

# Data isn't neutral, so what is it? Standpoint theory, positionality, and data in context

“*what one does both enables and limits what one can know.*” (Harding, 1992) | “*the personal is political*” (Combahee River Collective, 1977)

## Overview:

Broadly, **Standpoint Theory** argues that *knowledge is situational* – one’s capacity to create or interpret knowledge is mediated by who they are and how they perceive the world.

**Positionality** takes into account all of the intersectional aspects of the researcher and returns the ‘position’ or standpoint from which the researcher approaches their research. “All persons, including the theorist, have fleshy, material identity that will influence and pass judgment on all political claims.” (Alcoff, 1988) pg 432.

The **Positionality Statement** states this position in the research output and can include gender, race, class, education, geographical region, institutional affiliation, and anything else that might impact the research at hand.

Sarah Holsapple;  
Email: [sarahwh@ibiblio.org](mailto:sarahwh@ibiblio.org)

References and examples on [my website](http://www.ibiblio.org/sarahwh/standpoint.html)  
[www.ibiblio.org/sarahwh/standpoint.html](http://www.ibiblio.org/sarahwh/standpoint.html)

acknowledgement to [Data Feminism \(D'Ignazio & Klein, 2020\)](#) where I originally learned about Standpoint Theory and positionality

## Data in Context:

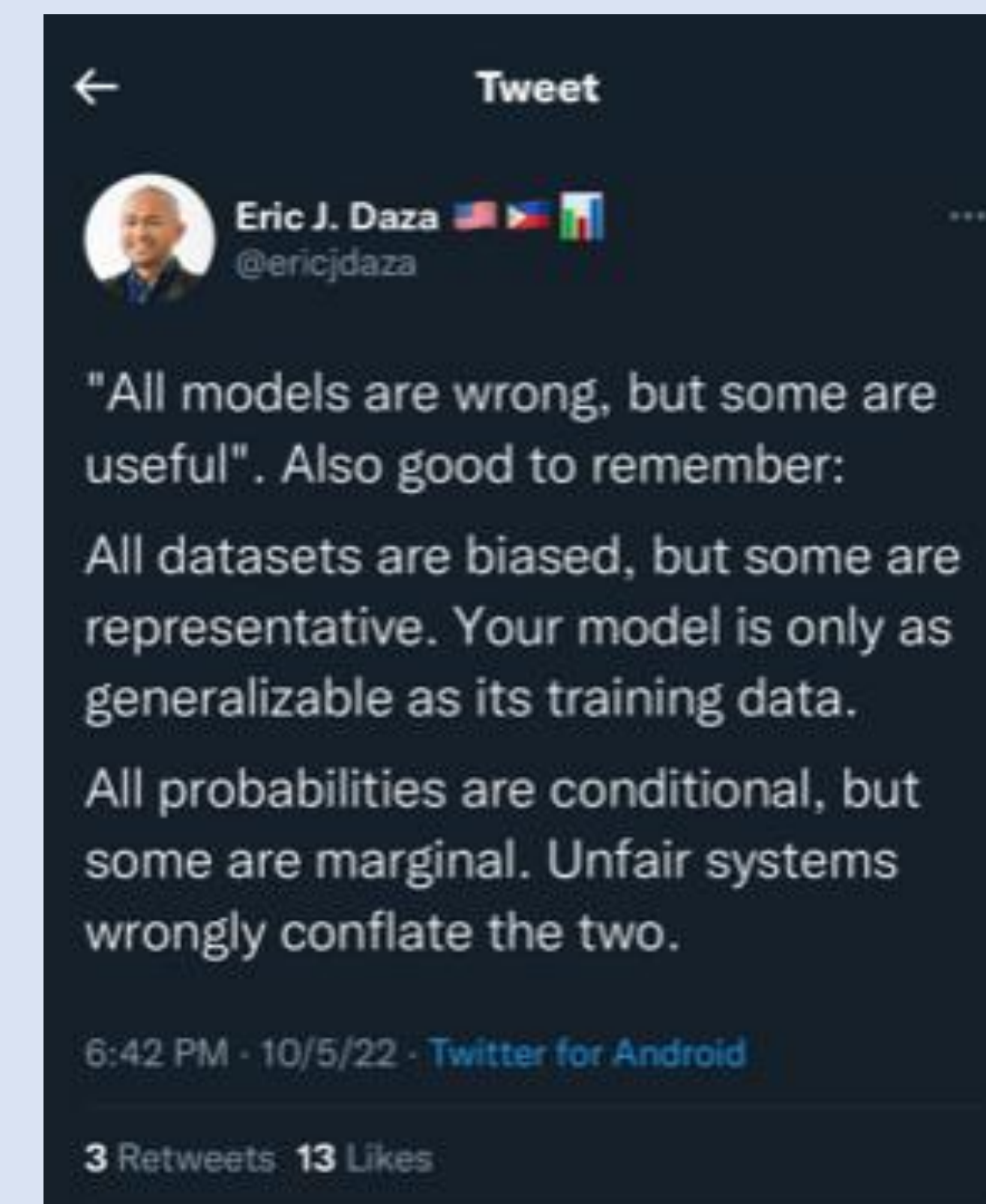
I propose that data is not objective truth and that all research output is mediated by the standpoint of the researcher and the researched. Data in a repository should be accompanied by a data positionality statement that states the **standpoint of the data** – where it was collected, who it was collected from, etc.. That context matters.

## Re-use of data:

Researchers collect data for primary research and share that data in repositories for other researchers to use for secondary analysis – expanding on the original research question, asking new questions, or reproducing original studies. Datasets are sometimes combined to create new, larger datasets. Data without proper context is ripe for misinterpretation and improper generalization ([Boté & Termens, 2019](#))

## Machine Learning:

“Positionality, and how it embeds itself in standards, ontologies, and data collection, is the root for bias in our data and algorithms. Every perspective has its limits - there is no view from nowhere.” [Kaeser-Chen et al.](#)  
“The AI systems learn what they’re taught, and if they are not taught with robust and diverse datasets, accuracy and fairness could be at risk.” ([IBM](#))  
A positionality statement with the data could allow researchers to construct more representative datasets.



## Bigger picture context of all this data.

**Data Equity** - Who gets data collected from them for what purposes, and how does that create equity issues? “big data are producing digital divides as a consequence of the lack of participation, accentuating existing inequalities. ([Ferreira & Vale, 2022](#)) quoting (Benjamin, 2001)”

## More questions:

What should data positionality statement include?  
Can this information be used to harm?  
How to get positionality info into metadata?  
Positionality for all works, in metadata?