

Decolonizing Content Moderation: Does Uniform Global Community Standard Resemble Utopian Equality or Western Power Hegemony?



Farhana Shahid



Aditya Vashistha

Background



"The cows and cars cross the river."

- Rabindranath Tagore (Poem: Our Small River)

Our standards on harassment and bullying

We define bullying and harassment as targeting someone with threats or insulting someone's appearance or character. This includes:

- Comparisons to animals or insects to make them feel inferior (for example: cow, worm)
- Comparisons to objects to make them feel inferior (for example: potato)

“পার হয়ে যায় গরু, পার হয় গাড়ি”



You can't post or comment for 3 days. This is because you previously posted something that didn't follow our Community Standards.

Image Source: STL Today

Background

TECH / FACEBOOK / POLICY

Facebook reportedly bracing for US election chaos with tools designed for 'at-risk' countries

Analysis

Facebook's role in Myanmar and Ethiopia under new scrutiny

Emmanuel Akinwotu

Whistleblower Frances Haugen adds to long-held concerns that social media site is fuelling violence and instability

Background

Facebook while black: Users call it getting 'Zucked,' say talking about racism is censored as hate speech

Jessica Guynn USA TODAY

Published 7:26 a.m. ET April 24, 2019 | Updated 6:17 p.m. ET July 9, 2020

Will Instagram's new 'sensitive content' filter censor Black users?

Instagram recently announced a new content filter to limit what people see in their Explore pages. Critics say it will negatively affect Black users.

future tense

TikTok Admits It Suppressed Videos by Disabled, Queer, and Fat Creators



Image Source: USA Today, NBC News, Slate

Prior Work

Harms

Disproportionate Removals and Differing Content Moderation Experiences for Conservative, Transgender, and Black Social Media Users: Marginalization and Moderation Gray Areas

Authors:  [Oliver L. Haimson](#),  [Daniel Delmonaco](#),  [Peipei Nie](#),  [Andrea Wegner](#) [Authors Info & Claims](#)

 Open access |    | Research article | First published online July 22, 2020

Content moderation: Social media's sexist assemblages

[Ysabel Gerrard](#)   and [Helen Thornham](#) [View all authors and affiliations](#)

[Volume 22, Issue 7](#) | <https://doi.org/10.1177/1461444820912540>

#thyghgapp: Instagram Content Moderation and Lexical Variation in Pro-Eating Disorder Communities

Authors:  [Stevie Chancellor](#),  [Jessica Annette Pater](#),  [Trustin Clear](#),  [Eric Gilbert](#),  [Munmun De Choudhury](#)

[Authors Info & Claims](#)

Frameworks

Drawing from justice theories to support targets of online harassment



[Sarita Schoenebeck](#)  , [Oliver L Haimson](#), and [Lisa Nakamura](#) [View all authors and affiliations](#)

[Volume 23, Issue 5](#) | <https://doi.org/10.1177/1461444820913122>

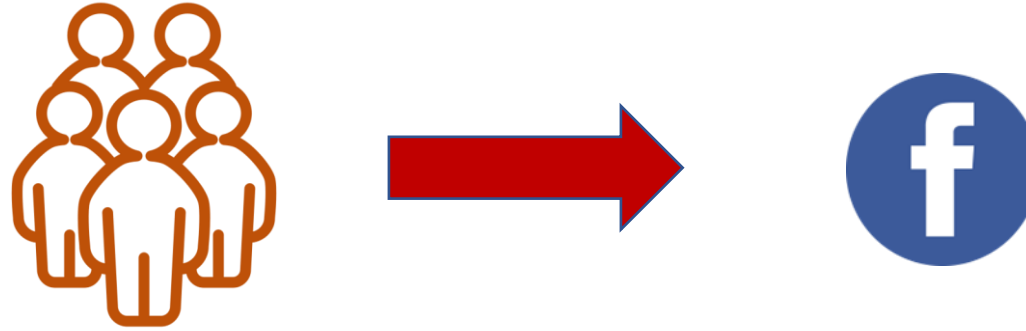
Democratising online content moderation: A constitutional framework

[Giovanni De Gregorio](#)  

Digital Juries: A Civics-Oriented Approach to Platform Governance

Authors:  [Jenny Fan](#),  [Amy X. Zhang](#) [Authors Info & Claims](#)

Methods



Semi-structured interviews with
19 Bangladeshi Facebook users

Violation of Facebook's
community standards

Research Questions

RQ1. How might content moderation systems propagate historical power relations and patterns of oppression?

