

Explainer: Weaponisation of social media against Hindus in UK



Centre for Integrated and Holistic Studies (CIHS) is a non-partisan, independent, research think tank headquartered in New Delhi, India. CIHS is dedicated to enriching individual decision making by presenting innovative ideas, fostering informed public debate, and advancing effective policy and programme development to advance humanity. Aspiring to positively shape the future of society, CIHS works to share knowledge on pressing global challenges and opportunities by fostering a 'culture of scholarship' and advancing informed public engagement.

Every effort is made to ensure that the information contained in these publicly available reports, factsheets, briefings is correct at the time of publication. If you have any comments on our briefings, email us on info@cihs.org.in

Disclaimer and Limitation of Liability: The report is for public distribution and has been furnished solely for information and must not be reproduced or redistributed to others without written consent. None can use the report as a base for any claim, demand or cause of action and, also none is responsible for any loss incurred based upon the report.

Weaponisation of social media against Hindus in UK

Artificial Intelligence models quantify the weaponisation of social media sites to spread hate and false narratives against Hindus that incited anti-Hindu violence in Leicester.

Leicester, the heartland of the Hindu community in the United Kingdom, is home to the second-largest Hindu populace in Europe. Widely known for its vibrant Hindu culture and stunningly beautiful Hindu festivities throughout the world. Since late August 2022 until September 2022, the city's minority British Hindu community had been the target of deliberate, organised, and systematic acts of violence including vandalisation of a local Hindu temple by South-Asian origin Muslims in the UK.

For nearly a month, Hindus were overwhelming victims of violent crimes like attempted stabbings, looting, organising to commit crimes, vandalism, and hate speech. In addition, despite assurances of safety from the local authorities and Leicestershire police, British Hindu homes and business establishments in Leicester were targeted by organised, violent Islamist mobs. Extreme Islamist groups in the UK waged a systematic terror campaign against British Hindus while they were in Leicester, subjecting them to extreme threats, assaults, and trauma.

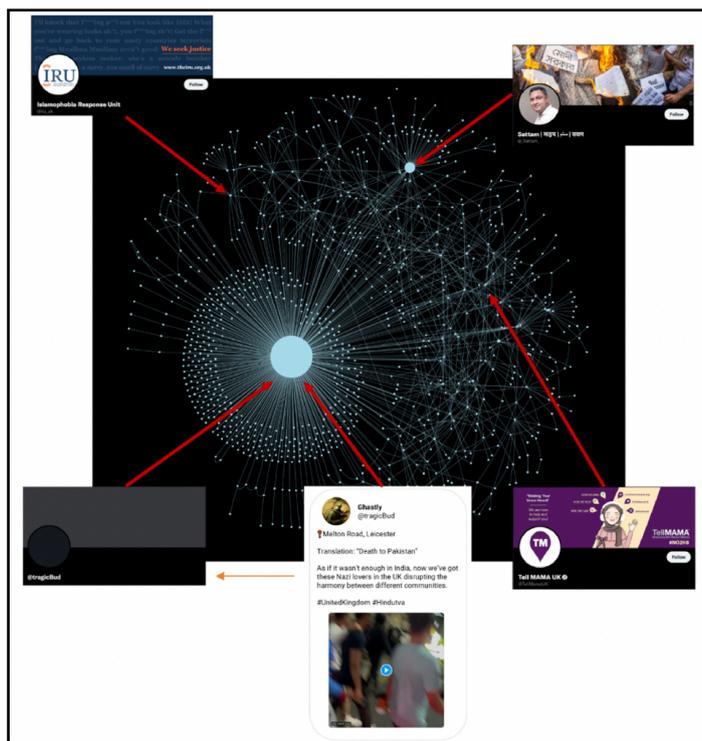
A large-scale mobilisation of violent extremist Islamists and criminals to Leicester with the intention of damaging the city's Hindu population was made possible by online hate campaigns that were planned both within and outside of the UK to incite hatred and spread misinformation against Hindus.

The media primarily characterised the anti-Hindu violence in Leicester as being sparked by racial and aggressive anti-Pakistan slogans; nevertheless, the planned, systematic, and organised attack on Hindus was never brought up or acknowledged. Statements condemning the horrific anti-Hindu violence were casually reported by the media.

