# Globalizing Fairness Attributes in Machine Learning: A Case Study on Health in Africa

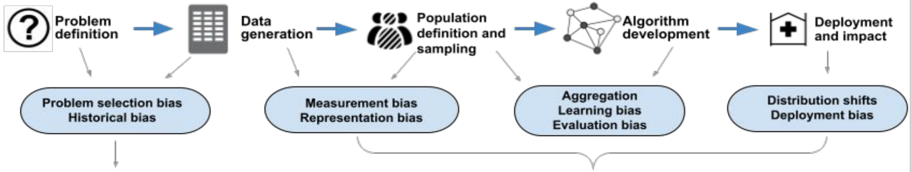
Mercy Nyamewaa Asiedu[1]*, Awa Dieng[1]*, Abigail Oppong[2], Maria Nagawa[3], Sanmi Koyejo[1,5], Katherine Heller[1]
* = equal contribution
1. Google Research, 2. Ghana NLP, 3. Duke University, 4. Ghana NLP , 5. Stanford University

## Summary

**Global Responsible Machine Learning for Health Practices**

- The field of Algorithmic Fairness had been majoritarily contextualized to Western context, raising the question of the meaning of fairness in the Global South.
- To develop fair machine learning models, one needs first to understand what the fairness attributes of a given context are and where to apply them.
- We identify axes of disparities for fairness between African and non-African countries, as well as provide a contextual understanding of globally applicable fairness attributes such as race and religion.
- We underline different limitations to the development of machine learning for health tools in Africa and delineate where fairness attributes need to be considered.
- Finally, we highlight important open challenges to developing fairness-aware methods in Africa including representative data collection and mitigating distribution shifts.

## African-contextualized fairness attributes along the ML pipeline



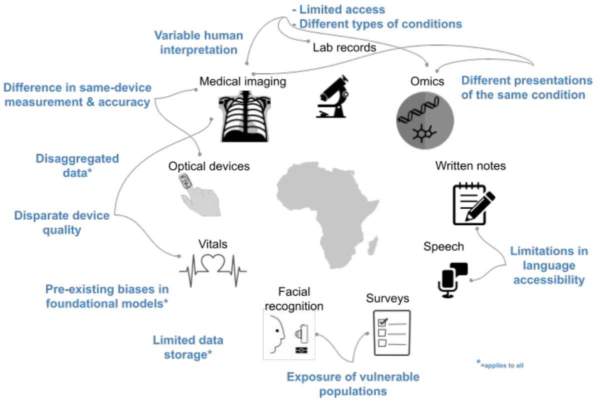

**Africa-specific axes**

Colonial History
National Income level
Country of Origin

**Global axes in the African context**

| Social attributes | Socio-economic attributes | Health attributes |
|---|---|---|
| Race | Literacy and education | Phenotype |
| Ethnicity | Rural-urban divide | Genetics |
| Skin tone | Family income | Pre-existing conditions |
| Religion | Disability | Co-existing conditions |
| Language | | |
| Gender | | |
| Sexual orientation | | |

## African-contextualized barriers to ML4H by health modality




There are several Africa-specific barriers to the development of fair machine learning for health (ML4H) in Africa. For each data modality, we identify these barriers and discuss the fairness attributes that need to be considered by fairness-aware methods.

- **Medical imaging:** race, skin tone, NIL.
- **Medical speech:** language, ethnicity, literacy, education
- **Survey data:** socio-economic attributes, gender, sexual orientation, ethnicity
- **Unstructured written health notes:** language, gender, ethnicity, race
- **Optical sensor devices:** Race, skin tone
- **Omics:** Country of origin, phenotype, genetics, race, gender
- **Lab values:** NIL, rural-urban divide

## Implications for machine learning

Whether training from scratch or finetuning a pretrained model, it is crucial to understand how **African-contextualized fairness attributes affect each step of the ML development pipeline.** We highlight important considerations for health-specific applications in Africa.

→ **Contextualize the fairness criteria** to ensure that fairness definitions are aligned with local laws and cultural beliefs

→ **Align incentives between stakeholders, researchers and practitioners** during problem selection to address contextually relevant problems

→ **Caution around using pretrained models** given there is no guarantees that fairness properties will transfer under distribution shifts.

Corresponding authors: {masiedu, awadieng}@google.com