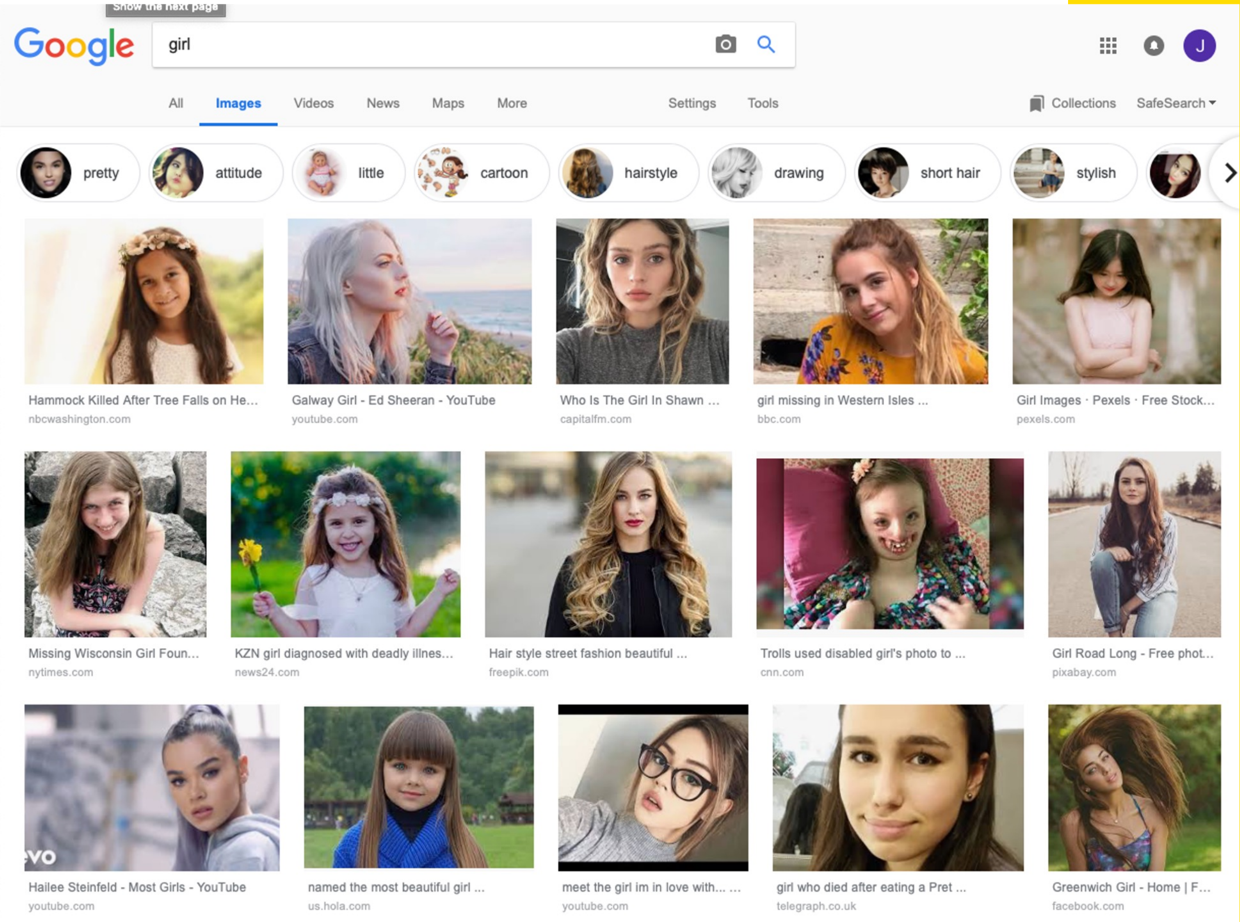
Algorithms of Opression



Noble, S. U. (2018). Algorithms of oppression: How search engines reinforce racism. New York University Press.

Search Engine Bias

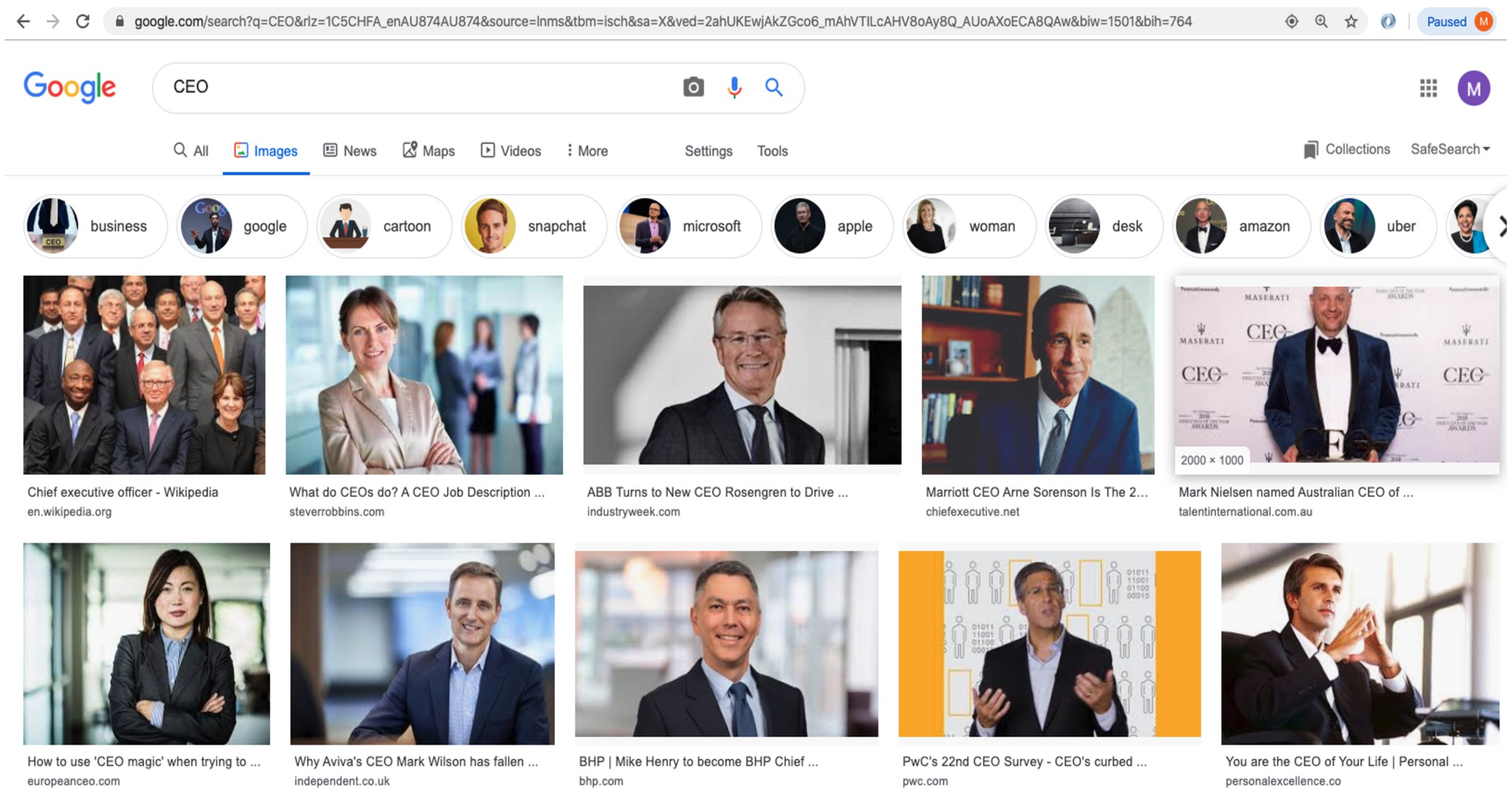




Source from https://scroll.in/article/921305/google-search-results-reflect-its-algorithms-bias-against-women-and-people-of-colour

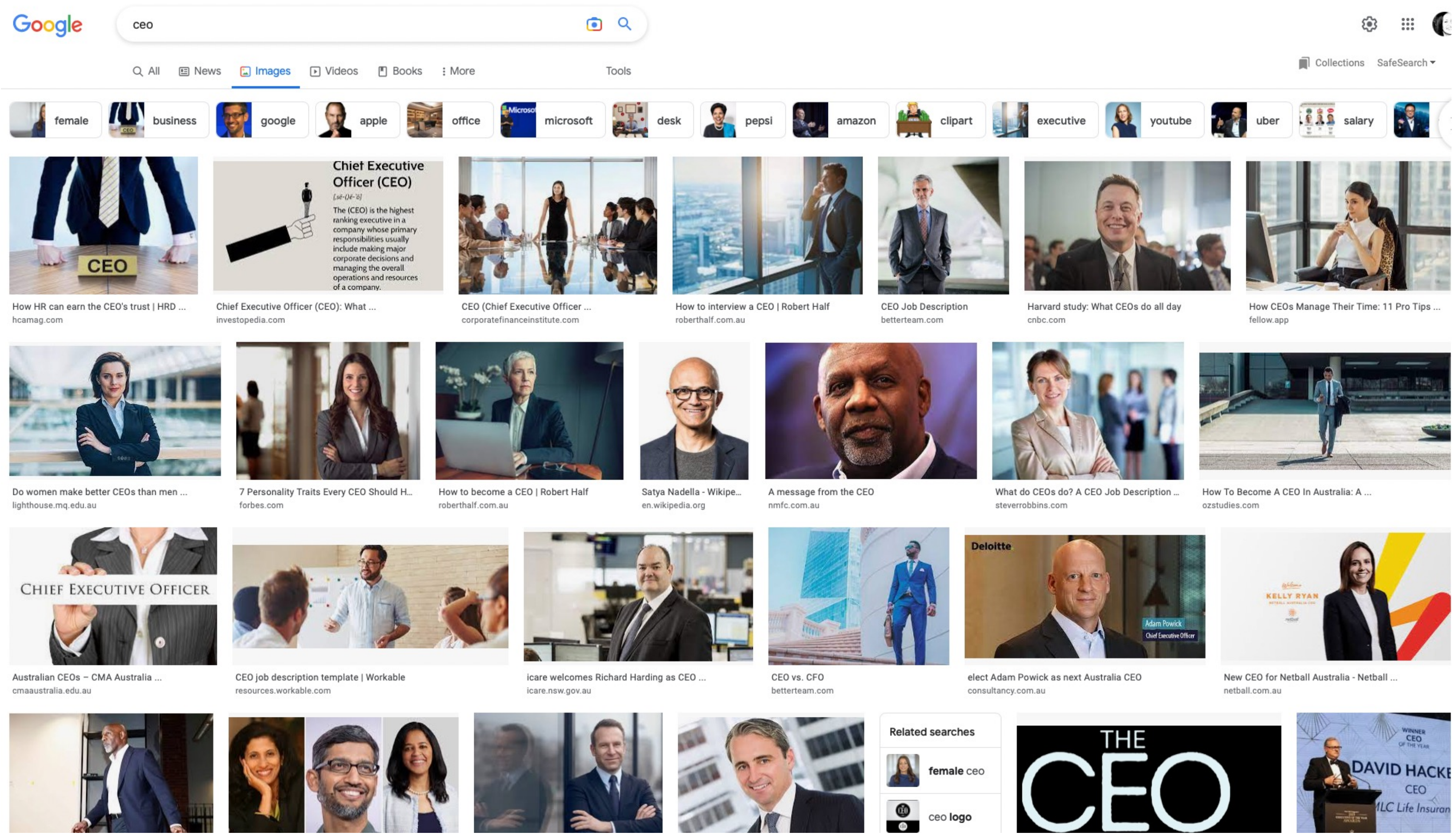
UNSW SYDNEY

Search Engine result on CEO (3 years ago)



