Google AI

Reimagining ML Fairness in India and Beyond

Nithya Sambasivan

Feb 2021

Image source: Universal History Archive/Getty Images

Team

NITHYA
SAMBASIVAN

ERIN
ARNESEN

BEN
HUTCHINSON

TULSEE
DOSHI

VINODKUMAR
PRABHAKARAN

Motivation

Facial recognition bans

Availability of resources like datasets, benchmarks, APIs, and information acts

Functional, neo-liberal institutions

A healthy civil society

A technically-literate government

A free media

Sources: Corsight (Israel21); Boulamwini and Gebru, 2018; Perpetual Lineup Report, 2016; @jackyalcine Twitter; Facial Recognition and Biometric Technology Moratorium Act of 2020; NYT

Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*

Joy Buolamwini JOYAB@MIT.EDU
MIT Media Lab 75 Amherst St. Cambridge, MA 02139

Timnit G
Microsoft

Amazon Pauses Police Use of Its Facial Recognition Software

The company said it hoped the moratorium "might give Congress enough time to put in place appropriate rules" for the technology.

Civil liberties advocates began calling for a ban on the use of facial recognition by law enforcement in 2018. Elaine Thompson/Associated Press

San Francisco Bans Facial Recognition Technology

Attendees interacting with a facial recognition demonstration at this year's CES in Las Vegas. Joe Buglewicz for The New York Times

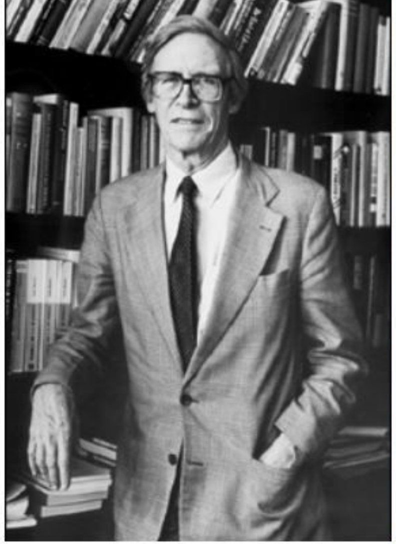
Motivation

Conventional Fair-ML is Western

Fair-ML research is premised on the United States

- **Structural injustices** race and gender
- **Datasets** imagenet and wordnet
- **Measurement** fitzpatrick
- **Laws** civil rights laws
- **Values** enlightenment ideals

Troublingly used in India, Tunisia, Mexico, Uruguay etc.



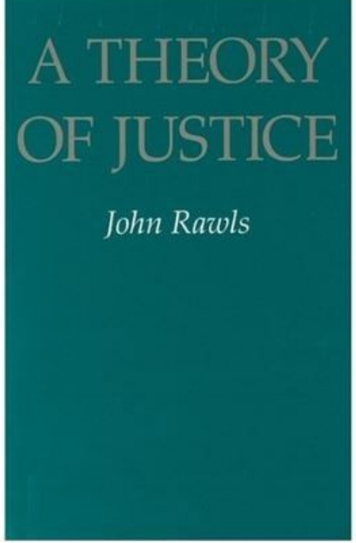


Image source: Harvard

Motivation

Fairness is context-specific

Conventional fairness can be a tokenism or pernicious

Fair-ML should identify defaults, biases, and blindspots to avoid exacerbating harms

Avoid a general theory of algorithmic fairness based on the West.

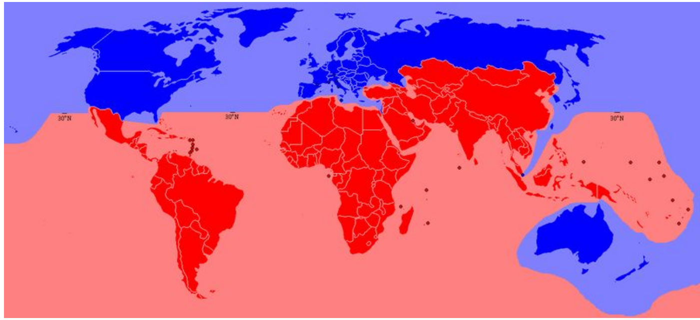


Image source: Global South blog

Motivation

Research questions

Could algorithmic fairness have a structurally different meaning in non-Western contexts?

Is there anything in conventional fairness that could actually be counterproductive in the non-West?

How do social, economic, and infrastructural factors interplay with algorithmic fairness?

