



Post-elections briefing: India

Introduction

The world's biggest general election kicked off on 19 April 2024 in India, where nearly one billion people were eligible to cast their ballots.¹² The vote was preceded by a polarising campaign, during which the ruling Hindu nationalist Bharatiya Janata Party (BJP) sought to secure a third consecutive term.³ Spread over a seven-phase voting period, it concluded on 1 June 2024.⁴⁵

In power for a decade,⁶ BJP, led by Prime Minister Narendra Modi, was hoping to win by a landslide, thanks to welfare payouts to the most needy and amid a campaign steeped in violent rhetoric against Muslims.⁷ But the election results, announced on 4 June 2024, shocked the world as BJP lost its single-party majority in the Lok Sabha (the lower house of Parliament).⁸

Months after the ballot, our analysis of evidence from a myriad of investigations and media reports, reveals that despite being lauded as a "successful" exercise in democracy in which the BJP failed to win its expected landslide, Muslim citizens lost their lives and election laws were broken in a context of Big Tech compliance and platform safety failures.

India's ballot and the post-election period were steeped in offline and online hate speech and accompanied by violence against Muslim voters. Such violence, which social media platforms have helped normalise, is not a hallmark of democracies.

Neither India's electoral laws, nor the tech companies' own policies appear to have mattered throughout. Even though India's laws prohibit references to religion and creating "communal disharmony" during election campaigning, this was fully ignored by the politicians and the Big Tech companies. Social media platforms have operated with impunity in deciding which democratic processes they want to follow, and which they want to ignore – it is clear India's election wasn't a priority.

¹ <https://restofworld.org/2024/exporter-india-deepfake-trolls/>

² <https://www.bbc.com/news/world-asia-india-68678594>

³ <https://www.bbc.com/news/world-asia-india-68872835>

⁴ <https://restofworld.org/2024/exporter-india-deepfake-trolls/>

⁵ <https://www.bbc.com/news/world-asia-india-68678594>

⁶ <https://apnews.com/article/india-modi-cabinet-nda-bjp-election-a607a3ac480c2ccc24c79cbde43fb61f>

⁷ <https://www.aljazeera.com/news/2024/5/3/vote-jihad-as-modi-raises-anti-muslim-india-election-pitch-whats-next>

⁸ <https://www.economist.com/asia/2024/06/04/a-shock-election-result-in-india-humbles-narendra-modi>

Platforms also failed to meaningfully engage with civil society groups throughout the election period and made next to no efforts at being transparent – there were no announcements on how election policies were being implemented and no stock taking after the ballot concluded.

These systemic failures are at the heart of tech harms to democracy and human rights across the globe. They're also at the core of the platforms' toxic business model, which incentivises harmful and hate-filled content and prioritises ad revenue over democratic integrity and human lives. Unless and until that model changes, the already devastating impacts of social media platforms on people and elections will only become more severe.

Social media landscape

Home to over 460 million social media users, India is among the largest and most significant social media markets globally.⁹ Platforms have played a key role in political campaigning since at least 2019, when a third of the country's population had access to social media for the first time.¹⁰

Meta-owned WhatsApp is the most often used social network and messaging app in the country, followed by YouTube, Facebook and Instagram.¹¹ With almost half a billion active users in 2024, India is the encrypted app's largest market.¹³ While all platforms have been employed by bad actors to spread falsehoods and incite violence, WhatsApp and YouTube were at the forefront of electoral disinformation and hate speech.¹⁵

Some ascribe YouTube's prominence in India to the fact that the authorities banned TikTok in 2020, which until then had 200 million users.¹⁷ This prompted many younger social media users to turn to YouTube and Instagram.¹⁸

India has also emerged as a petri dish for tech harms and their often life-ruining consequences – the social media trends and adverse impacts of tech platform failures to secure information integrity in India have been harbingers of things to come elsewhere in the world. This has been the case partly due to a number of factors, including 1) India's ever growing online population, 2)

