Anti-Muslim Hatred and Discrimination
Submission from Dia Kayyali
Associate Director of Advocacy at Mnemonic, independent consultant on technology and human rights, and co-chair of the Christchurch Call Advisory Network1
dia@mnemonic.org

Summary: This submission focuses on the online aspect of anti-Muslim hatred and discrimination (Islamophobia). Content that incites violence against Muslims is too often left up on major social media platforms, while important content produced by Muslims is often removed. This leads to diminishing opportunities for justice in conflict zones such as Syria, while facilitating increased violence against Muslim communities around the world. Furthermore, online content is never solely online. Rather, it is intimately linked to violence and discrimination against Muslims in a harmful feedback loop.

How the online feeds into the offline and back again, creating a vicious cycle

The online component of Islamophobia has deadly consequences for Muslims around the world. At the same time, predominantly Muslim communities see content they post online regularly removed by major social media companies.

Islamophobia, like other social ills, is stuck in a dangerous feedback loop. Offline discrimination and violence lead to online hate speech and dangerous speech. This content then worsens discrimination, and sometimes directly incites offline violence and other negative consequences. The cycle is self-perpetuating, and it is deadly. Muslim lives have already been lost as a result, and Islamophobia threatens other essential human rights including freedoms of expression and religion.

Who is Muslim?
Islamophobia doesn’t just impact Muslims. As outlined in this submission, it also impacts people who are secular or practice other religions but are in Muslim majority countries or communities. Islamophobia is not solely about religion.

Definitions:

Islamophobia: The term Islamophobia is justifiably contested. This submission doesn’t attempt to define it, but uses Islamophobia to refer to prejudice and discrimination against Muslims and people associated with Islam such as Christian Arabs. Islamophobia can be

1 See more online: https://mnemonic.org/; https://www.christchurchcall.com/advisory-network.html
conscious or unconscious, individual or structural. It is often linked very closely to xenophobia, even when directed against longstanding Muslim minority communities.

**Content moderation:** Content moderation is the process by which companies, especially social media platforms, take down or leave up content - posts, accounts, groups, and more. It is usually done through a mix of human content moderators and automation.

**Automated content moderation:** Automated content moderation is when all or part of decisions about content are made by artificial intelligence. This can happen at the time content is uploaded or later. The use of automation has greatly increased since the beginning of the COVID-19 pandemic.²

“**Terrorist and violent extremist content**” (TVEC): this phrase is increasingly used by companies and policymakers to refer to a poorly-defined swath of content related to terrorism and violent extremism, terms which are themselves contested.

**Dangerous speech:** Coined by Susan Benesch, dangerous speech refers to “any form of expression (speech, text, or images) that can increase the risk that its audience will condone or participate in violence against members of another group.”³ In this submission, it is used as shorthand for the phrase “content that leads to offline violence.”

**Who is a terrorist?**

The term terrorism is generally defined as an unlawful use of force, particularly against civilians, with an aim of influencing politics. How the term is used, and which groups and individuals are officially designated as terrorist, varies widely from context to context. Violent extremism, another term commonly thrown about by social media platforms and lawmakers, is even less well-defined, but often used interchangeably with the word terrorism. These terms are officially used by states, and in some contexts by social media platforms, almost exclusively to refer to Islamic groups.

When it comes to social media platforms and other technology companies, the most influential lists of officially designated terrorist organizations are the United Nations Security Council Consolidated List and United States government definitions.⁴ It’s important to note that "[t]here is no legal mechanism in the United States for labeling purely domestic organizations as terrorist

---


groups,” meaning it’s nearly impossible to apply that label to the far-right groups widely associated with violent attacks, including those that target Muslims.⁵

Recently the United States added its first and only far-right group to its list of designated foreign terrorist organizations, and the United Nations Security Council list does not have any of these groups listed.⁶ In fact, after the Christchurch Massacre, when an extremist murdered 51 Muslims and livestreamed the act, only one Security Council member raised the need to consider how to address international far-right terrorism.⁷ This is slowly changing, but the fact remains that terrorism and violent extremism are terms largely associated with Islam.⁸

**Islamophobic terrorism?**

In addition to a lack of officially designated white supremacist groups, other types of extremism that target Muslims are also not officially listed. Particularly notable, groups like Ma Ba Tha (influenced by extremist monk Wirathu to engage in violence against the Rohingya in Myanmar), and various identifiable Hindu extremist groups such as the Rashtriya Swayamsevak Sangh that are paramilitary in nature and linked to violence, are not officially designated as terrorist or extremist groups, nor are their most militant factions. That is despite the fact that these groups employ violence to obtain the political goal of forcing Muslims out of Myanmar and India, respectively.