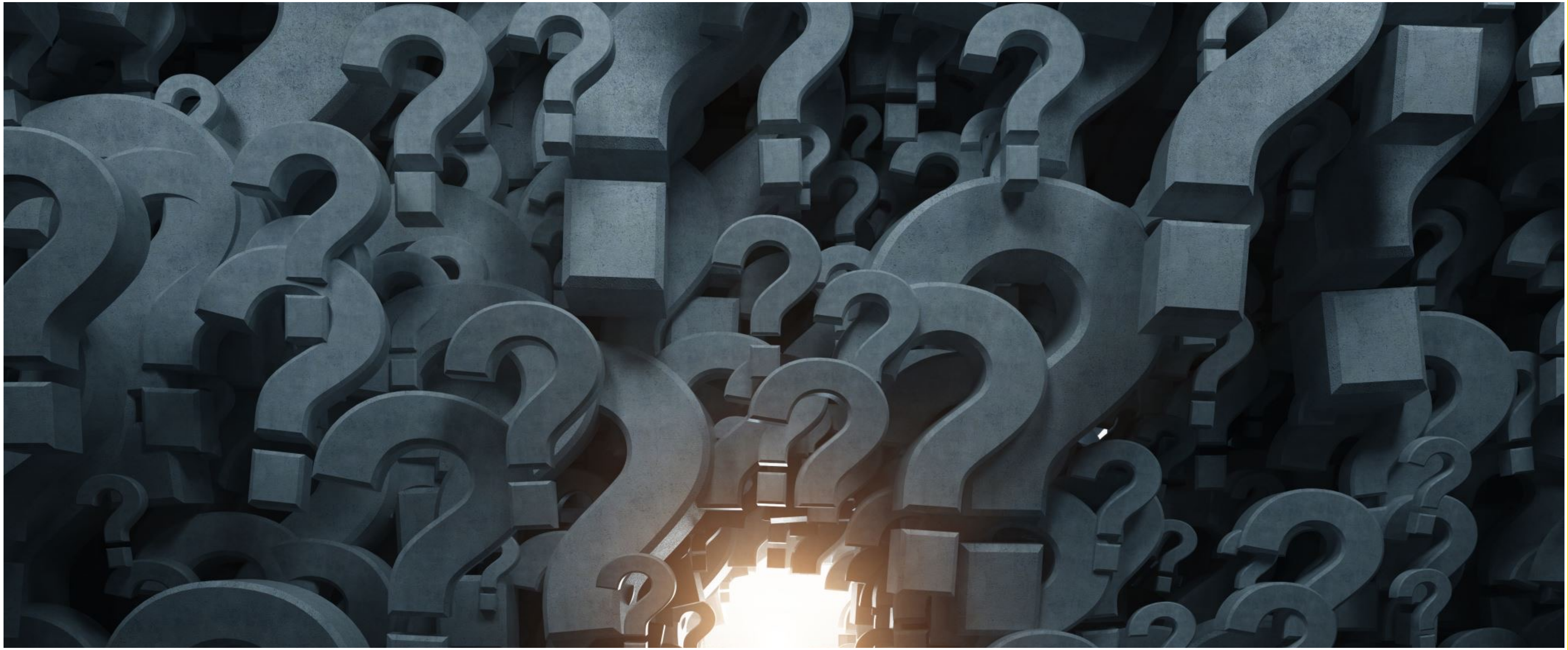
UNSW SYDNEY

Search Engine result on CEO (today)

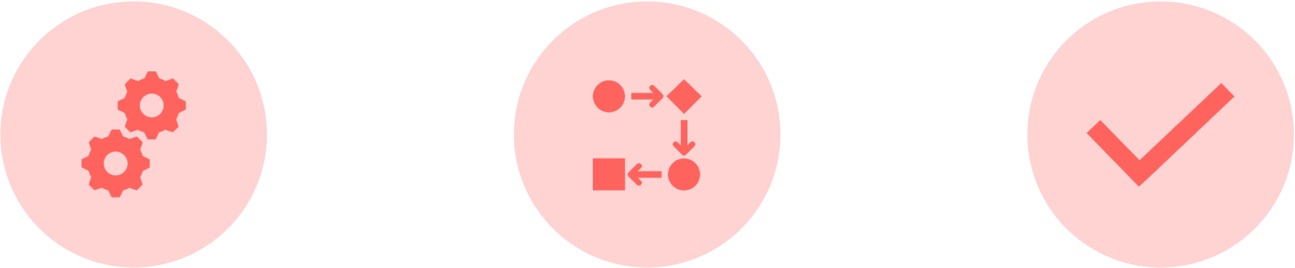


UNSW SYDNEY

The Problem of Bias

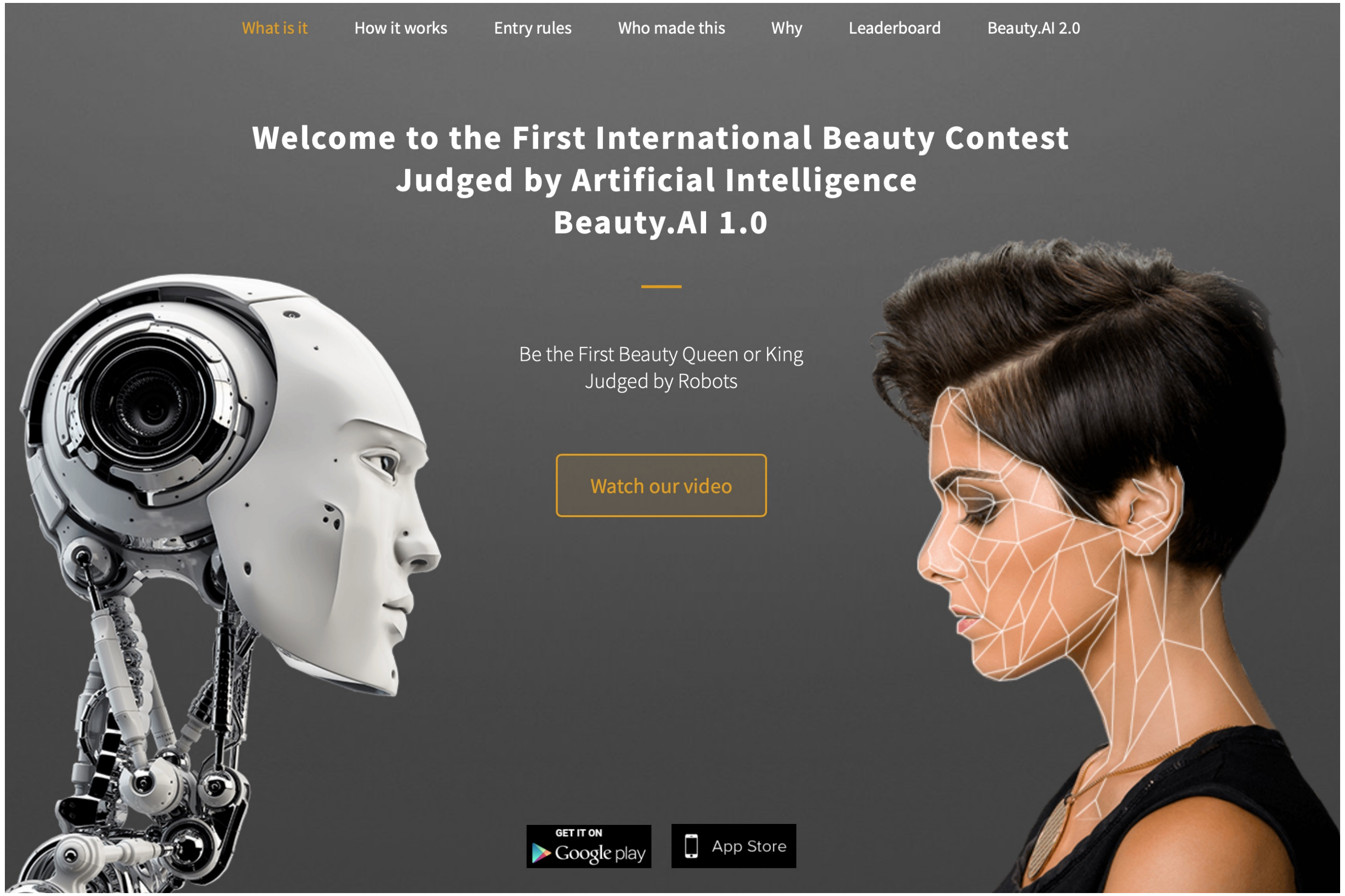


Bias – where do they come from?



INPUT PROCESS OUTPUT

A case study: Beauty.ai

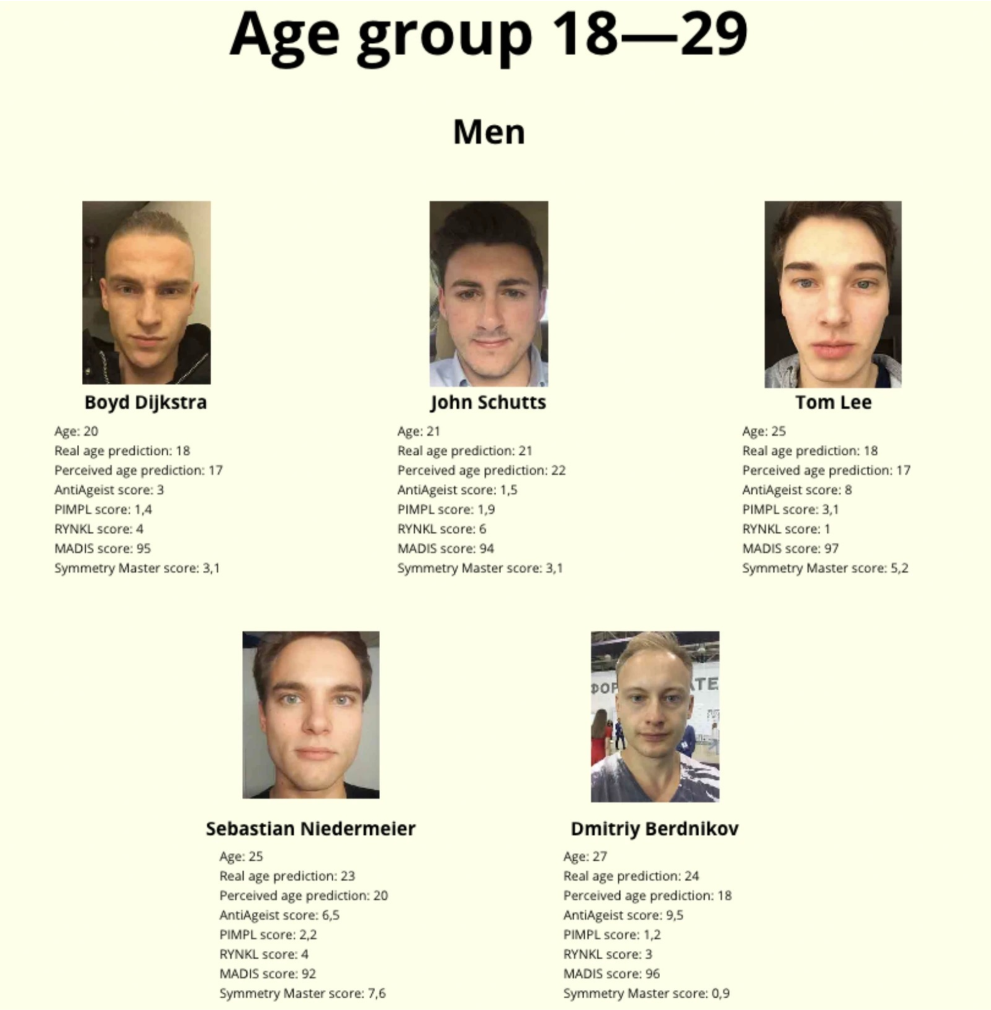


Why An AI-Judged Beauty Contest Picked Nearly All White Winners

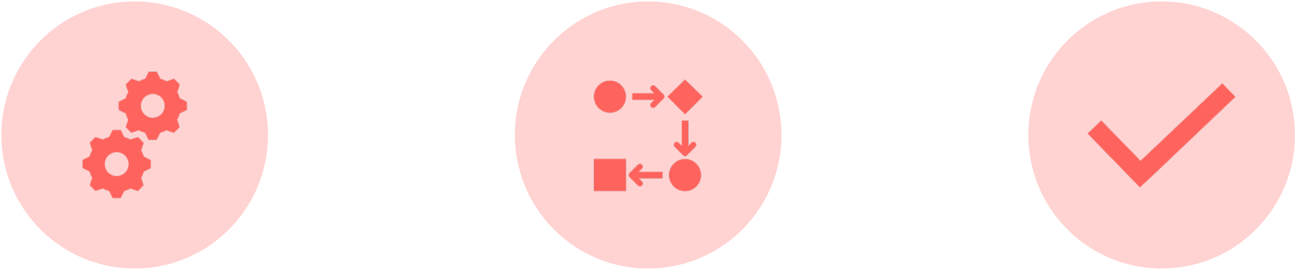
Beauty.ai, an initiative by the Russia and Hong Kong-based Youth Laboratories and supported by Microsoft and Nvidia, ran a beauty contest with 600,000 entrants, who sent in selfies from around the world—India, China, all over Africa, and the US. They let a set of three algorithms judge them based on their face's symmetry, their wrinkles, and how young or old they looked for their age. The algorithms did not evaluate skin color.

The results, released in August, were shocking: Out of the 44 people that the algorithms judged to be the most "attractive," all of the finalists were white except for six who were Asian. Only one finalist had visibly dark skin.

https://www.vice.com/en/article/78k7de/why-an-ai-judged-beauty-contest-picked-nearly-all-white-winners



Bias – where do they come from?



