



Global Network on Extremism & Technology

Polarising Narratives and Deepening Fault Lines: Social Media, Intolerance and Extremism in Four Asian Nations

Jordan Newton, Yasmira Moner, Kyaw Nyi Nyi and Hari Prasad

GNET is a special project delivered by the International Centre for the Study of Radicalisation, King's College London.

The authors of this report are Jordan Newton, Yasmira Moner, Kyaw Nyi Nyi and Hari Prasad, with contributions by Joseph Franco, Nazneen Mohsina and Cameron Sumpter.

The Indonesia country section was written by Jordan Newton, an independent research consultant and former security analyst for the Australian government. The Philippines country section was written by Yasmira Moner, acting director of IPDM. The Myanmar country section was written by Kyaw Nyi Nyi, a visiting fellow at Yusof Ishak-ISEAS. The India section was written by Hari Prasad, a research associate at Critica Research. CENS researchers Joseph Franco, Nazneen Mohsina and Cameron Sumpter conceptualised the project and provided editorial assistance.

The Global Network on Extremism and Technology (GNET) is an academic research initiative backed by the Global Internet Forum to Counter Terrorism (GIFCT), an independent but industry-funded initiative for better understanding, and counteracting, terrorist use of technology. GNET is convened and led by the International Centre for the Study of Radicalisation (ICSR), an academic research centre based within the Department of War Studies at King's College London. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing those, either expressed or implied, of GIFCT, GNET or ICSR.

CONTACT DETAILS

For questions, queries and additional copies of this report, please contact:

ICSR
King's College London
Strand
London WC2R 2LS
United Kingdom

T. **+44 20 7848 2098**
E. **mail@gnet-research.org**

Twitter: **@GNET_research**

Like all other GNET publications, this report can be downloaded free of charge from the GNET website at www.gnet-research.org.

© GNET

Executive Summary

The use of social media platforms and chat applications in Asia has grown exponentially in recent years. Throughout the 2010s, violent extremists (VEs) in different parts of the continent exploited this growing access to audiences, disseminating their divisive messages broadly, while targeting individuals in fringe online groups. Technology companies and governments eventually imposed relatively effective measures to moderate overtly terrorist content, remove accounts and limit reach. However, the dynamics of broader communication on platforms that reward contentious engagement is continuing to inflame domestic political polarisation and societal division.

Indonesia, the Philippines, Myanmar, and India are four Asian nations with unique but comparable experiences regarding the impact of online communications on social fault lines, extremism and violence. This report outlines and analyses these respective contexts.

Indonesian VEs on social media, particularly Islamic State (IS) supporters, are facing intense government and commercial pressure, placing more strain on their online recruitment efforts than ever before. Effective law enforcement in Indonesia and IS' decline in the Middle East have combined to weaken Indonesian IS supporters' reach and precipitated a decline in the quantity and quality of their online propaganda. But increasing government crackdowns on non-violent Islamist groups could yet provide VEs with new, disaffected audiences to tap for recruitment.

In the Philippines, VE propaganda is rooted in highly localised grievances involving perceptions of exclusion, discrimination and underdevelopment. Leading up to the Battle of Marawi city in 2017, Filipino-speaking VE groups created original content focusing on tactical dispatches, militant training activities, the celebration of militant leaders, and small victories against the military. Encouraging democratic engagement and highlighting positive narratives could dissipate ongoing exclusionary tendencies.

Anti-Muslim sentiment in Myanmar erupted following communal violence in the early 2010s, facilitated and accelerated by the concurrent rapid growth of Facebook, which for many became the default platform for accessing the Internet. What initially began as a 'buy Buddhist' campaign developed into the Ma Ba Tha movement in 2013, involving hundreds of Facebook accounts as nodes to spread hate. By 2018, Facebook had begun to moderate content and remove belligerent accounts in collaboration with Myanmar civil society organisations.

Social media platforms such as WhatsApp, Facebook and Twitter play important roles in intensifying Hindutva narratives in India. Complementing grassroots activities, the Hindutva movement exploits the online environment in three ways. First, social media serves to amplify narratives that Hindu identity is under threat from liberals and pluralists. Second, campaigns of online falsehoods demonise

perceived 'others', particularly Muslim communities. Finally, legions of online trolls seek to divert public attention from shortcomings in governance and manipulate narratives.

The four countries demonstrate the fallout when malign actors exploit social media and messaging platforms. Content moderation is the primary response from tech companies to deny space to extremist organisations, but its success depends largely on the domestic political environment. From Mindanao to Mumbai, various groups are able to circumvent terms of service restrictions. No technical intervention can resolve messaging campaigns against targeted out-groups. Polarising narratives, delivered either by memes or manifestos, are symptomatic of underlying divisions. Rather than looking at malign activities in the online space as a problem distinct from the offline world, it is more prudent to look at them as an extension of human experience.

Contents

Executive Summary	1
1 Introduction	5
2 Indonesia	7
The Rise and Decline of ISIL's Social Media Empire	7
Stubborn Resistance, New Potential Audiences	9
Repression and the Seeds of New Networks Online	9
3 The Philippines	13
DIL Propaganda Before and After the Battle of Marawi	13
The Broader Online Jihadist Milieu	15
Online Jihad: Recruitment via Facebook	15
The Way Forwards	16
4 Myanmar	17
Anti-Islam/Muslim Extremism	17
Anti-Rohingya Extremism	18
What to Do about Technologically Mediated Online Extremism	20
5 India	21
Creating a Hindutva-Friendly Environment	21
Amplification	22
Demeaning the Other	22
Distortion and Diversion	24
Challenges to Content Moderation	25
6 Conclusion	27
Policy Landscape	29

1 Introduction

Asia accounted for almost 60% of the world's online social networking users in 2020.¹ As in other parts of the world, different groups and actors throughout the continent have exploited these contemporary media methods to realise their extremist ambitions. Whether based on religious bigotry, sectarian contempt or ethnic hatred, social media accounts are deployed to recruit followers, exacerbate social fault lines and organise violence. Global technology companies are increasingly working with national governments in South and Southeast Asia to curb some of the most blatant abuse, but as some problems are managed, fresh issues emerge in the quickly evolving landscape.

This report aims to provide a concise update on the developing threats posed by actors with different extremist profiles in Indonesia, the Philippines, Myanmar and India. In recent years, these four nations have experienced violent extremism unfolding in separate forms but elevated and impacted by the same dramatic increase in social media uptake, particularly on mobile devices. India, Indonesia and the Philippines now comprise three of the top six Facebook user populations in the world, while Myanmar is in 19th place, despite ranking 82nd in global Internet penetration.² Hundreds of millions of new social media accounts have been generated in these four countries alone over the past few years.

While violent extremists have profited from this unprecedented audience expansion, content moderation mechanisms have eventually caught up following a range of attacks and organised violence closely associated with technology platforms. There is still much work to be done to keep extremists away from user accounts while limiting the impact on legitimate communications. But an even more pressing challenge is the tendency for social media platforms to encourage interactions and posts that foment societal polarisation and the subsequent demonisation of perceived 'others'. While terrorist propaganda may have disintegrated, pushed into the margins of online discourse, disinformation and toxic arguments in the mainstream could prove just as effective for violent extremist recruitment and radicalisation.

1 Clement, J. (15 July 2020), "Distribution of worldwide social media users in 2020, by region", *Statista*. <https://www.statista.com/remotexs.ntu.edu.sg/statistics/295619/regional-distribution-of-social-media-users-worldwide/>.

2 Clement, J. (24 November 2020), "Leading countries based on Facebook audience size as of October 2020", *Statista*. <https://www.statista.com/remotexs.ntu.edu.sg/statistics/268136/top-15-countries-based-on-number-of-facebook-users/>.

2 Indonesia

Increasing polarisation in Indonesian society is offering local violent extremists, particularly supporters of Islamic State (IS), a potential recruitment lifeline. Reflecting global trends, Indonesian social media is becoming more hotly contested between polarised communities – in this case, opposition Islamists and a coalition of pro-government forces – which is ratcheting up tensions. As increasing repression reduces the space for criticism of the government, members of some Islamist opposition movements could become more receptive to the argument offered by extremists that violence is the only path left for their struggle.