Universal Ethics: Centering Western Norms

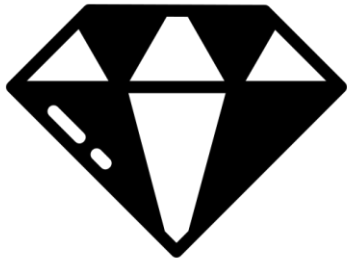


Facebook removed an image of
village kids swimming in the pond for
child nudity



Facebook group was deleted as **emergency**
blood donation requests containing **blood**
group, contact number violated **privacy**

Universal Ethics: Ignores Local Sensitivities



Grandmother's affectionate remark
"black diamond" for grandson
removed for racist hate speech



Islamophobic removal of "No religion
teaches to disrespect the holy script of
another religion" –without explanation

Inaccessible Standards



*“All users do not know English well. Rickshaw pullers [drivers of local passenger cart] also **use Facebook and they won’t be able to read the English guideline.**”*

- Rashed

*“If we want to establish a **global community standard**, are we saying that we want to move towards one world one order? Is it the **old socialist utopia for equity?**”*

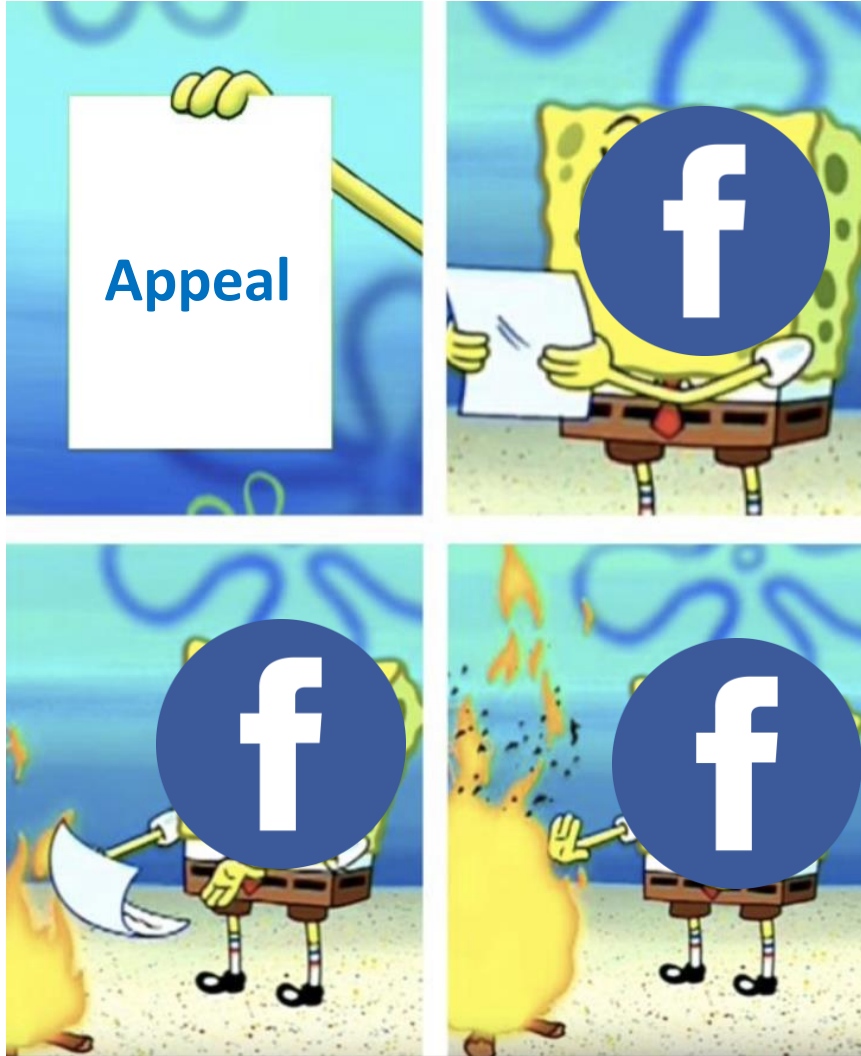
- Shafique



Threadbare Appeal Process

“... I reached out to my acquaintance working at Facebook, gave my account details, and within one week my account was restored. Even though I received official email from Facebook that my account was banned, I didn’t receive official notification when my account was up again.”

