

facebook INDIA

**TOWARDS THE TIPPING POINT OF VIOLENCE:
CASTE AND RELIGIOUS HATE SPEECH**



EQUALITY LABS

EQUALITY LABS

Equality Labs works at the intersection of community research, education, technology, and art to support South Asian religious, ethnic, and cultural minority movements dealing with intractable systems of oppression. Equality Labs is the first South Asian American human rights and technology start-up whose leadership is centered around women, gender non-conforming, trans, and queer people.

ISBN: 978-0-578-49021-2

First published in 2019 by Equality Labs.

www.equalitylabs.org

equalitylabs@protonmail.com

Suggested citation: Soundararajan, T., Kumar, A., Nair, P., Greely, J.

Except where otherwise noted, content in this document is licensed under a Creative Commons (attribution, non-commercial, no derivatives, international 4.0) license.

<https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>

Facebook India: Towards The Tipping Point of Violence Caste and Religious Hate Speech. Equality Labs, USA.

Facebook India

**TOWARDS THE TIPPING POINT
OF VIOLENCE:
CASTE AND RELIGIOUS HATE SPEECH**

Equality Labs 2019

Table of Contents

Exexecutive Summary	6
Recommendations.....	8
Introduction	10
The Limits of Community Standards.....	12
Failure in Reporting.....	19
Hate Speech Analysis.....	23
Conclusion	64

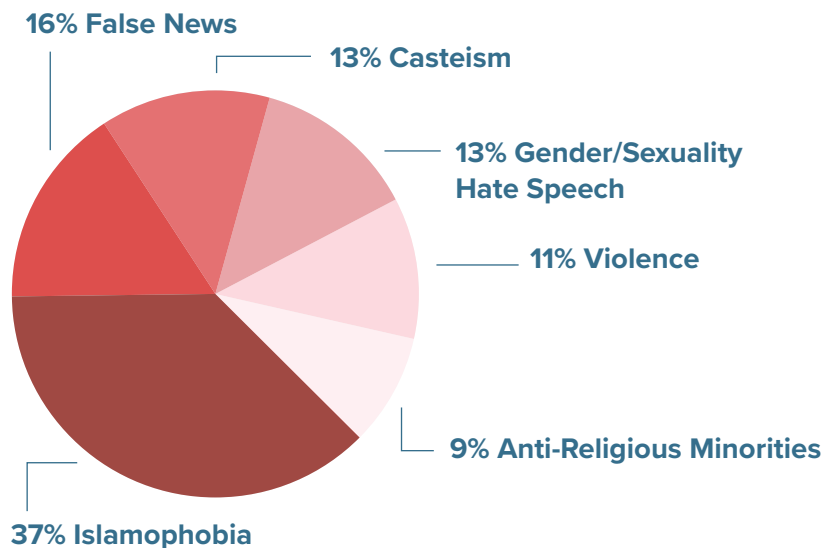
Executive Summary

This report summarizes Equality Labs' advocacy concerning hate speech on Facebook during 2018. It reveals disturbingly real threats that Facebook content presents to an estimated 300 million Indian caste, religious, gender, and queer minorities both in India and abroad.

We urge Facebook, the Indian public, and the global Internet Freedom community to act quickly and appropriately based on our findings. Without urgent intervention, we fear we will see hate speech weaponized into a trigger for large-scale communal violence. We further urge Facebook to take responsibility for its platform, and to cooperate with representatives of Indian civil society and Internet Freedom activists. Facebook must act to mitigate specific risks, and get more aggressive about the platform's impact on the physical safety and fundamental human rights of India's most vulnerable communities.

KEY FINDINGS

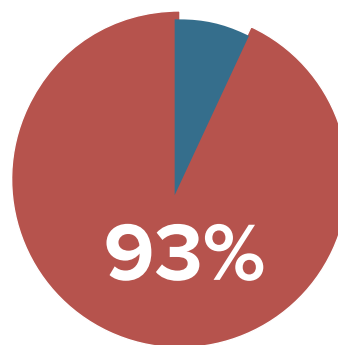
Hate speech targeting Indian caste, religious, gender, and queer minorities is rampant across Facebook.



Facebook staff lacks the cultural competency needed to recognize, respect, and serve caste, religious, gender, and queer minorities. Hiring of Indian staff alone does not ensure cultural competence across India's multitude of marginalized communities. Minorities require meaningful representation across Facebook's staff and contractor relationships. Collaboration with civil society and greater transparency of staffing diversity strengthens hate speech mitigation mechanisms like content moderation.

There is widespread doxxing occurring on the platform, threatening activists, journalists, and others who speak on behalf of the vulnerable.

93% of all hate speech posts reported to Facebook remain on Facebook. This includes content advocating violence, bullying and use of offensive slurs, and other forms of Tier 1 hate speech, reflecting a near total failure of the content moderation process.



Procedures for reporting these activities to Facebook are opaque, increasing people's vulnerability, and safety concerns for the persons affected.

Most hate speech violations on Facebook India are Islamophobic. 6% of Islamophobic posts were specifically anti-Rohingya, with calls to violence similar to content that led to the Rohingya genocide in Myanmar.

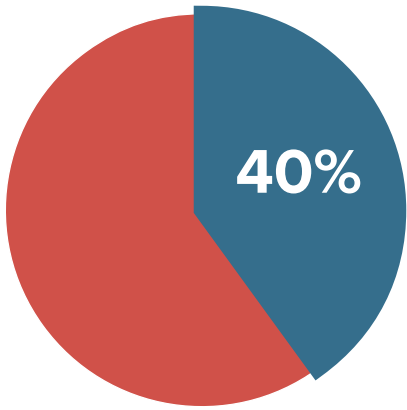
1 out of 4 posts containing gender/sexuality hate speech were transphobic or queerphobic in content. Half of these called for or glorified rape and other forms of sexual violence.

While hate speech almost completely remains online or is reinstated by moderators on Facebook, an increasing number of minority user accounts are being banned or removed entirely.



Nearly half of the hate speech Facebook initially removed was later restored.

100% of these restored posts were Islamophobic in nature.



40% of all casteist hate speech on Facebook India is preoccupied with India's affirmative action policy, known as "reservation."

Remaining posts include caste slurs, derision of caste-based menial labor, anti-Ambedkar messages, and anti inter-caste personal relationships content.

While hate speech almost completely remains up or is reinstated by moderators on Facebook, an increasing number of minority user accounts are being banned or removed entirely.

Facebook Community Standards are not yet localized for many of India's major languages. Despite our engagement with Facebook, there has been no progress to date on this front.

Facebook lacks clear user hate speech reporting mechanisms for Indian caste-oppressed minorities. After a year of advocacy with Facebook, we are deeply concerned that there has been little to no response from the company.

Facebook is unable to track caste and religious hate speech disaggregated by category. Such data is critical to identifying hate speech by scope and scale. Data transparency on this front will allow experts in civil and public society to effectively track and contribute to mitigating hate speech.

RECOMMENDATIONS

Content moderation of hate speech in India is complex, requiring relevant context and collaborative expertise from civil society and advocates. Facebook cannot solve this problem in a vacuum. Transparency, accountability, and resource allocation are the keystones required for an effective and durable solution to the exponential expansion of hate speech enabled by social media platforms like Facebook. The result will be a safer and more welcoming platform for users in a growing market. We believe the following steps are necessary in order to achieve this outcome.

Facebook must prioritize an aggressive and thorough human rights audit of the following:

- The human rights impact of Facebook policy and programs and how the platform has been used by hate groups, political entities, and other public figures to stoke casteist and religious animosity or violence

- What risk assessments, if any, were conducted to improve understanding of the threats faced by Indian minorities on the platform
- Facebook’s hiring practices, especially with respect to safety, policy, and content moderation teams.
- Content Moderation, which should include hiring practices, contractor demographics, and slur lists. These lists should be open and transparent to the public.
- User Privacy
- Targeted Advertisements
- Security Policies
- Facebook’s elections and government unit’s work from elections 2014 to elections 2019.

The audit should follow guidelines similar to the US civil rights audit and should include:

- Empowering an independent audit team that is approved and monitored by both civil society and Internet Freedom advocates as well as by Facebook.
- This audit team must have clear competencies in caste, religious, and gender/queer minorities and includes members of Indian minorities in its composition.
- Research comparing impacts across India’s discrete language markets for analysis of implementations.
- Determinations regarding risk prevention, mitigation, and remediation plans for vulnerable communities.
- Revision of policies and practices to address the human rights risks identified.

Facebook must engage in more transparent dialogues with Internet Freedom advocates and Indian civil society to create more efficient conduits for advocacy, intervention, education, and training:

- We recommend a regularly convened working group of Indian Internet freedom and civil society groups that work on the issues of caste, religious, and gender/queer minorities. This group could work actively to counter casteist and religious bigotry while also helping provide input into Facebook’s policies and processes.

Facebook must establish a remediation and diversity fund for survivors of violence that result from the platform.

So many of the problems of hate speech are amplified by the lack of transparency and engagement between Facebook and organizations that could help with frontline experience.

The situation is grave enough that we must see an overhaul of these processes—nothing less than the integrity of India, the world’s largest democracy, is at stake.

Introduction

India is now the largest market for Facebook and its related companies. In a country of 1.3 billion people with over 294 million active accounts, Facebook India's user base will soon exceed the entire population of the United States. And, unlike the U.S., Facebook India has an unprecedented growth potential of over 400% because a whole new generation of users has yet to come online. The future of Facebook then is the Indian market.

Facebook operates as a global public square, providing crucial community information to millions of people each day. For many first-time internet users, it is a primary source of news and has become the de facto internet for a growing number of people in the global south. As such, Facebook plays a crucial role in shaping public opinion, community dynamics, and relationships.

For Indian minorities, Facebook India has become a critical platform for building community and seeking new audiences. It provides a unique window into the lives of Indian Dalits, women, religious and gender/queer minorities, lifting up voices that might not be heard otherwise. It also creates powerful opportunities for dialogue, engagement, and global connection. This is why, as representatives of these communities, we are deeply concerned that Facebook products have also become the central avenue for spreading hate speech and disinformation. Equally concerning is that Facebook's own community management policies in India further undermine user safety and privacy on the platform.

Facebook has repeatedly come under fire for insensitivity to the concerns of Indian minorities. This includes suspending or banning the accounts of minority leaders and press outlets, and a general lack of transparency regarding its collaborations with law enforcement. This reckless disregard is coupled with Facebook's frustrating opacity around content moderation, data stewardship and crucially, elections operations in the country.

This report is part of an extended advocacy campaign conducted by Equality Labs throughout 2018 to address these issues and promote public conversation concerning very real threats that the Facebook presents to Indian minorities in India and abroad. The report features data and analysis from a six-month research project monitoring caste, religious, and gendered hate speech on Facebook India. The writers of this report Thenmozhi Soundararajan, Abishek Kumar, Priya Nair, and Josh Greely supervised an inter-faith, inter-caste, and multi-racial research team composed of researchers from over 15 countries. Many researchers were Indian minorities who experienced this violence not just in India but in their new countries of residence. Our team was global because the problem of hate speech in Facebook India is not limited to the region which is why this research merits an international response.

Our findings reveal a disturbing volume of direct hate speech, disinformation, and calls to violence against Indian caste, religious, and gender/queer minorities as well as the free

press. Such calls to violence have led to dangerous attacks on civil society, including the murder of award-winning journalist Gauri Lankesh and the detention of numerous Dalits.

As a civil society organization committed to South Asian minority communities, we disseminate materials and training on digital literacy, and safety tools relevant to Facebook and other social media platforms, to counter the disinformation underlying caste and religious tensions. However, without Facebook as a transparent and accountable partner in the India market, the scope of the problem far outweighs our capacity to assist all those in need. We cannot do this alone. And more importantly the burden of user safety is not on the users but on Facebook itself.

We urgently raise our concerns to Facebook, the Indian public, and the global Internet Freedom community. Without swift and aggressive intervention, we fear the lack of necessary caste and religious competence at Facebook India enables the weaponization of hate speech to instigate large-scale communal violence. The Rohingya Genocide, the Christchurch mass murders, and the epidemic of Indian Whatsapp-mob lynchings are clear evidence that online hate speech can quickly erupt into severe violence. We must act now and swiftly to prevent further societal harm.

We ask that Facebook take responsibility and cooperate with Indian Internet Freedom activists and Indian civil society in order to mitigate further risk and harm. Facebook must commit to addressing its impact on the human rights of India's most vulnerable peoples according to the UN guiding principles for business and human rights.

Also examined in this report are the limitations of community management on the platform with regard to caste and religion. Additionally, our preliminary findings from the hate speech monitoring project are presented, shedding light on the range and types of hate speech found on Facebook. Finally wherever other reporting and research has been done to help shed light on problems like banning of user accounts and doxing we have included those reports in order to amplify those findings but also to help provide a comprehensive picture of the problem. Our hopes are this report opens an urgently needed conduit for both advocacy and corporate accountability.