In fact, mainstream UK media outlets like the Guardian, the BBC etc. made sure that fake news disseminators and supporters of Islamist extremist groups like Majid Freeman, who spread false information to incite violence against Hindus both online and offline, or roving reporters like Sunny Hundal, who stoked Hindu hate and sparked violence in the area by tweeting inciting messages, were given a platform to further spread Hinduphobia and extremist views.

Most of these observations were confirmed in a research released on November 17, 2022 by the National Contagion Research Institute (NCRI), a cyber threat intelligence organisation. In the report entitled "In a Cyber Social Swarming Precedes Real World Riots in Leicester: How Social Media Became a Weapon for Violence"¹. NCRI collected data about incidents in Leicester from Instagram, YouTube, TikTok, and Twitter between August 27, 2022, and September 19, 2022. The report's findings are based on a variety of technological tools, machine learning, open source intelligence, natural language processing (NLP) models, network analysis, and linguistic models like Conflibert which are used to map political conflict and violence.

The findings of the reports show that bots used cyberswarming² to encourage violence against Hindus online, which eventually spread to the real world. Real-world violence against Hindus, hate speech and slurs, vandalism of Hindu temples, and intimidation were manifestations in the socio-physical domain. The first social media incident that has been



attributed to inciting violence in Leicester was a video that went viral and was shared by the unknown account @tragicbud. In the video, cricket fans (who were not necessarily Hindu) can be heard chanting phrases like "Death to Pakistan."

The user @tragicbud equated the sloganeering in the video with Hindutva, thereby giving India's victory in the cricket match an ethno-religious connotation. In order to create a network analysis, this tweet was examined by NCRI. This has been highlighted in Figure 1.

Fig.1. Network analysis of the first tweet by @tragicbud.
Source: NCRI

¹ Cyber social swarming precedes real world riots in Leicester: How social media became a weapon for violence(2022) Network Contagion Research Institute. Available at: <https://networkcontagion.us/reports/11-16-22-cyber-social-swarming-precedes-real-world-riots-in-leicester-how-social-media-became-a-weapon-for-violence/> (Accessed: November 17, 2022).

² Definition of cyberswarming by Network Contagion - "Even as law enforcement and intelligence begin to map how social media can rapidly radicalise individuals to commit acts of domestic terror, they remain challenged to understand how social media empowers entirely new groups to self-organise radicalised militant cells and incite violence. The ability of extremist groups to self-organise creates a new and poorly understood thereafter for emerging threats in the cyber domain."

Following this, Pakistan inclined twitter handles in used cyberswarming and began notifying their followers and urging a call to action in both the cyber and real worlds. In Figure 2; Three of these tweets are displayed.

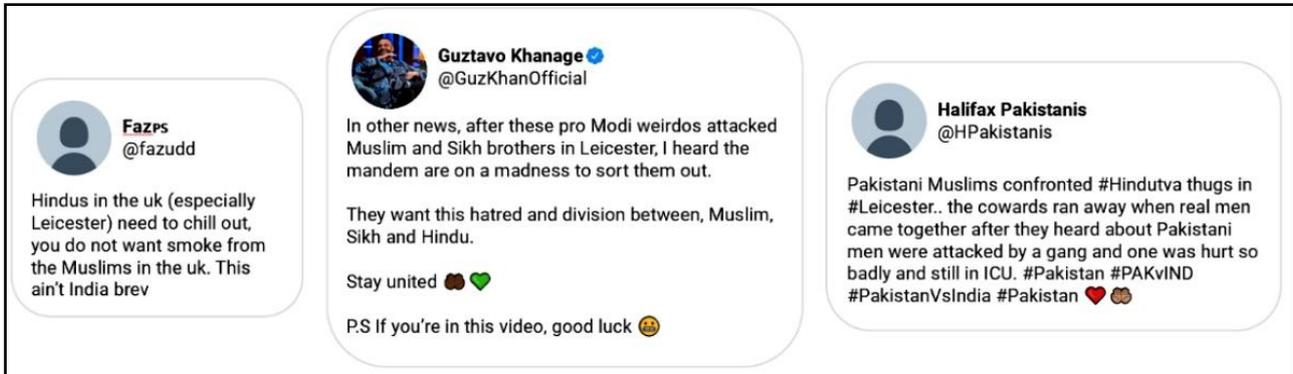


Fig. 2. Call to action raised by British Pakistanis and their organisations following the viral video. | Source: NCRI

In the days following the altercation, the narratives being spread by these Pakistan-leaning twitter handles and entities grew more intense, making the terms "Hindu," "Hindutva," and "Abuse" recurrent themes. Figure 3 illustrates network analysis and statistical association between these themes as reported by NCRI. Additionally, it demonstrates the "Hindu" term's cosine similarity³ to 10 additional terms that were adopted by the South-Asian origin Muslims to support Islamophobic narratives.