INPUT PROCESS OUTPUT

Bias – where do they come from?

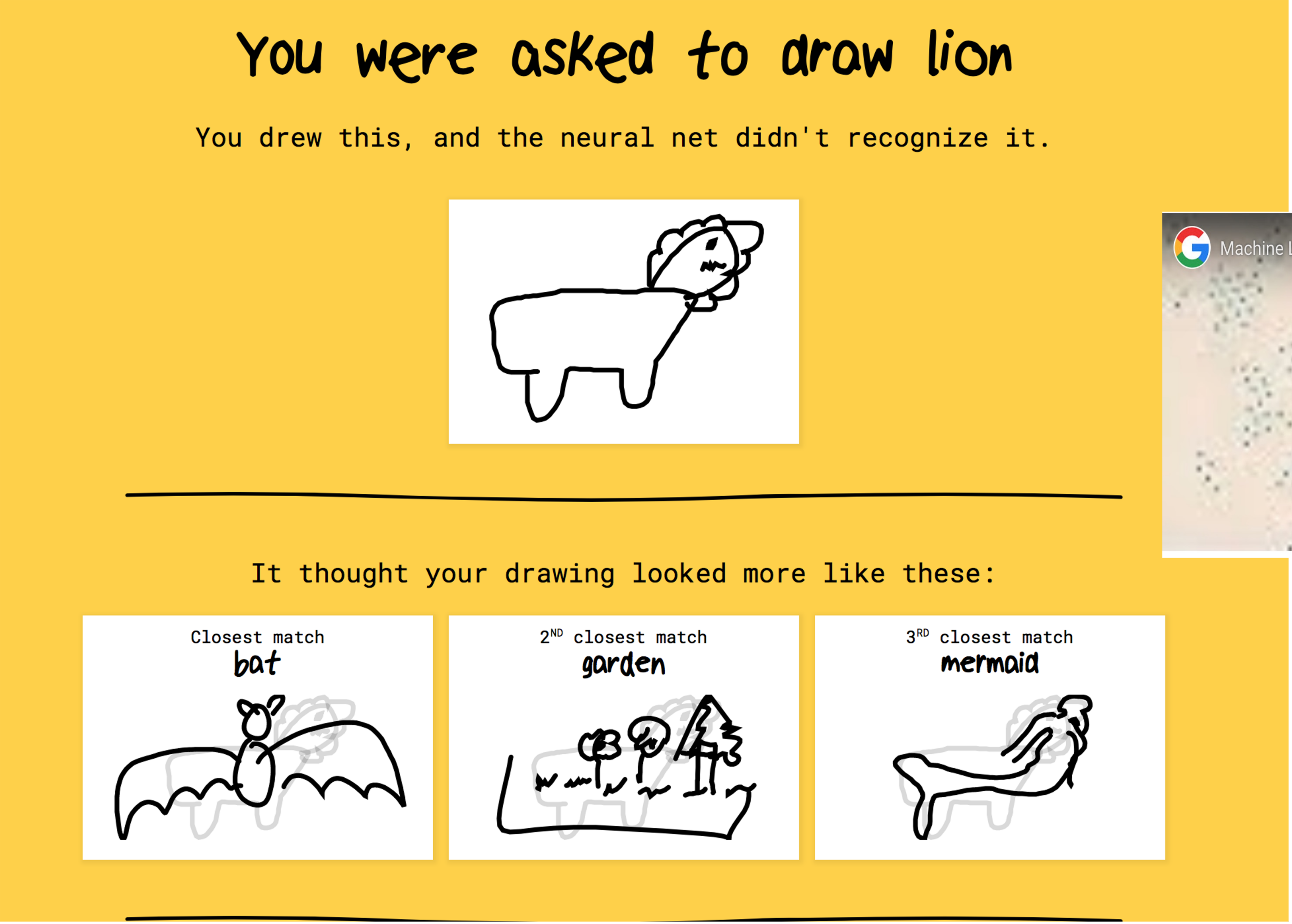
INPUT PROCESS OUTPUT

Bias – where do they come from?

INPUT

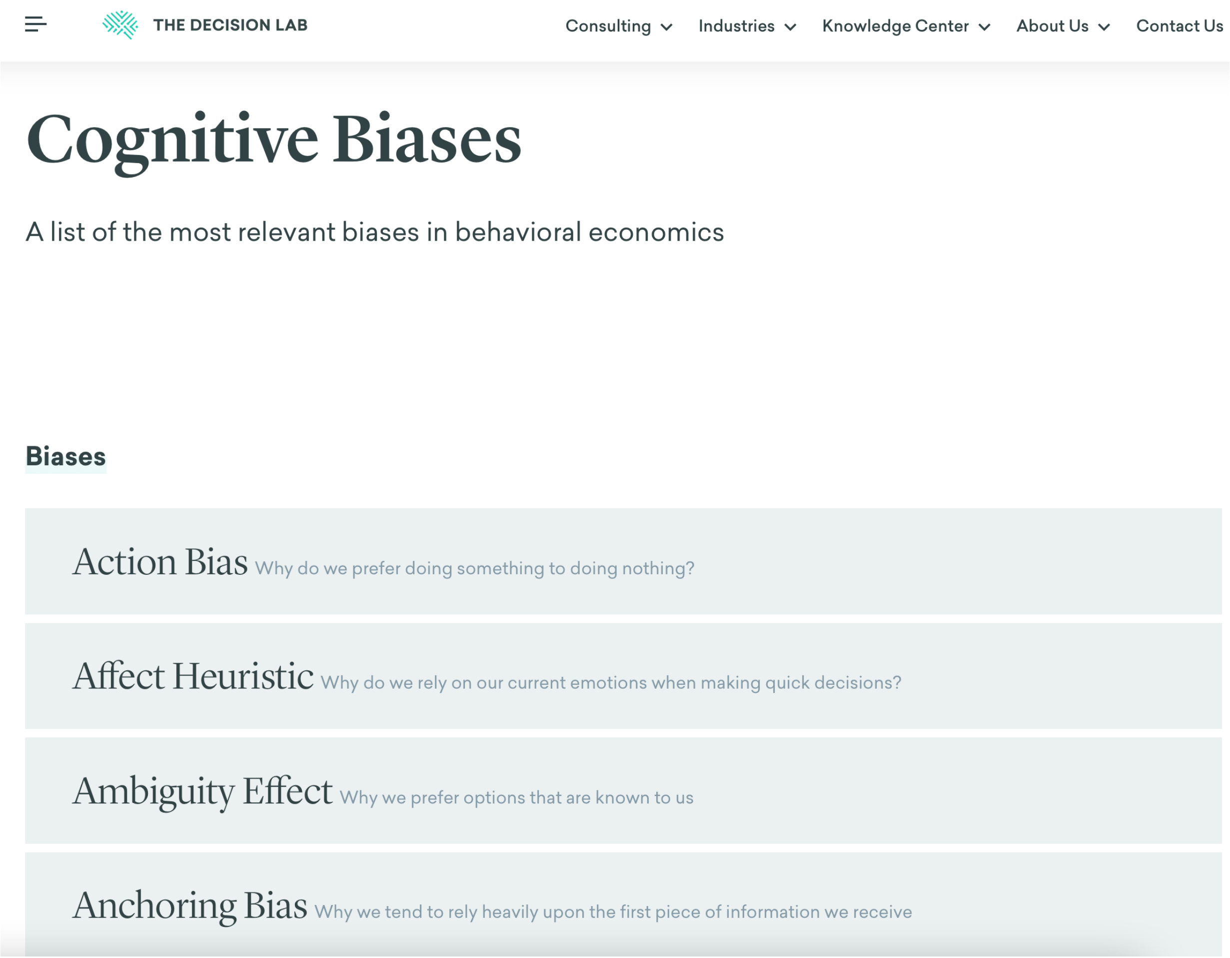
PROCESS

OUTPUT


You were asked to draw lion
You drew this, and the neural net didn't recognize it.
It thought your drawing looked more like these:
Closest match
bat
2ND closest match
garden
3RD closest match
mermaid


Machine Learning and Human Bias

Cognitive Biases (>100 of them)



https://thedecisionlab.com/biases ; https://en.wikipedia.org/wiki/List_of_cognitive_biases

Anchoring Bias

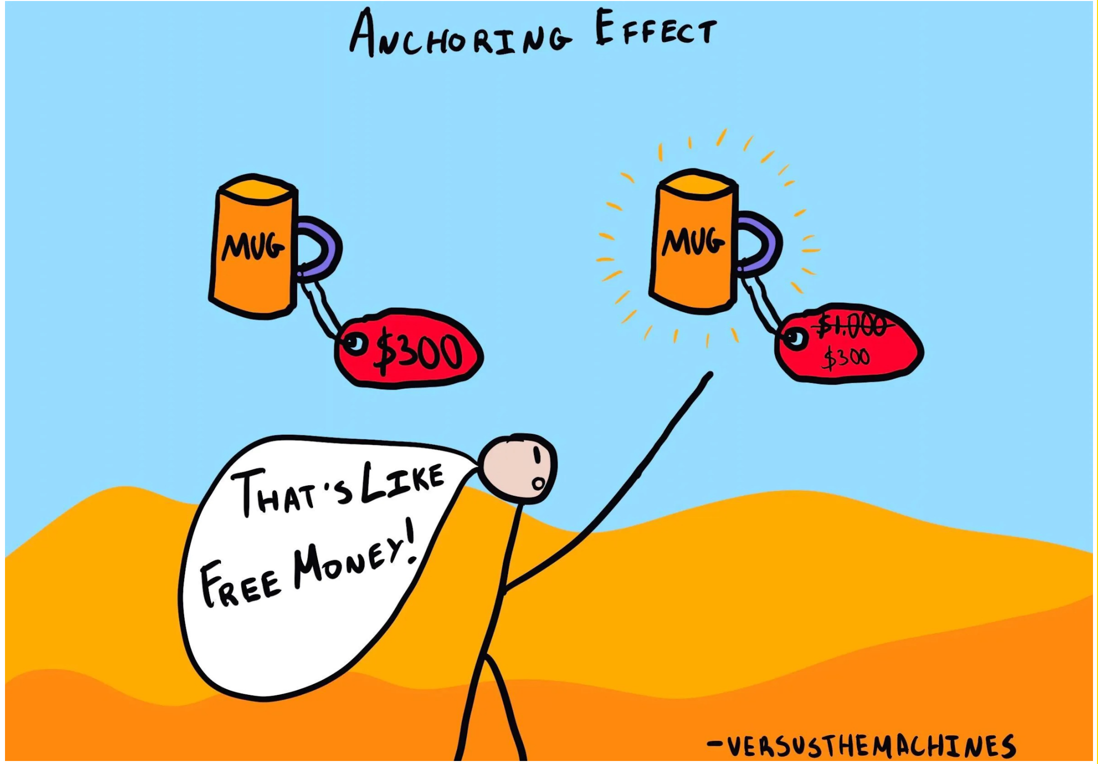
Anchoring bias is a pervasive cognitive bias that causes us to rely too heavily on information that we received early on in the decision making process. Because we use this “anchoring” information as a point of reference, our perception of the situation can become skewed.

https://thedecisionlab.com/biases/anchoring-bias

Anchoring Bias

Anchoring bias is a pervasive cognitive bias that causes us to rely too heavily on information that we received early on in the decision making process. Because we use this "anchoring" information as a point of reference, our perception of the situation can become skewed.

https://thedecisionlab.com/biases/anchoring-bias



Implicit Bias

Implicit bias *occurs when assumptions are made based on one's own mental models and personal experiences that do not necessarily apply more generally.*

https://developers.google.com/machine-learning/crash-course/fairness/types-of-bias

Implicit Bias

Implicit bias *occurs when assumptions are made based on one's own mental models and personal experiences that do not necessarily apply more generally.*

EXAMPLE: An engineer training a gesture-recognition model uses a head shake as a feature to indicate a person is communicating the word "no." However, in some regions of the world, a head shake actually signifies “yes”

https://developers.google.com/machine-learning/crash-course/fairness/types-of-bias

A technology case study: COMPAS

COMPAS (Correctional Offender Management Profiling for Alternative Sanctions)

- Developed by Northpointe, Inc in 1990s
- a statistically-based algorithm designed to assess the risk that a given defendant will commit a crime after release
- Used in court by judges to make sentencing decision

COMPAS (Correctional Offender Management Profiling for Alternative Sanctions)

An ADM recidivism predicition software to forecast which criminals are most likely to reoffend.