The Rise and Decline of ISIL's Social Media Empire

IS's embrace of social media, including in Indonesia, enabled the group to reach a larger and more mainstream audience than any extremist group ever had before. Historically, violent extremism online was largely corralled behind password-protected forums for carefully vetted supporters.³ By comparison, IS and its Indonesian supporters developed a prolific and vibrant presence across Facebook, Twitter and such messaging apps as Telegram and WhatsApp, posting propaganda videos and religious tracts, and engaging potential recruits in one-on-one chats. Online contact did not displace traditional face-to-face recruitment but supplemented it and provided a means for more intense, continuous and rapid radicalisation.

But IS supporters' online dominance was relatively short-lived as social media companies gradually made life difficult for extremists online. Twitter had some of the earliest and most comprehensive successes in 2015 and 2016,⁴ largely driving ISIL supporters off its platform. This precipitated an extremist migration to more secure options, notably Telegram. But even on Telegram, respite was only temporary, with a large-scale campaign by Europol 'trashing' IS's presence on the platform in 2019.⁵ IS in Indonesia also felt the heat, with vibrant groups that hosted as many as 8,000 members in 2017 dwindling to less than 200 by 2020.⁶ Supporters' inability to maintain an effective presence on one platform has forced them into a more dispersed, disjointed existence, spread out across various platforms.⁷

The symbiotic relationship between online and offline worlds meant that real-world setbacks for IS in Iraq and Syria were also a blow for its virtual presence, starving the group's supporters of new and engaging content to share online and attract recruits. Indonesian supporters

3 IPAC Report no.24 (30 October 2015), "Online Activism and Social Media Usage among Indonesian Extremists", *Institute for Policy Analysis of Conflict*. http://file.understandingconflict.org/file/2015/10/IPAC_24_Online_Activism_Social_Media.pdf.

4 Lakhani, Suraj, and Maura Conway (21 August 2017), *The Conversation*. <https://theconversation.com/islamic-states-twitter-network-is-decimated-but-other-extremists-face-much-less-disruption-82657>.

5 Winter, Charlie, and Amarnath Amarasingam (2 December 2019), "The decimation of Isis on Telegram is big, but it has consequences", *Wired*. <https://www.wired.co.uk/article/isis-telegram-security>.

6 IPAC Report no.48 (27 July 2018), "Indonesia and the Tech Giants Vs Isis Supporters: Combating Violent Extremism Online", *Institute for Policy Analysis of Conflict*. http://file.understandingconflict.org/file/2018/07/IPAC_Report_48.pdf.

7 Krona, Michael (21 October 2020), "Revisiting the Ecosystem of Islamic State's 'Virtual Caliphate'", *Global Network on Extremism and Technology (GNET)*. <https://gnet-research.org/2020/10/21/revisiting-the-ecosystem-of-islamic-states-virtual-caliphate/>.

online fared no better. Indonesia-focused materials from IS's central propaganda arm dried up as fighters from the archipelago fell in key battles in Mosul⁸ and Raqqa⁹, and official Indonesian-language translations of other IS material all but ceased. In place of IS's flashy propaganda videos, Indonesians now rely on a mix of homemade propaganda posters by 'unofficial' outlets, such as Ash Shaff Media, and grainy smartphone-recorded videos of pro-IS Eastern Indonesia Mujahidin militants holed up in the mountains of Central Sulawesi. Even these are scarcer, as frequent police counter-terrorism raids and generally inept attacks by IS supporters have left local militants with little to crow about in recent years.

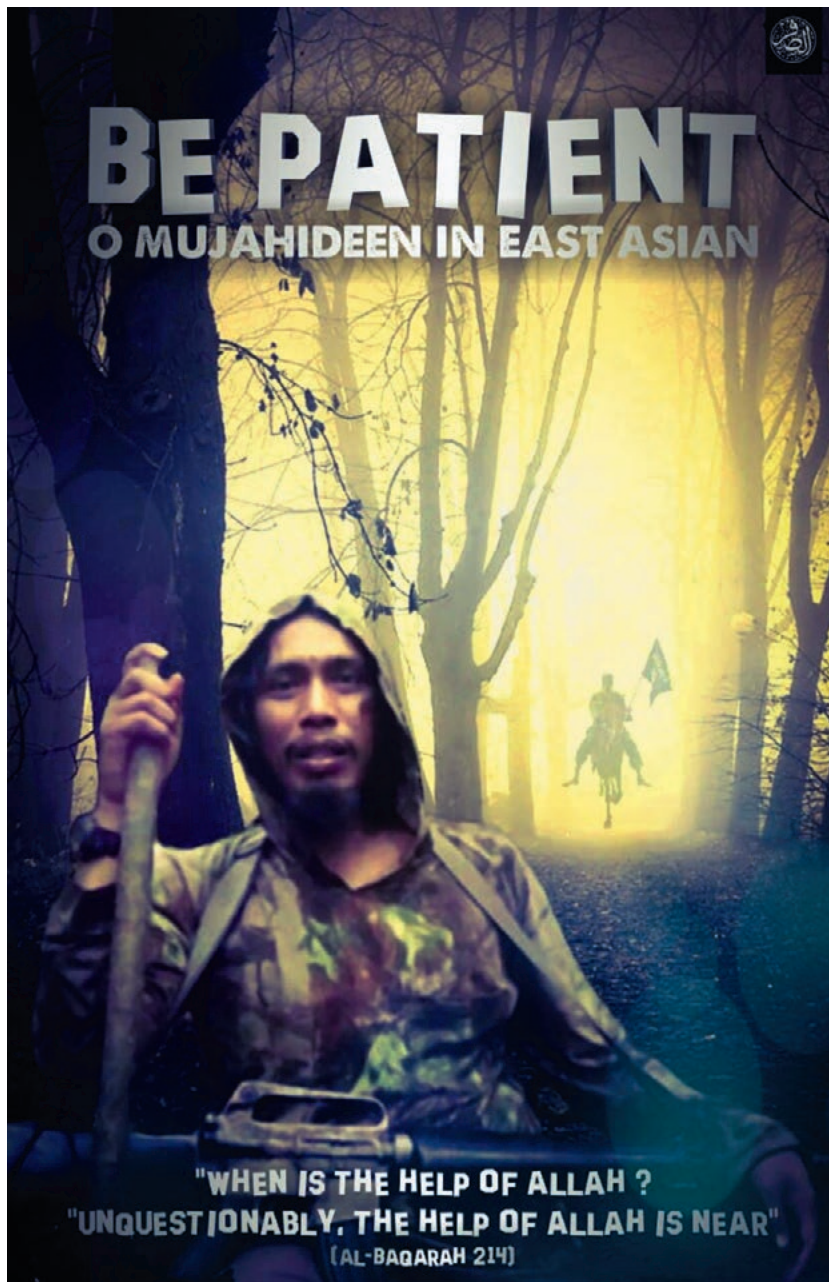


Figure 1: Poster produced by Ash Shaff Media (Telegram: 8 October 2020)

⁸ Rahayu, Heny, and Tia Asmara (9 November 2016), "Keluarga Benarkan Abu Jandal Tewas di Mosul", *Benar News*. <https://www.benarnews.org/indonesian/berita/abu-jandal-11092016165829.html>

⁹ BBC News (27 April 2018), "Benarkah Bahurmsyah, komandan tertinggi ISIS asal Indonesia tewas di Suriah?", *BBC News Indonesia*. <https://www.bbc.com/indonesia/indonesia-43889637>.

Stubborn Resistance, New Potential Audiences

Still, IS supporters continue to offer stubborn resistance online, employing a combination of deception and sheer determination consistently to subvert tech companies' best efforts to eliminate their virtual presence. One prominent Indonesian IS supporter currently in Syria's Al Hol refugee camp has been repeatedly banned by Facebook but stages regular comebacks, immediately welcomed by hundreds of friend requests from old acquaintances and new fans alike.¹⁰ Others are utilising relatively simple tactics such as Internet 'leet speak' (replacing letters with numbers and symbols) to disguise controversial terms and new platform features, such as time-limited 'stories' (image posts), to dodge both automated and human content-censors.

Though at their lowest ebb, both online and offline, in years, violent extremists may yet be able to regain some momentum by taking advantage of the increasing polarisation gripping Indonesia. Since 2017, political and social battle lines have been drawn between the opposition movement, spearheaded by conservative Islamist groups like the Islamic Defenders Front (FPI), and a pro-government 'militant pluralist' coalition, consisting of more secular-minded political parties, such as the Indonesian Democratic Party of Struggle, and mainstream Muslim groups, including Nahdlatul Ulama.¹¹ Much like other polarised polities elsewhere in the world, both sides have come to view the other as an existential threat to the state and themselves.

