

AI (Machine Learning): Bias & Garbage In, Bias & Garbage Out

John Rochford



INDEX Program Director & Faculty Member

Eunice Kennedy Shriver Center



Invited Expert

World Wide Web Consortium



About Me

I am:

- legally blind.
- I trade privacy for utility. The 14 Alexas in my home enable me to control devices with interfaces I can't see.
- (Alexa records to the cloud everything I say.)
- an AI Researcher focused on ML text simplification; ML fairness & empowerment.
- a 35-year developer of technology for the disabled.



INDEX Staff: Representing the Diversity of our Community



Creating Easy-To-Understand COVID-19 Info Worldwide



INDEX ML engineers and disabled staff are developing our ML model and using crowdsourcing to simplify COVID-19 text from the websites of every country's governments.

Simple COVID-19 text will help huge populations of people with cognitive disabilities and/or low literacy; non-native language speakers; and seniors make the world safe & healthy.



Presentation Targets

- ML Social Bias
- Why we should care.
- What we can do about it.
- How we can do it.
- I hope to persuade you that, for your ML efforts to be successful, their development and data must include disabled people.



But First, A Shout Out To Black Lives Matter

“University of Toronto and MIT researchers found that every facial recognition system they tested performed better on lighter-skinned faces. That includes a 1-in-3 failure rate with identifying darker-skinned females.”

Joy Buolamwini, et al., Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification



How I'm
fighting bias in
algorithms,
Joy Buolamwini



Why Care? Profound Responsibility

“There’s nothing artificial about AI. It’s inspired by people, it’s created by people, and – most importantly – it impacts people. It is a powerful tool we are only just beginning to understand, and that is a profound responsibility.”

Fei-Fei Li, Bias in, bias out: the Stanford scientist out to make AI less white and male



Why Care About ML Fairness and Disability?

- AI/ML Fairness is our most-significant challenge.
- If it does not work for all, nobody will trust it.
- The disabled are part of every segment of society.
- Thus, if we can solve the problem for the disabled,

we can solve the problem for everyone.



Why Care? What About You?

- “Bias in ML has been almost ubiquitous when the application is involved in people and it has already hurt the benefit of people in minority groups or historically disadvantageous groups.
- Not only people in minority groups but everyone should care about the bias in AI.
- If no one cares, it is highly likely that the next person who suffers from biased treatment is one of us.”



- Ziyuan Zhong, [A Tutorial on Fairness in Machine Learning](#)



Issue: Being Smart Can Be A Deficit in ML



- “People who perform better on a test of pattern detection—a measure of cognitive ability—are also quicker to form and apply stereotypes.”
- “In other words, being smart might put you at a greater risk of prejudice, but you can still fight against those instincts by challenging your thinking and getting to know people who aren’t like you.”
- Olga Khazan, Smart People Are More Likely to Stereotype



How can you get to know people who aren't like you?
Include The Disabled in ML Development

“To ensure AI-based systems are treating people with disabilities fairly, it is essential to include them in the development process.”

Shari Trewin, [How to tackle AI bias for people with disabilities](#) 

MIT ML Failure to Include the Disabled

- An MIT project claimed to translate American Sign Language with ML and sign language gloves.
- It failed because sign language is much more than communicating with hands. It's body language and facial expressions.
- MIT would have known that was essential had it involved the deaf community.