⁹ <https://datareportal.com/reports/digital-2024-india>

¹⁰ <https://in.boell.org/en/elections-2024-media>

¹¹ <https://datareportal.com/reports/digital-2024-india>

¹² <https://restofworld.org/2024/bjp-whatsapp-modi/>

¹³ <https://restofworld.org/2024/bjp-whatsapp-modi/>

¹⁴ <https://backlinko.com/whatsapp-users>

¹⁵ <https://restofworld.org/2024/bjp-whatsapp-modi/>

¹⁶ <https://www.boomlive.in/decode/sudarshan-news-youtube-channel-islamophobia-hate-speech-misinformation-25591>

¹⁷ <https://www.nytimes.com/2024/03/22/business/tiktok-india-ban.html>

¹⁸ <https://www.nytimes.com/2024/03/22/business/tiktok-india-ban.html>

immense investment into 'digital' by the Indian government and both domestic and foreign companies, 3) India, being a low rights jurisdiction with poor on-ground protection of human rights, and 4) absence of robust regulatory hurdles/controls such as data protection law, online harms regulation and health-tech and financial credit related regulations.

Social media trends: Online hate, real-world violence

BJP politicians and other Hindu nationalist hardliners have been inciting violence against India's Muslims offline and online for years, but their hate speech intensified in the leadup and during the election.¹⁹ The problems of online hate in India had pre-dated the 2024 elections and in some cases led to real-life violence. In 2013, for example, a misleading video in northern India incited riots and five years later a spate of mob lynchings were linked to messages that circulated on WhatsApp groups.²⁰²¹

Additionally Big Tech companies have been a part of India's political landscape for almost a decade and in that time have gained awareness and understanding of the track record of the ruling BJP – a supremacist party with a documented history of anti-Muslim hate speech – in stoking social polarisation and weaponising vitriol to win votes.

Despite this knowledge, tech platforms routinely failed to enforce their own content policies. They neither stymied nor removed posts fomenting enmity or hatred between different classes of citizens on grounds of religion, race, caste, community, or language – which violated India's election laws.²² These failures have meant that videos, ads and other social media posts attacking and calling for violence against India's Muslims have become a fixture, fostering an environment of hate and societal approval for the dehumanisation and violence against Muslims. This in turn has led to Muslim men and women being violently attacked, and in some cases tortured and killed.²³ Such violence is not a hallmark of democracies.

The impacts of Big Tech failures have been long-lasting, continuing well after the election concluded on 1 June 2024, including in states ruled by the opposition Congress party, as part of a purported retaliation for BJP's loss of majority.²⁴ Families had their homes razed to the ground

¹⁹ <https://www.aljazeera.com/news/2024/5/3/vote-jihad-as-modi-raises-anti-muslim-india-election-pitch-whats-next>

²⁰ <https://indianexpress.com/article/news-archive/latest-news/muzaffarnagar-rioters-used-whatsapp-to-fan-flames-find-police/>

²¹ <https://www.bbc.com/news/world-asia-india-44897714>

²² https://thelondonstory.org/wp-content/uploads/Slander-Lies-and-Incitement_Report-2-compressed.pdf

²³ <https://www.aljazeera.com/features/2024/6/1/india-election-fairness-concerns-amid-muslim-vote-suppression-allegations>

²⁴ <https://www.aljazeera.com/features/2024/6/25/eid-means-mourning-muslims-lynched-in-india-after-shock-election-result>

and at least three men were beaten to death after being tortured.²⁵ Like with past attacks against Muslims, the violence was often recorded and shared on social media, where it went viral.^{26,27}

While much of the anti-Muslim violence appears to be a result of a years-long environment of hate, offline and online, there is a growing body of evidence linking social media to specific harms perpetrated by those espousing supremacist Hindutva views. The attacks are often documented by perpetrators or onlookers and shared online, where they cause further distress and suffering to members of already marginalised communities.