This sense of existential threat is triggering intense competition online and conditioning more Indonesians to see intimidation and violence as acceptable tactics for dealing with opponents.¹² 'Hashtag wars' between proponents of the two sides are common on Twitter,¹³ providing trigger points for online debates and offline violence. Islamist groups led a months-long campaign, fuelled by propaganda and disinformation online,¹⁴ to discredit the 2019 national election results, which they had lost. The campaign created a febrile environment and contributed to riots that broke out in Jakarta in May last year.

Repression and the Seeds of New Networks Online

Heavy-handed repression is ramping up pressure on Islamist groups but also increasing the risk that some may be pushed into the arms of extremists. Most notably, the deaths of six FPI members in an alleged shoot-out with police in December 2020 prompted fury from FPI supporters online, with at least some appearing to be more willing to resort to violence.¹⁵ One FPI supporter group posted a meme to almost 50,000 followers on Telegram calling on them to ready their weapons to fight back against the police.¹⁶

10 Facebook (November 2020), Ummu Azzam (pro-IS Indonesian woman in Syrian refugee camp).

11 Warburton, Eve (16 April 2019), "Polarisation in Indonesia: what if perception is reality?", *New Mandala*. <https://www.newmandala.org/how-polarised-is-indonesia/>.

12 Hui, Jennifer Yang, and Jordan Newton (2 October 2020), Global Network of Extremism and Technology (GNET). <https://gnet-research.org/2020/10/02/toxic-cybercampaigns-the-muslim-cyber-army-versus-militant-pluralism/>.

13 CNN Indonesia (21 May 2019), "KPU Umumkan Jokowi Menang Pilpres, Netizen 'Perang Tagar'", *CNN*. <https://www.cnnindonesia.com/teknologi/20190521072914-192-396717/kpu-umumkan-jokowi-menang-pilpres-netizen-perang-tagar>.

14 Potkin, Fanny (22 May 2019), "Indonesia curbs social media, blaming hoaxes for inflaming unrest", *Reuters*. <https://www.reuters.com/article/indonesia-election-socialmedia/indonesia-curbs-social-media-blaming-hoaxes-for-inflaming-unrest-idUSL4N22Y2UP>.

15 Chadajah, Rina (7 December 2020), "Indonesian Police Kill 6 Supporters of Firebrand Muslim Cleric in Road Encounter", *Benar News*. <https://www.benarnews.org/english/news/indonesian/id-rizieq-killings-12072020101536.html>.

16 Pro-FPI Telegram propaganda channel Angin Gunung (12 December 2020).



Figure 2: Pro-FPI poster encouraging supporters to ‘prepare your weapons, oh mujahid’ (Telegram: 11 December 2021)

To date, no violence has resulted from this incident. FPI followers have also refrained from retaliating over the government’s proscription of the organisation (prompted in part by the December shoot-out) and arrest of its leader, Rizieq Syihab. Still, IS supporters appear to be positioned to take advantage of any spike in militancy, with some musing in private chat groups that FPI supporters are now ‘ripe for recruitment’.¹⁷

A large-scale shift by the FPI or other Islamist groups to terrorism appears unlikely, but more frequent contact with IS supporters online increases the risk that at least some could see violence as the only path remaining for their struggle. Although they share some ideological touchstones, such as support for the establishment of an Islamic polity, FPI and IS supporters have traditionally been vehemently opposed to one another.¹⁸ Still, there is precedent for clusters of Islamists switching to terrorist networks: the entire FPI branch in Lamongan Regency, East Java, declared its support for IS in 2016 after a spiritual-leadership vacuum in the branch was filled by a renowned pro-IS preacher.¹⁹ IS supporters may seek to fill similar voids for disillusioned FPI members: some have already set up chat groups using Islamist naming conventions and iconography to encourage

¹⁷ IS sympathiser on Facebook (8 December 2020).

¹⁸ Some IS supporters even plotted to attack FPI leader Habib Rizieq Shihab. IPAC Report no.25 (1 February 2016), “Disunity among Indonesian ISIS supporters and the risk of more violence”, *Institute for Policy Analysis of Conflict*. http://file.understandingconflict.org/file/2016/04/IPAC_25_-_5.pdf.

¹⁹ IPAC Report no.18 (15 April 2015), “Indonesia’s Lamongan Network: How East Java, Poso and Syria are Linked”, *Institute for Policy Analysis of Conflict*. http://file.understandingconflict.org/file/2015/04/IPAC_18_Lamongan_Network.pdf.

more contact between themselves and FPI supporters.²⁰ If any FPI members do become radicalised, it is plausible that these kinds of chat groups will be an important gateway for that process.

The significant progress in recent years in reducing the reach of violent extremists could be undermined if a surge in polarisation provides a new crop of potential recruits online. Technological solutions will be vital for keeping a lid on any slide to violence: content moderation has played a key role in suppressing violent extremist voices and material advocating violence should continue to be the focus of efforts. But social media platforms' structure, which rewards sensationalism and virality – two things in which extremists of all stripes revel – and encourages ever greater fragmentation and the development of self-selected digital echo chambers, will also need to be addressed. More needs to be done to promote civil debate and shared spaces between competing views in society to avoid pushing people to the fringes, where violence becomes an attractive option.

²⁰ Pro-IS 'Alumni 411 & 212', Telegram chat group (13 December 2020).

3 The Philippines

In 2017, the centre of the Islamic City of Marawi in southern Philippines was razed to the ground by the self-proclaimed local Dawlah Islamiyah-Lanao (DIL) group, who pledged allegiance to IS. The so-called Battle of Marawi was foreshadowed by the online activity of DIL members and sympathisers. DIL propaganda is centred on the politics of identity, claiming that Christianised Filipinos marginalise Muslims. The othering of minority Muslim Moro populations can be traced back to the founding of the ethnonationalist liberation front, known as the Moro National Liberation Front (MNLF), in the 1970s.

The MNLF emerged in response to the politics of exclusion and economic marginalisation offered by the Manila-based national government, which enacted discriminatory land laws against the Moro and indigenous peoples, favouring Christian settlers from Luzon and the Visayas.²¹ The government of the Philippines entered into successive peace negotiations with the MNLF and its successor group, the Moro Islamic Liberation Front (MILF). It culminated in the creation of the expanded Bangsamoro political entity called the Bangsamoro Autonomous Region in Muslim Mindanao (BARMM) in 2019.

DIL Propaganda Before and After the Battle of Marawi

Extremist messaging in the Philippines is highly localised and connects with local grievances that often spring from the parochial interests of the community. Content expressing support for VE is commonly shared in Bangsamoro languages, particularly Maranao, Maguindanao and Tausug.²² Local sentiments are fuelled by the dissemination of viral messages of global Islamic channels capitalising on the 'Islam versus the West' discourse in global politics.

For one religious leader in Marawi City, the messaging of 'Islamic State Ranao', a moniker used by DIL in online recruitment, is this: 'We are now at war with the Khafirs [Infidels], hence the struggle of the Muslim nations is a battle between the "good" (Islamic world) and the "evil" (US-led democratic world order). "Jihad Fi Sabilillah" (War in the Cause of Islam) must be fought. It becomes a moral obligation for every believing men to fight the cause of Allah towards peace and the revival of the Khalifa Islamiyyah (Islamic State).'²³

DIL messaging on such apps as Telegram went beyond the repackaging of ideological content. Two channels moderated by Filipino-speaking users were active in producing original content between June and August 2016, with one focused on producing

21 Montiel, C. J., Rodil, R. B., & Guzman, J. M. d. (2012). "The Moro struggle and the challenge to peace building in Mindanao, Southern Philippines", in Landis, D. & Albert, R. D. (eds.), *International and cultural psychology. Handbook of ethnic conflict: International perspectives* (Springer Science + Business Media): pp.71–89. https://doi.org/10.1007/978-1-4614-0448-4_4.

22 The Asia Foundation and Rappler (2018). A Study on Understanding Violent Extremism: Messaging and Recruitment Strategies on Social Media in the Philippines.

23 Personal interview, Aleem Musa (alias) (5 September 2019), Barrio Salam, Mindanao State University-Marawi, Philippines.

tactical-level dispatches.²⁴ It was an indicator of preparations for the skirmishes between the Philippine military in late 2016 in Butig, a municipality south of Marawi. Content shared on the other channel focused on previous activities by the late Isnilon Hapilon before his pledge to IS in 2015.