See Julia Angwin et al., Machine Bias, ProPublica (May 23, 2016), https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing .

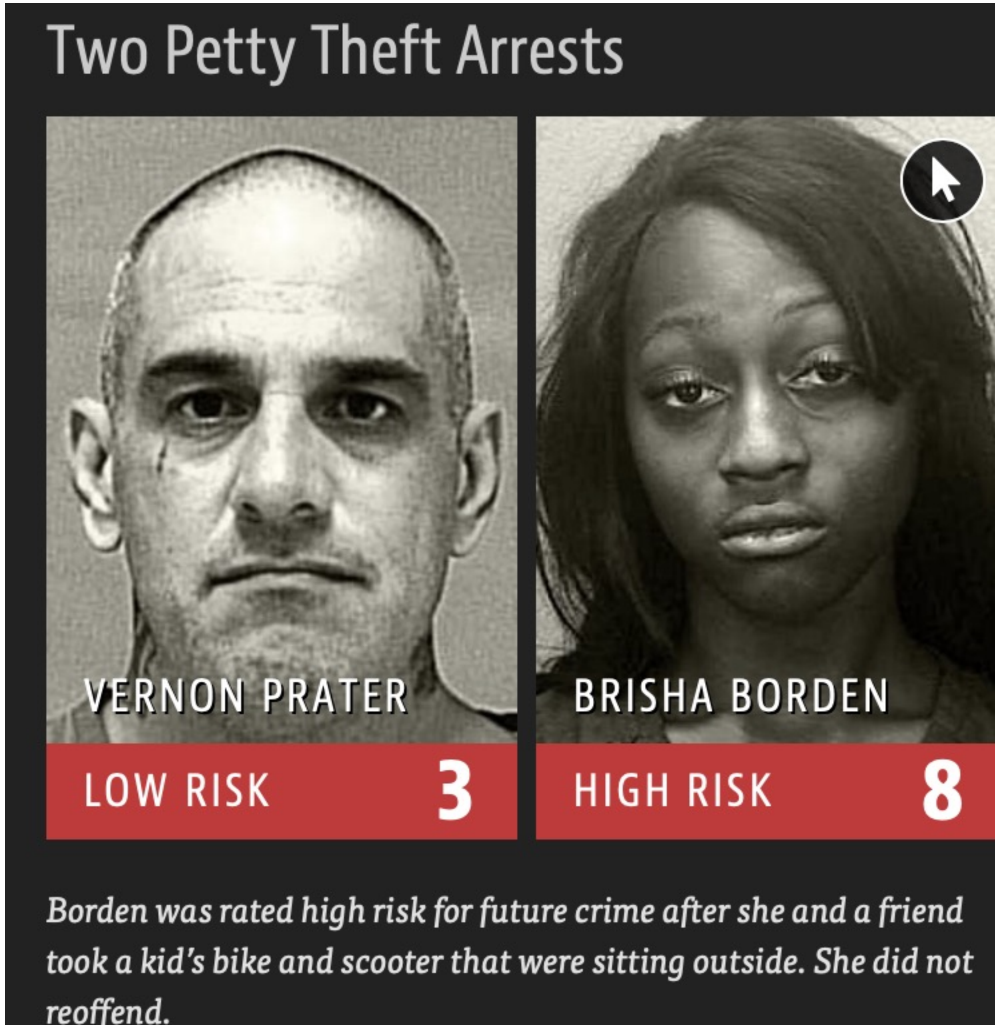


Image source from https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

COMPAS (Correctional Offender Management Profiling for Alternative Sanctions)

An ADM recidivism predicition software to forecast which criminals are most likely to reoffend.

See Julia Angwin et al., Machine Bias, ProPublica (May 23, 2016), https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

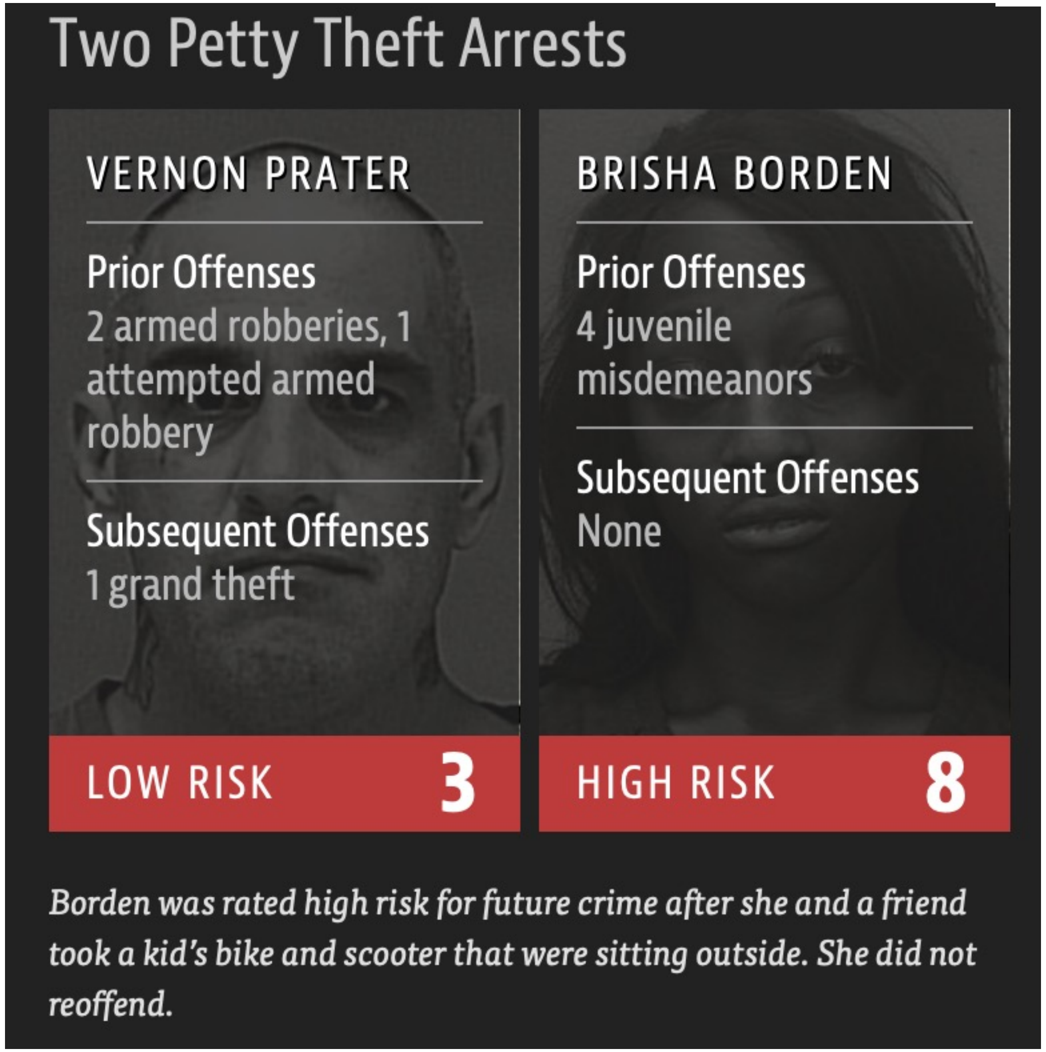


Image source from https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

COMPAS (Correctional Offender Management Profiling for Alternative Sanctions)

An ADM recidivism predicition software to forecast which criminals are most likely to reoffend.

See Julia Angwin et al., Machine Bias, ProPublica (May 23, 2016), https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Image source from https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

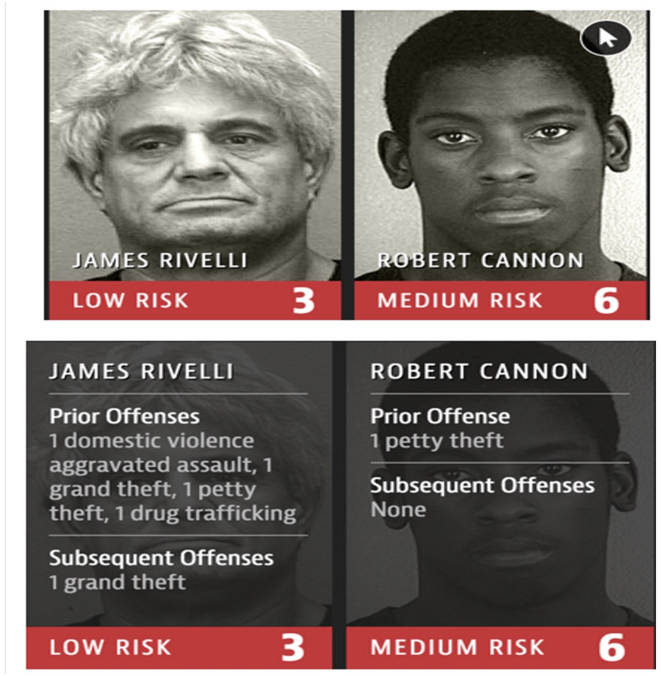
COMPAS (Correctional Offender Management Profiling for Alternative Sanctions)

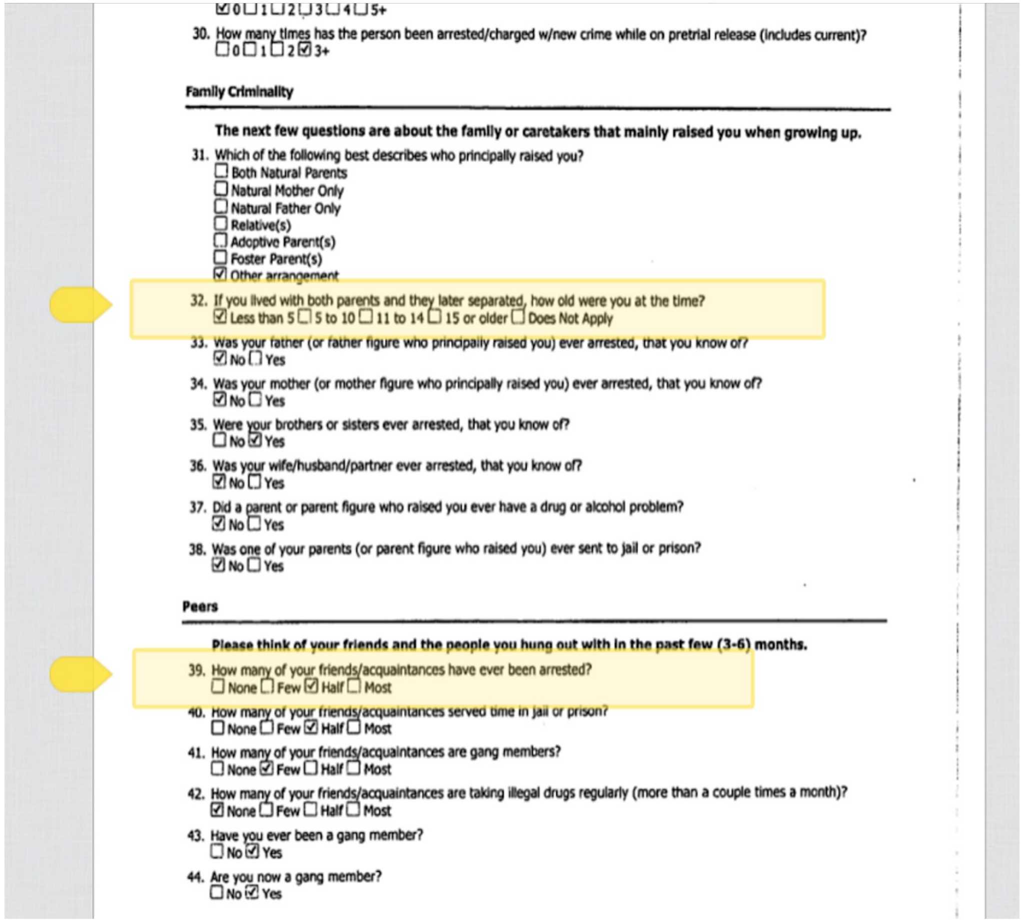
An ADM recidivism predicition software to forecast which criminals are most likely to reoffend.

See Julia Angwin et al., Machine Bias, ProPublica (May 23, 2016), https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Image source from https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

COMPAS (Correctional Offender Management Profiling for Alternative Sanctions)





☑ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5+

30. How many times has the person been arrested/charged w/new crime while on pretrial release (includes current)?
☐ 0 ☐ 1 ☐ 2 ☑ 3+

Family Criminality

The next few questions are about the family or caretakers that mainly raised you when growing up.