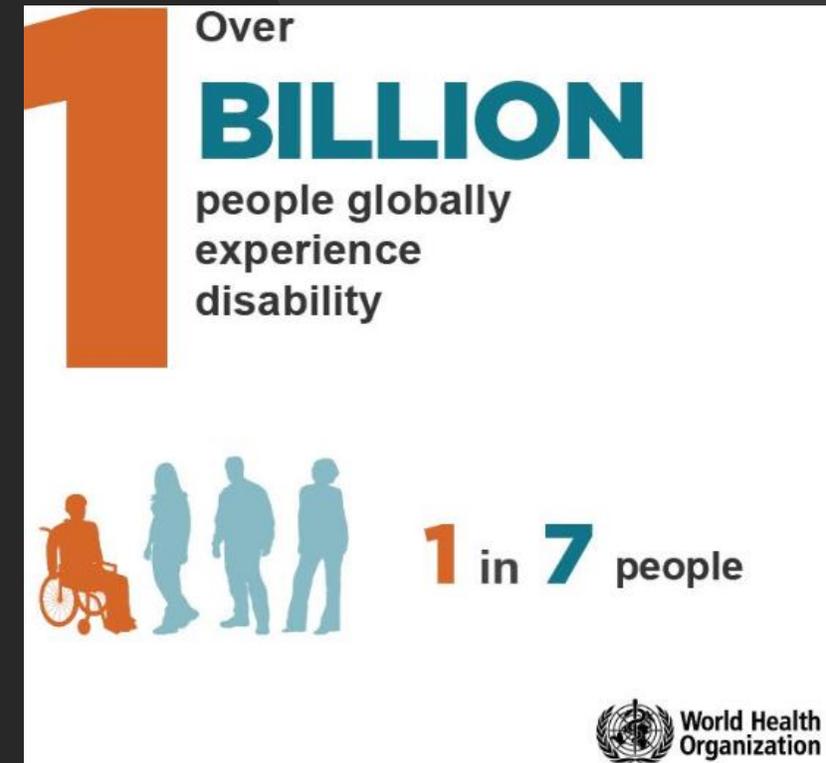


[Why Sign-Language Gloves Don't Help Deaf People](#)



Issue: Lack of Data from The Disabled Due to Privacy Concerns

- Many of us do not disclose our disabilities.
- When we do, we are denied employment, housing, more.
- Thus, despite being 15% of the population, we are **significantly** underrepresented in training data.



What We Can Do About A Lack of Data from The Disabled?

Accurate Analysis

“We are helping machine learning engineers figure out what questions to ask of their data to diagnose why their systems may be making unfair predictions.”

David Sontag, MIT researchers show how to detect and address AI bias without loss in accuracy



What Can We Do with Data We Have From the Disabled?

Fairness Through Unawareness

“No information about protected attributes (e.g. gender, age, disability) is gathered and used in the decision-making.”

Fairness Through Awareness

“Membership in a protected group is explicitly known, and fairness can be formally defined, tested and enforced algorithmically.”

Shari Trewin, [AI Fairness for People with Disabilities: Point of View](#)



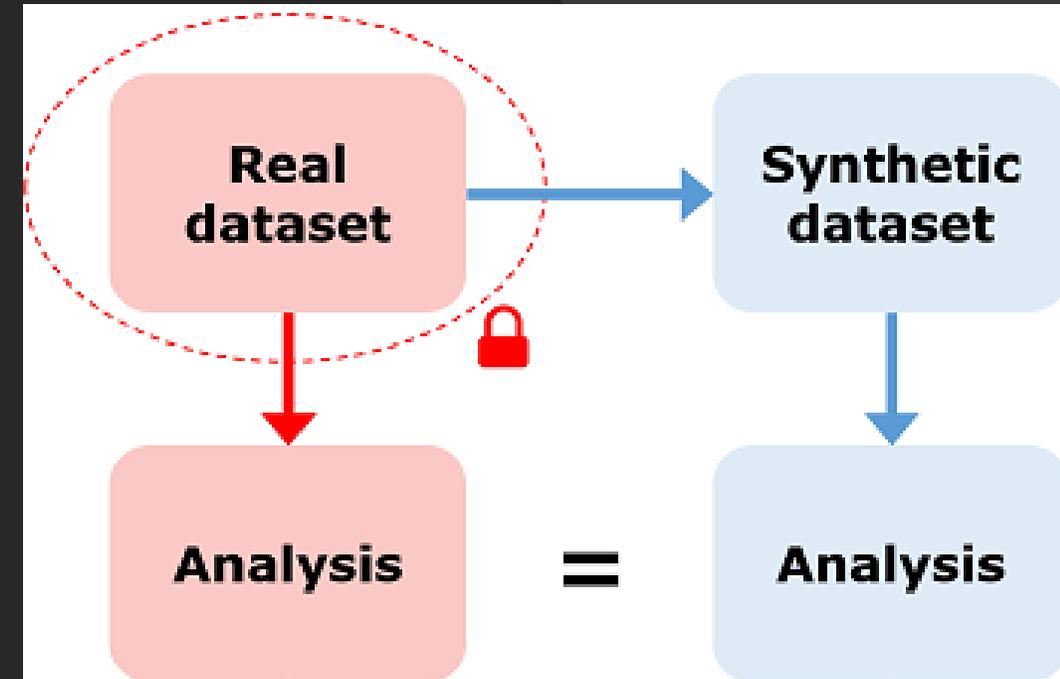
What Can We Do to Mitigate Our Lack of Data from the Disabled?