Telegram content mostly depicted training activities by 'IS Lanao' forces [Figure 3] and displayed *ghanimah* (spoils of war) seized from government troops [Figure 4].²⁵ Curiously, the expected deluge of content from inside the Marawi battle zone also referred to as the 'Most Affected Area' (MAA) did not materialise. Only a few DIL-supplied clips were redistributed via IS-linked outlets, such as Amaq. These clips mostly depict the initial days of fighting.²⁶



Figure 3: IS Lanao photo depicting marksmanship training



Figure 4: Photo of materiel captured from assassinated enlisted personnel

24 Franco, Joseph (2017), "Assessing the feasibility of a 'Wilayah Mindanao'", *Perspectives on Terrorism* vol.11 no.4: pp.29–38.

25 Screenshots from IS Lanao Telegram accounts adapted from Franco, "Assessing the Feasibility", p.32.

26 Knight, Charles, and Katja Theodorakis (31 July 2019), "The Marawi crisis – urban conflict and information operations", *Australian Strategic Policy Institute*. <https://www.aspi.org.au/report/marawi-crisis-urban-conflict-and-information-operations>.

While short lived, content from inside the MAA was useful propaganda for the IS leadership in Iraq and Syria. IS magazine *Rumiyah* published a special issue entitled 'The Jihad in East Asia', which was followed by 'Inside the Khilafah', a series of propaganda videos. The narrative placed blame for the destruction of the city with the Armed Forces of the Philippines (AFP) and emphasised how the jihad in Marawi was a defensive campaign.²⁷

The Broader Online Jihadist Milieu

Beyond propaganda, extremist groups increased their recruitment efforts in local universities and high schools, and on social media sites. The new group of young, zealous, violent extremists have even challenged the MILF leadership by declaring that the leadership has abandoned the true spirit of jihad by negotiating with the Philippine government, a government of infidels.²⁸ Information from participants suggests that the DIL is viewed as more than just an organisation; it represents an ideology that creates a heroic narrative of an idealised version of the Ummah. Among prospective recruits, this narrative is either shared online through social media or spread by indoctrination in spaces called *halaqa* (study circles).²⁹

While Internet-mediated communication is not solely responsible for radicalisation in Muslim Mindanao, social media strengthens the existing offline social network, which contains locally rooted, legitimate grievances of exclusion, discrimination and underdevelopment. Social media is employed as a launch pad for extremists' political agenda and psychological manipulation, under the guise of legitimate religious activities.

Online Jihad: Recruitment via Facebook

Facebook is instrumental in the recruitment process as extremist recruiters seek to influence and persuade. In one personal conversation from a former member of the DIL, it was found that prospective recruits are young, male university or college students who are living in economically depressed communities in Mindanao or the BARMM. Female recruits exist but are relegated to acting as the support network for the male-dominated VE groups. Some of the female recruits joined reluctantly, becoming de facto members as wives or sisters of group members.³⁰ These types of members usually do not undergo indoctrination.

Recruiters utilised a combination of online recruitment, through private messages in closed Facebook groups, and offline engagement.³¹ In order to attract more followers, recruiters used charismatic young members well versed in reciting verses from scripture.³² Central

27 Franco, Joseph (29 November 2017), "Marawi: Winning the War After the Battle", International Centre for Counter-Terrorism – The Hague (ICCT). <https://icct.nl/publication/marawi-winning-the-war-after-the-battle/>.

28 Discussions with participants of a listening session on 6 November 2016 point to an incident during the previous year when a group of young individuals, who are allegedly members of the Maute, labelled Commander Bravo 'murtad' after the latter refused their invitation to join the cause. Commander Bravo heads the Northwestern Front of the Moro Islamic Liberation Front.

29 An Arabic term that means religious gathering purposely to study the Koran and Islam.

30 Personal interview, female former member of IS Lanao (19 December 2019), Marawi City.

31 Pantao, Jovar (2019), "Ecological Influences to Violent Extremism: Basis for the Development of Global Citizenship Education Framework", unpublished dissertation, School of Graduate Studies, Mindanao State University, General Santos City.

32 Field notes, Peace Formation Workshop (25 November 2019), MSU-Marawi City.

to their recruitment strategy are the themes of 'Jihad Fi Sabilillah' (fighting for the cause of Allah) and the Ummah. Da'wah (an invitation to the religion) is conducted through closed groups or private Facebook pages:

Firstly, we are asked about our personal background. Then, the person on the other side of the screen will start his indoctrination by citing cases of Muslim persecutions around the globe being a sign that Islam is under siege, hence the need the Islamic world (Ummah) is living in a perpetual war against the infidels. Moreover, the recruiter will continue by striking fear intimating that those who fail to join them will regret in the end. Indoctrination happens online, especially among the young member recruits. We also felt like we know each other because we speak the same language, not English, not Arabic, but my own language which is Meranao.³³

This private, small-group process serves to solidify confirmation bias. It also emphasises parochial loyalty to the religious and subnational ethnic affiliations. DIL online recruitment is successful because its social media accounts are an extension of its offline network. Individuals recruited online are employed to work on sharing messages translated in the local language/dialect of the community.

The Way Forwards

In the Bangsamoro Watch,³⁴ a supposedly open Facebook group that was initially conceived as a citizen-led watchdog of the BARMM, what emerges is a microcosm of the contested identities in the Bangsamoro. Group members include religious conservatives, progressive activists, and cynics promoting their own brand of exclusionary politics online. The polemics on the Facebook page highlight the need to promote the spirit of democratic engagement in the BARMM both online and offline.

The MILF-led Bangsamoro Transition Authority is attempting to champion the policy of inclusive and gender-fair autonomous government in the BARMM. Engaging the youth and female-led households in the Bangsamoro is imperative, as these groups are central to countering hate speech online by providing alternative messaging on diversity, religious toleration and cultural pluralism in order to dispel the notion of Islamophobia and exclusionary politics.

³³ Field notes, Peace Formation Workshop (19 December 2019), Iligan City.

³⁴ See <https://web.facebook.com/groups/821542374608074>.

4 Myanmar

Social media is usually considered not just a medium³⁵ but a breeding ground for extremism.³⁶ But technologically mediated extremism on social media platforms such as Facebook and Twitter should not be treated as a self-contained problem. Online extremism often precedes, accompanies or follows offline extremism, and extremism on the two platforms feed each other.³⁷

This section explores the relationship between extremism, groups and technological mediation. Myanmar has undergone a transition from military or military-dominated authoritarianism to partial electoral democracy since 2010 and is now reverting back to military rule in the aftermath of another military coup on 1 February 2021. Online/offline extremism and violence have emerged as a serious problem in the past eight years. Anti-Muslim extremism and anti-Rohingya extremism are the two most serious forms of extremism in Myanmar in transition.

Anti-Islam/Muslim Extremism

Anti-Muslim extremism emerged in October 2012 with the establishment of the symbolic 'buy Buddhist' 969 movement, in response to interreligious and intercommunal violence in Rakhine State. It grew into the larger and more mobilisational Ma Ba Tha (Organisation for the Protection of Race and Religion) movement in June 2013. Ma Ba Tha produced and propagated an extremist version of anti-Muslim-cum-pro-Buddhist extremism.³⁸ Emotional, 'reasonable' and conspiratorial, Ma Ba Tha clamored that Muslims intentionally buy only from Muslim-run shops to enrich their own community, while polygamous Muslim men plot to attract Buddhist women of humbler socioeconomic origins in order to marry and convert them to Islam and eventually have bigger families.³⁹ This Muslim scheme, it was argued, would eventually result in a demographic catastrophe and a smaller Myanmar Buddhist community.⁴⁰

All these extremist anti-Muslim discourses with alleged, unproven or exaggerated stories of Buddhist women badly treated and converted by Muslim men were relentlessly spread, both offline and online, by Ma Ba Tha monks, lay propagandists and followers from 2013 to 2015. The Thein Sein administration, then in power, and parliament sided with them and helped pass four race protection laws by August 2015.⁴¹

35 Klausen, Jytte (2015), "Tweeting the Jihad: Social Media Networks of Western Foreign Fighters in Syria and Iraq", *Studies in Conflict & Terrorism* vol.1 no.38: pp.1–22.