The Limits of Community Standards

Facebook Community Standards set parameters for safe engagement on the platform by laying out user rights and processes for mitigation of harm. Ideally, these standards would be a key pillar in Facebook’s strategy for effective community management. In practice, however, the Facebook India’s standards have not evolved in tandem with the company’s expansion into markets in the Global South. Our research determined that they were formulated out of little to no collaboration with civil society experts or Internet Freedom advocates from Indian minority communities.

From Facebook’s guidelines on hate speech:

https://www.facebook.com/communitystandards/hate_speech.

We define hate speech as a direct attack on people based on what we call protected characteristics — race, ethnicity, national origin, religious affiliation, sexual orientation, caste, sex, gender, gender identity, and serious disease or disability. We also provide some protections for immigration status. We define attack as violent or dehumanizing speech, statements of inferiority, or calls for exclusion or segregation. We separate attacks into three tiers of severity, as described below.

Facebook now includes “caste” when describing protected classes in hate speech. Facebook deserves credit for this addition and other social media platforms would serve their users well to adopt it. However, Facebook fails in implementing protections for these classes on their platform in the Indian market, as we will describe later in the report.

Facebook ranks the severity of different violations of its standards using a three-tier system. Corresponding actions are provided to moderators according to the violation’s ranking. On the next page is the Facebook guidelines as shared in their community standards.

Tier 1	Tier 2	Tier 3
<p>Attacks which target a person or group of people who share one of the above-listed characteristics or immigration status.</p> <p>(including all subsets except those described as having carried out violent crimes or sexual offences), where an attack is defined as</p> <ul style="list-style-type: none"> ● Any violent speech or support in the written or visual form ● Dehumanizing speech such as reference or comparison to: <ul style="list-style-type: none"> ● Insects ● Animals that are culturally perceived as intellectually or physically inferior ● Filth, bacteria, disease and faeces ● Sexual predator ● Subhumanity ● Violent and sexual criminals ● Other criminals (including but not limited to “thieves,” “bank robbers,” or saying “all [protected characteristic or quasi-protected characteristic] are ‘criminals’”) ● Mocking the concept, events or victims of hate crimes even if no real person is depicted in an image ● Designated dehumanizing 	<p>Attacks, which target a person or group of people who share any of the above-listed characteristics, where an attack is defined as</p> <ul style="list-style-type: none"> ● Statements of inferiority or an image implying a person’s or a group’s physical, mental, or moral deficiency ● Physical (including but not limited to “deformed,” “undeveloped,” “hideous,” “ugly”) ● Mental (including but not limited to “retarded,” “cretin,” “low IQ,” “stupid,” “idiot”) ● Moral (including but not limited to “slutty,” “fraud,” “cheap,” “free riders”) ● Expressions of contempt or their visual equivalent, including (but not limited to) <ul style="list-style-type: none"> “I hate” “I don’t like” “X are the worst” ● Expressions of disgust or their visual equivalent, including (but not limited to) <ul style="list-style-type: none"> “Gross” “Vile” “Disgusting” ● Cursing at a person or group of people who 	<p>Attacks, which are calls to exclude or segregate a person or group of people based on the above-listed characteristics.</p> <p>We do allow criticism of immigration policies and arguments for restricting those policies.</p> <p>Content that describes or negatively targets people with slurs, where slurs are defined as words commonly used as insulting labels for the above-listed characteristics.</p>

At first glance, these standards can be daunting for the user to wade through, but if Indian Internet Freedom activists and civil society were brought to the table, our groups could serve as a crucial bridge to user training and education. Civil society represents a much needed missing link to improving and maintaining these standards, -- from giving insight into evolving forms of hate speech to training vulnerable users. Our input could better inform the distinctions between Tier 1, Tier 2, and Tier 3 hate speech.

Civil society and Internet Freedom advocates could also assist in creating culturally competent training that enables all Facebook users to understand how its standards, coupled with rigorous content moderation, can work together for a safer Facebook India.

This input is necessary because the community standards were designed primarily for the US and European markets. The Indian religious and socio-political contexts are complex enough to require their own review and co-design process to adequately address safety. It is also quite crucially, not a static context — the constantly evolving memes and narratives at play in the heart of hate speech must be met with an equally nimble process to contain them.

HATE SPEECH GUIDELINES ARE NOT LOCALIZED INTO ALL INDIAN LANGUAGES

Beyond the larger issue of the drafting process of Facebook's community standards, the major concern is that the standards are not available in all Indian languages, despite the fact that millions of users speak, read, and write in these languages.

This is simply unacceptable. Without nuanced and accurate translations of the guidelines and reporting mechanisms, users are completely uninformed of their rights on Facebook and the remedies available to deal with abuse.

How can Facebook guarantee the safety of all of its users if the basic community standards are not available for all to read? If Facebook has prioritized investments in the technical infrastructure to support these language markets, it must also allocate funds and attention to creating localized resources to protect users in these markets. Safety cannot be an afterthought — it must be central to the production workflow.

Currently, community standards pages for Indian languages often present only the headings in a regional language and the rest of the text in English. Most frustrating are Facebook's handling of languages like Urdu. Urdu is the native language of over 61 million people, most of whom are Indian and Pakistani. Over 100 million people speak Urdu worldwide. Like Arabic, Urdu is written right-to-left. The Urdu page for Community Standards reads in English, but it's English that is right-justified as if it were written right to left. Further, end punctuation is inexplicably moved from the right to the left side of the line. Such clumsiness is needlessly confusing at best.

Despite engaging with Facebook about localization, we have not seen progress in fixing any of these issues. We had hoped for progress before the Indian Lok Sabha elections. We implore Facebook again to consider the damage to users from participating without protection, and to invest in the localization, education, and dissemination of community standards in all of India's official languages.

English: Full translation.

English Curriculum - Google D... x FB Community standards lang... x Google Translate x translate.google.com x Community Standards

https://www.facebook.com/communitystandards/hate_speech

Community Standards Search the Community Standards

Introduction

I. **Violence and Criminal Behavior**

II. **Safety**

III. **Objectionable Content**

12. Hate Speech

13. **Violence and Graphic Content**

14. **Adult Nudity and Sexual Activity**

12. Hate Speech

Policy Rationale

We do not allow hate speech on Facebook because it creates an environment of intimidation and exclusion and in some cases may promote real-world violence.

We define hate speech as a direct attack on people based on what we call protected characteristics — race, ethnicity, national origin, religious affiliation, sexual orientation, caste, sex, gender, gender identity, and serious disease or disability. We also provide some protections for immigration status. We define attack as violent or dehumanizing speech, statements of inferiority, or calls for...

Hindi: Full translation.

Facebook x Facebook x समुदाय मानक | Facebook x Facebook x Community Standards | Facebook x +

https://hi-in.facebook.com/communitystandards/hate_speech

समुदाय मानक समुदाय मानक खोजें

भूमिका

I. **हिंसा और अपराधिक व्यवहार**

II. **सुरक्षा**

III. **आपत्तिजनक सामग्री**

12. अभद्र भाषा

13. **हिंसा और ग्राफिक सामग्री**

14. **व्यस्क नरता और यौन मतिविधि**

15. **यौन आग्रह**

12. अभद्र भाषा

नीति की तर्कसंगतता

हम Facebook पर अभद्र भाषा की अनुमति नहीं देते हैं क्योंकि इससे डर और बहिष्कार जैसा माहौल बन जाता है और कुछ मामलों में वास्तविक दुनिया में हिंसा को बढ़ावा मिल सकता है।

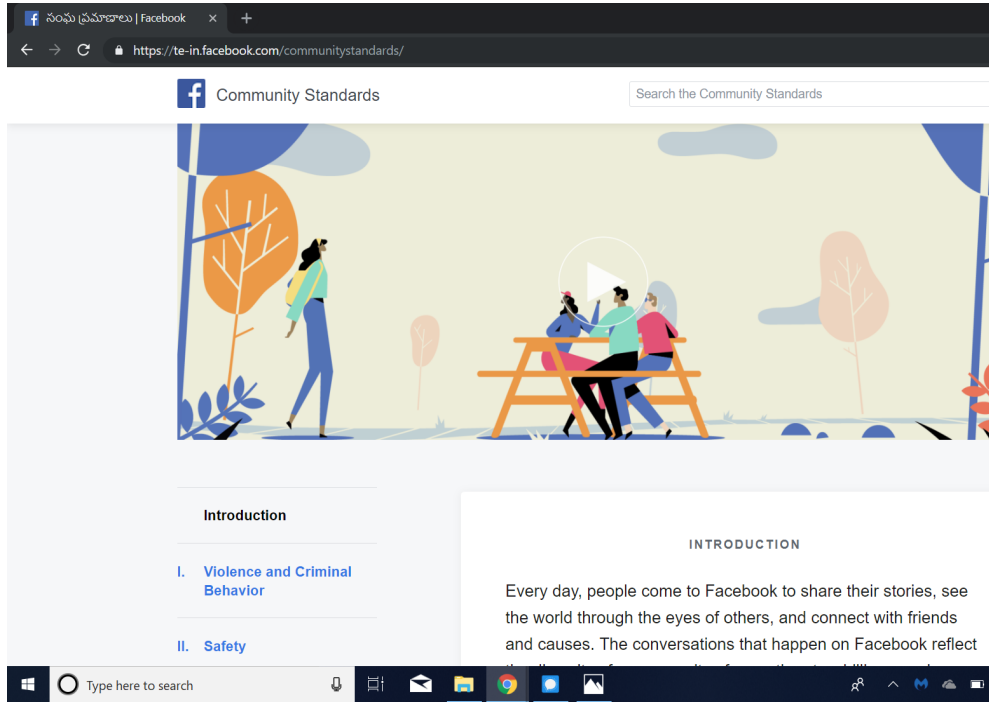
हम अभद्र भाषा को लोगों की सुरक्षित विशेषताओं — नस्ल, मूल राष्ट्रियता, धार्मिक मान्यता, यौन अभिविन्यास, जाति, लिंग या लैंगिक पहचान और बीमारी या अक्षमता के आधार पर उन पर सीधे आक्रमण के रूप में परिभाषित करते हैं। हम अप्रवासन स्थिति के लिए भी कुछ सुरक्षा उपाय प्रदान करते हैं। हम हमले को हिंसक या वहशी बनाने वाला भाषण, हीनता के कथन अथवा अपवर्जन या अलगाव की माँग के रूप में परिभाषित करते हैं। हम हमलों को नीचे बताए गए अनुसार गंभीरता के तीन हिस्सों में अलग-अलग वर्गीकृत करते हैं।

कभी कभी लोग ऐसी सामग्री साझा करते हैं, जिसमें किसी अलग व्यक्ति की अभद्र...

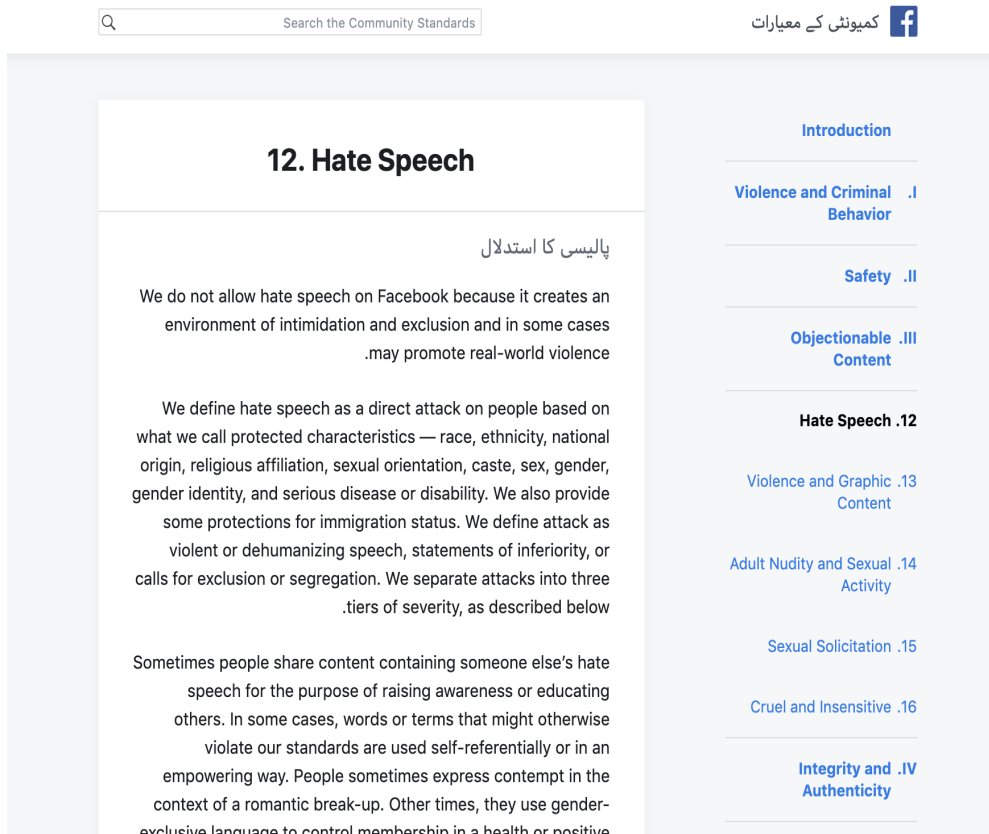
Type here to search

ENG 06:18 26-10-2018

Telugu: No translation offered, only English text but right-justified.