Use Synthetic Data

“Given a dataset involving a protected attribute with a privileged and unprivileged group, we create an “ideal world” dataset:

for every data sample, we create a new sample having the same features (except the protected attribute(s)) and label as the original sample but with the opposite protected attribute value.”

Shubham Sharma et al., [Data Augmentation for Discrimination Prevention and Bias Disambiguation](#)



What Tools Will Help Us Implement ML Fairness?

Commercial Tools

Many companies, such as Microsoft, Amazon, and IBM, offer bias-mitigation services in their ML platforms (Azure, SageMaker, IBM Cloud).

Opensource Tools

There are an increasing number of opensource tools on GitHub. The following are the best of the current ones.



plots Compute distance Show nearest different classification: L1 L2 ⓘ

PERFORMANCE + FAIRNESS DATAPPOINT EDITOR FEATURES

Binning | X-Axis age Co. 10 Binning | Y-Axis marital-stat

17-23 24-29 30-35

Select a datapoint to begin exploring features and values. →

Clicking on a datapoint in the visualization will load all the features and values associated with that example. Here are some of the things you can do:

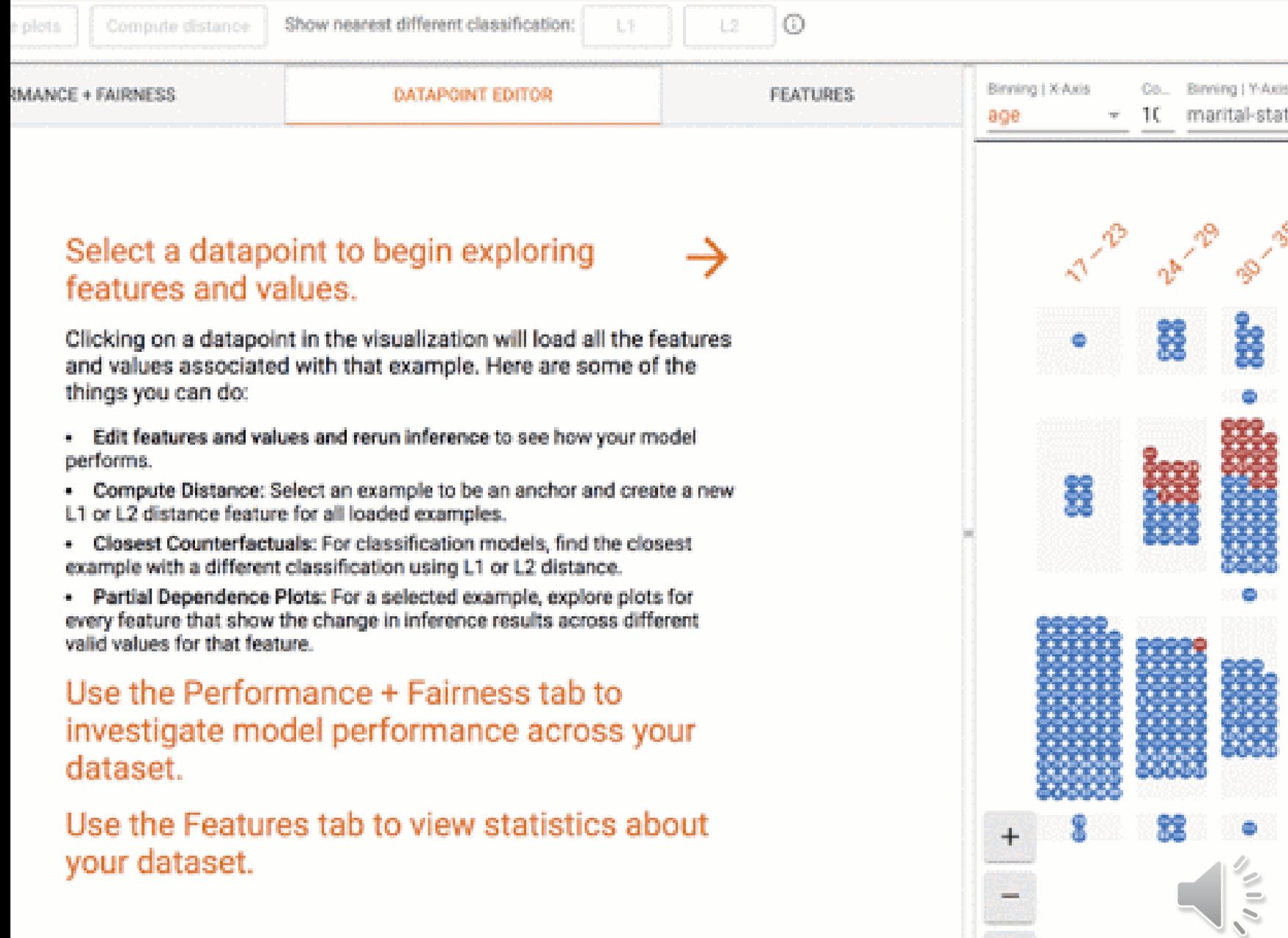
- Edit features and values and rerun inference to see how your model performs.
- Compute Distance: Select an example to be an anchor and create a new L1 or L2 distance feature for all loaded examples.
- Closest Counterfactuals: For classification models, find the closest example with a different classification using L1 or L2 distance.
- Partial Dependence Plots: For a selected example, explore plots for every feature that show the change in inference results across different valid values for that feature.