Contents

Method

Methodology

1. Semi-structured interviews

- Interviews with 36 scholars, activists, practitioners with expertise on social justice in India, working at the grassroots
- Across law, computer science, economics, sociology, journalism, STS, and political sciences.
- Disability, caste, gender, gender identity, privacy & surveillance, health, constitutional rights, languages.
- 25 male, 10 female, 1 non-binary

2. Analysis of algorithmic deployments

- News publications, policy documents, community media

Image sources: Factor Daily

Method

Feminist, decolonial, and anti-caste lenses

South Asian feminism: equality for all, especially with a focus on caste, religion, class

Decolonialism: undoing of colonialism and imperialism, in knowledge, language, values

Anti-caste emancipation: towards annihilation of caste, undoing brahminical patriarchy


The Times of India

BOMBAY: FRIDAY, AUGUST 15, 1947

BIRTH OF INDIA'S FREEDOM

NATION WAKES TO NEW LIFE

Mr. Nehru Calls For Big Effort From People

"INCESSANT STRIVING TASK OF FUTURE"

Assembly Members Take Solemn Pledge

WILD SCENES OF JUBILATION IN DELHI

STATE VISIT TO KARACHI

LORD MOUNTBATTEN GREETS PAKISTAN

Mr. Jinnah Re-Affirms Firm Friendship With Britain

FRENZIED ENTHUSIASM IN BOMBAY

Crowds In Festive Mood

"MAY BOMBAY PROSPER"

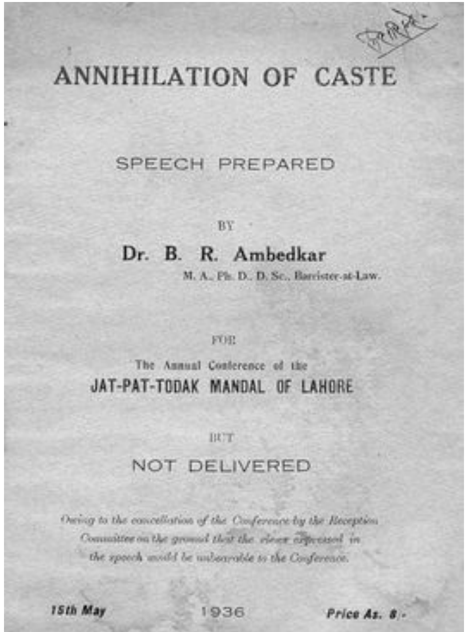
Governor's Message

GOOD WISHES TO FREE INDIA

NEW CABINET OF INDIA

Fourteen Members

PANDIT NEHRU TO BE PREMIER


ANNIHILATION OF CASTE

SPEECH PREPARED

BY

Dr. B. R. Ambedkar

M. A., Ph. D., D. Sc., Barrister-at-Law.

FOR

The Annual Conference of the

JAT-PAT-TODAK MANDAL OF LAHORE

BUT

NOT DELIVERED

Owing to the cancellation of the Conference by the Reception Committee on the ground that the views expressed in the speech would be unbearable to the Conference.