36 Vicario, Michela Del, Gianna Vivaldo, Alessandro Bessi, Fabiana Zollo, Antonio Scala, Guido Caldarelli, and Walter Quattrociocchi (29 June 2016), "Echo Chambers: Emotional Contagion and Group Polarization on Facebook", *Scientific Reports*. <https://arxiv.org/pdf/1607.01032v1.pdf>.

37 Wojcieszak, Magdalena (2010), "'Don't Talk to Me': Effects of Ideologically Homogeneous Online Groups and Politically Dissimilar Offline Ties on Extremism", *New Media & Society* vol.4 no.12: pp.637–55.

38 Kyaw, Nyi Nyi (2016), "Islamophobia in Buddhist Myanmar: The 969 Movement and Anti-Muslim Violence", in Crouch, M. (ed.), *Islam and the State in Myanmar: Muslim-Buddhist Relations and the Politics of Belonging* (Delhi: Oxford University Press).

39 *ibid.*

40 *ibid.*

41 The four laws ban polygamy, restrict interfaith marriage between Buddhist women and non-Buddhist men, hamper religious conversion and reduce family size and population growth, in response to alleged Muslim behaviour that threatens Buddhists and Buddhism.

Ma Ba Tha's extremism was mediated by Facebook, which became the most popular social media platform in Myanmar.

Hundreds of Facebook accounts belonging to Ma Ba Tha monks, such as U Wirathu, and his followers were involved in the extremist hate campaign,⁴² which was fed by offline activities. The extremist hate campaign materials they used included but were not limited to signature campaigns, journals, magazines, pamphlets, statements, books, street protests, monks' sermons, lay people's talks, conferences, public consultative workshops, press conferences, pictures, songs, life-story sessions and films — all delivered both offline and online.

Anti-Rohingya Extremism

Ma Ba Tha is a pro-Buddhist religious movement. Therefore, its extremism did not really distinguish between non-Rohingya Muslims and Rohingya Muslims. For the organisation, Islam and most Muslims, if not all of them, are a threat to Buddhists and Buddhism. However, there is a stark difference between non-Rohingya Muslims and Rohingya Muslims. While non-Rohingya Muslims (who account for approximately 2% of the population of Myanmar) are largely accepted as Myanmar citizens, Rohingya Muslims (2.3%) are almost universally perceived not as citizens but as 'illegal' Bengali migrants or infiltrators from Bangladesh.

Due to these doubly unacceptable qualities – being both Muslim and illegal – Ma Ba Tha was also extremely anti-Rohingya. Rohingya's alleged illegality has helped to recruit additional believers to the group, as well as producers and propagators of anti-Rohingya extremism. In general, 'Rohingya-phobia' has a range of versions of varying severity. Strong Rohingya-phobia is expressed by parties, such as the ethnonationalist Rakhine Buddhists and Ma Ba Tha, that 'totally' reject the Rohingya. The milder version is evidenced by people in Myanmar who question the veracity of Rohingya's claims to Myanmar citizenship or the sociocultural qualifications to be considered part of the nation of Myanmar due to ethnoreligious differences between the Muslim minority and the Buddhist majority.⁴³

Anti-Rohingya extremism in Myanmar emerged in the aftermath of communal violence in Rakhine State in June 2012 and became more widespread than ever after the Arakan Rohingya Salvation Army (ARSA) – formerly the Harakah al-Yaqin (Faith Movement) – attacked northern Rakhine State on 9 October 2016 and 25 August 2017, killing nine and twelve members of the Myanmar security forces on the respective occasions. The ARSA has been designated a terrorist organisation by the Myanmar government since the day of the second attack. Rohingya-phobia was mostly confined to Ma Ba Tha, Rakhine Buddhists and the Myanmar military until 2016. The ARSA's emergence and attacks during the rule of the National League for Democracy (NLD) government (March 2016–January 2021) has drawn in supporters of the chair of the NLD, Aung San Suu Kyi.

⁴² Based on the author's observation of such accounts from 2012 to 2016. Most if not all accounts have been either removed by Facebook or become inactive, as of late 2019.

⁴³ Cheesman, Nick (2017), "How in Myanmar 'National Races' Came to Surpass Citizenship and Exclude Rohingya", *Journal of Contemporary Asia* vol.3 no.47: pp.461–83.

Aung San Suu Kyi's supporters visibly and loudly defended her and the NLD government on Facebook and Twitter.⁴⁴ Eventually, such 'well-meaning' defence coalesced into a larger anti-Rohingya discourse. The discourse subsided in 2018, only to rise again in November and December 2019,⁴⁵ when the Republic of the Gambia sued Myanmar under genocide charges at the International Court of Justice (ICJ) and Aung San Suu Kyi had to appear at the court. Her supporters are in the majority in Myanmar and mostly believe that the military's heavy-handed actions against the ARSA resulted in the ongoing ICJ proceedings, whether or not those actions amount to genocide, war crimes or crimes against humanity.

Facebook, which was lamentably slow in responding to the tide of anti-Muslim extremism in 2012–15, found itself implicated in mounting international criticism of Myanmar's oppression of the Rohingya and the role its platform played in inciting and condoning violence and extremism against the Rohingya.⁴⁶ Facebook eventually admitted in November 2018 that it could have done better to monitor and remove online extremism and promised that it would.⁴⁷ Facebook removed 18 Facebook accounts, 52 pages and one Instagram account, including those of Myanmar Commander-in-Chief Min Aung Hlaing and the military-owned Myawady television network on 28 August 2018; 13 Facebook pages and ten accounts on 15 October 2018; 135 Facebook accounts, 425 pages, 17 Groups and 15 Instagram accounts on 18 December 2018;⁴⁸ four ethnic armed groups, including the Arakan Army, on 5 February 2019;⁴⁹ 89 Facebook accounts, 107 pages, 15 groups and five Instagram accounts on 21 August 2019;⁵⁰ and 13 Facebook accounts and ten pages on 12 February 2020.⁵¹

By early 2020, all these measures seem to have created a Facebook clientele in Myanmar that is more careful about what it posts and shares, especially if it is in relation to the Rohingya. Though not directly related to the Rohingya or their extremist adversaries, Facebook announced in August 2020⁵² that it would be on the look-out for election-related hate speech and fake news. Mostly in relation to elections, Facebook continued to remove Facebook accounts, pages, groups and Instagram accounts through 2020,⁵³ and worked with Myanmar civil society.⁵⁴

Due to Ma Ba Tha and likeminded groups' loss of influence, at least since 2018, the non-emergence of the Rohingya and the ICJ proceedings as an electoral issue, and the largely online nature of the electoral campaign from 8 September to 6 November 2020, extremist

44 The author's observation of NLD and Aung San Suu Kyi supporters' Facebook and Twitter accounts between August 2017 and December 2019.

45 *ibid.*

46 Miles, T. (12 March 2018), "U.N. Investigators Cite Facebook Role in Myanmar Crisis", *Reuters*. <https://www.reuters.com/article/us-myanmar-rohingya-facebook-idUSKCN1GO2PN>.

47 Warofka, A. (5 November 2018), "An Independent Assessment of the Human Rights Impact of Facebook in Myanmar", Facebook. <https://about.fb.com/news/2018/11/myanmar-hria/>.

48 Facebook (28 August 2018), "Removing Myanmar Military Officials from Facebook". <https://about.fb.com/news/2018/08/removing-myanmar-officials/>.

49 Facebook (5 February 2019), "Banning More Dangerous Organizations from Facebook in Myanmar". <https://about.fb.com/news/2019/02/dangerous-organizations-in-myanmar/>.

50 Facebook (21 August 2019), "Taking Down More Coordinated Inauthentic Behavior in Myanmar". <https://about.fb.com/news/2019/08/more-cib-myanmar/>.

51 Gleicher, N. (12 February 2020), "Removing Coordinated Inauthentic Behavior From Russia, Iran, Vietnam and Myanmar", Facebook. <https://about.fb.com/news/2020/02/removing-coordinated-inauthentic-behavior/>.

52 Frankel, R. (31 August 2020), "How Facebook Is Preparing for Myanmar's 2020 Election". <https://about.fb.com/news/2020/08/preparing-for-myanmars-2020-election/>.

53 Facebook (5 November 2020), "October 2020 Coordinated Inauthentic Behavior Report". <https://about.fb.com/news/2020/11/october-2020-cib-report/>.