In August 2024, vigilantes from the right-wing group Hindu Raksha Dal (HRD) assaulted Muslim families in two separate incidents in northern India – both recorded and shared on HRD’s social media accounts.²⁸ “*They were so angry, went to everyone’s house, and asked if they were Muslims or Hindus,*” a witness of one of the attacks reportedly said. “Hindus were spared, and Muslims were beaten brutally.”²⁹

The violence took place after BJP politicians and other hardliners called for revenge on social media for violence against Hindus in Bangladesh³⁰ following the escape of Bangladesh’s Prime Minister from the capital Dhaka on 5 August 2024. At the time a wave of deadly anti-government protests had swept through Bangladesh.³¹

²⁵ <https://www.aljazeera.com/features/2024/6/25/eid-means-mourning-muslims-lynched-in-india-after-shock-election-result>

²⁶ <https://edition.cnn.com/2024/04/30/india/india-elections-varanasi-hindu-muslim-tensions-intl-hnk-dst/index.html>

²⁷ https://thewire.in/communalism/attack-on-elderly-muslim-man-over-beef-suspicion-court-cancels-bail-granted-to-cow-vigilantes/?mid_related_new

²⁸ <https://m.thewire.in/article/communalism/hindutva-group-attacks-muslims-in-ghaziabads-shanties-to-avenge-anti-hindu-violence-in-bangladesh/amp>

²⁹ <https://maktoobmedia.com/india/they-went-to-every-house-asking-if-residents-were-muslims-or-hindus-eyewitness-of-ghaziabad-attack-on-muslims/>

³⁰ <https://maktoobmedia.com/india/they-went-to-every-house-asking-if-residents-were-muslims-or-hindus-eyewitness-of-ghaziabad-attack-on-muslims/>

³¹ <https://www.bbc.com/news/articles/clywww69p2vo>



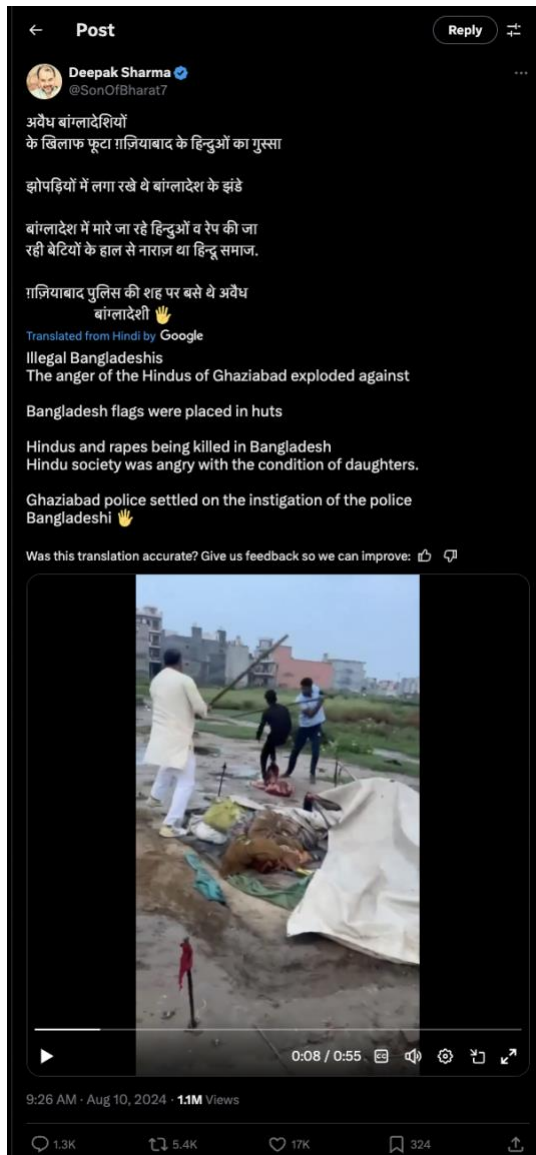
BJP’s Nitish Rane reportedly called for outright murder, tweeting: “*If Hindus are targeted and killed in Bangladesh, why should we allow even one Bangladeshi to breathe here. We will also target and kill.*”³² Though at the time of writing the tweet appears to have been deleted.