31. Which of the following best describes who principally raised you?
☐ Both Natural Parents
☐ Natural Mother Only
☐ Natural Father Only
☐ Relative(s)
☐ Adoptive Parent(s)
☐ Foster Parent(s)
☑ Other arrangement

32. If you lived with both parents and they later separated, how old were you at the time?
☑ Less than 5 ☐ 5 to 10 ☐ 11 to 14 ☐ 15 or older ☐ Does Not Apply

33. Was your father (or father figure who principally raised you) ever arrested, that you know of?
☑ No ☐ Yes

34. Was your mother (or mother figure who principally raised you) ever arrested, that you know of?
☑ No ☐ Yes

35. Were your brothers or sisters ever arrested, that you know of?
☐ No ☑ Yes

36. Was your wife/husband/partner ever arrested, that you know of?
☑ No ☐ Yes

37. Did a parent or parent figure who raised you ever have a drug or alcohol problem?
☑ No ☐ Yes

38. Was one of your parents (or parent figure who raised you) ever sent to jail or prison?
☑ No ☐ Yes

Peers

Please think of your friends and the people you hung out with in the past few (3-6) months.

39. How many of your friends/acquaintances have ever been arrested?
☐ None ☐ Few ☑ Half ☐ Most

40. How many of your friends/acquaintances served time in jail or prison?
☐ None ☐ Few ☑ Half ☐ Most

41. How many of your friends/acquaintances are gang members?
☐ None ☑ Few ☐ Half ☐ Most

42. How many of your friends/acquaintances are taking illegal drugs regularly (more than a couple times a month)?
☑ None ☐ Few ☐ Half ☐ Most

43. Have you ever been a gang member?
☐ No ☑ Yes

44. Are you now a gang member?
☐ No ☑ Yes

Image source from https://medium.com/thoughts-and-reflections/racial-bias-and-gender-bias-examples-in-ai-systems-7211e4c166a1

Bias vs Fairness

Bias focuses more on the representation

Fairness focuses more on the decision outcome

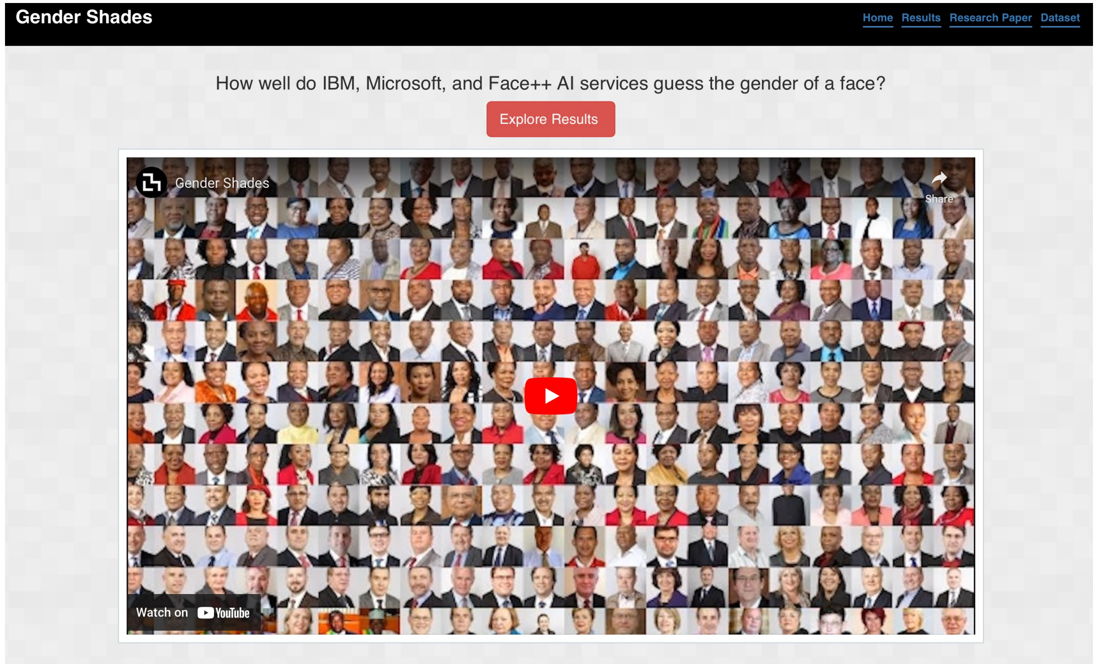
Fairness – two different *Worldviews*

What you see is what you get (WYSIWYG) worldview

We are all equal (WAE) worldview

Friedler, S.A., Scheidegger, C. and Venkatasubramanian, S., 2016. On the (im) possibility of fairness. *arXiv preprint arXiv:1609.07236*.

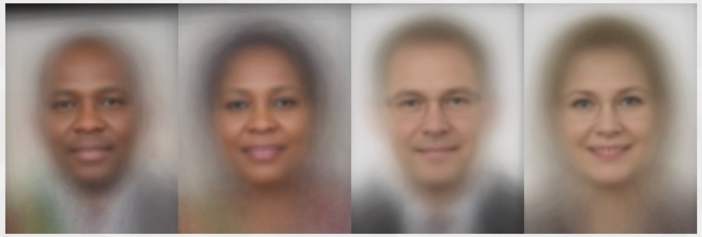
Gender Shades



When we analyze the results by intersectional subgroups - darker males, darker females, lighter males, lighter females - we see that all companies perform worst on darker females.

IBM and Microsoft perform best on lighter males. Face++ performs best on darker males.

Gender Classifier	Darker Male	Darker Female	Lighter Male	Lighter Female	Largest Gap
Microsoft	94.0%	79.2%	100%	98.3%	20.8%
FACE++	99.3%	65.5%	99.2%	94.0%	33.8%
IBM	88.0%	65.3%	99.7%	92.9%	34.4%



gendershades.org;

http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf

Bias

Bias in AI & Automated Decision Making (ADM) System

Types of Bias in Machine Learning (Mehrabi et al. 2019)

>23 types of bias !!

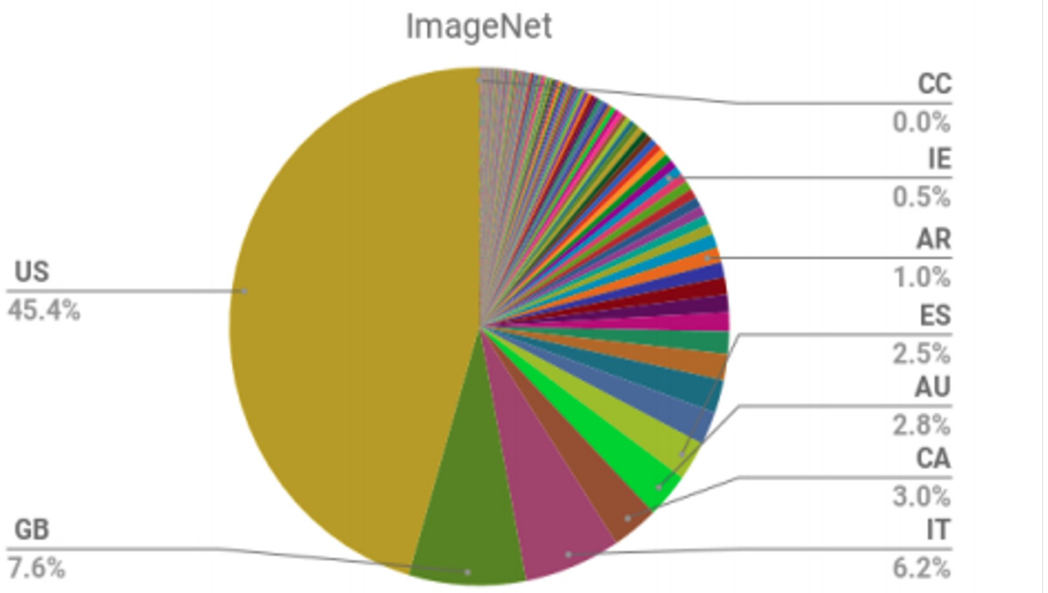
- Historical Bias
- Representation Bias
- Measurement Bias
- Evaluation Bias
- Aggregation Bias
- Population Bias
- Simpson's Paradox
- Longitudinal Data Fallacy
- Sampling Bias
- Behavioral Bias
- Content Production Bias
- Linking Bias
- Temporal Bias
- Popularity Bias
- Algorithmic Bias
- User Interaction Bias
- Social Bias
- Emergent Bias
- Self-Selection Bias
- Omitted Variable Bias
- Cause-Effect Bias
- Observer Bias
- Funding Bias

https://arxiv.org/pdf/1908.09635.pdf

UNSW

Common Types of Bias

- Representation bias
  - Representation bias happens from the way we define and sample from a population feature selection.
  - Comes from the way we define and sample from a population, e.g. ImageNet



Source from https://arxiv.org/pdf/1711.08536.pdf

Representation bias

- **Representation bias** occurs when building datasets for training a model, if those datasets poorly represent the people that the model will serve.

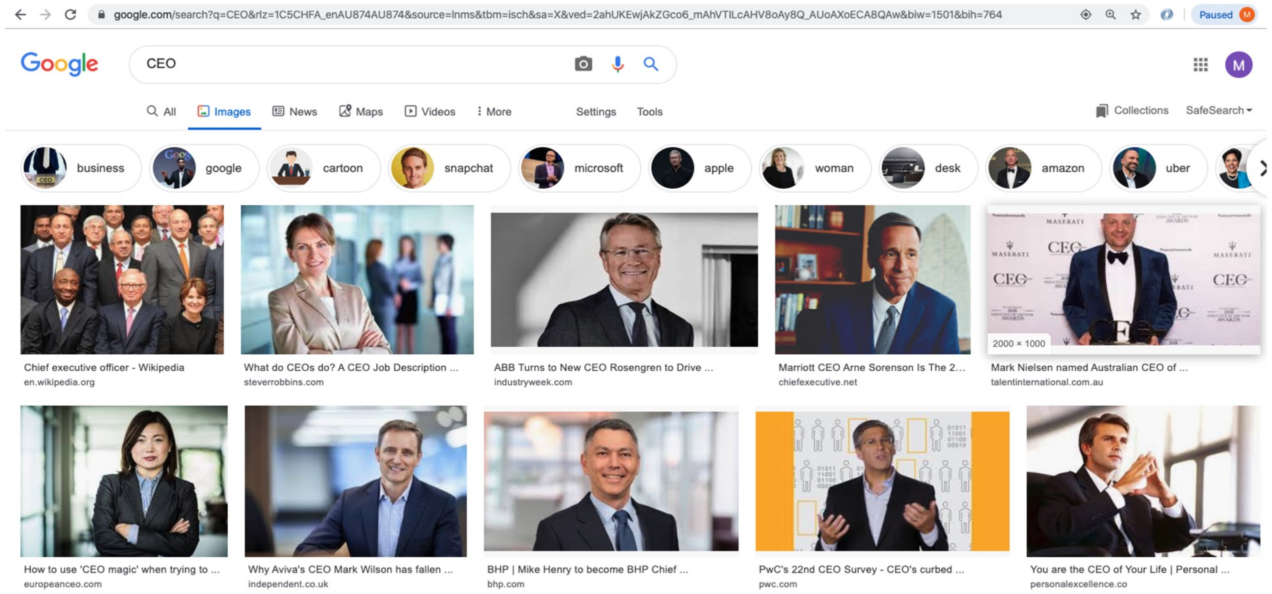
- Data collected through smartphone apps will under-represent groups that are less likely to own smartphones.

- Eg. In US, individuals > 65 y.o. will be under-represented.
- What if the data used to design US transportation system https://www.bloomberg.com/news/articles/2017-08-04/why-aging-americans-need-better-transit

UNSW

Common Types of Bias

- Historical bias
  - Historical bias is the already existing bias and socio-technical issues in the world and can seep into from the data generation process even given a perfect sampling and feature selection.



Source from https://arxiv.org/pdf/1908.09635.pdf

Historical bias

- **Historical bias** occurs when the state of the world in which the data was generated is flawed.

- As of 2020, only 7.4% of Fortune 500 CEOs are women. Research has shown that companies with female CEOs or CFOs are generally more profitable than companies with men in the same position, suggesting that women are held to higher hiring standards than men.