In fact, there is little research on such extremism, but what little there is urges policymakers to expand their understand of far-right terrorism and violent extremism. Writing in February of this year, researcher Dr. Eviane Leidig said:

> Earlier this month, there was a far-right terrorist shooting at Jamia Millia Islamia (JMI) University in New Delhi, India. The perpetrator broadcast the attack live on Facebook, shouting Hindu nationalist slogans whilst opening fire. Significantly, the perpetrator targeted a crowd that had gathered to the mark the 72nd

---

anniversary of Gandhi’s assassination. Gandhi was murdered by a Hindu nationalist who believed Gandhi to be too ‘secular’ and accommodating to India’s Muslims [and who was] a member of the paramilitary Hindu nationalist organization, the Rashtriya Swayamsevak Sangh (RSS), which advocates for an ethno-nationalist Hindu state.9

Dr. Leidig points out that the media has not taken Hindu nationalism seriously because the violence is largely confined to India’s borders, and because India is an ally to many Western countries.10 This is perhaps why discussions about this kind of violence are not happening in policy spheres, either. High profile events like the Jamia Millia shooting are, unfortunately, not isolated.11

Definitions of terrorism and extremism are flawed in myriad ways, and perhaps in their current form UN and US lists can’t include far-right groups—but if that’s the case, social media companies and others need to rethink reliance on these lists and explain how they’re addressing these major gaps in definitions when searching for the dangerous speech that can cost Muslim lives.

**Counter terrorism through content moderation**

Counter-terrorism is not a field that is known for its respect of human rights.12 It has often led to the implementation of repressive measures. The United Nations Office of Counter-Terrorism notes “the shrinking space for human rights defenders and civil society actors to exercise their freedoms as a consequence of counter-terrorism measures that are not human rights-compliant.”13 In a very current and disturbing example, in the wake of terrorist attacks, France has forced Muslim civil rights organization Collectif contre l'islamophobie en France (CCIF), to disband without any judicial intervention.14

---


10 Id.


Civil society space is also rapidly disappearing online, and Muslims are particularly hard hit.

Deleting our content
There is little empirical evidence on how moderation of “TVEC” impacts radicalization and violence, but there is evidence suggesting that automated moderation based on biased definitions of terrorism and violent extremism is removing huge swathes of human rights documentation that could be valuable in achieving justice for human rights abuses in conflict zones like Syria.15

Syria and Yemen
Justice for the atrocities committed in Syria relies on the ability to find witnesses willing to testify—and on human rights documentation.16 Much of that documentation lives on social media platforms, where groups like Mnemonic are able to find, verify, archive, and deploy it. In October, Mnemonic, the Justice Initiative, and the Syrian Center for Media and Freedom of Expression submitted a criminal complaint to the German federal public prosecutor over sarin gas attacks in Syria.17 Mnemonic is also currently assembling further evidence for the Prosecutor’s office, raising hopes that the Assad regime may actually face justice for its crimes against humanity.18

Unfortunately, this documentation is at risk, especially on YouTube where machine learning-powered automated flagging and takedowns have led to termination of thousands of Syrian


18 Christoph Reuter et al., German Prosecutors Are Collecting Evidence on Chemical Weapons Attacks in Syria, 27 Nov. 2020, Spiegel
https://www.spiegel.de/international/world/german-prosecutors-are-collecting-evidence-on-chemical-weapons-attacks-in-syria-a-74fec59c-9461-42f6-a6a4-dc89e1a08692; Al Badry et al., How Germany could indict Syria's Assad for war crimes. Deutsche Welle, 27 Nov. 2020.
YouTube channels that published videos documenting human rights violations. Takedowns have ranged from documentation of protests in Syria to non-traditional media reporting on violent attacks.