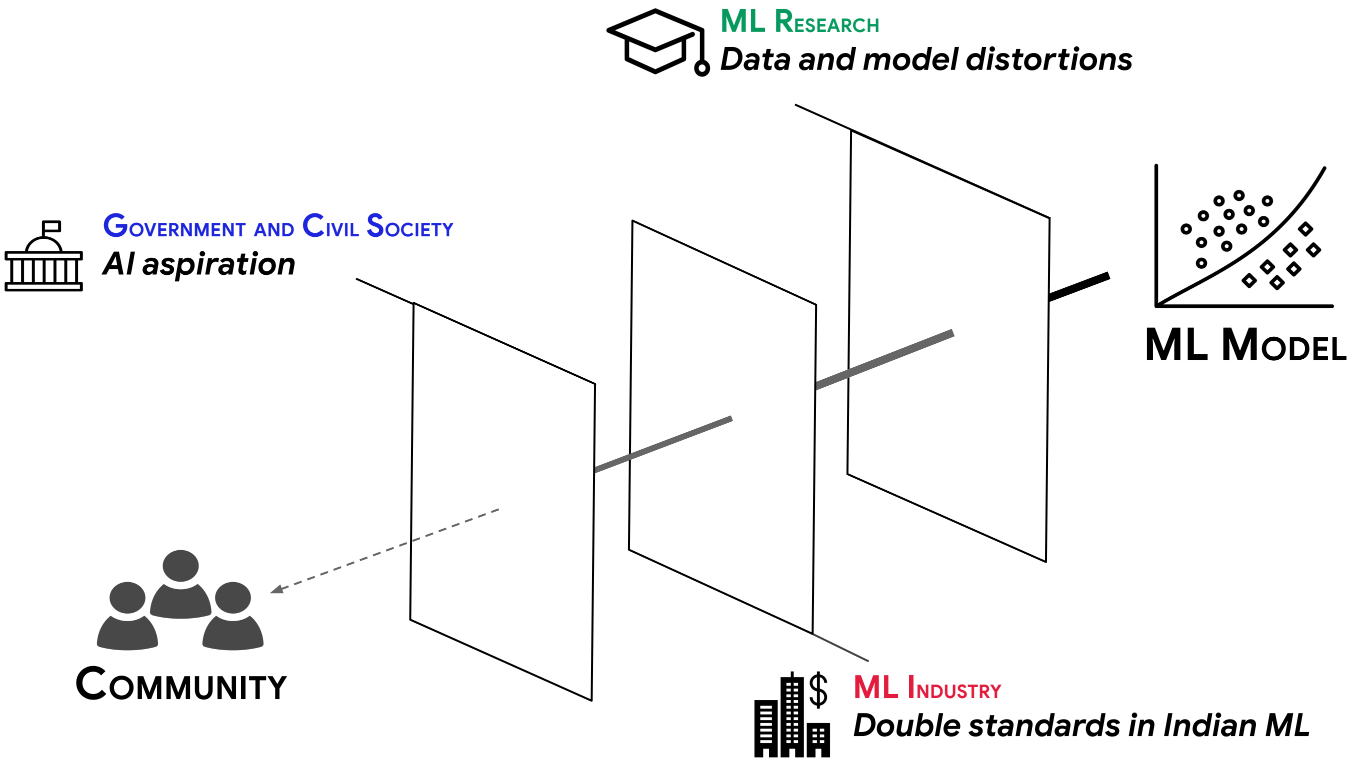
15th May 1936 Price As. 8/-

Image sources: Outlook India

Contents

Findings

Large distance between models and disempowered communities



Data and model distortions

Findings

Missing data and humans

Missing data and invisible humans

- Digital divide
- Caste, gender, or class margins
- ‘Off data’ practices and ‘confusing’ algorithms
- Human infrastructures in data

Data non-transparency and non-availability

Image sources: Indian Express

Mis-recorded identities

One user, one account can fail. e.g., women tend to share their phones and apps.

Only 29% of Internet users were women in 2018.

Mobility and transience: SIM cards change often. Migrant locations change.

Image sources: Google Research

Different sub-groups and proxy implementations

Caste, Religion, Gender, Gender Identity & Sexual Orientation, Income, Ethnicity, and Ability

Pluralistic country

Proxies do not generalize within

Names: Most semantically meaningful

Zip code: Heterogenous SEC types

Mobility: inverse of gender and disability

Oppression is under-reported in discourse

Image source: NDTV

Overfitting models to the privileged

Good data and inferences privilege data-rich users

Typically middle-class men

Mobility, literacy, income

Women end up taking loans in men's names

Model retraining can create new biases

Image source: PARI

Indic social justice

How does India lead us to redefine fairness with its value systems, history, and philosophy?

Reservations and social justice

Collectives and plurality

Image sources: Newsminute

Double standards and distance from ML makers

Findings

‘Bottom billion’ petri dishes

India as a sandbox and playground for ML

Recourse is non-existent and insensitive

Dire for marginalized groups like Dalits

- e.g., Human efficiency tracking
- Dalit and muslim bodies as test subjects

Removal of street-level bureaucrats, officers and human infrastructures by apps and infrastructures

High tech illegibility

Image sources: Huff Po India and *Bureaucratics* by Jan Banning

Entrenched privilege in ML makers

Built by white or dominant caste Indian men

Not inclusive of women, dalits, muslims, PWDs, or the economically poor

AI labour from India thru data collectors and annotators

Caste and religion are eluded even in India

Image sources: Million needs

Findings

Unquestioning AI aspiration

Findings

AI euphoria

High confidence in high stakes AI
e.g., Delhi FR used to arrest 1,100 protestors

*"This is a **software**. It **does not see faith**. It does not see clothes. It **only sees the face** and through the face the person is caught."*

Undertrials: 52% Dalits & Adivasis (25% of pop.). 21% Muslims (14% of pop.)

'Smart cities' not inclusive of minorities

Fully inscrutable AI inputs and outputs

Challenging for researchers to interrogate

Lack of inclusion of stakeholders in high-stakes AI

Image sources: ToI, Indian Express

Questioning AI power

Indian tech journalism is business-oriented, not critical

95% of reporters come from privileged castes

Fairness issues require a healthy ecosystem of activists, media, and civil society

World Emoji Day: Apple, Google Tease New Emojis Coming to Android and iOS

The tech giants are set to roll out the new emoji options later this year.