Urdu: No translation offered, only English text but right-justified.



A LACK OF CULTURAL COMPETENCY

In our advocacy, we found that Facebook staff often lacked the necessary cultural competency and literacy in the needs of caste, religious, and gender/queer minorities. This can be the result of many factors including, most critically, a lack of diversity in the staff and contractors of Facebook India.

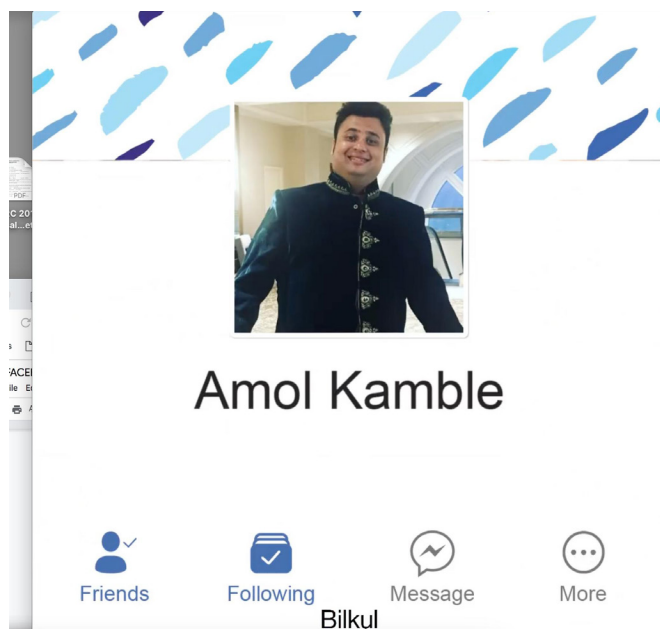
We want to emphasize that, in order to not further alienate and jeopardize already vulnerable Indian minorities, this competency is not optional.

One example where Facebook’s lack of cultural competency led to widespread dismay and outrage by caste-oppressed communities on the platform came in March 2019. Facebook deployed an advertisement for Marathi users to inform them about resources available to protect users against harassment and hate speech. The campaign backfired because of its employment of casteist tropes in its messaging. The troll in the scenario of the ad was a man named “Amol Kamble” — Kamble is a common last name associated with Dalit communities in Maharashtra.

Image of a female Facebook user being harassed by the ‘troll’, Amol Kamble



Profile of the name of the ‘troll’ with the recognizable Dalit last name



The resulting outrage led to hundreds of users reporting the ad as hate speech and many calls for a boycott of the platform. Caste-oppressed users left this episode feeling it was just the latest proof that Facebook simply does not care about either the representation of caste-oppressed communities or the violence and harassment these communities face on their platform. The ad was considered even more egregious in light of the consensus in Indian society that widespread harassment and disinformation are not a result of single individuals, but rather emerging from a well-organized political network of religious and casteist trolls.

This type of unintentional but reckless and significant harm is one of the main reasons we have repeatedly asked Facebook to directly engage with Indian Internet Freedom advocates and civil society organizations. We could help co-design more appropriate campaigns and training using our frontline experience. Our absence from the discussion table can only lead to more fumbles like this one when it comes to protecting Facebook India's most vulnerable users.

Failures of Reporting

Despite numerous requests by Indian minority activists for Facebook to address casteist and religious violence, there is still no actual consistent means of reporting the rampant casteist, Islamophobic, and other anti-minority hate speech and disinformation on the platform.

One essential step would be to establish clear, forward-facing reporting mechanisms. Equality Labs has been advocating for this change directly with the company for over a year, and we are concerned that there has been little or no movement on this vital plank of community management.

The existing process for reporting hate speech is as follows. Generally, each post has a small menu that is accessible at its right-hand corner. This menu allows users to report the post and, in some cases, to classify the type of hate speech. These types of forward-facing reporting mechanisms are often the only way users can report abuse, and they offer the platform an easy way to aggregate data on the types and frequency of hate speech reported in each market. Below is an example of one reporting workflow our users came across:

Workflow 1: Showing only Race, Sex, Orientation, and Ability.

Help Us Understand What's Happening ×

What's wrong with this post?

- It shows someone using drugs
Examples: people injecting or snorting drugs
- It harms or humiliates based on race, sex, orientation or ability
Examples: racism, sexism, anti-gay content
- It describes buying or selling drugs, guns or adult products
Examples: shotguns, marijuana, tobacco, adult items
- It shows someone harming themselves or planning to harm themselves
Examples: self-inflicted wound, weapon held to head, suicide note, promotes eating disorder
- I think it's an unauthorized use of my intellectual property
Examples: copyright or trademark infringements
- It's something else

Back
Continue

Each reporting mechanism offers more or less detail. It is hard to predict which reporting workflow you are about to see. This system can be very challenging to use, especially for first-time internet users. Given that the Indian market is now larger than the US market, it is especially puzzling that our protected minority categories are not clearly and distinctly listed for members of those communities to access when faced with hate speech, disinformation, and direct calls to violence.

When we started this work, we were shocked at the absence of caste as a protected category for the Indian market. Caste-oppressed communities represent at least 300 million people in India alone, and caste is not solely an Indian phenomenon. It affects millions across all the countries of South Asia, as well as the diasporas in the US and the global North. This is a major reason we demanded from the beginning that Facebook include caste as an explicitly protected category in its reporting mechanism.

Because of our continuous advocacy and monitoring, in late March 2019, we finally uncovered one sample workflow that offers a way to report casteist hate speech. However, this is still complicated by the fact that the user cannot predict which reporting screen they will receive when reporting posts. In fact, many members of our team of researchers have not been able to find this reporting screen again. So, while we would love to say that the inclusion of caste in the reporting workflow is progress, the fact that it is not consistently replicable/accessible by a team of over 10 researchers across multiple Indian languages is troubling and defeats the purpose of its inclusion. If technologically savvy and knowledgeable Dalits on our team cannot access this mechanism, what of the first time user who does not even know they have an option for action against casteist abuse?

Below are two other reporting workflows, including the new workflow that finally includes caste:

Workflow 2: This type allows for a clear listing of the protected category but does not mention caste.

Help Us Understand What's Happening ✕

What's going on?

- It's annoying or not interesting
- I think it shouldn't be on Facebook
- It's a false news story
- It's spam

[Continue](#)

Help Us Understand What's Happening
✕

What's wrong with this?

- It's rude, vulgar or uses bad language
- It's sexually explicit
- It's harassment or hate speech
- It's threatening, violent or suicidal
- Something else

Back
Continue


Help Us Understand What's Happening
✕

This is hateful towards

- A race or ethnicity
- A religious group
- A gender or orientation
- People with disability or disease
- An individual

Back
Continue

Workflow 3: This for the first time represents caste, but has not been replicated successfully by our team. This screen was first discovered in March 2019.

 We use your feedback to help us learn when something isn't right.

Nudity

Violence

Harassment

Suicide or self-injury

False news

Spam

Unauthorised sales

Hate speech

Terrorism

Incorrect voting information

🔍 Something else

Help us understand the problem.

Race or ethnicity

National origin

Religious affiliation

Social caste

Sexual orientation

Sex or gender identity

Disability or disease

Something else

We would recommend that Facebook cease in experimenting with workflows on vulnerable Indian minority communities. While we appreciate that software development practices require testing between user groups, hate speech is too grave a matter to allow for experimentation in real time. There is a well-established body of work documenting the history and ethics of experimentation on human subjects, particularly in the global South, raising both moral and legal questions about testing around such volatile issues in vulnerable communities.

This is the primary reason academic institutions have adopted rigorous standards for such research under the supervision of Institutional Review Boards (IRB) around the world. IRBs are administrative bodies established to protect the rights and welfare of human research subjects recruited to participate in research. We would ask Facebook whether its product teams conduct such cultural experiments under similarly rigorous protocols. If so, how can civil society participate to ensure experimentation on our communities does not lead to greater risk for us?

A more sensitive and reasonable approach would be to partner with Indian Internet Freedom and civil society partners to host co-design sessions addressing these problems, and then to roll out these solutions collectively with appropriate training.

The fact that Equality Labs was not informed about this rollout even though we were the first organization to advocate for the push for caste in Facebook's hate speech reporting mechanisms reveals how opaque their processes are with partners who could help with caste and religious competency they sorely need. Keeping civil society partners in the dark profoundly limits our ability to help train and educate vulnerable users and work collectively with Facebook to address this urgent problem. We urge Facebook to consider being more transparent and collaborative on these issues so we can all work to keep more users and their communities safe.

Inability to Track Hate Speech Disaggregated by Category

Even more concerning than the lack of consistency in reporting functions is the issue of how Facebook is tracking casteist and religiously discriminatory content. If these categories are not disaggregated in the reporting of hate speech and violence, how can Facebook track its effectiveness in assuring the safety of these communities on its platform?

This also leads us to our concern about the composition of the moderation team. We are concerned about their understanding of the nuances of how caste, gender, and religious stigma operate in Indian society. We have observed that, despite such posts being reported, the review committee often treats offensive content as mere humor without realizing how this humor rests on stigmas and violence against certain groups, namely caste, religious, and gender/queer minorities.

Hate Speech Analysis

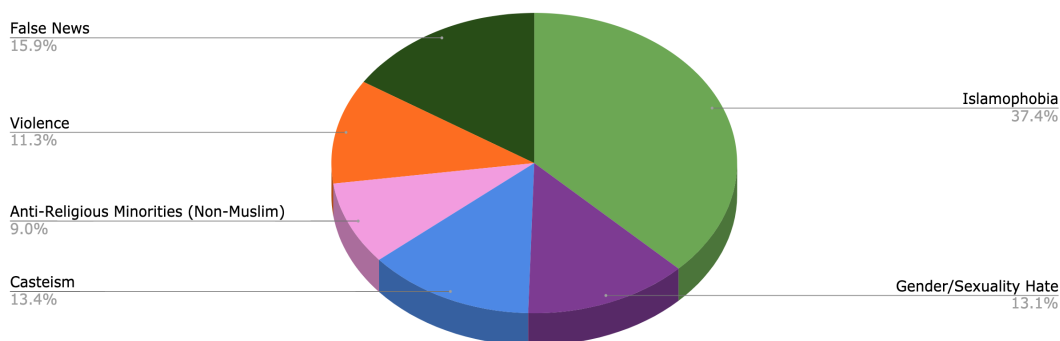
There are so many examples of the failures of content moderation on Facebook that our team had to strategize how best to start a productive conversation about the problem. However, we developed a methodology for reporting that allowed us to create one of the largest existing sample sets of hate speech on Facebook India, with help from multiple stakeholders in the field including the Next Billion Coalition, the Internet Freedom Festival, Media Matters, Access Now, and Dr. Joan Donovan, Director of the Technology and Social Change Research Project at the Harvard Kennedy School's Shorenstein Center. Over a four-month period, we gathered over 1,000 posts in six different languages including Hindi, English, Marathi, Telugu, Kannada, and Tamil.

We tracked their authorship, the reporting process, and Facebook's response to reporting. We disaggregated hate speech content into categories that were actually relevant to the Indian context as we analyzed content for both the level of hate and Facebook's response characteristics.

Our most important finding is that, without question, hate speech targeting Indian caste, religious, and gender/queer minorities is the most rampant across Facebook. Outside of direct slurs, we found a consistent set of disturbing content that includes discriminatory posts and clear escalation in calls for violence against minorities.

We broke down the instances of hate speech and disinformation into the following categories: Casteism, Islamophobia, Religious Minorities, Gender/Sexuality, Violence, and False News.

Frequency of Appearance of Different Types of Hate Speech



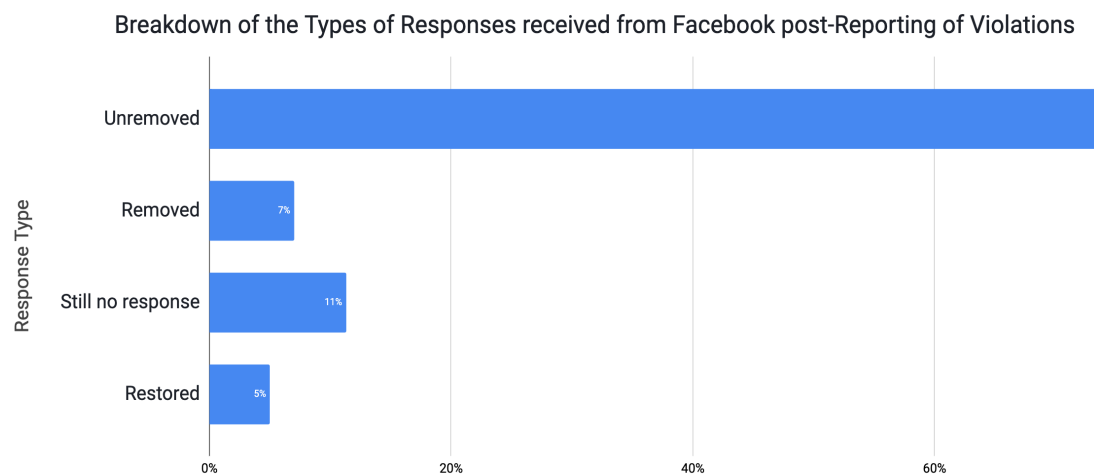
We originally included all hate speech against minority religions in one category, but Islamophobic hate speech was the largest single sample set in our report. We were compelled to re-categorize it separately for finer analysis.