Another BJP hardliner, Kangna Ranaut, in her X post drew comparisons between India and Israel, saying that “*now we are also covered by extremists. We must be ready to protect our people and our land.*”³³ “*Peace is not air or sunlight that you think is your birthright and will come to you for*

³² <https://tribune.com.pk/story/2487292/hindu-nationalist-vigilantes-attack-muslim-settlement-in-india>

³³ <https://x.com/KanganaTeam/status/1822122929242378665>

free...Pick your swords and keep them sharp, practise some combat form everyday,” said Ranaut’s post which at the time of writing had 1 million views.³⁴



Five days later HRD thugs, led by Daksh Chaudhary were seen abusing the Muslim residents of Delhi's Shastri Nagar district. “Go to your Bangladesh!” Chaudhary shouted in the video which was shared on social media.³⁵ (Chaudhary is known for hate speech and violence. In June 2024, he was arrested after the ballot for abusing the residents of the northern city of Ayodhya after BJP lost there.³⁶) The second assault by an HRD mob took place on 9 August 2020, when thirteen men attacked Muslim families in a shanty town of Ghaziabad, a city north of the capital Delhi.³⁷ This attack was also filmed and reportedly first shared on the group's social media accounts, and by others.³⁸ One video of the assault shared on X, garnered over 1 million views.³⁹

Other instances of violence include offline and online calls for the eradication of mosques by vigilante influencers like Preet Sirohi. Sirohi keeps filing complaints with local authorities, claiming mosques he wants demolished were built illegally. In June 2024, he was behind the campaign to demolish two mosques in Delhi in a week.⁴⁰

Sirohi’s hateful campaign has continued at the time of writing. In October 2024, Sirohi vowed to keep on demolishing mosques, claiming they are illegal and sharing posts with inflammatory language, calling for

³⁴ <https://x.com/KanganaTeam/status/1822122929242378665>

³⁵ https://x.com/alishan_jafri/status/1822188539158597744?t=FUIFQaWlbi0QxeYCCxjQA&s=19

³⁶ <https://www.livehindustan.com/videos/uttar-pradesh/daksh-chaudhary-who-slapped-kanhaiya-kumar-abused-the-voters-of-ayodhya-arrested-1-10175530>

³⁷ <https://m.thewire.in/article/communalism/hindutva-group-attacks-muslims-in-ghaziabads-shanties-to-avenge-anti-hindu-violence-in-bangladesh/amp>

³⁸ <https://m.thewire.in/article/communalism/hindutva-group-attacks-muslims-in-ghaziabads-shanties-to-avenge-anti-hindu-violence-in-bangladesh/amp>