54 Facebook (23 September 2020), "Additional Steps to Protect Myanmar's 2020 Election". <https://about.fb.com/news/2020/09/additional-steps-to-protect-myanmars-2020-election/>.

hate speech, whether anti-Muslim or anti-Rohingya, was scant and negligible, and some fake news about parties and candidates was visible on Facebook before, during and after the elections held on 8 November. Fake news did not spread as fast as it could have, in part because several fact-check programmes run by such Myanmar media outlets as Myanmar Now⁵⁵ and Mizzima,⁵⁶ as well as by civil society organisations, such as Myanmar ICT Development Organisations,⁵⁷ often in cooperation with Facebook, managed to check the spread of fake news and hate speech.

What to Do about Technologically Mediated Online Extremism

The case of Myanmar shows four aspects of technologically mediated extremism on social media. First, it occurs in the context of a broader political and social transformation, a democratic transition in the case of Myanmar. Second, online is not the only platform on which extremism manifests –online and offline realms are partners in crime. Third, it is not individuals but groups that engage in that project. Fourth, vigilance by social media platforms such as Facebook and by civil society and media organisations in countries play a significant role in reducing online extremism.

55 Myanmar Now (n.d.), "FACTCHECK". <https://www.myanmar-now.org/mm/topics/2705>.

56 Mizzima. (n.d.), "MISINFORMATION BUSTERS". <https://www.mizzimaburmese.com/news-category/32199>.

57 Hmue Angel (15 October 2020), "MIDO To Fact-Check Facebook News", *Myanmar Times*. <https://www.mmtimes.com/news/mido-fact-check-facebook-news.html>.

5 India

As India dealt with the coronavirus pandemic, two noticeable issues took place on Indian social media. First, a gathering of the Islamic group Tablighi Jamaat was blamed for spreading COVID-19, whereby pro-Hindu nationalist media and political parties escalated the incident into a fully fledged campaign⁵⁸ to demonise Indian Muslims.⁵⁹ The campaign resulted in the boycott of Muslim businesses and vendors,⁶⁰ the barring of Muslims from neighbourhoods⁶¹ and violence. Another fervent social media issue was a campaign to withdraw an advertisement showing an interfaith marriage between a Muslim and a Hindu. The advert was met with outrage among right-wing Hindu nationalists who alleged that this was an example of 'Love Jihad', a process in which Muslim men allegedly woo and trick Hindu women into marrying them to convert them to Islam.⁶² Subsequently, in several Indian People's Party-led (BJP) states, legislation to stop love jihad was introduced.⁶³

Both issues gained prominence partially due to a significant Hindu nationalist presence using both traditional and new media to spread misinformation.⁶⁴ Platforms like Facebook,⁶⁵ WhatsApp⁶⁶ and Twitter⁶⁷ have played important roles in amplifying Hindu nationalist narratives and misinformation across a wide domestic audience in India. This has complemented efforts by Hindutva organisations and political parties in the parliament and on the ground.⁶⁸

Creating a Hindutva-Friendly Environment

Hindutva, also known as Hindu nationalism, is a right-wing religious-nationalist ideology that maintains that India is and should be the homeland for Hindus above all religions. While Hindutva organisations have been present since pre-independence India, they operated at the fringe of Indian politics. However, Hindutva groups like the Rashtriya Swayamsevak Sangh (RSS), which formed in the 1920s, have built up strong grassroots support bases over the last

58 Banaji, S., and Ram Bhat (30 September 2020), "How anti-Muslim disinformation campaigns in India have surged during COVID-19", *LSE Blogs*. <https://blogs.lse.ac.uk/covid19/2020/09/30/how-anti-muslim-disinformation-campaigns-in-india-have-surged-during-covid-19/>.

59 Desai, S., and Amarnath Amarsingam (2020), "#CoronaJihad COVID-19, Misinformation, and Anti-Muslim Violence in India", *ISD*. <https://strongcitiesnetwork.org/en/wp-content/uploads/sites/5/2020/06/CoronaJihad.pdf>.

60 Al Jazeera (29 April 2020), "'Do not buy from Muslims': BJP leader in India calls for boycott", *Al Jazeera*. <https://www.aljazeera.com/news/2020/4/29/do-not-buy-from-muslims-bjp-leader-in-india-calls-for-boycott>.

61 Ellis-Petersen, H., and Shaikh Azizur Rahman (13 April 2020), "Coronavirus conspiracy theories targeting Muslims spread in India", *The Guardian*. <https://www.theguardian.com/world/2020/apr/13/coronavirus-conspiracy-theories-targeting-muslims-spread-in-india>.

62 Desai, S. (22 October 2020), "The Shadowy Online Hindutva World That Killed The Tanishq Ad", *Article 14*. <https://www.article-14.com/post/the-shadowy-online-hindutva-world-that-killed-the-tanishq-ad>.

63 Seth, M. (15 January 2021), "Explained: Uttar Pradesh's 'love jihad' law, and why it could be implemented vigorously", *The Indian Express*. <https://indianexpress.com/article/explained/explained-uttar-pradeshs-love-jihad-law-and-why-it-could-be-implemented-vigorously-7066156/>.

64 Poonam, S., and S. Bansal (1 April 2019), "Misinformation Is Endangering India's Election", *The Atlantic*. <https://www.theatlantic.com/international/archive/2019/04/india-misinformation-election-fake-news/586123/>.

65 Desai, S. (6 October 2020), "The Hateful Facebook Adventures Of Ragini Tiwari & Friends", *Article 14*. <https://www.article-14.com/post/the-hateful-facebook-adventures-of-ragini-tiwari-friends>.

66 McLaughlin, T. (12 December 2018), "How WhatsApp Fuels Fake News and Violence in India", *Wired*. <https://www.wired.com/story/how-whatsapp-fuels-fake-news-and-violence-in-india/>.

67 Chaturvedi, S. (2016), *I Am a Troll: Inside the Secret World of the BJP's Digital Army* (Juggernaut Books).

68 Ali, M. (14 April 2020), "The Rise of the Hindu Vigilante in the Age of WhatsApp and Modi", *Wired*. <https://www.wired.com/story/indias-frightening-descent-social-media-terror/>.

century. The RSS and other Hindutva organisations were propelled to national prominence in 1992 following the destruction of the Babri Masjid, an event that marked a turning point in Indian politics. Since the BJP took power in 2014, the party and its affiliated groups have tried to consolidate power and influence through local violent Hindutva organisations, pressure on the media and dominance in national politics.

Social media platforms have been a notable tool for the Hindutva movement. The BJP maintains an IT cell, a network of influencers and trolls to dominate Twitter and other social media platforms in which each platform serves specific purposes.⁶⁹ It is important to note that while the BJP maintains a significant, organised presence on these social media sites, they are just one part of the Hindutva ecosystem. Alongside its grassroots prowess, the online environment serves three important roles: the amplification of Hindutva, the demonisation of the 'other' and distortion in the public sphere.

Amplification

Social media is used to amplify narratives created and supported by Hindutva groups. It complements the overall message that the BJP already advocates at the grassroots level. For instance, there were at least 300,000 tweets with the hashtag #CoronaJihad, which relates to the conspiracy that Muslims were spreading COVID-19 to Hindus deliberately.⁷⁰ Another popular bogeyman is the love jihad phenomenon, which resurfaced after several BJP states promised to implement laws against interreligious marriages.⁷¹ Other popular topics of Hindu victimhood involve using the term 'Hinduphobia' to silence any criticism of Hinduism, Hindutva or the BJP.

This all combines to help push back against liberal and pluralistic values. Hindutva supporters fundamentally envision India as a majoritarian state, one in which Hindus are privileged. While independent India's foundations embraced pluralism and diversity, Hindutva supporters have seen this as disadvantaging the majority community and instead appeasing the minority religious communities. For them, India should have been the Hindu alternative to Islamic Pakistan; as such, the culture of the religious majority (as interpreted by them) should be enshrined. While it is doubtful that most Hindus will agree with the Hindutva ideal of a Hindu *rashtra* (nation), it has pushed the narrative that India must be a state that caters to the Hindu majority.

Demeaning the Other

One aspect of the Hindutva narrative that has gained international notoriety is its demonisation and violence towards the 'other'. Even though minority communities, such as Dalits and Christians, and political opponents are attacked on a regular basis, Muslims receive the brunt of the abuse from Hindutva supporters. Muslims who refuse to accept the Hindutva narrative are portrayed as fifth columnists.