Use the Performance + Fairness tab to investigate model performance across your dataset.

Use the Features tab to view statistics about your dataset.

+ 8 8 8

-



Google What-If Tool

Visually probe the behavior of trained machine learning models, with minimal coding.

- [GitHub](#)



Fairlearn

A Python package to assess and improve fairness of machine learning models.

• [GitHub](#)



Skater

- Python library designed to demystify the learned structures of a black box model both
 - globally (inference on the basis of a complete data set); and
 - locally (inference about an individual prediction).
- [GitHub](#)





Pymetrics

Detects demographic differences in the output of machine learning models or other assessments.

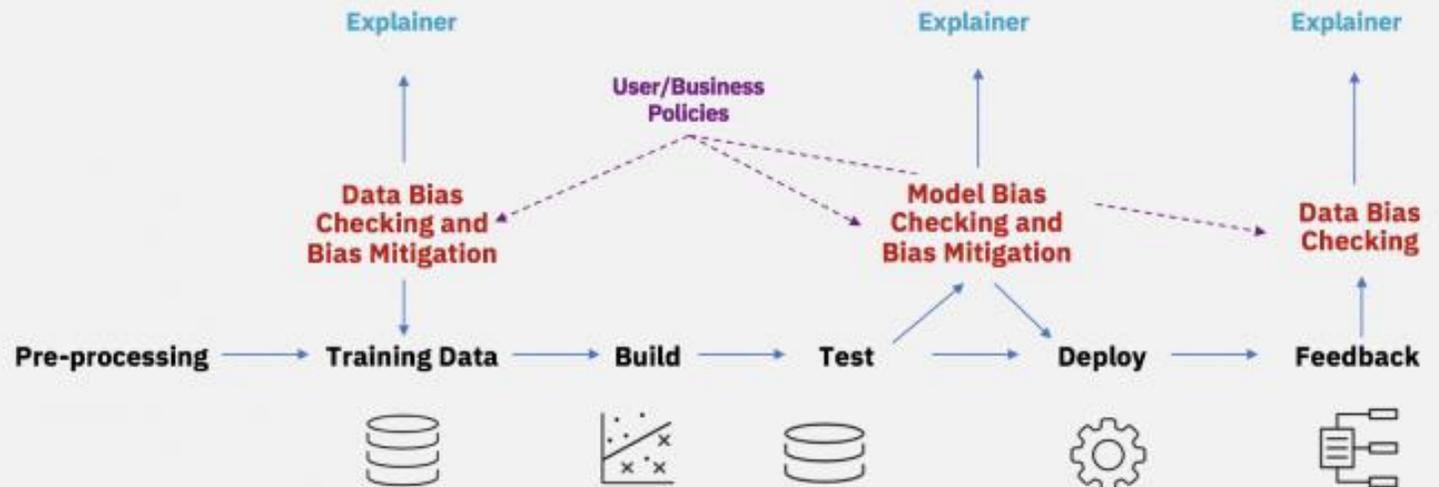
- [GitHub](#)