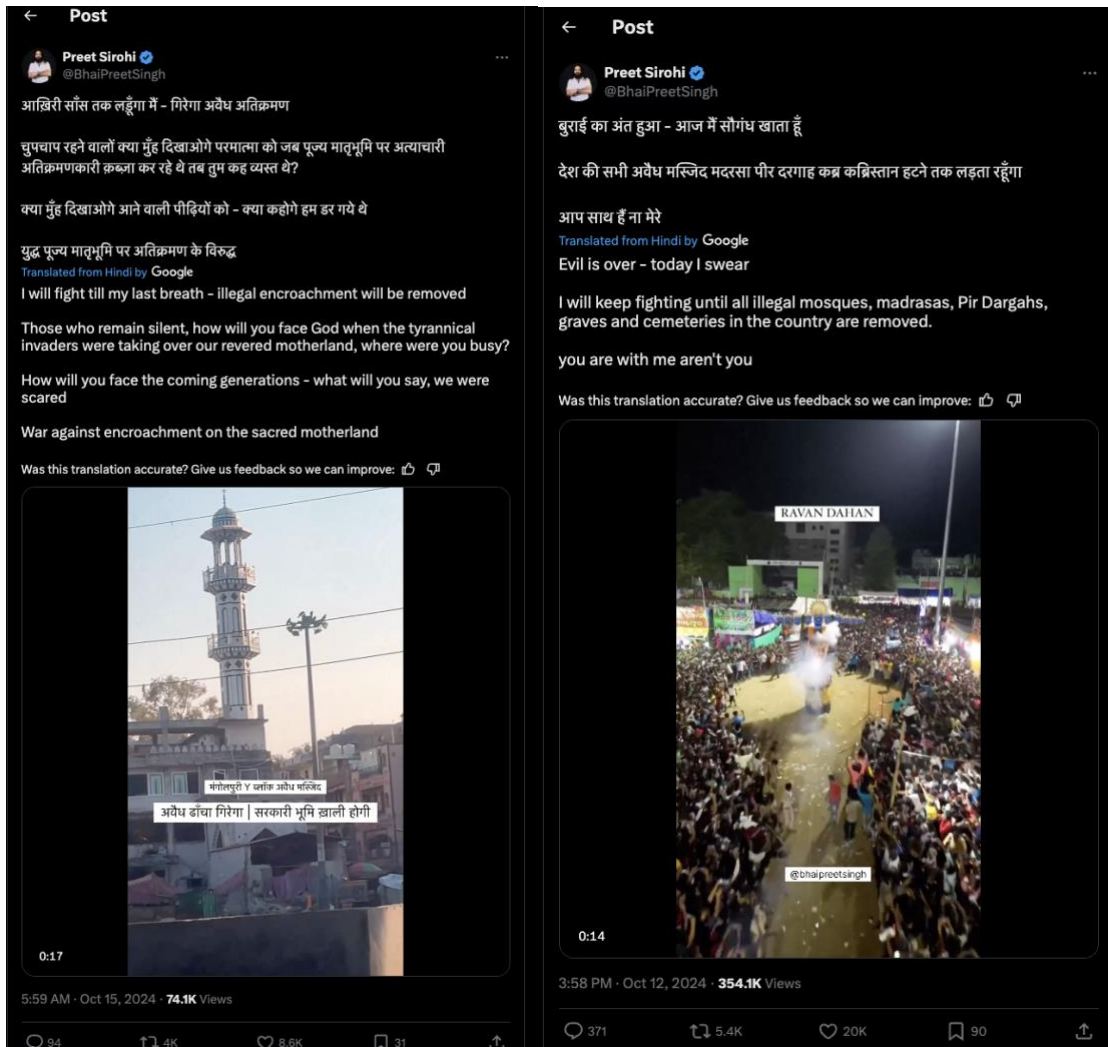
³⁹ <https://x.com/SonOfBharat7/status/1822172551067451633>

⁴⁰ <https://maktoobmedia.com/india/hindutva-hate-monger-claims-his-complaints-prompted-demolition-of-2-mosques-in-7-days-in-delhi/>

⁴¹ <https://x.com/theobserverpost/status/1803821493832622341>

⁴² <https://x.com/BhaiPreetSingh/status/1805534514451484776>

the destruction of Mosques.⁴³ “Those who remain silent, what face will you show to God, where were you busy when the tyrannical encroachers were occupying the revered motherland?”, his October 2024 tweet reads.⁴⁴



The details: How platforms helped foster a violent environment

Content moderation and policy enforcement

Modi peddled an islamophobic conspiracy theory around the so-called “vote jihad”, which echoed the alt-right “great replacement theory”⁴⁵ – implying Muslims were in India to replace

⁴³ <https://x.com/BhaiPreetSingh/status/1846038241675198744>

⁴⁴ <https://x.com/BhaiPreetSingh/status/1846038241675198744>

⁴⁵ <https://theconversation.com/modis-anti-muslim-rhetoric-taps-into-hindu-replacement-fears-that-trace-back-to-colonial-india-229570>

Hindus – further disenfranchising India’s Muslims and dissuading them from voting.⁴⁶ The narrative echoed across all social media platforms, with pro-BJP Facebook pages becoming a “ubiquitous part of Indian election campaigns as political parties and their leaders seek to directly connect with their followers,” according to Usha M. Rodrigues, a professor at Charles Stuart University.⁴⁷

Facebook was at the heart of a countrywide hate-filled campaign. Researchers examined posts shared in 812 Facebook pages and in 15 Facebook groups between 1 March and 10 May 2024.⁴⁸ In that sample alone, they found over 50 posts inciting hostility between Hindus and Muslims, with a large share of the content boosting Prime Minister Modi’s speeches.⁴⁹

"In most cases, it was only after large and blatant policy failures were made public through reports and press articles that companies took action. In multiple cases, that action came too late to protect the election," the diaspora organisation India Civil Watch International (ICWI) said.