By Abhik Sen

Twitter rolls out new interface for DMs on web

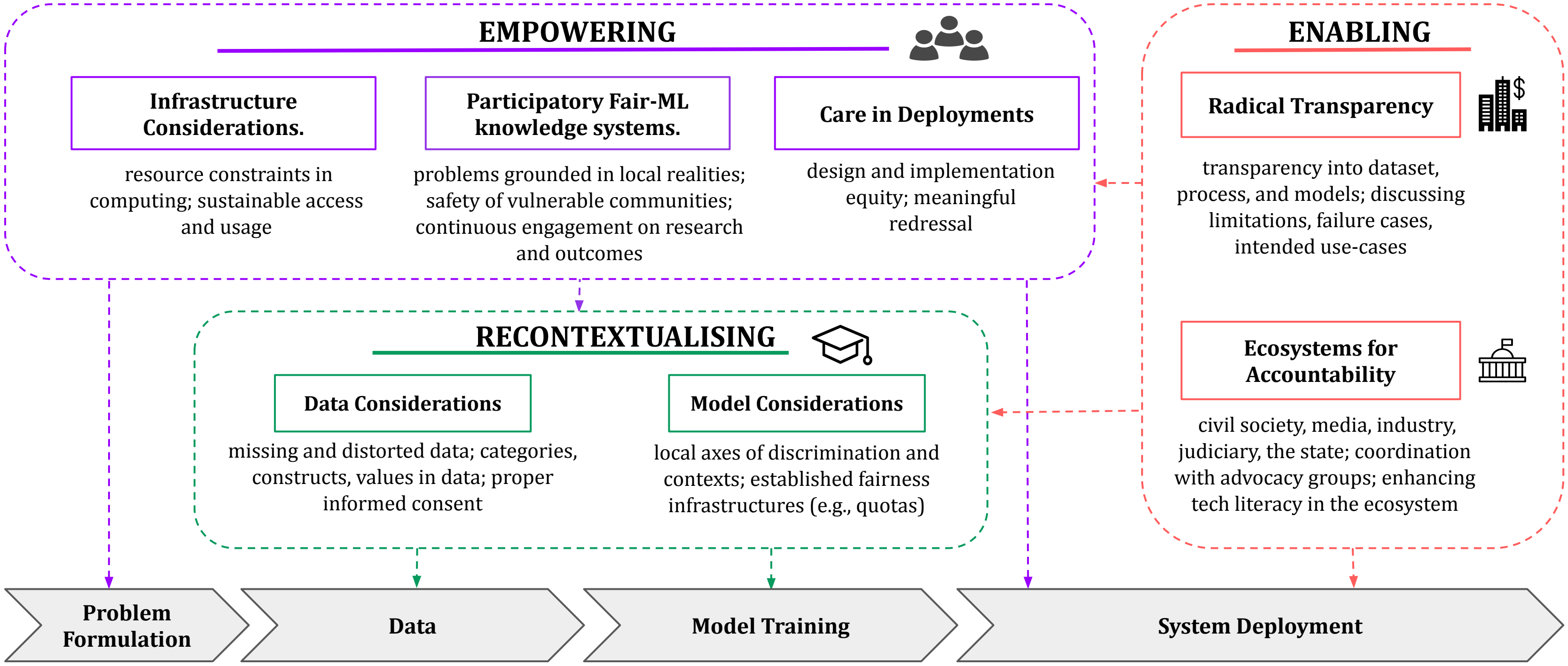
TIMESOFINDIA.COM | Jul 17, 2020, 05:13PM IST

Image sources: ToI, NDTV, Propublica

Contents

Discussion

Towards AI fairness in India



Discussion

Recontextualizing data and models

Combining observational research & dataset analysis

Data gathering & consent via community relationships

Social audits

Normative frameworks, e.g. ethics of care

Image sources: The Hindu

Discussion

Empowering communities

Participatory & assets-based research approaches

ICTD & HCI4D design approaches

First-world care in deployments

Diversity of ML makers

Image sources: Design Beku

Discussion

Enabling a Fair-ML ecosystem

Transparency on datasets, models and processes

Granting access to APIs, data & negative results

Project partnerships

Need investigative journalism on Indian AI

Discussion

Coda

Context matters

Considerations are not limited to India

Move towards a pluriverse of AI ethics

nithyasamba@google.com

Acknowledgements

A. Aneesh
Aishwarya Lakshmiratan
Ameen Jauhar
Amit Sethi
Anil Joshi
Arindrajit Basu
Avinash Kumar
Chiranjeeb Bhattacharya
Dhruv Lakra
George Sebastian
Jacki O Neill
Mainack Mondal
Maya Indira Ganesh
Murali Shanmugavelan
Nandana Sengupta
Neha Kumar
Rahul De
Rahul Matthan
Rajesh Veeraraghavan
Ranjit Singh
Ryan Joseph Figueiredo (Equal Asia Foundation)
Savita Bailur
Sayomdeb Mukerjee
Shanti Raghavan
Shyam Suri
Smita
Sriram Somanchi
Suraj Yengde
Vidushi Marda
Vivek Srinivasan
and other experts who wish to stay anonymous
Jose M. Faleiro
Daniel Russell
Jess Holbrook
Fernanda Viegas
Martin Wattenberg
Alex Hanna
Reena Jana
for their invaluable feedback