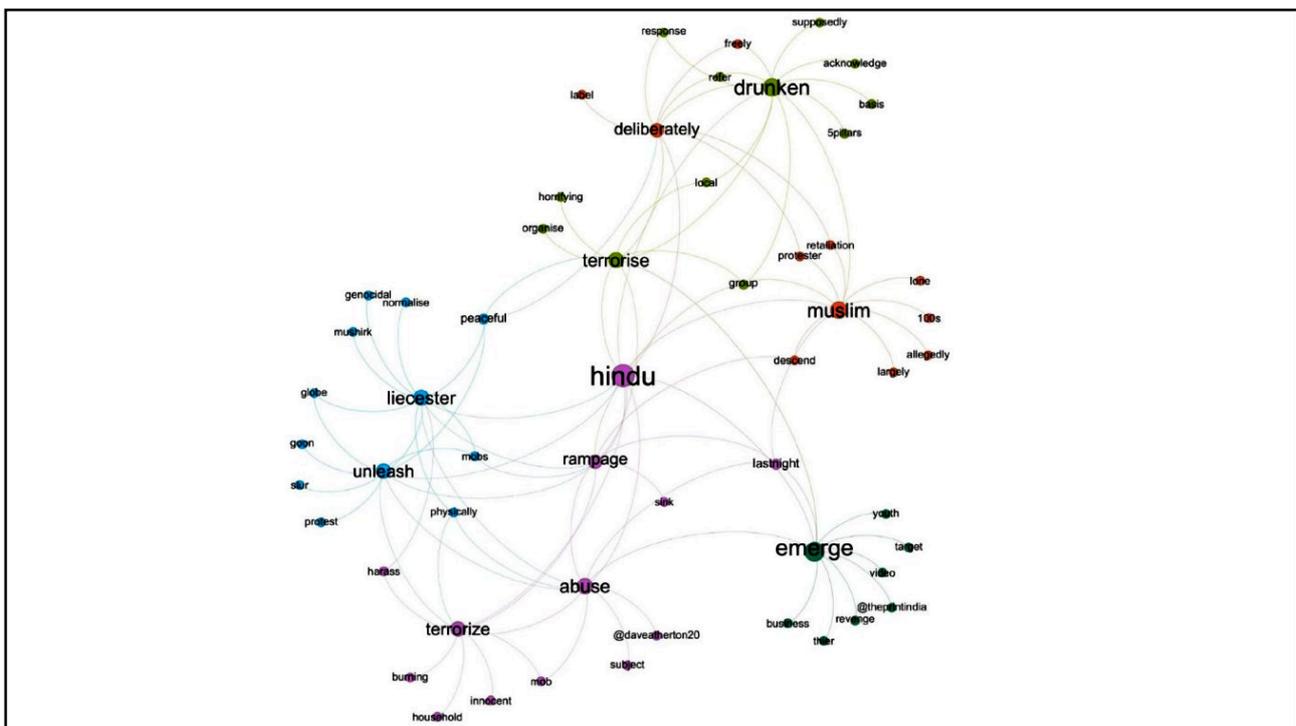


Fig. 3. Network analysis for the term "Hindu"

³ Cosine similarity is a measure of similarity between two sequences.

Word	Cosine Similarity	Word	Cosine Similarity
Deliberately	0.973	Terrorise	0.943
Abuse	0.961	Unleash	0.942
Liecester	0.953	Drunken	0.941
Emerge	0.953	Muslim	0.938
Rampage	0.951	Terrorize	0.937

Fig 3 Cont.. Hindu and its accompanying cosine similarity analysis. | Source: NCRI

The on-ground violence increased along with the amount of Hinduphobic content on Twitter (mainly) and other social media. The use of racial slurs like "Cow Piss" and other hinduphobic epithets increased abruptly. This was examined by NCRI and is depicted in Figure 4.