Before proceeding, we would like to provide a severe content warning. Readers may find the examples presented here deeply disturbing. For our part, we have decided to present all of the content as-is, even the most graphic, to be able to provide undiminished evidence for the scope of violence Indian minority communities encounter on Facebook.

Below are some of the most crucial findings.

Slow Response Times to Hate Speech Reports

Facebook has stated in multiple reports that they review the great majority of the content reported to them in under 24 hrs. However, we found that the median time for Facebook India to respond to a report was, in fact, 48 hours.



Once a post was reported, Facebook sent us a response informing us either of removal of the offending post or of its continued presence on the site. It is useful to note that this did not include the 11% of reported posts that received no response from Facebook whatsoever — neither a response indicating deletion, nor one indicating a non-violation. Several posts were not responded to until almost a week after reporting.

Restoring Removed Posts

43% of all initially removed posts were found to have been RESTORED after an average period of 90 days from the date of reporting. This phenomenon lowers the actual rate of removal of reported posts from 12% to 6.8%. Alarmingly, 100% of all restored posts were Islamophobic posts.

It is unclear what guidelines Facebook India followed for the eventual restoration of removed posts. It is especially confusing because, according to our observations, these posts were filled with hate-mongering and calls to violence.

Below, we provide a sample of the kinds of posts that were initially removed, then restored.

Example Post 1: This post is about validating a common Islamophobic trope that the government is offering too much welfare to Muslim communities. It does this by showing a Muslim man, sitting as a child would, on Gandhi's lap and receiving a bag of money from him. Meanwhile, Gandhi is holding down another Hindu man under his foot and his walking stick.

The caption chillingly makes reference to Nathuram Godse, the Hindu nationalist who assassinated Gandhi, and implies that Godse's choice to murder Gandhi is vindicated by such welfare.



Example Post 2: This post repurposes a common white supremacist character, Pepe the Frog. Here he is depicted as a Hindu nationalist, standing in front of the infamous image of the 1992 demolition of the Babri Masjid in Ayodhya by Hindu nationalist mobs. Beneath the image is the sardonic caption “Good evening Jai Shri Ram,” which again refers back to communal violence that broke out when the Babri Masjid was destroyed to make way for a Ram temple. This incident was preceded and followed by extreme acts of violence against Indian Muslims.

To glorify a pogrom that led to the desecration of a place of worship and the murder and rape of innocent Muslims is a vile act of hate and a clear violation of Facebook standards. The fact that this image was returned to the platform is just shocking.



It is particularly alarming since Huffingtonpost India Editor-in-chief Aman Sethi noted¹ in a report about his recent visit to Facebook India’s New Delhi office that the slogan “Mandir yahi banega” or “Build a temple here” was written on the office entrance. This slogan has been continually used by Hindu nationalists to push for a temple on the site of this destroyed mosque. The same slogan is now used in communal encounters across the country. Google faced a similar issue when nationalists attempted to edit the name of the Babri Masjid’s contested site with the same slogan in Google Maps. Google apologized and removed it quickly².

If Google can act quickly and with sensitivity to this same issue, we ask why Facebook lacks the caste and religious competency required not to allow such a provocative and violent phrase to be prominently displayed in their headquarters. Further, how does this lack of competence contribute to their allowing content like this post on the platform?

1 https://www.huffingtonpost.in/entry/can-facebook-fix-fake-news-in-time-for-indian-elections-dont-hold-your-breath_in_5c98c340e4b057f7330d5fdb

2 <https://www.thehindu.com/news/national/other-states/mandir-yahi-banega-controversial-marker-appears-near-ram-janmabhoomi-site-on-google-maps/article25635160.ece>



If Google can act quickly and with sensitivity to this same issue, we ask why Facebook lacks the caste and religious competency required not to allow such a provocative and violent phrase to be prominently displayed in their headquarters. Further, how does this lack of competence contribute to their allowing content like this post on the platform?

Example Post 3: This post has been reported, removed, and restored several times. It features a quote from the current Chief Minister of Uttar Pradesh, Yogi Adityanath. The quote says, “Those people whose mothers took their pants (salwars) off after seeing swords in the hands of Mughals today proudly claim to be Muslim.” It then asks the reader if they agree.

There are so many levels of offensiveness to this post. First, it makes the assumption that all Muslim rulers in India used the power of their swords to rape women and force them to bear their offspring. Second, it also implies that Indian Muslims are a result of “weak” women who obliged these rulers. It makes the assumption that it is through rape and the resultant birth that most of the Muslims today in India profess this faith, and that otherwise they would still be Hindus.

Yogi Adityanath himself is no stranger to religious controversy. Before his ascension to the chief minister of Uttar Pradesh he ran Hindu nationalist militia called the Hindu Yuva Vahini. There are videos showing the Chief Minister sharing a stage with his colleagues addressing hundreds of Hindus in the audience when one of his leaders asks his followers to dig up Muslim women from their graves and rape them.

The post below is an example of how much historical and political context is required to understand whether a post qualifies as hate speech. While we understand that Facebook gives certain concessions to public figures, we must also ask whether Facebook helps enable public figures who normalize hate speech that in any other context would be Tier 1 hate speech. What are the guidelines? Does notoriety outweigh the need to keep the public safe?

Moderators must consider overlapping violations of religious and gendered hate speech as well. That this post has been reported, removed, and restored reflects the need for urgent dialogue on these issues so that there can be a consensus based on safety.

Hindu Ram Singh shared a post.
24 November 2018



जिनकी अम्मी ने मुगलो के हाथ में तलवार देख कर अपनी सलवार उतार दी थी। उन्हीं की नाज़ायज औलादे आज गर्व से कहती है कि -

"हम मुस्लमान है"

क्या आप योगी जी की बात से सहमत है?

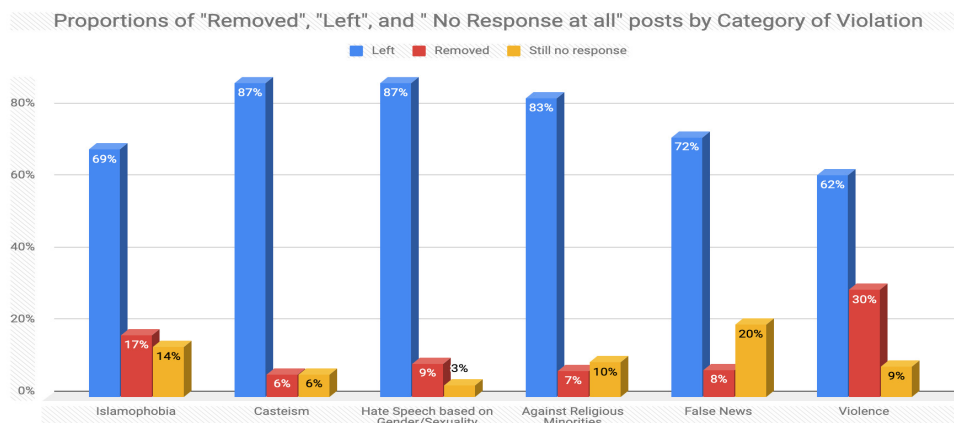
The Power of Hindu
24 January 2018

॥ ॐ ॥ सभी HINDU सगठन मे रहने के लिए पेज को जरूर LIKE करे अगर सच्चे HINDU है
॥ ॐ ॥

Removed, Left to Remain, or No Response at All: A Breakdown by Violation Type

The rate of removal for reported posts remains more or less comparable across categories. However, it is notable that the highest rate of removal was found in the "Violence" category and the lowest rate of removal in the "Casteism" category.

The highest rate of “No response at all” was found in the “False News” category. This is especially a cause for concern in this age of disinformation leading to real-life violence. Acting quickly and effectively to curb and debunk false news should have been a top priority. Yet while Facebook has been actively touting its response to fake news to journalists, it is clear that the reality is quite different.



We` have selected some samples of each type of violation to share. Our intention is not to require a post-by-post response from Facebook, but to demonstrate the pattern and intensity of violations these posts represent.

Islamophobia Dominates Indian Hate Speech

The vast majority of violations we observed and reported were Islamophobic in nature. Islamophobic content is foundational to the hate speech echo chamber in Facebook India, accounting for almost 40% of the content we reported. Islamophobic content is often the most violent, most threatening, and most gruesome. This is in alignment with the growing tide of violence against Indian Muslims with India Spend’s Hate Crime Watch reporting that Muslims were the victims in 58% cases of religiously motivated violence, between 2009 and 2019.³ This atmosphere of violence underscores the reality that hate speech in a cultural context that has already seen an exponential rise of violence is like adding gasoline to kindling. All that is needed is the spark, and as we have seen in Myanmar, social media can be that tragic spark.

Within the spectrum of Islamophobic content, there are some recurring themes.

Anti-Rohingya Posts

Almost 6% of all Islamophobic posts we examined were anti-Rohingya posts. These posts follow patterns of violation that are extremely similar to those seen in Myanmar just before and during the Rohingya genocide. This anti-Rohingya tone adopted and escalated by Rakhine state actors is in fact, faithfully replicated on Facebook India. Many of these messages involve dehumanizing Rohingya peoples with comparisons to cockroaches or insects, or to pigs, an animal considered impure in Islam. We are concerned because while Facebook assured the world “Never again” after its failure in Myanmar, it is clear that it is

3 <https://p.factchecker.in/>

again at a moral and political crossroads in complicity with this brand of hate speech.

Other Islamophobic posts target Rohingya Muslim refugees in India by spreading false news about these communities. Usually these stories claim that Rohingya refugees are swarming India; that they are swaying elections, obtaining false identification, or that they have married “local girls” who are giving birth to multiple children very quickly and swelling the population of Muslims in India. This narrative in particular, aggravates chronic Hindu nationalist anxieties in India.

It is a fascist narrative not unlike the ones used to create anxieties around Jews in Hitler’s Germany. It uses a pernicious combination of xenophobia, Islamophobia, and misinformation to target Rohingya Muslim refugees. This Anti-Rohingya rhetoric is then seamlessly coupled to existing hate against Indian Muslims.

Example Post 4: Anti-Rohingya post calling Rohingyas cockroaches needing to be “trampled.”




Example Post 5: This post is a call for boycott against Rohingya Muslims. It repeats the trope that they are illegal immigrants, radicalized, and terrorists. The open call to exclude them from jobs and places of residence is chilling.

November 11, 2018 · 🌐

Like Page

We all know what a threat the entry of Rohingya illegal immigrants is to our national security & demographics. This report sheds light on the facts. They aren't 'refugees' here, they are ILLEGAL IMMIGRANTS. Some of them are heavily radicalised & are Islamic terrorists. We can all do something in our individual strength to thwart this attempted demographic change.

NEVER hire 1 of them for any job that you're in control of. Don't hire them as maids, cooks, cleaners, sweeper, security, watchman, nanny etc. If you're staying in a housing society, get a resolution passed for the same. If you're on the hiring committee in your office, do the same. They're here for economic opportunities, if they don't get any, they'll disappear..



SWARAJYAMAG.COM

Ground Report: Rohingyas Are Refugees In Bangladesh; In India, They Are Just Illegal Immigrants Seeking A Better...

44 1 Comment 26 Shares

Example Post 6: This horrifying post features screenshots from a now debunked staged video that claimed Rohingyas were slaughtering and cannibalizing Hindus. The post translated from Hindi says, “This is Rohingyas Muslims. Watch how they are killing and eating Hindus and yet you keep crying about Petrol!”

The shocking nature of this post was not only that the video was staged^{4a} as reported by Boom live in 2018 but the post was disseminated both on Whatsapp and on Facebook in coordinated posts across both platforms. Additionally even when videos were taken down images of the videos were recirculated over and over again with some shares happening again during the 2019 Indian Lok Sabha elections. Shockingly if there had been no fact checks these alarming and inciting videos would have no counter on the platform as they linger on the platform a year after having been widely debunked. Clearly something is wrong with Facebook moderation when it comes to Rohingya centered hate speech and given the precarious conditions Rohingyas face in India and across South Asia this issue must be dealt with immediately.