While Meta’s blog post boasted having 20 local languages in India covered by content moderators – out of a total of 780⁵⁰ – that obscured the fact that the measures were neither new nor sufficient to protect the integrity of the elections.⁵¹ In fact, the tech companies’ election measures have remained largely the same over the years, “despite mounting evidence that they must be adjusted”, said corporate accountability watchdog The London Story.

Similarly, WhatsApp appears to have failed to implement its content moderation policies, allowing BJP-affiliated accounts to spread conspiracy theories, anti-Muslim rhetoric and hate speech free from public scrutiny⁵² (The party reportedly operates at least 5 million WhatsApp groups in the country.⁵³) This included messages glorifying police brutality, which spread like wildfire within minutes.⁵⁴

Failures of the tech platforms also resulted in electoral misinformation, with conspiracy theories and falsehoods about politicians and opposition candidates circulating on Facebook, WhatsApp YouTube, and X.

YouTube failed to remove policy-violating content, according to a report by ICWI, Dalit Solidarity

⁴⁶ <https://thelondonstory.org/wp-content/uploads/Vote-Jihad-Report-2024-V2.pdf>

⁴⁷ <https://time.com/6984947/india-election-disinformation-modi/>

⁴⁸ <https://thelondonstory.org/wp-content/uploads/Vote-Jihad-Report-2024-V2.pdf>

⁴⁹ <https://thelondonstory.org/wp-content/uploads/Vote-Jihad-Report-2024-V2.pdf>

⁵⁰ <https://www.bbc.com/news/world-asia-india-41718082>

⁵¹ <https://about.fb.com/news/2024/03/how-meta-is-preparing-for-indian-general-elections-2024/>

⁵² <https://restofworld.org/2024/bjp-whatsapp-modi/>

⁵³ <https://www.deccanherald.com/elections/india/political-theatre-bjp-on-social-media-2950186>

⁵⁴ <https://restofworld.org/2024/bjp-whatsapp-modi/>

Forum, Indian American Muslim Council, Hindus for Human Rights and Tech Justice Law Project.⁵⁵

The groups reported to YouTube that 26 videos, published between April and May 2024 by the far-right Sudarshan News YouTube channel, violated the platform’s hate speech and misinformation policies – they contained conspiracy theories, hate speech dehumanising Muslims and misinformation about the opposition Congress party.⁵⁶ Instead of taking the videos down, YouTube shared ad revenue with the channel.⁵⁷

The lack of meaningful engagement with civil society groups has been a hallmark of the 2024 election season in India and beyond. Multiple members of the Global Coalition for Tech Justice who have monitored the social media space in the leadup and during the ballot reached out to YouTube and Meta to act on violating and harmful content, and take it down. But in most cases, they said, platforms failed to take any action and when they did the hate speech or disinformation in question had already spread far and wide.

X’s crowdsourced fact-checking programme Community Notes, which allows X contributors to write context notes for misleading tweets, also proved to be a complete fiasco, leaving tweets with outright falsehoods without any notes or disclaimers.⁵⁸ This included a tweet by a verified X user, who falsely claimed that a non-existent Dubai-based Association of Sunni Muslims was financially supporting Muslims travelling to the southwestern state of Karnataka to unseat the BJP and Modi.⁵⁹ There was no Note accompanying the tweet, even though fact-checkers flagged it as containing false information.⁶⁰

Failure to enforce ad policies

Social media platforms, including YouTube, also failed to implement their own ad policies. Corporate accountability rights groups tested YouTube’s systems for ad review and ad approval mere days before the election kicked off, only to find out they didn’t work.⁶¹

Global Witness and Access Now submitted 48 ads in English, Hindi, and Telugu, all breaching YouTube’s advertising and election misinformation policies⁶² – the ads contained disinformation meant to suppress voter turnout among youth and women, and content inciting violence against the Muslim minority. (It was based on already existing falsehoods specific to India.) Although

⁵⁵<https://www.boomlive.in/decode/sudarshan-news-youtube-channel-islamophobia-hate-speech-misinformation-25591>

⁵⁶<https://www.boomlive.in/decode/sudarshan-news-youtube-channel-islamophobia-hate-speech-misinformation-25591>

⁵⁷<https://www.boomlive.in/decode/sudarshan-news-youtube-channel-islamophobia-hate-speech-misinformation-25591>