Fig. 4. Hinduphobic slurs as mapped by NCRI during Leicester violence in September 2022. | Source: NCRI

According to a sentiment analysis of the twitter database compiled by NCRI, accusations and blame were levelled at particular identities and gradually increased in the days following the anti-Hindu violence that broke out in Leicester. The researchers used Perspective API, which analyses text strings to determine how they might affect a discussion as a whole, to conduct the analysis on four key sentiments: "Toxicity," "Insult,"

"Identity Attack," and "Threat." Top 500 tweets, which received the highest engagement, were analysed for these sentiments, and a score was assigned to each category.

The outcomes are shown in Figure 5.

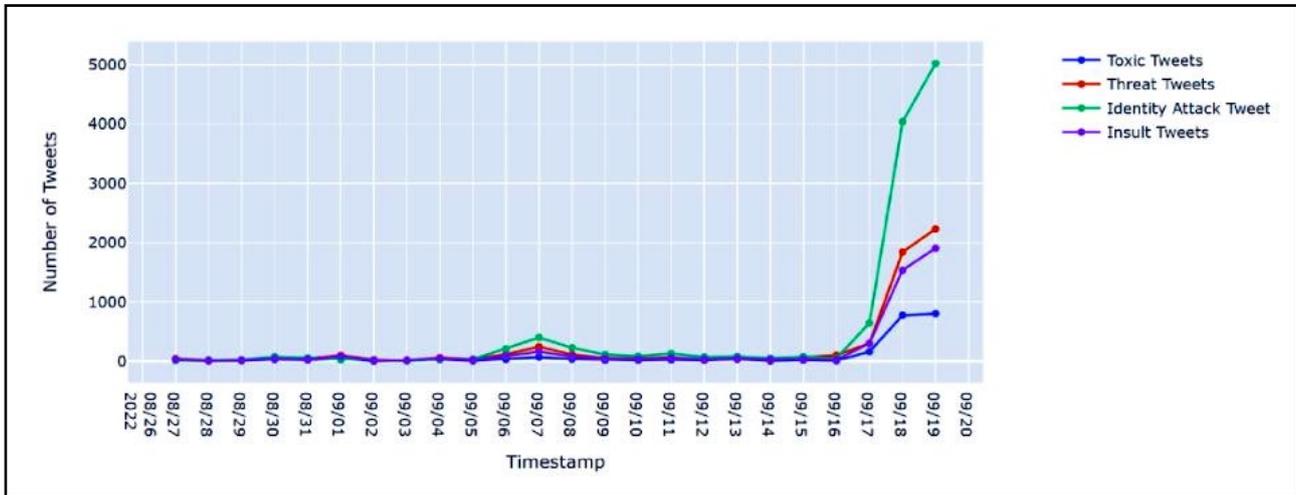
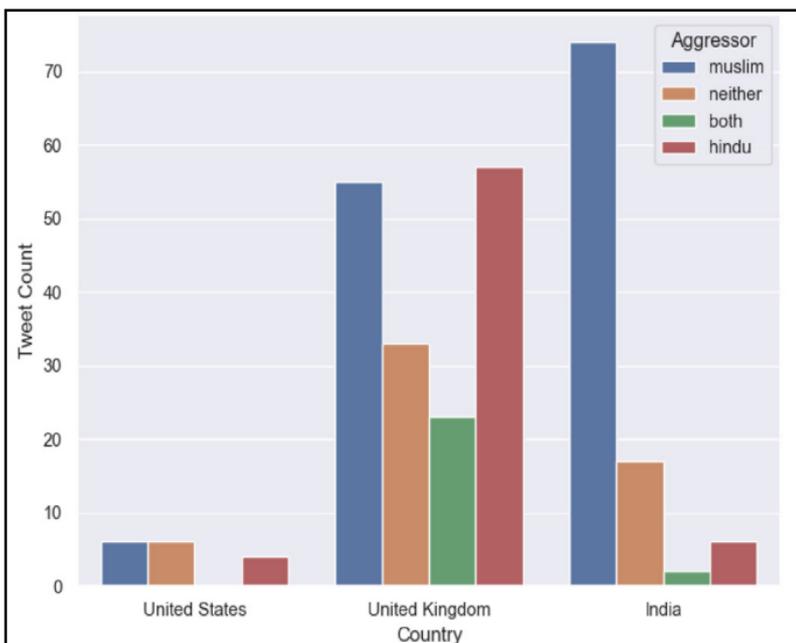