^{4a} <https://www.boomlive.in/gory-staged-video-being-shared-as-rohingya-slaughtering-hindus/>



हिन्दू हेमन्त गौड़ बिजौलिया

May 27, 2018 · 🌐

ये हैं रोहिंग्या मुस्लिम देखो कैसे खा रहे हिन्दूओ को...ओर हिन्दू तुम पेट्रोल पे रोते रहो...और काँग्रेस कह रही है वो सत्ता में आई तो इनको पक्के मकान बना कर देगी।

[See Translation](#)



👍👎🙄 55

11 Shares



Love Jihad

5% of all Islamophobic posts we examined relate to the Hindu nationalist conspiracy theory of “Love Jihad.” “Love Jihad” posts repeatedly emphasize the dangerous and fictitious idea of a Muslim plot to target and entrap Hindu girls into marriage with Muslim men with the intention to defile their (Hindu) honor and grow the Muslim community. Often there is a connected idea that they are using these marriages to recruit for ISIS.

Much work has been devoted to debunking “Love Jihad” but the conspiracy persists, abetted in part by a large Love Jihad disinformation ecosystem on Facebook and other social media platforms. Meanwhile, Muslim men in the real world face the consequences of being portrayed as bogeymen in such interfaith relationships and, in many cases, are subjected to extreme violence as a result. Ultimately the fear of “Love Jihad” is an externalization of an anxiety against interfaith relationships and the supposed assault on family, religious, and caste honor that these relationships represent.

Example Post 7: The post below is an example of how this disinformation is spread on Facebook and surprisingly allowed to remain despite the ramifications of violence that these posts caused for their target audience. It explains how a marriage between a Muslim man and a Hindu girl is actually a trap laid by the Muslim man to ensure that five new individuals (his wife whom he would successfully fool into converting and their resultant four children, a number randomly assigned to Muslim family size) would become Muslims, thereby contributing to the population growth of their community. This growth then is a threat to the Hindus whose numbers would dwindle in comparison.

This propaganda often spills over into violence offline as was seen in December 2017 in Rajasthan⁴, when a man named Shambhu Lal murdered a 45-year-old Muslim labourer, Mohammad Afrazul, by hacking him to death. He recorded the murder in a chilling video in which he proudly proclaimed, “This is what will happen to you if you do Love Jihad in our country.” Internet access had to be suspended temporarily in the area after he posted two of these videos on WhatsApp.⁵

4b <https://www.bbc.com/news/world-asia-india-42261797>

5 <https://www.ndtv.com/india-news/chilling-murder-in-rajasthans-rajsamand-on-video-man-hacks-labourer-burns-him-1784669>

October 4, 2018 · 🌐

LOVE JIHAD IS REAL!!

When a muslim man and non-muslim woman start dating, in most of the cases its very much observable that the muslim man's agenda is to make the girl fall in love with Islam more than him, It starts with little gestures like gifting a Quran to his girlfriend, then eventually graduating to explaining Islam in appealing(zakir naik) ways, telling false stories of Muhammad's greatness, Islam's role in empowering and liberating women, wrapped lollipop example(in case of very dumb girlfriend 😊)

All this brainwashing is done in order to make things easier for a smoother transition when it comes to proposing marriage, the girl is mentally a muslim before she even reads the shahda, because a muslim man very well knows its compulsory to convert a non-muslim woman before marrying her and as it is, converting a non-muslim has extra reward points in Islam.

A non-muslim woman who was born in a more liberal atmosphere, with no big time taboos regarding such issues will take the bait easily, and no serious issue of apostasy in her former faith, makes thing easier for the muslim man.


SIMPLE MATH:

One such marriage subtracts a member from a non-islamic faith and add it to Islamic faith, the kids who will be born to such couple will end up inheriting the father's faith, and on average, a muslim couple has 4 kids globally(approx), so basically one such marriage takes 1 member away from Non-Islamic faith and adds 5 to Islam.

How do you think Islam grew faster than all the other religions??? 😊

Because some morons never pay attention, simple!!

#LoveJihad #Islamophobia #Islam #Marriage



👍👎😊 104

9 Comments 51 Shares

Example Post 8: This post directed at Hindu girls says, "This is not love dear sister, they pretend to love you and convert you and use you for terrorist activities. The girl called Deepa Cherian became a terrorist and went to jail through this kind of love jihad." It ends with an advertisement for a "Hindu Helpline."

love jihad
In Time the Photos

ഇത് പ്രണയമല്ല പെങ്ങളെ
Love Jihad

ഇത് ശാന്തിയുടെ മതം

143 786

മതഭ്രാന്തിന്റെയും, തീവ്രവാദത്തിന്റെയും പുതിയ രൂപം ഇവർ പ്രണയം നടിച്ച് നിങ്ങളെ മതം മാറ്റുകയും, തീവ്രവാദ പ്രവർത്തനങ്ങൾക്ക് ഉപയോഗിക്കുകയും ചെയ്യുന്നു.

ദീപാചെറിയാൻ എന്ന ഷാഹിന 'ലൗ ജിഹാദി'ലൂടെയാണ് തീവ്രവാദി ആയതും ഇപ്പോൾ ജയിലിൽ കിടക്കുന്നതും

Call Hindu Helpline
9497545511

Like Comment Share Tag Photo Options Send in Messenger

Glorification of Violence by Alluding to Previous Incidents of Religious Violence Against Muslims in Ayodhya and Gujarat

10% of all Islamophobic posts are about the anti-Muslim attack in Ayodhya. In 1992, Hindu nationalist mobs converged on a mosque in the town of Ayodhya, purported to be the birthplace of the Hindu God Lord Ram, and destroyed the mosque. Since that time, there have been repeated calls from the nationalists to “re-build a temple to Lord Ram on the exact same spot.” Over the past quarter-century, this issue has torn the nation apart along religious lines. These posts aim to do the same: divide people along religious lines and instigate violence against Muslims and other caste-oppressed communities.

Example Post 9: In this post the image of bullets is accompanied by this alarming caption: “Either by words or guns, no matter who comes in between, the temple will be built in Ayodhya where Lord Ram was born, and it will be built by the Bajrang Dal of Hindustan (ed.note: India is being identified here as the Hindu homeland). My workers and my office bearers and all my respected members are all ready, we will not listen to the courts. You can build your taj mahal where you want we will build our Ram Temple Jai Shri Ram.”



Example Post 10:

The meme below was uploaded on March 29, 2019, three weeks after Babu Bajrangji, referred to here as the “Real Gangster” was granted bail by the Supreme Court of India on medical grounds. Babu Bajrangji was convicted and sentenced to life imprisonment for his role in the communal riots of Gujarat and for his role in the violence at the Naroda Patiya locality of Ahmedabad where 97 mostly Muslim residents were killed, one day after the Sabarmati express train (also in the meme), was burnt in Godhra. This was an incident that triggered riots in Gujarat, where thousands of people were killed. Babu Bajrangji has also openly bragged about how he slit open the stomach of a pregnant Muslim woman, pulled her fetus out and paraded it on a trishul (a kind of sharp weapon mostly depicted in Hindu

rituals). To glorify him as a “real gangster” is to glorify extreme violence.

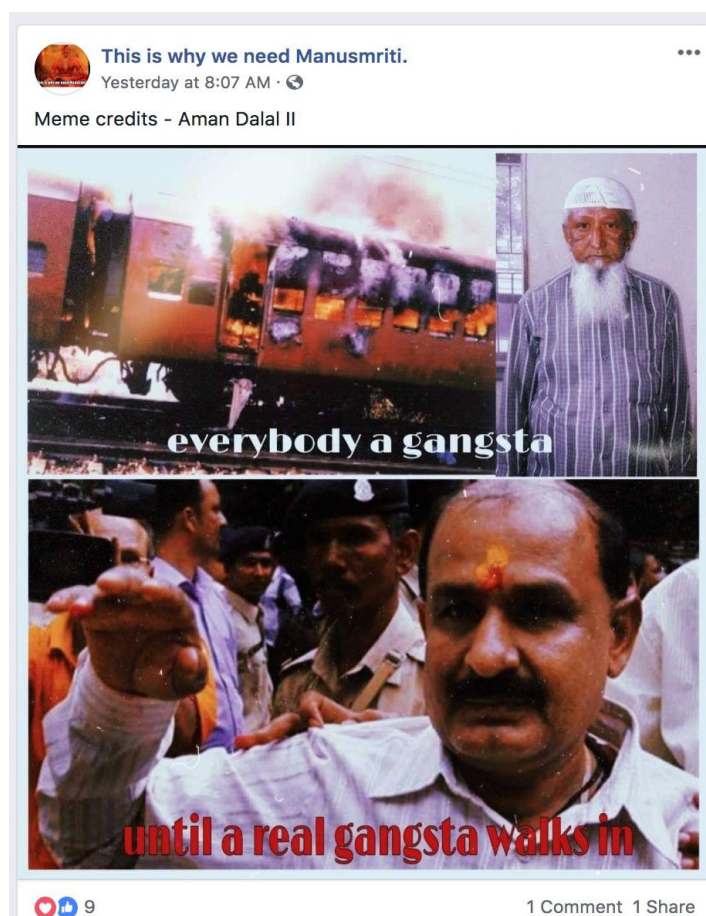
It is also significant that this page which shared the post is called “This is why we need Manusmriti” and provides a link: The Manusmriti refers to the Hindu scripture that established caste prohibitions and the enslavement of caste oppressed communities. It is a text that is disturbing to many communities and in fact a growing number of Hindus now disavow the text. However, this does not stop Facebook from allowing a proliferation of groups aligned with the Manusmriti, whose primary purpose is to issue violent casteist memes against caste oppressed communities.

The link to this page is <https://www.facebook.com/MlecchasBTFO/>.

Mleccha⁶ is a Sanskrit word meaning “non-Vedic,” or “**barbarian**,” derived from how Indo-Aryans described the language of other communities in the region. It is used in memes as a slur to target caste oppressed and religious minority communities.

They are also using the term “BTFO,” an acronym for “Blown The Fuck Out” because the original page was removed and they have created this second back-up page for their users. This is another behavior that moderators for safety must watch.

The complexities of this post reflect again how much historical awareness of Indian history and communal tensions is required to moderate in this space. Having seen many posts like this one, we would like to know what training Facebook India offers its moderators to address this cultural complexity and history.



Islamophobic Slurs and Profanities

Most of the rest of the Islamophobic posts focus on obscenities. These include words like “Terrorist,” “chopped dicks,” “Sulla,” “Tuluk,” “pigs,” “cockroaches,” and “chuslim,” all of which are some of the common slurs directed at Muslims in the Indian context.

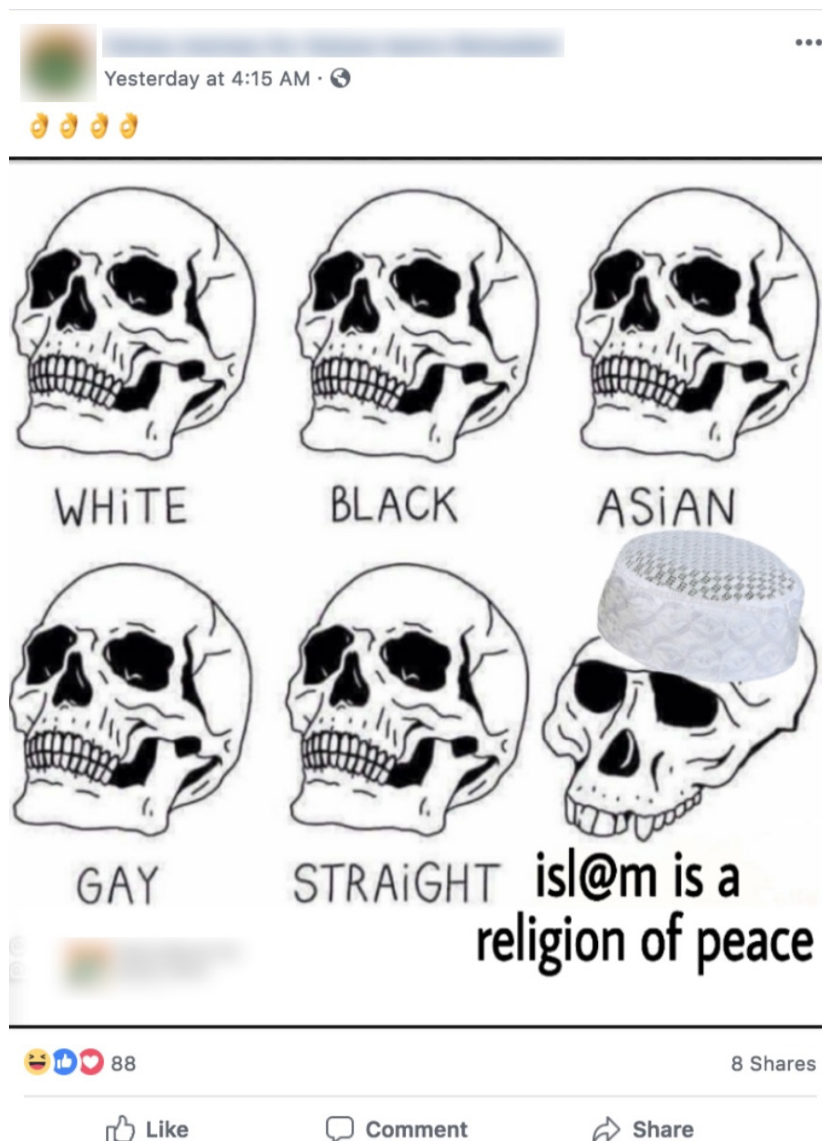
Example Post 11: Here Siddhu, a Sikh public personality, is accused of being a “Muslim.” Terms like “chopped dick” are created to equate Muslims with their religious practice of circumcision.