⁵⁸ <https://x.com/CommunityNotes/status/1775674773777100866>

⁵⁹<https://www.boomlive.in/fact-check/xs-community-notes-struggles-to-curb-misinformation-in-indian-election-25247>

⁶⁰<https://www.boomlive.in/fact-check/xs-community-notes-struggles-to-curb-misinformation-in-indian-election-25247>

⁶¹<https://www.globalwitness.org/en/campaigns/digital-threats/votes-will-not-be-counted-indian-election-disinformation-ads-and-youtube/>

⁶²<https://www.globalwitness.org/en/campaigns/digital-threats/votes-will-not-be-counted-indian-election-disinformation-ads-and-youtube/>

YouTube purportedly reviews all ads prior to publication, it approved all the ads, which the groups withdrew post approval by the platform.

YouTube was at pains to defend itself, saying the ads would have been reviewed by a moderator before going live.⁶³ But when Global Witness tested YouTube’s ad approval system in English and Spanish ahead of the 2022 US elections, the video-sharing platform rejected all the ads at the first stage and suspended the host channel.⁶⁴ Such glaring disparities in the treatment of harmful content by YouTube proves that the company has the means to enforce its policies, but it chooses not to do so in Global Majority countries, even in major markets like India – something civil society groups have been flagging for years.⁶⁵

Breaches of campaign financing laws

India’s campaign financing laws may have been violated due to lack of transparency across social media platforms.⁶⁶

Political advertising in India, including on social media, is subject to a number of content and financing regulations.⁶⁷ To skirt around the rules and evade the scrutiny of the Election Commission, political parties and actors allegedly resorted to surrogate and shadow advertising, thus undermining the integrity of India's ballot.⁶⁸

Surrogate and shadow advertising involves an individual or a group placing social media ads on behalf of a candidate or a political party without disclosing that the ads are directly funded by said party.⁶⁹ Surrogate advertisers are typically identifiable by their tax number but shadow advertisers are less transparent and further obfuscate who’s behind purchased advertisements.⁷⁰

Corporate accountability rights groups Eko, The London Story and ICWI reviewed Meta’s Ad library in the leadup to India’s vote and uncovered a network of 22 far-right shadow advertisers, who purchased ads with memes and videos attacking the opposition, fomenting violence against the Muslim minority and promoting the BJP⁷¹. The adverts amassed more than 10 million interactions in over 90 days and appeared to breach Meta’s own content and transparency

⁶³ <https://www.aljazeera.com/news/2024/5/31/in-shadows-meta-youtube-black-market-clouds-india-election-integrity>

⁶⁴ <https://www.globalwitness.org/en/campaigns/digital-threats/votes-will-not-be-counted-indian-election-disinformation-ads-and-youtube/>

⁶⁵ <https://www.globalwitness.org/en/campaigns/digital-threats/votes-will-not-be-counted-indian-election-disinformation-ads-and-youtube/>

⁶⁶ https://thelondonstory.org/wp-content/uploads/Slander-Lies-and-Incitement_Report-2-compressed.pdf

⁶⁷ <https://indianexpress.com/article/explained/explained-govts-new-guidelines-banning-surrogate-ads-7974028/>

⁶⁸ <https://indianexpress.com/article/explained/explained-govts-new-guidelines-banning-surrogate-ads-7974028/>

⁶⁹ <https://indianexpress.com/article/explained/explained-govts-new-guidelines-banning-surrogate-ads-7974028/>

⁷⁰ <https://indianexpress.com/article/explained/explained-govts-new-guidelines-banning-surrogate-ads-7974028/>

⁷¹ <https://www.boomlive.in/fact-check/xs-community-notes-struggles-to-curb-misinformation-in-indian-election-25247>

policies.⁷²

The 22 shadow pages spent over \$1 million on advertisements for BJP and in particular Prime Minister Modi, which accounted for almost 22% of the total sum of election advertisements during a 90 day period before the ballot.⁷³

The researchers couldn't establish who owned the pages due to lack of verifiable information.⁷⁴ The pages had no active phone lines and the researchers' emails to addresses indicated on the shadow pages remained unanswered.⁷⁵