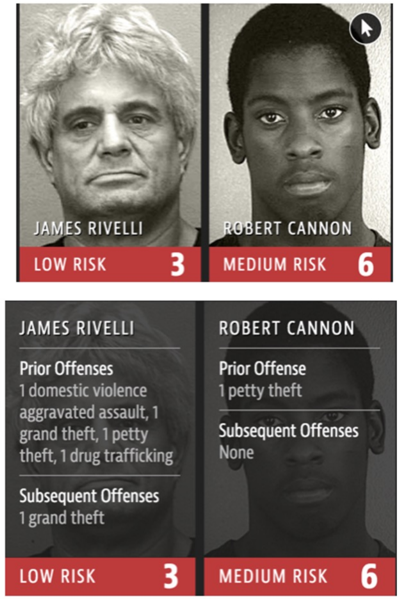
- If we are using AI to make the hiring process more equitable, how to fix this?

from https://arxiv.org/pdf/1901.10002.pdf

UNSW

Common Types of Bias

- Measurement bias
  - Measurement bias happens from the way we choose, utilize, and measure a particular feature.
  - Choosing and measuring the particular features of interest, e.g. COMPAS



Source from https://arxiv.org/pdf/1908.09635.pdf

UNSW

Measurement bias

Measurement bias occurs when:

- the accuracy of the data varies across groups.
- This can happen when working with proxy variables (variables that take the place of a variable that cannot be directly measured), if the quality of the proxy varies in different groups.

Measurement bias

Your local hospital uses a model to identify high-risk patients before they develop serious conditions, based on information like past diagnoses, medications, and demographic data. The model uses this information to predict health care costs, the idea being that patients with higher costs likely correspond to high-risk patients. Despite the fact that the model specifically excludes race, it seems to demonstrate racial discrimination: the algorithm is less likely to select eligible Black patients. How can this be the case?

Common Types of Bias

- Aggregation bias

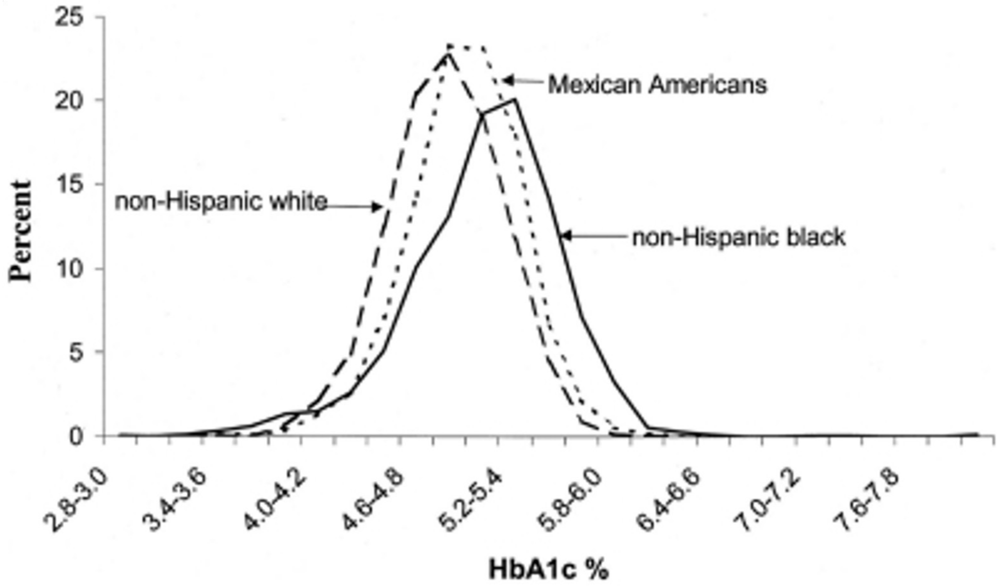
Aggregation bias arises when a one-size-fit-all model is used for groups with different conditional distributions, $p(Y|X)$.

False conclusions are drawn for a subgroup based on observing other different subgroups or generally when false assumptions about a population affect the model's outcome and definition.

Aggregation bias

Hispanics have higher rates of diabetes and diabetes-related complications than non-Hispanic whites.
If building AI to diagnose or monitor diabetes, how to mitigate this?

Race/Ethnicity	Type 1 Diabetes Prevalence (per 1,000)		Type 2 Diabetes Prevalence (per 1,000)		Reference
	0-9 years	10-19 years	0-9 years*	10-19 years	
Non-Hispanic White	1.03	2.89	0.0046	0.18	[10]
Non-Hispanic Black	0.57	2.04	0.0005	1.06	[11]
Hispanic American	0.44	1.59	0.0003	0.46	[12]
Asian and Pacific Islanders	0.26	0.77	0.014	0.52	[13]
Native American	0.08	0.28	0.021	1.45	[14]

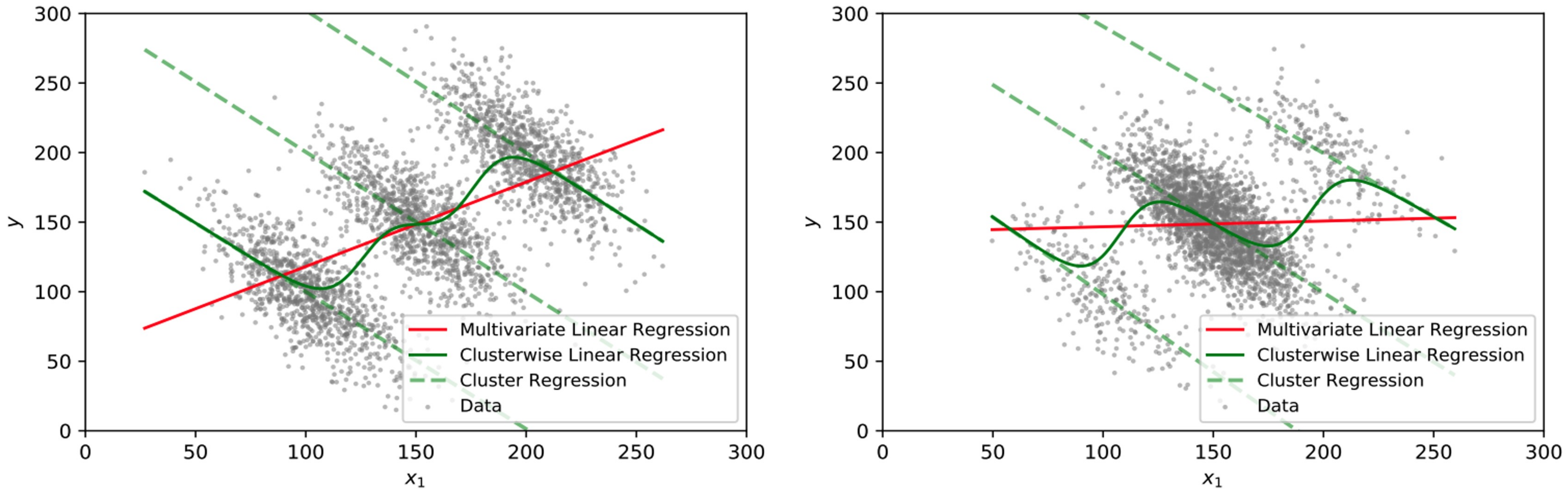


Source from https://arxiv.org/pdf/1908.09635.pdf and https://www.ncbi.nlm.nih.gov/pubmed/22238408

Bias in Data

Simpsons Paradox

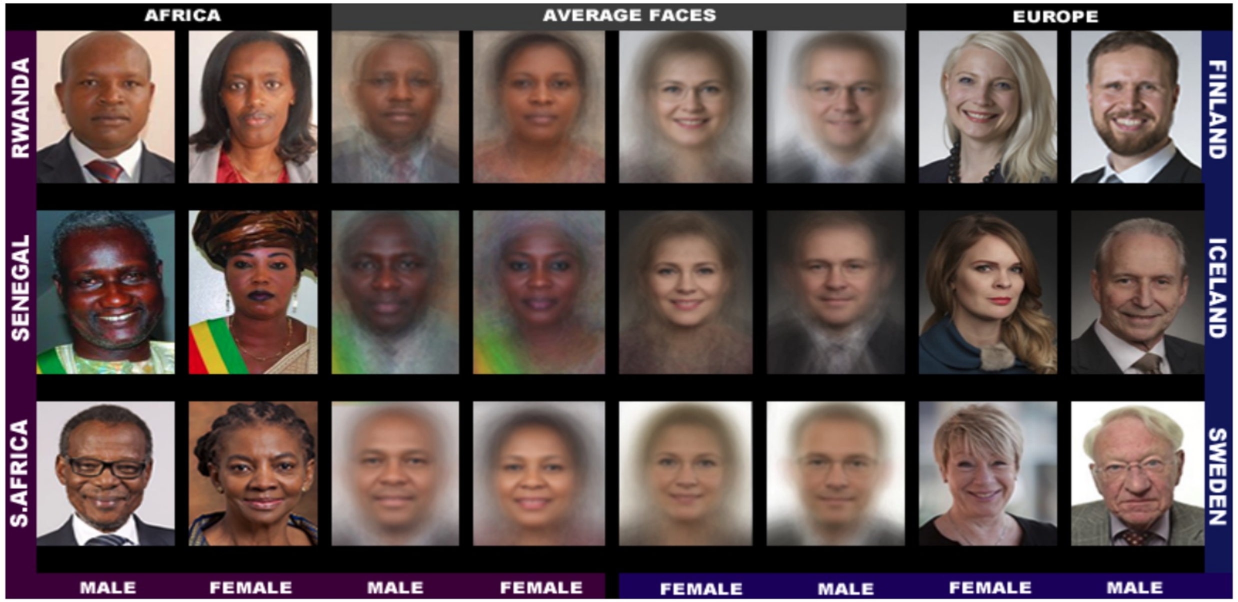
A hypothetical nutrition study which measured how the outcome, body mass index (BMI), changes as a function of daily pasta calorie intake (Figure 1).

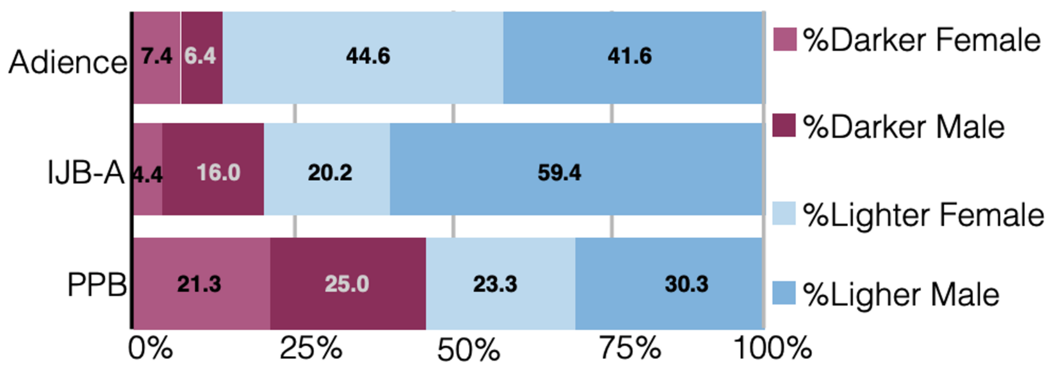


Source from https://arxiv.org/pdf/1908.09635.pdf

Common Types of Bias

- Evaluation bias
  - Evaluation bias happens during model evaluation. This includes the use of inappropriate and disproportionate benchmarks
  - occurs during model iteration and evaluation, e.g., Gender shades





Source from http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf

UNSW

Deployment bias

- **Deployment bias** occurs when the problem the model is intended to solve is different from the way it is actually used. If the end users don’t use the model in the way it is intended, there is no guarantee that the model will perform well.

Other Common Types of Bias

- **Population Bias.** Population bias arises when statistics, demographics, representatives, and user characteristics are different in the user population represented in the dataset or platform from the original target population.
- **Sampling Bias.** Sampling bias arises due to the non-random sampling of subgroups. As a consequence of sampling bias, the trends estimated for one population may not generalize to data collected from a new population.
- **Temporal Bias.** Temporal bias arises from differences in populations and behaviors over time.
- **Social Bias**. Social bias happens when other people's actions or content coming from them affects our judgment.