- Kabir



Research Questions

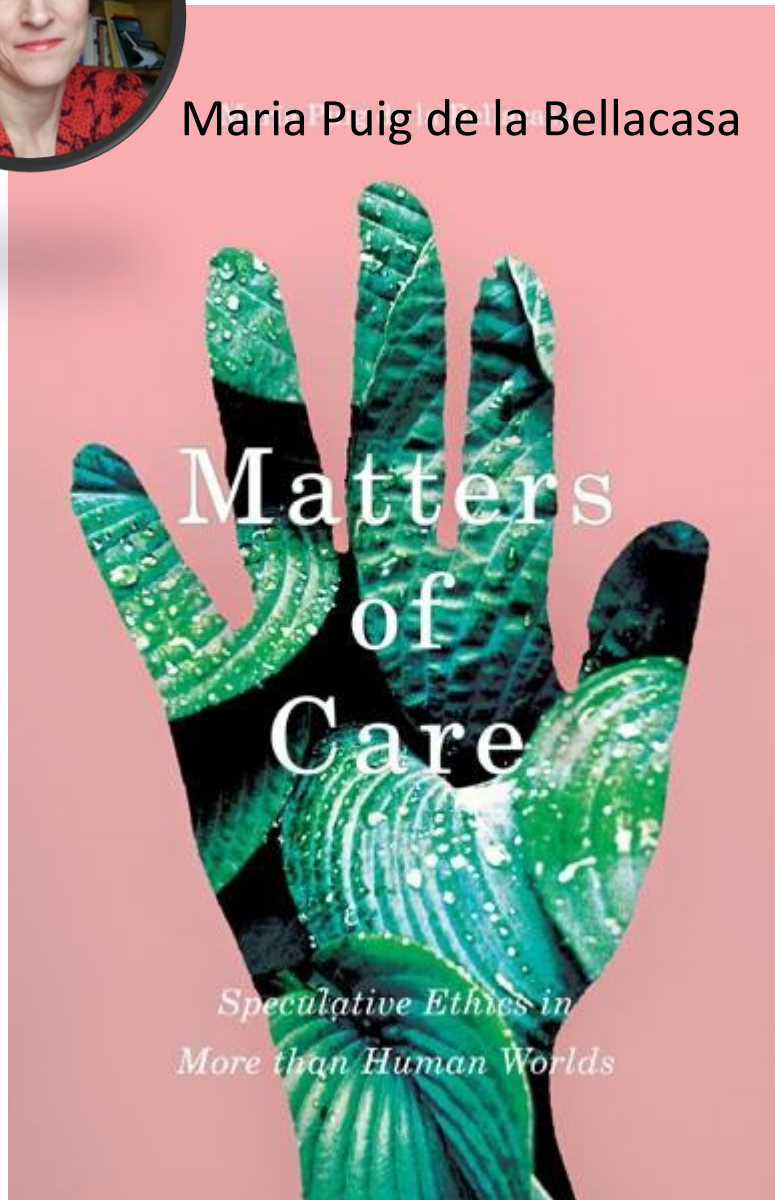
RQ2. What steps can we take towards a fundamentally decolonial content moderation system?

Multidimensional Framework of *Care*

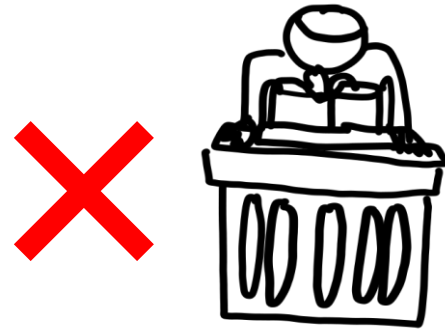
- Ethics/ Politics
- Work/ Labor
- Affect/ Affections



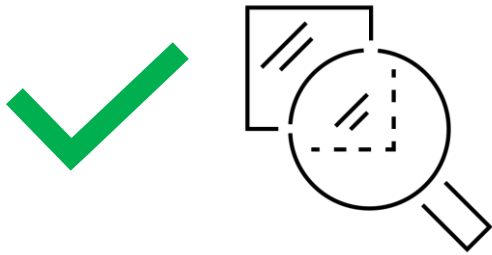
Maria Puig de la Bellacasa



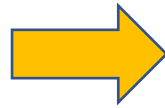
Towards Decolonial Moderation: Ethics/ Politics



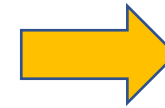
Paternalistic attitude



Increase transparency



Power imbalance



Explainable community standards

Towards Decolonial Moderation: Work/ Labor



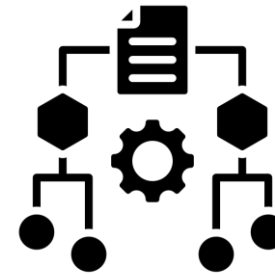
Linguistic discrimination



Objectified care

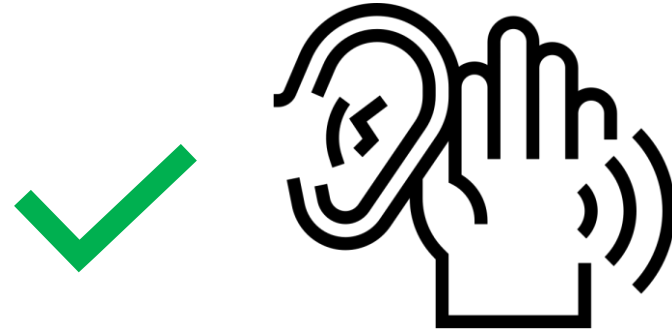


Fairness aware computing



Algorithmic transparency

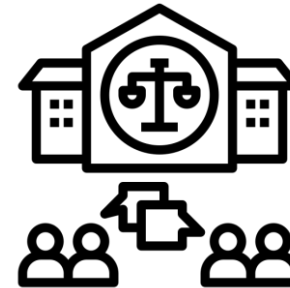
Towards Decolonial Moderation: Affect/ Affection



Listen with care



Restorative justice



Resources to appeal

Dankeschön (Thank You!)



https://bit.ly/decolonial_mod

✉ fs468@cornell.edu

🐦 @FarhanaShahid_