69 *ibid.*

70 Perrigo, W. (3 April 2020), "It Was Already Dangerous to Be Muslim in India. Then Came the Coronavirus", *Time*. <https://time.com/5815264/coronavirus-india-islamophobia-coronajihad/>.

71 Srivastava, P. (22 November 2020), "Draft UP law on 'love jihad' proposes 5-yr jail, marriage annulment for 'forced conversion'", *The Print*. <https://theprint.in/judiciary/draft-up-law-on-love-jihad-proposes-5-yr-jail-marriage-annulment-for-forced-conversion/549637/>.

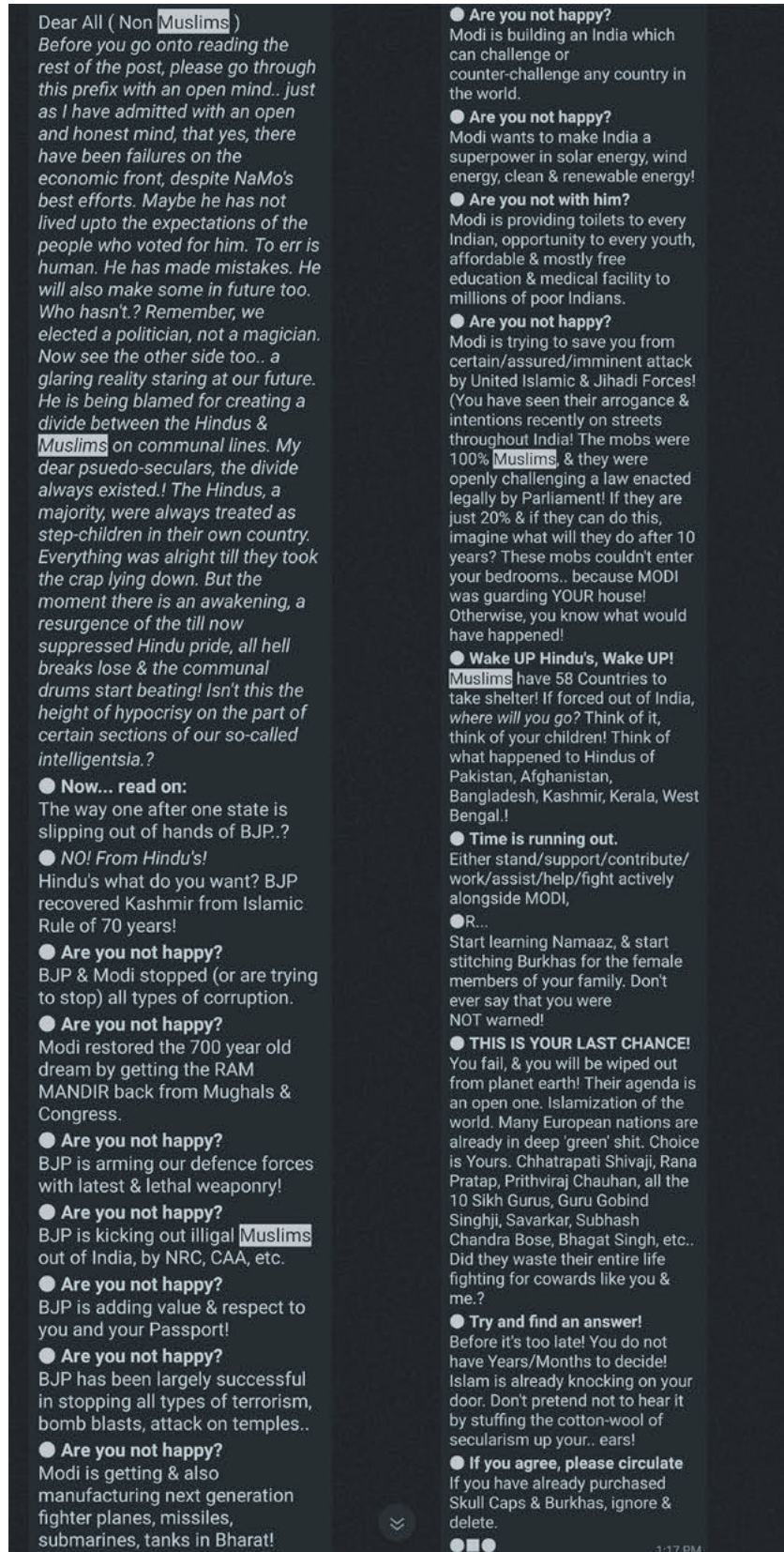


Figure 5: WhatsApp message during the anti-CAA protests.

On social media, false stories are spread about Muslim intentions towards Hindus. Recently, as COVID-19 spread in India, fact-checking websites have documented several instances of fake videos of Muslims allegedly spitting on Hindu food in an attempt to spread the virus.⁷² Even before the pandemic, pictures of weapons supposedly discovered in mosques were spread around social media to suggest that Muslims have violent intentions.⁷³ Although the videos of Muslims spitting, weapons in mosques and related to love jihad have been repeatedly debunked, social media has allowed these narratives to prosper.⁷⁴ An example of this can be seen in the screencap below of a WhatsApp message during the anti-CAA protests in early 2020.

These narratives put forward two messages. First, they emphasised the apparent danger that Muslims and other minorities pose to society. Second, they implied a more obscure fault, that pluralists and liberals have refused to address this, instead appeased Muslims. The latter message is particularly important, as it is a jab at the BJP's various opponent political parties that are often attacked by Hindutva organisations for being 'pseudo-secular' or appeasing minorities.

Distortion and Diversion

Using a sizeable social media presence allows Hindu nationalists to dominate or at the very least contest the public space. This happens in three ways: creating a hostile online environment for critical journalists and academics, spreading disinformation and the diversion of attention.⁷⁵ All of this is aided by a sizeable troll army that works to try to control the narrative on social media platforms.⁷⁶ The troll army can threaten and harass journalists and academics critical of the BJP-led government.

Trolls spread disinformation and pro-Hindutva messages on various social media platforms. This often follows the themes and topics discussed above. This is noticeably prevalent when certain issues are receiving increased attention. Take for example the disinformation around love jihad. As legislation to combat 'love jihad' was being considered, pro-Hindutva Internet users shared fake stories, videos, or images of Muslim men allegedly wooing Hindu women, forcibly converting them, or killing them.⁷⁷ As fact checkers and journalists worked to debunk this disinformation,⁷⁸ these fake incidents were used to spread the 'necessity' of the Love Jihad legislation.

Along with distorting the public sphere, the online environment also offers an opportunity to divert attention. An illustration of this during the COVID-19 crisis can be seen in the suicide of a prominent

72 Patel, J. (29 April 2020), "Old video falsely shared as Muslims spitting on relief food during lockdown", *Alt News*. <https://www.altnews.in/old-videos-falsely-shared-as-muslims-spitting-on-relief-food-during-coronavirus-lockdown/>.

73 Sidharth, A. (24 July 2019), "Old photos, false claim: Weapons recovered from a mosque in Gujarat." *Alt News*. <https://www.altnews.in/old-photos-false-claim-weapons-recovered-from-a-mosque-in-gujarat/>.

74 The Federal (24 November 2020), "'Love jihad' doesn't exist, says UP SIT after two months of investigation". <https://thefederal.com/states/north/uttar-pradesh/love-jihad-doesnt-exist-says-up-sit-after-two-months-of-investigation/>.

75 Sundaram, R. (2020), "Hindu nationalism's crisis machine", *HAU: Journal of Ethnographic Theory* vol.3 no.11: pp.734–41. https://www.journals.uchicago.edu/doi/10.1086/712222#_i3.

76 Guinee, S. (21 May 2019), "Journalists fighting fake news during Indian election face threats, abuse", *Committee to Protect Journalists*. <https://cpj.org/2019/05/india-elections-online-harassment-female-journalists/>.

77 Dhawan, Himanshi and Anam Ajmal. (8 November 2020), "How the Myth of Love Jihad is Going Viral" *Times of India*. <https://timesofindia.indiatimes.com/home/sunday-times/how-the-myth-of-love-jihad-is-going-viral/articleshow/79111670.cms>

78 Archit. (20 November 2020), "Explicit Video of a Gym Workout Viral with False 'Love Jihad' Spin". *Alt News*. <https://www.altnews.in/explicit-video-of-a-gym-workout-viral-with-false-love-jihad-spin/>.