It is pertinent to note that during religious riots and massacres in India, Muslim men have been asked to remove their clothing to be checked for circumcised penises. This violent obsession with Muslim men’s genitalia is replicated online and should be deemed absolutely unacceptable.



Example Post 12: Here a primitive hominid or monkey’s skull is used to represent a Muslim. This is offensive, dehumanizing and racist. It represents nearly everything about hate speech that Facebook says it discourages on its platform.

Also significant is that casteist and religious trolls on Facebook India know how to dodge filters designed to identify obvious hate speech. So in this post Islam is spelled Isl@m. We have seen many instances of this gimmick in our research which is reminder that AI cannot replace intentional human centered moderation.



Casteism

Casteist posts are another area of serious concern. Caste discrimination is pervasive throughout South Asia, and Indian casteist hate speech is part of an ecosystem of violence designed to shame, intimidate, and keep caste oppressed communities from asserting their rights and participating as equals in society.

Some posts are overtly against Dalit Castes:

Example Post 13: This is from a group called Anti-Chamaar Group. Chamaar is a Dalit caste found throughout North India who work with leather. Both the group and the image are explicitly against this caste. The existence of this group was named as early as 2016 by Soibal Dasgupta⁷ and yet the group continues to exist despite numerous attempts to report it to Facebook.



We additionally found in our research that 40% of all casteist posts were anti-reservation posts. Reservation is a policy in India that can be compared to Affirmative Action in the United States. It was put into place as a form of reparations required to address the disparity of access that Dalits and Adivasis peoples face in Indian society resulting from centuries of oppression. It provides for quotas in government education and employment positions for people from historically oppressed communities.

The main form of hate for Dalits and Adivasi people that we observed came through rhetoric of anti-reservationism. Additional casteist posts included caste-based slurs, derogatory references to caste-based occupations such as manual scavenging, anti-Ambedkar posts (such as photoshopping Ambedkar's face onto memes as an echo of real-world vandalism), and anti-inter-caste love unions posts.

⁷ <https://indiaresists.com/casteist-facebook-facebook-removing-posts-selectively/>

Example Post 14: Note that this post opens with a racist term which is a variation of the N-word. Authors have photoshopped a picture of Dr. B.R. Ambedkar, a Dalit anti-caste liberation leader whose stature is similar to that of Dr. Martin Luther King Jr. in America. The caption reads, “**When you clean all the toilets and then fuck up the education system with reservation.**”



To understand why this post is so deeply offensive, one must understand the work conditions of Dalits. These castes have historically been forced into forms of slavery requiring them to do the filthiest jobs, like handling dead bodies and cleaning toilets. Even in modern-day India, this includes the practice of manual scavenging, where Dalit workers clean out the sewage from sewers with their bare hands and bodies. Thousands of Dalits die every year as a result of inhaling sewer gases or by simply drowning in sewage. This dreadful exploitation of a caste-oppressed people is made the brunt of jokes that target Dalits including the one in this caption.

The social context is also important here. Statues of Dr. Ambedkar are regularly vandalized in India to intimidate and shame caste oppressed communities. Many cities and Dalit neighborhoods have to protect their statues of Dr. Ambedkar with cages to prevent “upper-caste” vandalism. Such vandalism includes statues being knocked down, splashed with red paint to look like blood, garlanded with slippers, or dirtied with cow dung.

This vandalism is echoed online. Spaces like Wikipedia see constant vandalism of the page on Dr. Ambedkar, where trolls descend to insert caste slurs into the text. On Facebook, the

act of photoshopping Dr. Ambedkar onto memes is a well-known extension of this practice.

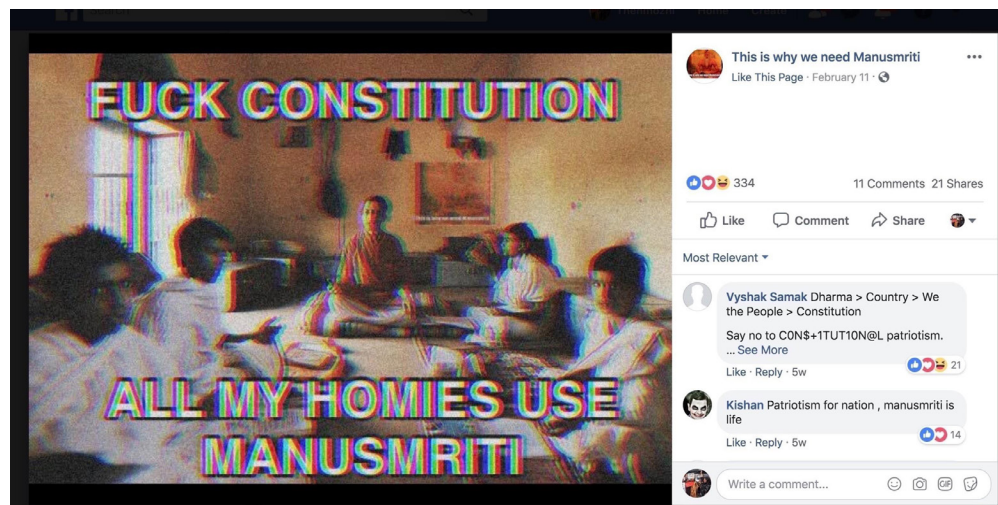
Real-life vandalism can be punished under the SC/ST Prevention of Atrocities Act, which provides special protections to Dalit and tribal peoples. In online spaces like Facebook however, such offenses put caste oppressed communities at the mercy of corporations who sorely lack the legally required caste competency to understand, much less remove, such inflammatory material.

Caste Slurs and Derogatory Occupational References

Example Post 15: This post is both anti-reservationist and makes a derogatory reference to the caste-based occupation of cleaning sewage. It implies that Dalits are good for only jobs involving sewers and not fit to compete with “savarnas” (upper caste people) for jobs with dignity. This page was made by yet another group referencing the Manusmriti called Make Manusmriti Constitution again.



Example Post 16: While this post is not hate speech per Facebook’s community standards, it is a post that is incredibly distressing for many because of its position asserting that Hindu scripture is above the constitution of India — and not just any scripture, but one that enforces caste oppression and violence. Again, we would like to know how Facebook India addresses instances like this. How can they respond to the normalization of what would have been seen as controversial even two or three years ago?



Example Post 17: In this post, another meme of Dr. Ambedkar shows him jumping off a bridge with the epithet “Bhangi,” making a pun on “bungee jumping.” Terms like “Bhangi” are extremely derogatory and comparable to the N-word in Indian usage. While Bhangi is the name of caste that does manual scavenging, or the cleaning of sewers, it is used as a slur in upper caste communities, as we see here. The laughing emoticon and the further caption distorting the term “Dalit” as “Da-lit” are evolved mechanisms intended to evade tracking of anti-Dalit hate speech and content.



Example Post 18: Below we see a deeply offensive manipulated image. The face of Dalit leader Behenji Mayawati has been photoshopped onto the body of a Ms. India beauty pageant winner. The caption below again references Indian affirmative action policies saying that in 2030 there will be Reservation (affirmative action) for beauty Pageants for Dalits and Adivasis (SC/ST)

Beyond the repetition and disregard of affirmative action, this image is meant to further mock and demean Mayawati with the underlying insinuation that she is unattractive. The idea is to trivialize Dalit people and Dalit women in particular, who are often stereotyped as dark, ugly, like demons, witches, unattractive, and so on.



Like This Page · November 9, 2018 ·

180 10 Comments 54 Shares

Like Comment Share

Most Relevant

Like · Reply · 8w

Hahahaaha

Like · Reply · 16w

Write a comment...

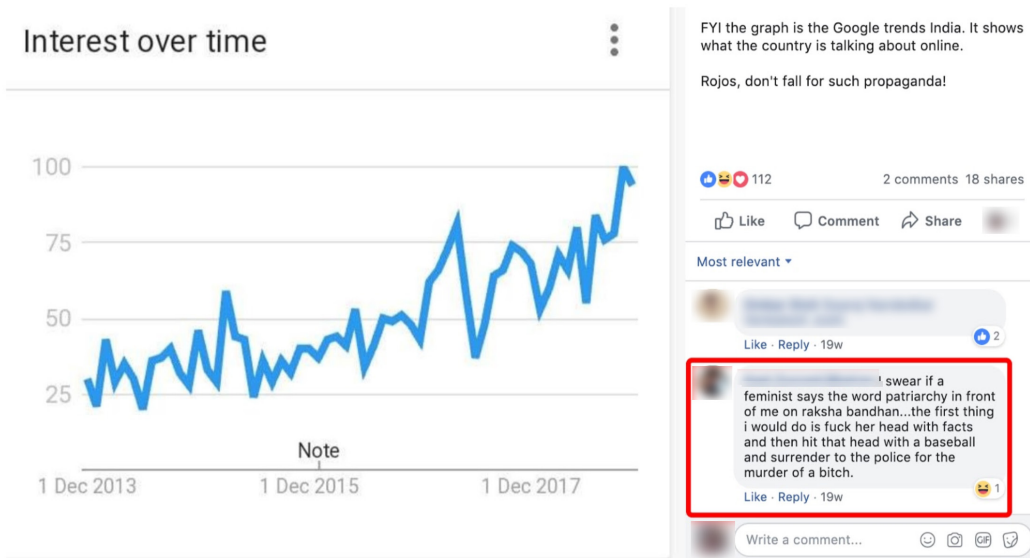
Hate Speech Based on Gender or Sexuality

While a vast majority of posts that were classified under Hate Speech Based on Gender or Sexuality displayed outright misogyny, over 25% of posts in this category were transphobic or queerphobic in content. 12% of the posts in this category made direct reference to violent rape, either calling for rape or glorifying or trivializing rape.

Example Post 19: This post is hosted by a casteist meme group called Normien't Memes for Chamaar Teens. The use of "Chamaar" is used as a slur here and is meant to provoke Dalit communities. The meme itself is a call for domestic violence implying that a baseball bat is an education tool for wives.



Example Post 20: Violent misogynist threats such as this one are unfortunately very common, not only as posts but also in the related comments section where they can go unchecked. Few users report comments but the misogyny here is clear. The comment says, **“If a feminist says the word patriarchy in front of me on raksha bandhan..the first thing I would do is fuck her head with facts and then hit that head with a baseball and surrender to the police for the murder of a bitch.”**



Example Post 21: Homophobic and transphobic references are often violent and frequently imply that a particular man is not “man enough,” or threaten anal rape. This post portrays Owaisi, a popular Indian Muslim politician, saying that he has got his “ass fucked” five times, comparing the act to the five prayers or namaz/salah that Muslims read during a day. The picture below it reads “Your prayer or Namaz has been accepted by God.” There is little more to say about this content other than that it should not be allowed on any platform.



Example Post 22: The text for this post below reads: “All those who are anti-Hindutva on Facebook, please put this dessert in your mouth and taste it.”



Example Post 23: The text reads: “Trans people are not those who wear a saree and roam around in markets. Trans people are those who despite being Hindus, oppose Hindutva.” This post is using trans-ness as an insult to those who oppose Hindu nationalism, or Hindutva.



Against Religious Minorities

Our findings show that 60% of all posts in this category were anti-Christian and 25% were anti-Buddhist. The rest were a mix of anti-Sikh and others. Note that Islamophobic content was so prevalent that we tracked it as its own category.

Anti-Christian

India's 29 million Christians constitute only 2.3 percent of the country's population, making them the second largest religious minority in the country after Muslims. On Facebook many posts depict Christianity as a religion that has come from outside of the country, but Christianity has been on the subcontinent as early as 2 C.E. with the establishment of the the Syrian church in Kerala, long predating the English arrival in India.

Despite that history, there is an insidious attempt by Hindu nationalists to depict Indian Christians as foreigners. This is exemplified by the the frequency of “No conversion” memes that showcawse the harassment of Indian Christians at the hands of Hindu Extremists. These posts often depict verbal and physical bullying, burning of Bibles, and exhortations to violence.

Example Post 24: This post is one example of a common anti-Christian meme. The post features an image of four Christians arrested for spreading their religion. The text accompanying this image in the post appears non-defamatory: “Two priests and two nuns arrested in Gorakhpur for religious conversion acts. All belong to the Church of Kerala.”