Meta's failures to implement its own policies on political advertising go beyond the national election. A group of rights groups, including Global Coalition for Tech Justice member ICWJ, documented how Meta has allowed BJP and its shadow pages to violate India's electoral law and Meta's own policies ahead of the November 2024 local elections in the north-eastern state of Jharkhand.⁷⁶ The researchers identified at least 87 shadow pages purchasing ads spreading BJP narratives in Meta.⁷⁷ These pages have been churning out nearly five times more ads and garnering four times more impressions than the official BJP Jharkhand page.⁷⁸

AI-generative content and system failures

It should be noted that, despite widespread fears that AI-generated content would inundate social media with falsehoods, in India deepfakes were predominantly used to troll rather than to launch information warfare.⁷⁹

This was a blessing for tech platforms, which despite their pledges to prioritise and fight generative AI misinformation failed a simple stress test by civil society groups, which proved that some platforms can't detect AI-manipulated content.

In May 2024, corporate accountability groups Ekō and ICWI created inflammatory and Islamophobic political ads and submitted them to Meta's Ad library – the repository for ads published on Facebook and Instagram. Some of the ads contained AI-manipulated images⁸⁰ and

⁷² https://thelondonstory.org/wp-content/uploads/Slander-Lies-and-Incitement_Report-2-compressed.pdf

⁷³ https://thelondonstory.org/wp-content/uploads/Slander-Lies-and-Incitement_Report-2-compressed.pdf

⁷⁴ https://thelondonstory.org/wp-content/uploads/Slander-Lies-and-Incitement_Report-2-compressed.pdf

⁷⁵ https://thelondonstory.org/wp-content/uploads/Slander-Lies-and-Incitement_Report-2-compressed.pdf

⁷⁶ https://techjusticelaw.org/wp-content/uploads/2024/11/Report_Final_Web_2.pdf

⁷⁷ https://techjusticelaw.org/wp-content/uploads/2024/11/Report_Final_Web_2.pdf

⁷⁸ https://techjusticelaw.org/wp-content/uploads/2024/11/Report_Final_Web_2.pdf

⁷⁹ <https://restofworld.org/2024/exporter-india-deepfake-trolls/>

⁸⁰ <https://www.theguardian.com/world/article/2024/may/20/revealed-meta-approved-political-ads-in-india-that-incited-violence#:~:text=Revealed%3A%20Meta%20approved%20political%20ads%20in%20India%20that%20incited%20violence,-Exclusive%3A%20Ads%20containing&text=The%20Facebook%20and%20Instagram%20owner,shared%20exclusively%20with%20the%20Guardian.>

all of them were created based on real hate speech and disinformation prevalent in India.⁸¹ One ad, for example, used inflammatory language targeting Muslims and read: “Hindu blood is spilling, these invaders must be burned”.⁸²

Out of the 22 ads the groups submitted in English, Hindi, Bengali, Gujarati and Kannada, Meta approved fourteen – all containing generative AI images.⁸³ The ads were approved despite containing harmful content, violative of the platform policies, and despite the fact that Big Tech giants pledged to prioritise tackling harmful generative AI content. (The adverts were submitted halfway through the six-week voting period and withdrawn post approval, so that they would not be circulated.)

Researchers monitoring social media throughout the six-week voting period have also raised concerns about the lack of clarity around the fact-checking of political advertising across Meta’s platforms. Some ads that were flagged by fact-checkers, they said, remained on the platforms, while others did not contain labels indicating they had been fact-checked, creating further confusion.

⁸¹ https://aks3.eko.org/pdf/Meta_AI_ads_investigation.pdf

⁸² https://aks3.eko.org/pdf/Meta_AI_ads_investigation.pdf

⁸³ <https://www.theguardian.com/world/article/2024/may/20/revealed-meta-approved-political-ads-in-india-that-incited-violence#:~:text=Revealed%3A%20Meta%20approved%20political%20ads%20in%20India%20that%20incited%20violence,-Exclusive%3A%20Ads%20containing&text=The%20Facebook%20and%20Instagram%20owner,shared%20exclusively%20with%20the%20Guardian.>