Mnemonic compares archived collections of videos with what is still available online to track takedowns. While some videos are made unavailable by users, many of them have been deleted due to faulty content moderation. Mnemonic’s most recent check of content from Syria shows that 23% of 1,772,992 YouTube videos and 11% of 1,172,202 tweets failed to download.

Mnemonic’s Yemeni Archive, like the Syrian Archive, includes collections of verified videos documenting attacks on civilian infrastructure, including medical facilities and bridges committed by all parties to the conflict. The collection is younger than the Syrian Archive, but just as vital. Many experts, including UN officials, are calling Yemen “the worst humanitarian crisis.”

Justice for human rights abuses in Syria has taken years of advocacy, but Yemen may be at a great disadvantage, as content is taken down ever more quickly. Our most recent check of content from Yemen shows that 15% of 445,777 YouTube videos and 10% of 251,143 tweets failed to download.

Allowing dangerous speech
Not only have social media platforms removed large swathes of content documenting human rights abuses in Muslim majority countries and communities, they continue to leave up dangerous speech. In the worst cases, they are employed as tools in atrocities, including genocide.

The true extent of dangerous Islamophobic speech on social media is staggering. Facebook is the easiest to point to, because it has been the focus of extensive advocacy and press reports. For example, a recent report from Muslim Advocates and Global Project Against Hate and Extremism points out:

Facebook has been used to orchestrate the Rohingya genocide in Myanmar, mass murders of Muslims in India, and riots and murders in Sri Lanka that targeted Muslims for death. Anti-Muslim hate groups and hate speech run rampant on Facebook with anti-Muslim posts, ads, private groups, and other content. Armed, anti-Muslim protests in the U.S. have been coordinated from Facebook event

---

pages. The Christchurch, New Zealand, mosque massacres were live-streamed on the site and the videos shared by untold numbers worldwide.\textsuperscript{21}

However, Facebook shouldn’t serve as a scapegoat for other companies. Twitter is also culpable, and YouTube in particular bears a lot of responsibility for hosting dangerous speech while investing minimal resources engaging with civil society to address it. What’s more, as research on misinformation and disinformation has made clear, content travels from platform to platform, often helped along by sophisticated groups, including state-sponsored “cyber armies.”\textsuperscript{22}

\textit{Myanmar}

Any submission on online Islamophobia wouldn’t be complete without pointing to what happened in Myanmar, which has been well documented by the Independent International Fact-Finding Mission on Myanmar. The IIIMM explicitly criticized Facebook’s role in the genocide of the Rohingya (a majority Muslim community) in Myanmar, calling it a “beast.”\textsuperscript{23} The role of content posted on Facebook in facilitating the violence is clear; “hundreds of military personnel created troll accounts and news and celebrity pages on Facebook and then flooded them with incendiary comments and posts timed for peak viewership.\textsuperscript{24} They also trolled posts critical of the government and “posted sham photos of corpses that they said were evidence of Rohingya-perpetrated massacres.”

\textit{India}

India provides one of the most blatant and disturbing examples of how dangerous speech spreads on social media platforms—and is inextricably intertwined with offline violence and discrimination.

In August, \textit{The Wall Street Journal} published two articles about Facebook India, with evidence that Ankhi Das, the head of Public Policy at Facebook, India, was allowing dangerous speech to

\begin{itemize}
\end{itemize}
thrive on the platform.\textsuperscript{25} Ankhi Das is gone, but the problems continue. As a civil society letter to Facebook about the report notes, dangerous speech targeting Muslims in India has been thriving on social media since at least 2013, when “a mislabeled video on social media was instrumental in stoking the horrific 2013 Muzaffarnagar riots in which 62 people were killed.”\textsuperscript{26}

Lynching is a pandemic in India, and while they can be linked directly to WhatsApp groups, they also “appear to have been sparked by online misinformation campaigns that often carry a strong Islamophobic bent.”\textsuperscript{27} In one case, perpetrators spread misinformation on Facebook and WhatsApp and even posted videos of the lynching, including victims begging for their life.\textsuperscript{28}