However the totality of the post is meant to incite violence as the whole ecosystem of the post is the mechanism for how hate speech is delivered, in this case tempting commenters for hate speech. Some users make horrific calls to violence like, “**Shave their heads, blacken their face, garland these pigs with old sleepers, strip them, get them paraded and make them apologize.**”

Another says, “If they are let off with a slight slap on the wrist. They’ll keep repeating these malicious acts. **Better is to kill them in encounter.** They are the unwanted scum of the society and are pollution in the country.”

“Encounter killings” describe killings by the police or the armed forces. The usage here as a call for state violence against Christians is utterly unacceptable.

Statements like these don’t occur in isolation. Indian Christians are facing severe levels of religious atrocity. In 2008 the Kandhamal riots saw Hindu mobs set fire to Christian settlements. At least 59 people were killed by arson with more than 50,000 left homeless. and thousands of houses and churches destroyed. Since then anti-Christian violence has become a part of Indian life. Christian pastors are routinely murdered, nuns are raped, Bibles are burned, and churches set on fire.

In 2017, the Evangelical Fellowship of India documented [351 instances](#) of violence against Christians, but activists and scholars believe that this is only a fraction of the actual violence as many cases go unreported.⁸

Example Post 25: Anti-Christian posts like this one are common and missionaries are often at the receiving end of this kind of rhetoric. This leads to harassment and mob violence against “visibly Christian” people, especially nuns and priests in their formal attire.



8 <https://thediplomat.com/2018/08/violence-against-christians-in-india-a-decade-after-kandhamal/>

Example Post 26: This post is a video of a Christian family in Tamil Nadu, whose home was invaded by a group of Hindu men off the streets. The men warn the family not to practice their religion or bear “consequences.” The female head of the family tries to negotiate peacefully with the intimidators while men who invaded her house threaten her and her family. These videos were widely shared and readers were encouraged to replicate such acts on their own.

Such videos provoke mob violence against minorities, as happened in 2013 when a fake video posted on Facebook and other platforms sparked communal riots in Muzaffarnagar, Uttar Pradesh, killing hundreds of Muslims.⁹ The fact that Facebook already had a tragic example of how content on its platform led to communal violence as early as 2013 should have been a red flag warning to its teams to invest proactively in cultural competence on caste and religious minority issues. Our team is unable to find a rational explanation in such cases for the disparity between obvious safety and risk scenarios and a complete lack of response by Facebook India.



False News and Disinformation

Across the Facebook product ecosystem, we are seeing an alarming number of fake news links and posts. These posts look like they originate from legitimate outlets, but they use manufactured screen grabs to incite violent dialogue around minority communities and issues. The sophistication of these posts can make it difficult for the user to identify the misinformation, and many users believe the incendiary content of these posts. Further, these posts are amplified in the Facebook algorithm by a sophisticated system of inter-connected pages. These pages have audiences in the millions and their traffic and interactions further lift the post in Facebook’s algorithms. The virality of the posts adds to the perceived credibility of the dangerous and violent content.

9 <http://archive.indianexpress.com/news/mla-who--shared-video--hops-parties-delights-in-guards/1168617/0>

Example Post 27: Pages like [Rightlog.in](https://www.rightlog.in) routinely present news-like articles that provoke emotional responses from readers based on religion. Their posts are widely shared and the page is followed by many users. Here we see a piece of false news about Barkha Dutt, a celebrated Indian journalist who is often under attack online. Hindu nationalist trolls express hate against her claiming she inspired the deeply misogynist and anti-free press term “presstitute,” an odious portmanteau of “press” and “prostitute.”

The comment above the post says, “Are you not seeing too many problems with Hindus these days? You have stooped so low in pleasing Mullas (ed note: referring to Muslims) that you want to destroy temples? Imagine had you said all this about Muslims or Christians **your body would have been found separately in a suitcase.**”

Referring to her as “Burkha,” after the garment worn by some Muslim women, it threatens consequences that she may face if she spreads the same “vile” against Muslims which, according to the commenter, she is spreading against Hindus in order to please Muslims. In invoking that violence it is a veiled threat against her to intimidate her from reporting.



6 hrs · 🌐

वाह रे दुरखा, हिन्दुओं में तुझे कुछ ज्यादा कमी नजर नहीं आ रही है, मुल्लों की चमचागिरी में इतना गिर गई कि अब तू मंदिरों को तबाह करने पर तुल गई? सोच अगर यह सब बकवास मुल्लों या इसाइयों के बारे में करती तो तो तेरी बाँडी अलग पीस में किसी सूटकेस में पड़ा मिलता.

[See Translation](#)



RIGHTLOG.IN

"Hindus are not victims, they are hypocrites and law breakers" Barkha Dutt's shocking new commentary

👍👎 25

12 Comments 46 Shares

Example Post 28: While 50% of posts classified as Fake News attempted to spread misinformation about a current event, it is worthwhile to note that over 20% of the posts in this category were about spreading a piece of falsified South Asian history.

This is an extension of offline attempts to edit textbooks across India to insert a Hindu nationalist version of events. Here we see depictions of seals from the Indus Valley Civilization age. They appear to depict people crossing sticks or swords in some sort of ritual. This is immediately identified with the modern day “Dandiya,” a Gujarati Hindu dance celebration. This is held up as evidence of the unproven idea that the Indus Valley Civilization was a “proto-Hindu” civilization. This is long-held Hindu nationalist propaganda, in support of the idea that Hindus have a legitimate and unbroken claim to the lands of the subcontinent.



During the Dandiya Raas dance, men and women wearing vibrant costumes dance around the Garbo, holding colorful sticks. They move in circles and rhythmically strike their own sticks and those of the dancers next to them. The dancers generally form two concentric circles, one revolving in a clockwise direction and the other anticlockwise. They are accompanied by a drum-player who provides the beats. In form and spirit, the Dandiya Raas is very similar to the Salla dance of Chhattisgarh.

The image on the Kalibangan seal can, therefore, also be interpreted as the performance of the Dandiya Raas with spears in honor of the tiger-riding goddess Amba Mata, during the Navratri festival.

This is evident that Hinduism was practiced during Indus Valley Civilization, 5000 years ago.

Example Post 29: Contemporary disinformation campaigns focus on creating violence and instilling fear among minorities. Here, a direct call for organized violence is made based on a piece of fake news, again about Rohingyas. The post says, “The report of Intelligence agencies in the Indian Railways has issued an alert that a large number of Rohingya trains are going to Kerala. **A request to all my brothers, if you meet any Rohingya in the train, throw him down.**”

The image shows a screenshot of a Facebook post. At the top, there is a profile picture and a timestamp: "October 4 at 11:02 PM". The post text is in Hindi: "सुफिया एजेंसियों की रिपोर्ट भारतीय रेलवे ने अलर्ट जारी किया है कि बहुत बड़ी संख्या में रोहिंग्या ट्रेनों से केरल जा रहे हैं, मेरा सभी भाइयों से निवेदन अगर ट्रेन में कोई रोहिंग्या मिलेता उसे नीचे फेंक दें". Below this is an English translation: "The report of intelligence agencies in the Indian Railways has issued an alert that a large number of Rohingya trains are going to Kerala, request to all my brothers if any rohingya meet in the train, throw it down." There is a "Rate this translation" link. Below the text are icons for Like, Comment, and Share, with a count of 89 likes and 6 comments/8 shares. Below the main post, there are three comments. The first comment says "Mera b" with a red heart icon and a '1' in a circle. The second comment says "Yes.Jai.Shri.Ram" with a red heart icon and a '1' in a circle. The third comment says "रवियार सुबह का जय श्री राम". Each comment has "Like · Reply · See Translation · 3w" and a "replied · 1 Reply" link.

October 4 at 11:02 PM · 🌐

सुफिया एजेंसियों की रिपोर्ट भारतीय रेलवे ने अलर्ट जारी किया है कि बहुत बड़ी संख्या में रोहिंग्या ट्रेनों से केरल जा रहे हैं, मेरा सभी भाइयों से निवेदन अगर ट्रेन में कोई रोहिंग्या मिलेता उसे नीचे फेंक दें

The report of intelligence agencies in the Indian Railways has issued an alert that a large number of Rohingya trains are going to Kerala, request to all my brothers if any rohingya meet in the train, throw it down.

⚙️ · Rate this translation

👍❤️👍 89 6 Comments 8 Shares

👍 Like 🗨 Comment ➦ Share

View 3 more comments

👤 Mera b 1
Like · Reply · See Translation · 3w
↳ replied · 1 Reply

👤 Yes.Jai.Shri.Ram 1
Like · Reply · See Translation · 3w
↳ replied · 1 Reply

👤 रवियार सुबह का जय श्री राम
Like · Reply · See Translation · 3w

Inciting Violence

Example Post 30: Violent posts are common and are often not removed promptly. Such posts not only incite physical violence but cause deep emotional trauma to minority communities, and should not be acceptable anywhere.

In this post the author is saying that red hot chilli powder should be **smear**ed into the vaginas of Muslims (the followers of Allah).

15 October at 18:48 · 🌐

अल्लाह को मानने वालों की मां की
चूत में M D H का टीका लाल मिर्ची
पाउडर

****Translation in English:
MDH hot red chilli powder in the
vaginas of the mothers of the
believers of Allah**

👍👎❤️ 69 17 Comments 2 shares

👍 Like 💬 Comment ➦ Share

👤 Jai shree ram

Example Post 31: The post reads, “If this time during Garbha [a regional Gujarati dance], any pigs [a slur referring to Muslims] try to enter, we will cleanse him with holy water from the Ganges. I am in communication with every state.”



See Translation

👍❤️😱 2.2K

207 comments 186 shares

False Reporting Leading to Banning of News Outlets, Journalists, and Minority Community pages

Another worrisome trend has been the growing number of accounts and pages that have been banned and removed from the platform. Many of these pages are banned without notice. News outlets banned in this way face a sudden loss of their audience, as Facebook is the primary place where they distribute their content. What concerns our team is how rapidly and how frequently these pages run by Indian minority people’s groups are targeted and removed from the website, while actual hate speech remains.

Caravan Daily wrote that in a ten day period Facebook disabled personal accounts of more than a dozen leading journalists in India without any warning or notice and without offering any credible reasons.¹⁰ This includes editors of BoltaHindustan.com, Mumtaz Alam, Editor

¹⁰ <http://caravandaily.com/mr-zuckerberg-whys-facebook-disabling-accounts-of-independent-journalists-in-india/>

of Caravan Daily, Syed Ghazanfar Abbas, National Correspondent of Caravan Daily. Ajay Prakash (News Editor, Dainik Bhaskar), Prerna Negi (Editor, Janjwar.com), Rifat Jawaid (Editor, JanataKaRepoter.com and former Editor of BBC) and Aijaz Zaka Syed, an award-winning Indian journalist and columnist based in the Gulf and former Opinion editor, Khaleej Times. Syed has also been associated with Caravan Daily as a columnist.

This also includes bans against prominent Dalit Bahujan news outlets and pages like National Dastak and Ambedkar's Caravan. The National Dastak Facebook page has 406,000 followers, making it one of the most followed Facebook pages raising issues related to Dalit-Bahujans and minority communities. Ambedkar's Caravan, with 220,00 followers, fills a similar niche and its owner no longer even pretends these bans are not casteist and openly challenges "the Manusmriti rules of Facebook." These bans have led Velivada another Dalit outlet to ask, "Who is afraid of Digital Dalits?"¹¹

Also significant is that numerous pages related to Kashmir by Kashmiri activists, journalists, and press outlets with the outlet Freepress Kashmir were taken down from Facebook for over 48 hours.¹² Even journalists who mention Kashmir face censorship.

This practice is so widespread on the platform that it merits its own report to further examine how such bans chill free speech in an increasingly polarized space.

Ironically one journalist wrote about his experience on Twitter since he had now been removed from Facebook.

Rifat Jawaid  [Follow](#)

[@RifatJawaid](#)

Great. [@facebook](#) has disabled my personal account [Facebook.com/RifatJawaid](#) because it thought I was impersonating my self. Just read the reasons why it disabled it. There's never a dull moment in social media user's life.

Why was my account disabled?

Your account was disabled for violating the Facebook Terms.

Our Policies

- Your account must list your authentic name.
- Personal accounts must represent individual people only. It's a violation of our policies to use a personal profile to represent anything other than yourself (ex: celebrities, pets, ideas, objects, etc.).
- Impersonating anyone or anything is not allowed.
- Maintaining multiple accounts is a violation of our policies.
- Accounts created for the purpose of spamming or harassing others are strictly prohibited.