Fig. 5. Sentiment analysis for the terms Toxicity, Insult, Identity Attack, and Threat. | Source: NCRI

The score given for the sentiment “Identity Attack” for these 500 tweets was assigned a score of >0.5. Further, the tweets were classified based on the aggressor mentioned therein and then these were further categorised based on location. This has been shown in Figure 6.

Activities linked to bots were identified as both Hinduphobic and Islamophobic content put



the onus of the violence on the other community. It was deduced that bots were deployed for quicker and amplified dissemination of content during the entire period of the violence. Time-series analysis of both streams of content shows that Islamophobic content got much more differential attention compared to Hinduphobic content. The results of this analysis are shown in Fig. 6.

Fig. 6. Locations for top 500 retweets attacking identity. Source: NCRI

Here, It is crucial to note that the dataset only includes the top 500 tweets with the most activity (which only takes into account retweets) based on a specific threshold. It doesn't represent the required number of retweets. Pakistan was not discovered since it did not satisfy the minimum criteria for retweets because per capita Twitter users in India will always be much larger than those in Pakistan and other smaller nations.

The NCRI also tracked the calls for violence on social media using two tools: ConflIBERT and HateBERT. HateBERT is a retrained BERT model for abusive language detection in English, while ConflIBERT is a domain-specific pre-trained language model for conflict and political violence. According to data gathered by NCRI, during the course of the events, roughly 70% of all calls for violent action were directed toward Hindus. The outcomes are displayed in Fig. 7.

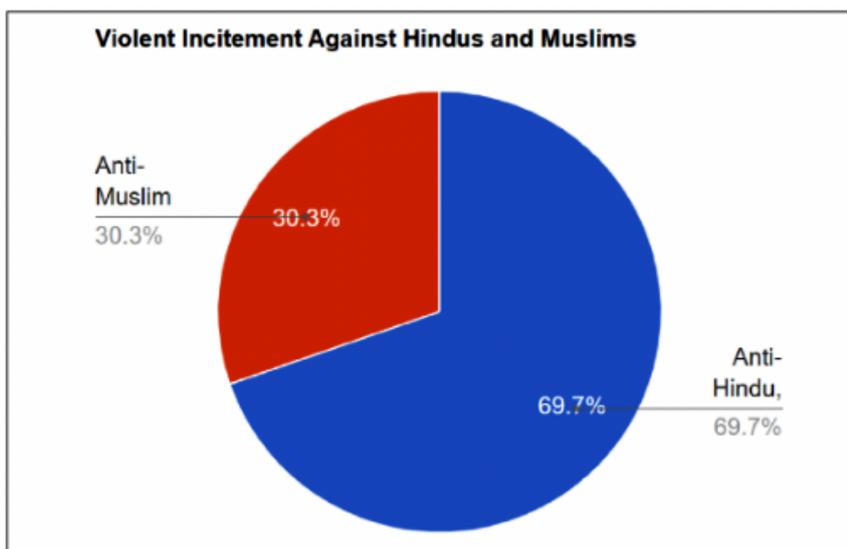


Fig. 7. Calls to violence targeting Hindus and Muslims during the Leicester violence earlier this year.

The report's findings imply that the recent violence in Leicester was not driven purely by organic outrage. The use of social media platforms as tools to spread malicious narratives appears to be a crucial aspect of how this violence was organised and amplified. Furthermore, according to NCRI study, religious animosity and disinformation about

Hindus as bloodthirsty and genocidal perpetrators influenced when and where attacks occurred by enlisting online support for real life attacks on Hindus. The BBC, The Guardian, The New York Times, and other mainstream media sources failed to do their due diligence on Majid Freeman and amplified the voice of a conspiracy theorist and extremist sympathiser who was a key instigator in the Leicester events.

Overall, NCRI research indicates that violence provoked by social media platforms and the use of social media platforms like weapons is frequently leading to real life violence and posing a threat to vulnerable populations like Hindus in the UK and around the world.