Reporting Bias

Reporting bias *occurs when the frequency of events, properties, and/or outcomes captured in a data set does not accurately reflect their real-world frequency.*

This bias can arise because people tend to focus on documenting circumstances that are unusual or especially memorable.

https://developers.google.com/machine-learning/crash-course/fairness/types-of-bias

Reporting Bias

Reporting bias *occurs when the frequency of events, properties, and/or outcomes captured in a data set does not accurately reflect their real-world frequency.*

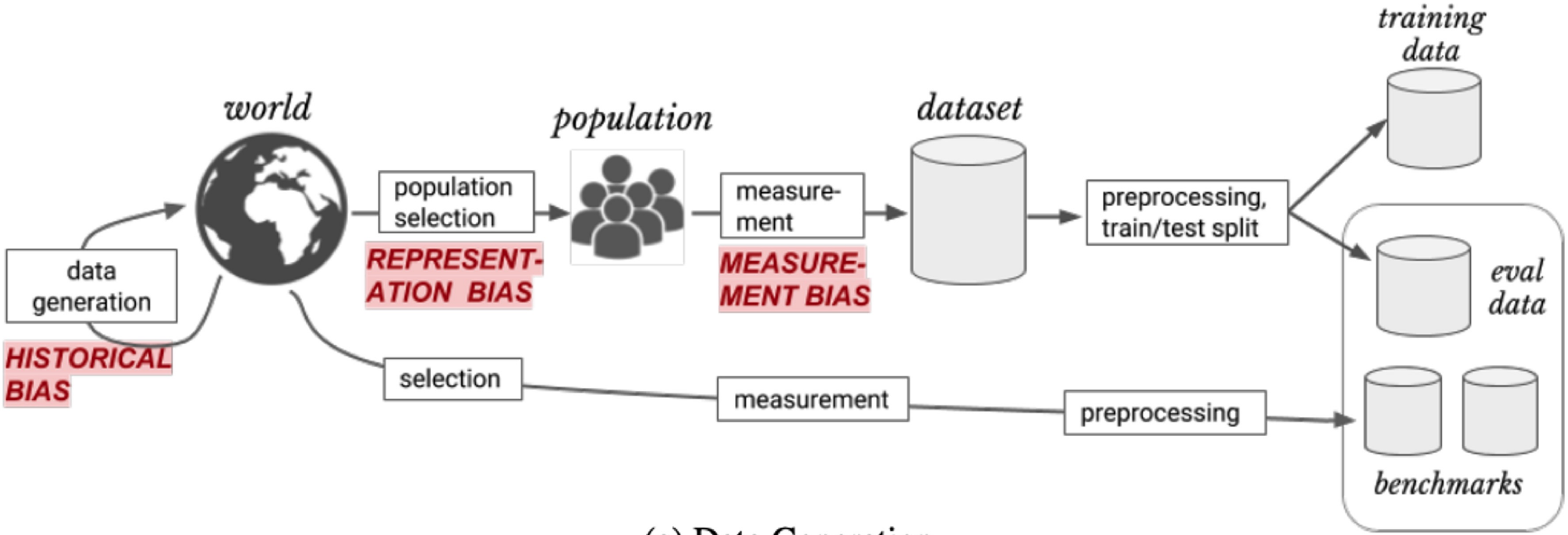
This bias can arise because people tend to focus on documenting circumstances that are unusual or especially memorable.

EXAMPLE: A sentiment-analysis model is trained to predict whether book reviews are positive or negative based on a corpus of user submissions to a popular website. The majority of reviews in the training data set reflect extreme opinions (reviewers who either loved or hated a book), because people were less likely to submit a review of a book if they did not respond to it strongly. As a result, the model is less able to correctly predict sentiment of reviews that use more subtle language to describe a book.

https://developers.google.com/machine-learning/crash-course/fairness/types-of-bias

UNSW SYDNEY

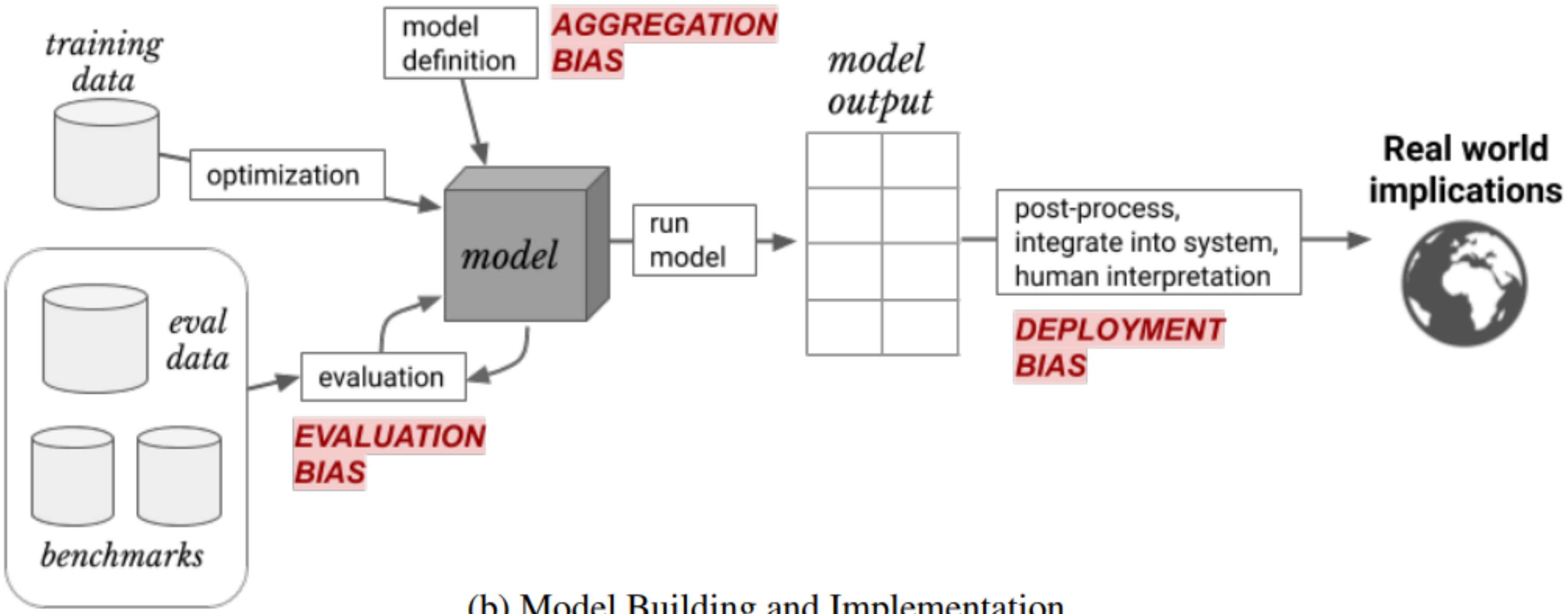
Where do they exist?



(a) Data Generation

Source from https://arxiv.org/pdf/1901.10002.pdf

Where do they exist?



(b) Model Building and Implementation

Source from https://arxiv.org/pdf/1901.10002.pdf

UNSW

Other classification and types of bias

- Population bias
- Simpson's Paradox
- Longitudinal data fallacy
- Sampling bias
- ……

[Optional Reading]

Olteanu A. et al. 2019, Social Data: Biases, Methodological Pitfalls, and Ethical Boundaries
https://www.frontiersin.org/articles/10.3389/fdata.2019.00013/full

UNSW

Algorithmic Fairness

Fairness – two different *Worldviews*

What you see is what you get (WYSIWYG) worldview

We are all equal (WAE) worldview

Friedler, S.A., Scheidegger, C. and Venkatasubramanian, S., 2016. On the (im) possibility of fairness. *arXiv preprint arXiv:1609.07236*.

Algorithmic Fairness

- Types of Discrimination
  - Direct vs. Indirect discrimination
  - Systemic discrimination
  - Explainable vs. unexplainable discrimination
  - ……

Source from https://arxiv.org/pdf/1901.10002.pdf

[Discussion]

Amazon's same-day delivery

A couple years ago Amazon rolled out same-day delivery across a select group of American cities. However, this service was only extended to neighborhoods with a high number of current Amazon users. As a result, predominantly non-white neighborhoods were largely excluded from the service.

Algorithmic Fairness

- Definitions of Fairness
  - Equalised odds
  - Equal opportunity
  - Demographic parity
  - Fairness through awareness
  - Fairness through unawareness
  - Treatment equality
  - Test fairness
  - ……

Source from https://arxiv.org/pdf/2006.16745.pdf

Algorithmic Fairness

- Categories of Fairness
    - Individual Fairness
    - Group Fairness
    - Subgroup Fairness

Source from https://arxiv.org/pdf/1710.03184.pdf

Algorithmic Fairness

- Categories of Fairness
  - Individual Fairness
  - Group Fairness
  - Subgroup Fairness

Name	Group	Individual
Demographic parity	✓	
Conditional statistical parity	✓	
Equalized odds	✓	
Equal opportunity	✓	
Fairness through unawareness		✓
Fairness through awareness		✓

Source from https://arxiv.org/pdf/1710.03184.pdf

Fairness criteria: Demographic parity /statistical parity

- **Demographic parity** says the model is fair if the composition of people who are selected by the model matches the group membership percentages of the applicants.

Fairness criteria: Demographic parity /statistical parity

- **Demographic parity** says the model is fair if the composition of people who are selected by the model matches the group membership percentages of the applicants.

- *A nonprofit is organizing an international conference, and 20,000 people have signed up to attend. The organizers write a ML model to select 100 attendees who could potentially give interesting talks at the conference. Since 50% of the attendees will be women (10,000 out of 20,000), they design the model so that 50% of the selected speaker candidates are women.*

Fairness criteria: Equal opportunity

- **Equal opportunity** fairness ensures that the proportion of people who should be selected by the model ("positives") that are correctly selected by the model is the same for each group. We refer to this proportion as the **true positive rate** (TPR) or **sensitivity** of the model.

Fairness criteria: Equal accuracy

- Alternatively, we could check that the model has **equal accuracy** for each group. That is, the percentage of correct classifications (people who should be denied and are denied, and people who should be approved who are approved) should be the same for each group. If the model is 98% accurate for individuals in one group, it should be 98% accurate for other groups.

Fairness criteria: Group unaware / "Fairness through unawareness"

- **Group unaware** fairness removes all group membership information from the dataset. For instance, we can remove gender data to try to make the model fair to different gender groups. Similarly, we can remove information about race or age.

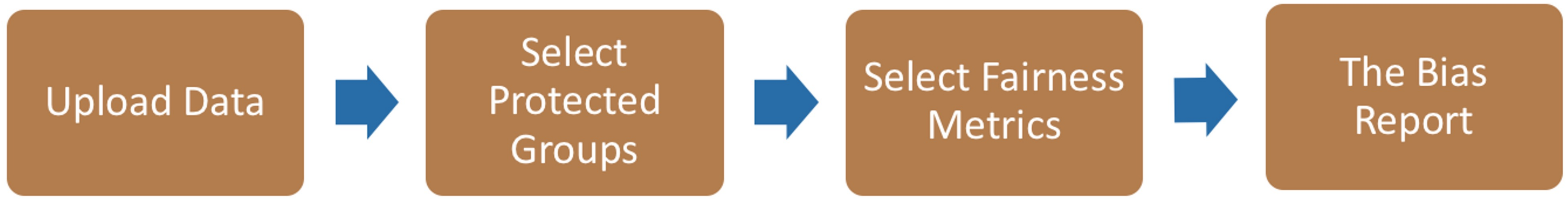
Fairness criteria: Group unaware / "Fairness through unawareness"

- One difficulty of applying this approach in practice is that one has to be careful to identify and remove proxies for the group membership data.
- Eg In cities that are racially segregated, zip code is a strong proxy for race. That is, when the race data is removed, the zip code data should also be removed, or else the ML application may still be able to infer an individual's race from the data. Additionally, group unaware fairness is unlikely to be a good solution for historical bias.

What can we do about it?

Assessing Fairness

Aequitas: Bias and Fairness Audit Toolkit



There is an example report on COMPASS http://aequitas.dssg.io/example.html#audit-results-details-by-fairness-measures

Source from http://aequitas.dssg.io

UNSW

Assessing Fairness

The AI Fairness 360 (AIF360 - IBM)

AI Fairness 360 - Demo

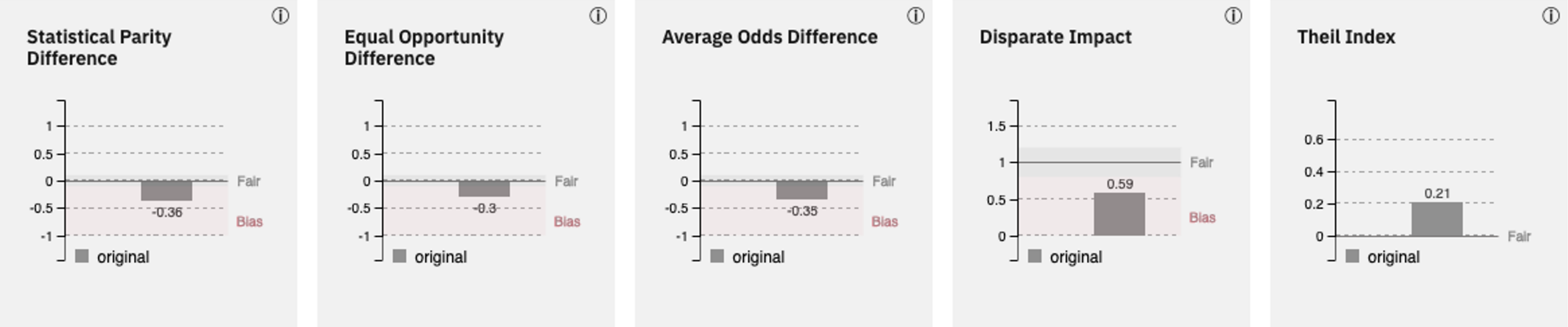
Data Check Mitigate Compare

2. Check bias metrics

Dataset: Compas (ProPublica recidivism)
Mitigation: none

Protected Attribute: Sex

Privileged Group: ***Female***, Unprivileged Group: ***Male***
Accuracy with no mitigation applied is 66%
With default thresholds, bias against unprivileged group detected in 4 out of 5 metrics



Source from https://www.ibm.com/blogs/research/2018/09/ai-fairness-360/

UNSW

Mitigation

Optimized Pre-processing

Use to mitigate bias in training data. Modifies training data features and labels.