\textit{Love Jihad, CoronaJihad, and more}

CoronaJihad is another particularly viral Islamophobic campaign. At the start of the pandemic, Muslims were accused of spreading the virus intentionally with the purpose of infecting non-Muslims.\textsuperscript{29} For weeks the hashtag #CoronaJihad trended on social media platforms, starting with a focus on the Tablighi jamaat, a Muslim missionary movement. The group held a conference in Delhi in March, a week before India declared a national lockdown.\textsuperscript{30} The attendees were blamed

\begin{footnotesize}
\begin{itemize}
\item Global Project Against Hate and Extremism et al, \textit{Major Human Rights and Internet Watchdog Organizations Sign On to Demands for #AuditFBIndia}, 9 Sep. 2020, \url{https://www.globalextremism.org/post/facebookindia}
\item Sukanya Shantha, \textit{COVID, Communal Reporting and Centre's Attempt to Use Independent Media as Alibi for Inaction}, The Wire, 18 Nov. 2020, \url{https://thewire.in/communalism/tabligi-jamaat-communal-reporting-ib-ministry-coronavirus}
\end{itemize}
\end{footnotesize}
for the spike in the number of cases, although other non-Muslim religious gatherings took place after this gathering.

The CoronaJihad hashtag included out of context videos that purported to show Muslim vendors spitting on food to spread COVID.\(^{31}\) It was accompanied by calls to boycott Muslim owned business, resulting in loss of basic income for some. It also occurred only months after Islamophobic riots in Delhi resulted in 53 deaths, over 200 injuries and 2,000 arrests.\(^{32}\)

In August Twitter told the Telangana High Court that it removed CoronaJihad content after being ordered to do so in June.\(^{33}\) Unfortunately, CoronaJihad had already become a tool in the Islamophobic toolkit.\(^{34}\)

Love Jihad is another viral Islamophobic conspiracy theory spread on social media. Proponents claim that Muslim men woo women from non-Muslim communities with the sole purpose of converting them to Islam, thereby increasing the Muslim population. This conspiracy theory has been helped along by viral social media content, including dangerous speech that targets specific individuals. In one case, names and personal details of Hindu-Muslim interfaith couples were published.\(^{35}\) One of the targets reported that men showed up outside his house, referenced the post, threatened him, and beat him up.\(^{36}\) The concept of Love Jihad was also part of what stoked the Muzaffarnagar riots.

---


It's not just violence: love jihad legislation in India

Just this month, the states of Uttar Pradesh and Madhya Pradesh passed or are in the process of passing legislation based on this theory, with the purpose of criminalizing Muslims.\textsuperscript{37} Although love jihad is not mentioned in the legal text, Muslims are clearly the target of this legislation as evidenced from speeches made by BJP ministers.\textsuperscript{38} Under these laws, punishable with jail term and fine, if someone wants to marry after converting into any other religion, they will need to take prior permission from the District Magistrate at least two months before the wedding.

Recommendations

- Work with the Security Council and the Special Rapporteur on the promotion and protection of human rights and fundamental freedoms to review the Security Council’s role in counter terrorism efforts, including its Consolidated List.
- Engage with the UN Special Rapporteur on minority issues, given that he is currently focusing on online hate speech and minorities, to produce a joint set of recommendations to social media platforms. Platforms should at a minimum:
  - Conduct a detailed, global audit of Islamophobia and country-specific human rights audits in countries with a high risk of Islamicphobic violence and share results of the audit publicly;
  - Increase transparency around “TVEC” designations and takedowns, including publicly providing rationale behind selection of organizations designated as terrorists; and
  - Engage with civil society activists and human rights organizations on content takedowns of human rights documentation, and to establish what constitutes Islamophobia and how to deal with it in online spaces.
- Dedicate resources specifically to addressing Islamophobia from a gendered lens, especially in countries where Muslim women are a minority and are targets of severe online harassment and sexual violence as a result of Islamophobia.

\textsuperscript{38} \textit{Id.}