IBM AI Fairness 360 Opensource Toolkit

A comprehensive set of fairness metrics for datasets and machine learning models, explanations for these metrics, and algorithms to mitigate bias in datasets and models.

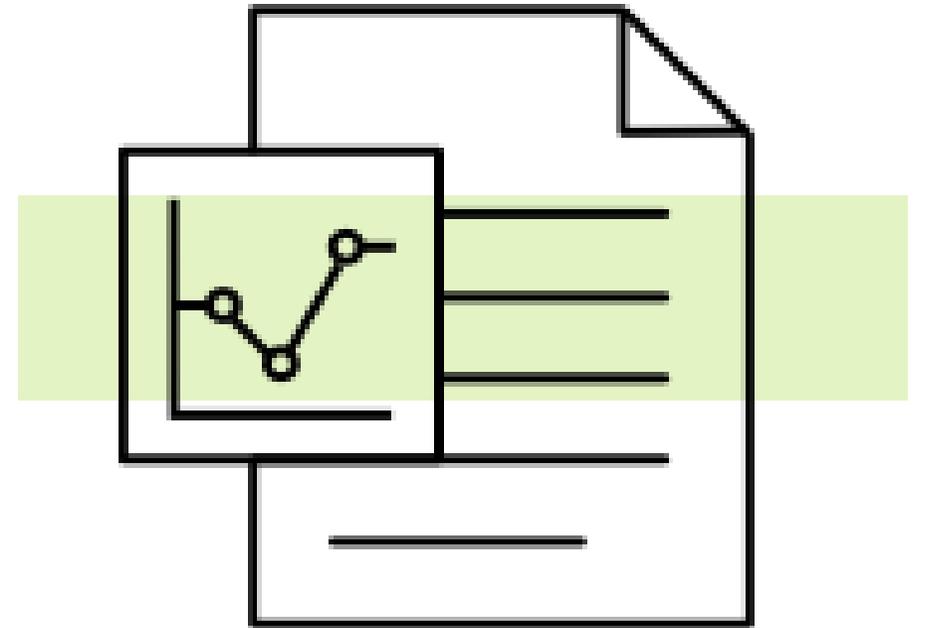
- [GitHub](#)



IBM AI FactSheets 360

A research effort to foster trust in AI by increasing transparency and enabling governance.

- Factsheets for AI Services
- Documentation



Final Thought About Disability

There is a we shout from the rooftops for all to hear. (I hope you do.)

“Nothing About Us Without Us!”



Final Thought For You

“You need to be woke if you want your AI to be woke.”

Noah Blier, [Bias in AI and Machine Learning: Sources and Solutions](#)



Thank You!

John.Rochford@umassmed.edu

Twitter: [@ClearHelper](https://twitter.com/ClearHelper)

LinkedIn: [john-rochford](https://www.linkedin.com/in/john-rochford)

Presentation: bit.ly/AIMLFairness

[AI Fairness Resources - EasyText.AI](#)

This is an [accessible presentation](#).