Reweighing

Use to mitgate bias in training data. Modifies the weights of different training examples.

Adversarial Debiasing

Use to mitigate bias in classifiers. Uses adversarial techniques to maximize accuracy and reduce evidence of protected attributes in predictions.

Reject Option Classification

Use to mitigate bias in predictions. Changes predictions from a classifier to make them fairer.

Disparate Impact Remover

Use to mitigate bias in training data. Edits feature values to improve group fairness.

Learning Fair Representations

Use to mitigate bias in training data. Learns fair representations by obfuscating information about protected attributes.

Prejudice Remover

Use to mitigate bias in classifiers. Adds a discrimination-aware regularization term to the learning objective.

Calibrated Equalized Odds Post-processing

Use to mitigate bias in predictions. Optimizes over calibrated classifier score outputs that lead to fair output labels.

Equalized Odds Post-processing

Use to mitigate bias in predictions. Modifies the predicted labels using an optimization scheme to make predictions fairer.

Meta Fair Classifier

Use to mitigate bias in classifier. Meta algorithm that takes the fairness metric as part of the input and returns a classifier optimized for that metric.

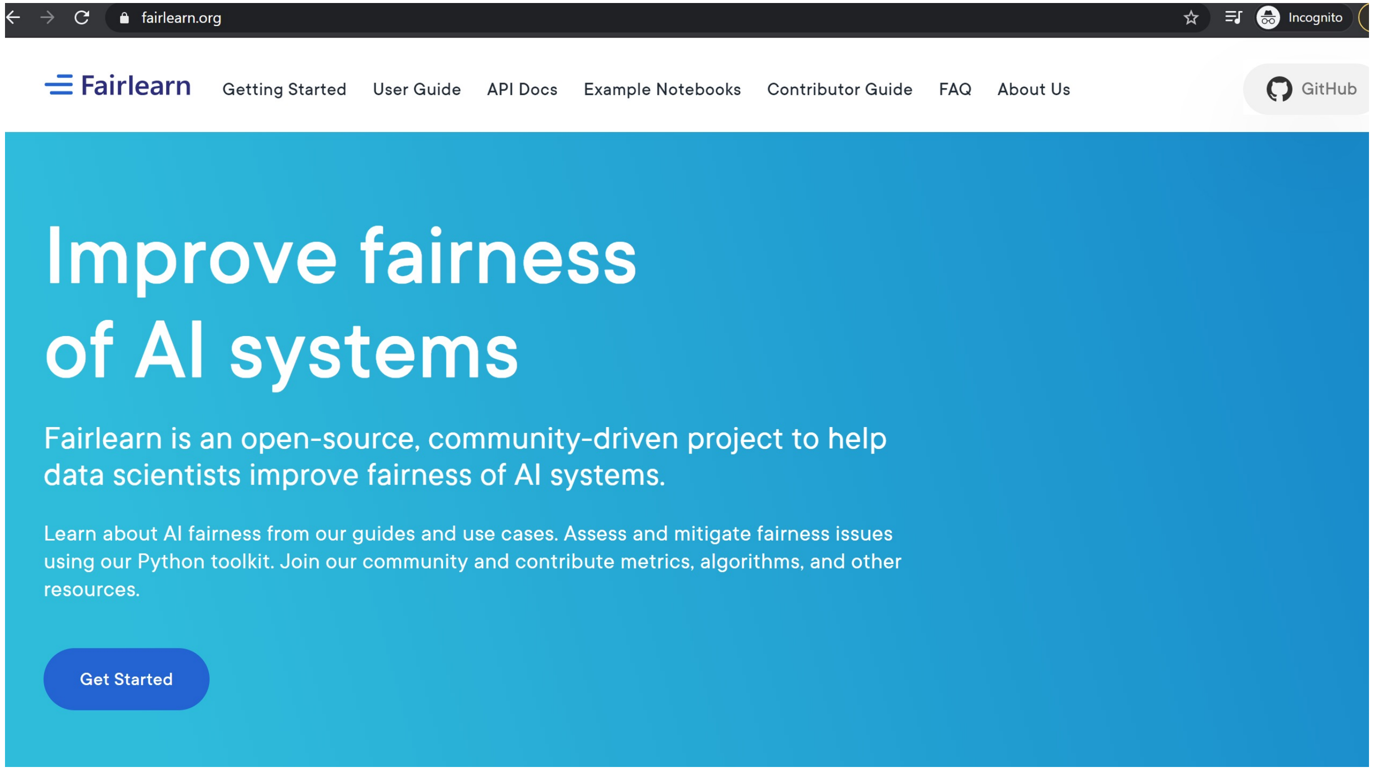
Source from https://aif360.mybluemix.net/

Mitigation Strategies

- Pre-processing
- In-processing
- Post-processing

Small E, Shao W, Zhang Z, Liu P, Chan J, Sokol K, Salim F. How Robust is your Fair Model? Exploring the Robustness of Diverse Fairness Strategies. arXiv preprint arXiv:2207.04581. 2022 Jul 11.

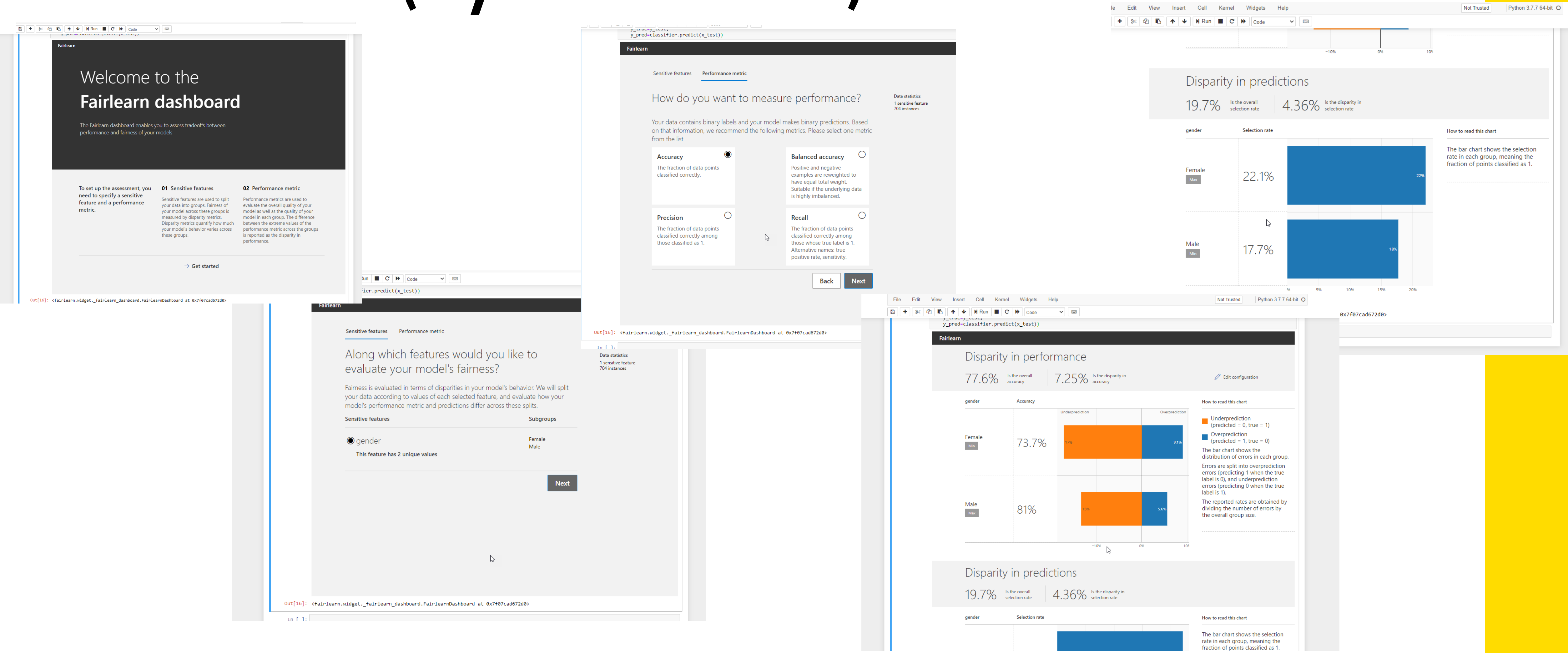
Fairlearn (by Microsoft)



https://fairlearn.org/

https://www.microsoft.com/en-us/research/publication/fairlearn-a-toolkit-for-assessing-and-improving-fairness-in-ai/

Fairlearn (by Microsoft)



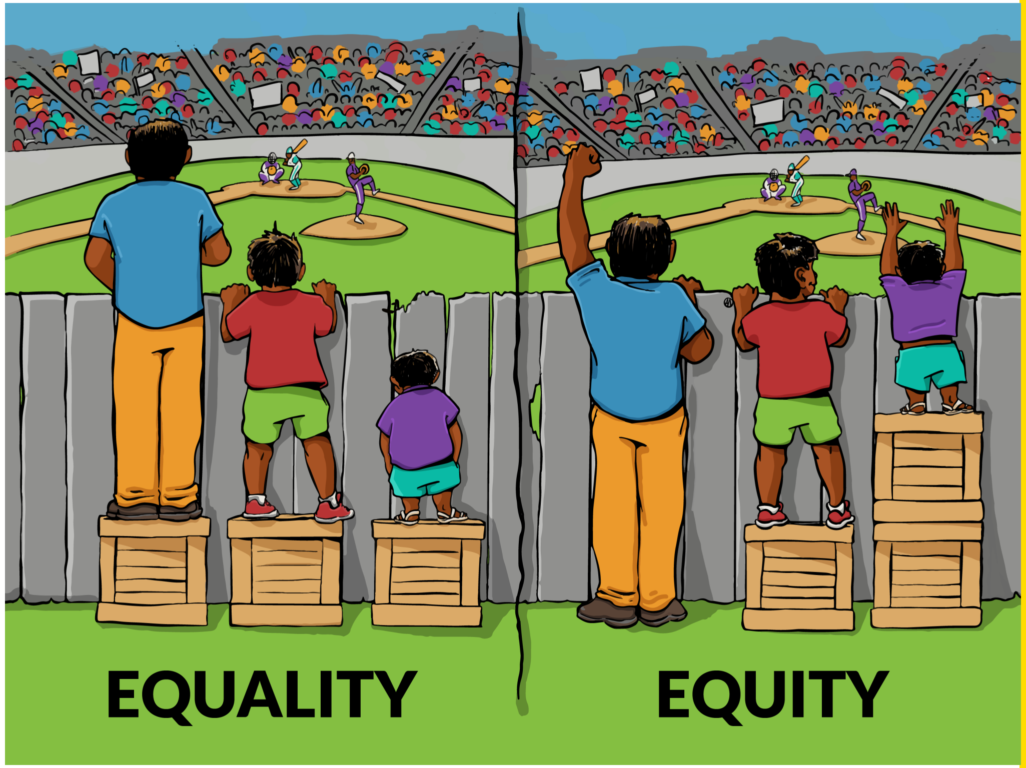
https://fairlearn.org/

Credit: https://github.com/wmeints/fairlearn-demo

Challenges?

Challenges

- Synthesizing a definition of fairness.
- From Equality to Equity
- Searching for Unfairness



https://interactioninstitute.org/illustrating-equality-vs-equity/

Additional Materials

- Friedler, S.A., Scheidegger, C. and Venkatasubramanian, S., 2016. On the (im) possibility of fairness. arXiv preprint arXiv:1609.07236.

- Must watch NIPS 2017 Guest Lecture by Kate Crawford

[The Trouble with Bias - NIPS2017](https://youtu.be/fMym_BKWQzk)
https://youtu.be/fMym_BKWQzk

UNSW