[Learn More](#)

4:13 AM - 27 Sep 2018

35 Retweets 80 Likes

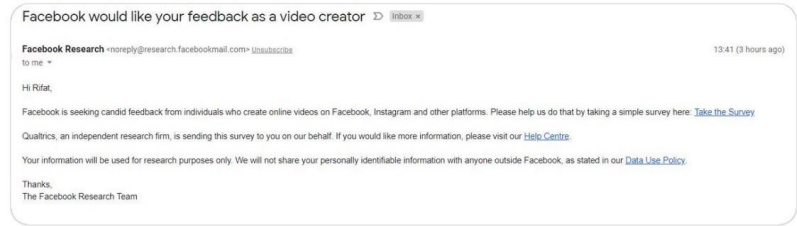
¹¹ <http://velivada.com/2017/07/05/manusmriti-rules-facebook-bans-national-dastak-posts/>

¹² <https://freepresskashmir.com/2018/12/29/look-out-for-the-phrase-free-kashmir-facebook-tells-its-moderators/>



Rifat Jawaid @RifatJawaid · 27 Sep 2018

At 1.41 pm today @facebook sends me a mail asking for my feedback as a video creator. At 4.40 pm it disables my account, minutes after I post my views on Ayodhya verdict. Coincidence?



7 30 57



Rifat Jawaid @RifatJawaid · 27 Sep 2018

So @Facebook must know that my last few posts, around 10-12 in last 24 hours, included photos with my parents, few news stories and comments on Ayodhya verdict. You don't need to be a rocket scientist which post would have prompted @Facebook to disable my account.

6 19 54

Multiple minority community news pages were removed and then returned after protest. They include the National Dastak a popular Bahujan (caste oppressed) news outlet.

National Dastak Banned



Political Cartoon about this issue from Free Press Kashmir:



Widespread Doxing and Targeted Threats Against Individuals

Many of the minority leaders we spoke to mentioned receiving a crippling number of direct threats of violence on the platform. These included journalists, activists, academics students and others who dared to speak out.

These personal attacks have been particularly worrisome when they include tactics like doxing where an individuals personal information such as photos, profile links, or phone numbers, are weaponized to further harm and make them a target of real world violence.

Facebook's procedures for reporting these threats are opaque, and many people struggle with such invasive intimidation.

Three Cases of Note

Sujatha Surepally: The first is the case of threats and violence against Ms. Sujatha Surepally, a Dalit activist, scholar, and leader. Because of her rising profile and growing activism in the state, local news channels in Telegana and Andra Pradesh with right-wing agendas falsely accused Sujatha of burning pictures of “Bharat Mata” (Mother India).

She was immediately swarmed by trolls who dug out her picture and added to the disinformation campaign by alleging she had Naxalite affiliations and was an “anti-national.”



One troll writes in the above post, “Lakshman cut off Surpanakha’s ears and nose then, what should be done with this wretch now?”

This refers to Lakshman and Surpanakha, two characters in the Ramayana. In this tale, Lakshman attacks and severs the nose and breasts of Surpanakha, who is portrayed as an “ugly demon-woman.” There is a deeply casteist underpinning to this story. Dalits, Adivasis, and caste-oppressed people are in many folk and other knowledges said to be descended from “Asuras” (demons). Dalit and Adivasi women, in particular, are subjected to extreme levels of physical and sexual violence based on these age-old tropes. The comparison of a Dalit woman activist and scholar to Surpanakha is therefore deeply casteist and anti-indigenous, as well as being a call for violence against her.

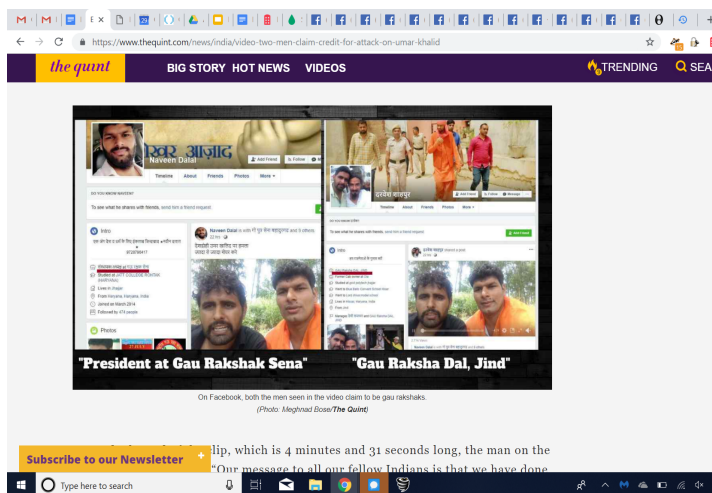
Sujatha endured a period of fear while these online attacks threatened to spread into her real life. Because she is a bold person who is committed to justice for her people, she has made it through this harrowing experience. But despite multiple escalations of her case to Facebook this content remained on the platform. There has to be more rapid response channels to protect vulnerable leaders in the face of clear incitement to violence. While Sujatha was lucky this time, Facebook should not rely on luck to keep its users safe.

Umar Khalid: The next case is an attack on the Muslim student activist Umar Khalid. Umar was shot at while speaking at the Constitution Club in Delhi on August 13. Fortunately Umar survived the attack unharmed. However, in a video widely circulated on Facebook and WhatsApp, two men, Naveen Dalal and Darvesh Shahpur openly confess to the attempted murder!¹³

The full video states, “We respect the Supreme Court and the Constitution, but we are angry

13 <https://thewire.in/rights/police-searching-for-two-men-who-claim-attack-on-umar-khalid-in-video>

that there is no proper provision in the Constitution to punish these mad dogs, like the JNU gang. Our elders in Haryana have taught us that mad dogs should be done away with, without delay. We, Darvesh Shahpur and Naveen Dalal Mandothi, are responsible for the attack on Umar Khalid in Delhi. This attack was a gift from us this Independence Day. We will surrender to the police on 17 August at 12 noon at patriot Kartar Singh Sarabha's village, at his home."¹⁴



Despite this page and video being reported numerous times, both of these users still have active Facebook pages as of June 1, 2019.^{15,16}

Clean the Nation: The last case we present to show widespread doxing campaigns are on Facebook is the case of the vigilante Facebook group Clean the Nation. As reported by Quartz¹⁷, Ashutosh Vashishtha — one of the group's founders — said in a note that the group's sole function was, "for cleaning and weeding out anti nationals who wear the tag of Indian but leave no stones unturned to insult and mock their own people, culture and the army that is one of the best in the world,"¹⁸ The note added: "**We ensure you, that if our esteemed Army is doing an external surgical strike in our neighbouring enemy, we will bloody well do an internal one against these enemies within the nation.**"

This kind of widespread doxing of Indians has been escalating since the 2014 elections but this campaign was one of the largest public doxing attempts recently on the platform. In this campaign Clean the Nation contributed with a network of other groups to the termination of 50 individuals from their jobs and schools. Further some individuals were even charged with sedition! Facebook was slow to reach in the wake of this crisis and when it finally remove

14 <https://www.thequint.com/news/india/video-two-men-claim-credit-for-attack-on-umar-khalid>

15 <https://www.facebook.com/gauhatyaroka.dusman>

16 <https://www.facebook.com/darvesh8572002916/>

17 <https://qz.com/india/1553592/after-pulwama-a-facebook-group-targeted-anti-national-indians/>

this page the founders simply created a new one!¹⁹

Clearly the issue of user safety in this volatile political climate must take priority at Facebook immediately. Facebook must not be a platform where violence and active doxxing campaigns can be organized and carried out in real life to affect people's jobs, academic life, and even their very freedom. The lives of our people and our leaders, should not be taken for granted on any platform. And it is clear that, even in the case of incidents as serious as the ones listed here, Facebook is slow to act and this must change.

Chowkidar Ayush Gupta
February 18 · 🌐

Attention 🇮🇳

Hi friends, our facebook group 'Clean the Nation' has been taken down by Facebook without any prior warning and notice. We are trying our best to resolve this issue and will get back to you with some updates very soon.

Meanwhile keep doing R&D and keep finding the traitors living inside India. You can send such details to us on cleanthenationmission@gmail.com

On the jayanti of Chhatrapati Shivaji Maharaj today, we pledge that we will follow his footsteps and do our best to make India free of traitors.

Vande Mataram 🙏

CLEAN THE NATION

#ExposeTheTraitors

93 12 Comments 8 Shares

¹⁹ https://www.facebook.com/cleanthenation1/?_tn_=_K-R&eid=ARAsS-GnySNG0-N_pOvsZP65J2pdw8lJwkmnhPL5rGquBzwfnrbDIzVbs8Frm-bLOA0WemOozx_0hXU7Z0J&fref=mentions



7 others.

February 21 · 🌐

चुनौतियों से सरकार निपट लेगी लेकिन अपने बीच बैठे चूतियों से निपटना सीखो दोस्तों।

#Attention

We at #CleanIndiaMission promise you that we will never take law in our hands but we will definitely make those traitors meet thier fate who talks against India or our army on social media but this isn't possible without your support.

I again request you all to share profiles, screenshot or anything or everything that is said against India by Indians on social media.

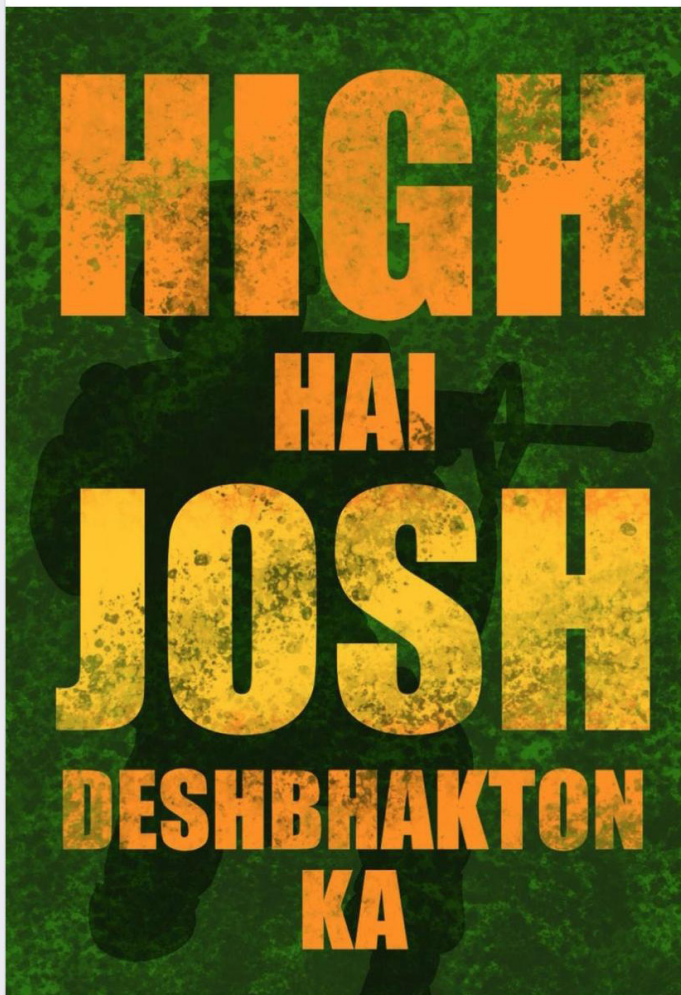
What you have to do...

1. Take screenshots of the post,comment.
2. Copy thier profile link.
3. Collect informations like college/university/companies they are working in.

Send the details to our mail id- cleanthenationmission@gmail.com

DM on twitter @Cleanthenation1

Team Clean the Nation!!!



👍❤️👤 190

21 Comments 12 Shares

Conclusion

We acknowledge that the problem of hate speech and content moderation is a difficult one, and it is not one that Facebook alone can solve. This requires a commitment of resources, transparency, and accountability. Civil society groups must be allowed to participate as welcomed and informed partners so we can collectively achieve a safer platform for all. To this end, we are requesting that Facebook conduct an independent, aggressive, and thorough independent human rights audit of its hiring practices, content moderation, privacy and security policies, and the scope of its elections and government unit's work since 2014.

We recommend that Facebook especially focus on the human rights impact of those policies on Indian minority communities, with input from key stakeholders in the Caste, Religious, and Gender/Queer minority communities. This audit should also be disaggregated across our discrete language markets so we can compare implementation across the many Indian languages.

Facebook's audit in the wake of the Cambridge Analytica scandal in the US was the bare minimum to address the harm the platform caused to vulnerable minority communities and to the integrity of the 2016 elections. This audit was fought for and won by a coalition led by the Color of Change, Muslim Advocates, Center for Media Justice and other civil rights groups. Audits work, and with the dire state of hate speech on India's Facebook platform, we must act now. With India now the country with the largest population on Facebook, don't Indian citizens and India's minorities deserve the same treatment as our American counterparts?

We also ask that Facebook meet with our team to immediately engage in product-level discussions of the findings in these briefings, and to discuss how we could work directly with engineers to solve some of these problems. Our colleagues in other Global South markets have established such conduits and provide instructive examples of what can work well.

We ask urgently that you treat the concerns we have raised with the same gravity in which we have shared them. This is a time of great division and violence in India, with growing caste and religious atrocities, including the creation of registries and detention centers for Indian Muslims. We must act now to protect Indian minorities, as we cannot afford for India to become the next Rakhine State. We already have ample evidence of how violent hate speech translates to physical violence. We implore Facebook to take action before a greater tragedy occurs across India.

We hope we can work together to unleash Facebook's potential for protecting human safety before it's too late.