Bollywood actor.⁷⁹ The story gained prominence as India faced criticism over its handling of the COVID-19 response and with China's intrusion into Indian territory.⁸⁰ Instead of focusing on these issues, the story of the actor's suicide was promoted, with pro-BJP outlets focusing their prime time slots on running a media trial against the late actor's actress girlfriend for allegedly abetting his suicide.⁸¹ Famously, a commentator on a news panel was chastised when he brought up that this Bollywood drama came on the day when India's GDP had contracted and China has intruded was silenced by the anchor of the news show.⁸² At a time when the BJP could have faced harsh criticism over a variety of issues, its online media environment and aligned news agencies instead discussed an actor's suicide as the "biggest story" of our times. Pushing the story complemented the messaging of the BJP, which sought to divert the public's attention from governance issues.

Challenges to Content Moderation

Moderation of content by social media companies in India is particularly difficult. First, the sheer size and diversity of India means content is difficult to monitor. With 22 official and many more unofficial languages, moderation itself remains a serious challenge. This makes it easier for Hindutva messaging to flourish across platforms, including Twitter, Facebook and WhatsApp.

Second, social media corporations face their own internal struggles in India. The strength of the BJP in India has also influenced the moderating policies of social media executives. Facebook India's former top policy official refused to crack down on hate speech by BJP politicians due to her close connections to government officials,⁸³ concerns over backlash from local hate groups⁸⁴ and a fear of damaging the company's business prospects in India.

Signs even point to an unruly troll mini-ecosystem that is now turning on symbols of Hindutva. While it is not as organised or prevalent as the BJP-led troll armies, there are aspects of the Hindutva ecosystem that view Modi as a Muslim appeaser, sometimes even mockingly calling him 'Maulana Modi'. While marginal on social media, it is clear that figures that represent the extreme right of the Hindutva movement are starting to take more prominent roles. This includes Indian chief ministers, including Yogi Adityanath, and parliament member Pragya Thakur, whom other BJP members have previously criticised.⁸⁵

79 Ramachandran, N. (7 October 2020), "Is India's Media Frenzy Over Sushant Singh Rajput's Death a Bollywood Circus or a Political Distraction?" *Variety*. <https://variety.com/2020/biz/news/india-media-sushant-singh-rajput-death-1234795033/>.

80 Bhagat, S. V. (3 October 2020), "Despite Crises, India Fixes Its Gaze on a Bollywood Tragedy", *The New York Times*. <https://www.nytimes.com/2020/10/03/world/asia/india-bollywood-sushant-singh-rajput.html>.

81 Pandey, G. (10 September 2020), "Rhea Chakraborty: Why is Indian TV Obsessed with Sushant Singh Rajput's Death?" *BBC News*. <https://www.bbc.com/news/world-asia-india-54098615>.

82 Ibid.

83 Horwitz, Jeff, and Newley Purnell (30 August 2020), "Facebook Executive Supported India's Modi, Disparaged Opposition in Internal Messages." *The Wall Street Journal*. <https://www.wsj.com/articles/facebook-executive-supported-indias-modi-disparaged-opposition-in-internal-messages-11598809348>.

84 Horwitz, Jeff, and Newley Purnell (13 December 2020), "In India, Facebook Fears Crackdown on Hate Groups Could Backfire on its Staff." *The Wall Street Journal*. <https://www.wsj.com/articles/in-india-facebook-fears-crackdown-on-hate-groups-could-backfire-on-its-staff-11607871600>.

85 The Times of India (28 November 2019), "BJP condemns Pragya Thakur's remarks on Godse, drops her from defence panel". <https://timesofindia.indiatimes.com/india/godse-remarks-pragya-thakur-dropped-from-defence-panel-stopped-from-attending-meetings-of-bjp-parliamentary-party/articleshow/72271531.cms>.

As Hindutva social media continues to play a significant role in the Indian social media sphere, this opens the door for further extremists who view their counterparts as too moderate. Considering the messages being spread online, this is a cause of great worry for anyone interested in social harmony, especially in the world's largest democracy.

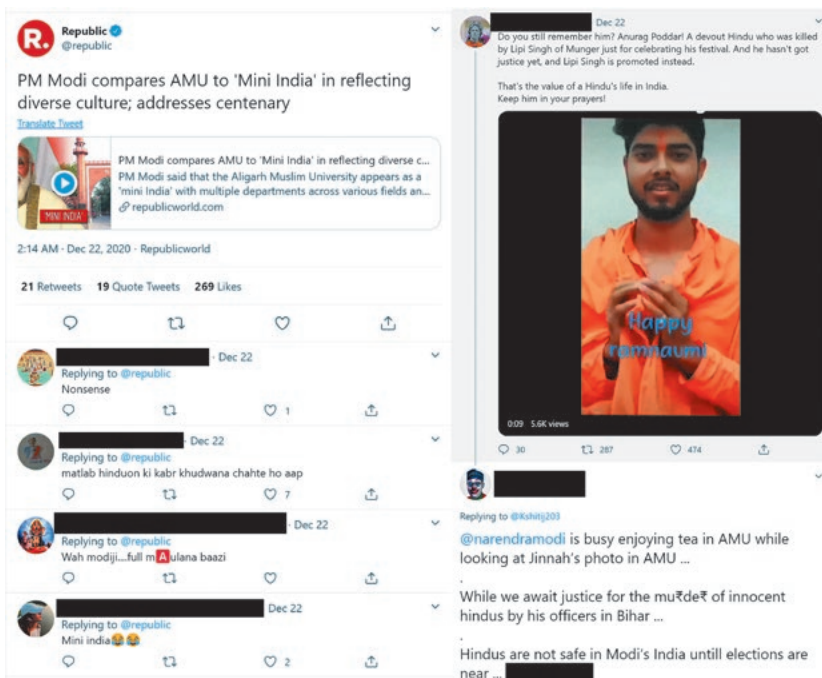


Figure 6: Twitter comments gathered on 23 December 2020. Republic is an Indian news station known for its pro-BJP bias. AMU is a predominantly Muslim university located in Aligarh, India.

6 Conclusion

The four countries featured in this policy paper demonstrate the spectrum of potential fallout when malicious actors exploit social media and instant messaging platforms. Indonesia, the Philippines, Myanmar and India may each have their distinct conflict histories but they exhibit similar patterns around how technology can exacerbate existing polarisation and violence.

Content moderation is often used by tech companies to deny space to both violent and non-violent extremist organisations, but its success depends largely on the domestic political environment. Extremist content pushed by organisations and individuals closely affiliated with governments or pro-government partisans are handled differently from anti-state extremist content. Hesitation among tech companies around whether to remove pro-BJP, pro-FPI or pro-NLD content is emblematic of the tension between applying global moderation policies and discerning what passes as permitted speech in specific national jurisdictions. From Mindanao to Mumbai, various groups are able to circumvent terms of service restrictions, whether they are violent extremists belonging to IS Ranao or political anti-BJP trolls. Restrictions are not made in a vacuum but are imposed against agile adversaries.

No amount of technical intervention can resolve the underlying issue of othering and demonisation against out-groups. Polarising narratives, whether through memes or lengthy ideological manifestos, are symptomatic of underlying ethnic, religious and class divisions. Electoral contests at all levels provide the incentive for political activists to use dehumanising speech in a deliberate but often cynical strategy to mobilise votes. In such a context, interventions, such as content removal, can serve only as fodder for further polarisation or even a justification for offline violence. Allowing a certain amount of online acrimony to exist may serve a cathartic function to forestall real-world harm.

Distinguishing between offline and online may ultimately be considered a pedantic exercise. Social media and instant communications platforms continue to be woven into the fabric of everyday life. The medium has become the message. Rather than looking at malign activities in the online space as a problem, it is more prudent to look at them as an extension of human experience.

Policy Landscape

This section is authored by Armida van Rij and Lucy Thomas, both Research Associates at the Policy Institute based at King's College London. It provides an overview of the relevant policy landscape for this report.

Introduction

Terrorism is not a new phenomenon. While it is almost impossible to verify when the first terrorist act took place in the history of world, some have traced one of the first acts of terror to 68 BC, when pirates attacked and set alight Ostia, Rome's port.⁸⁶ The use of violence for ideological and/or political gain is clearly not new, but the means by which terrorists recruit, manifest and organise are. The onset of social media and online platforms is increasingly cause for concern. Over the past decade, terrorists have utilised social media to pursue their own goals, activities that are now becoming increasingly restricted and illegal. Policymakers are struggling to keep up with technological developments that allow terrorists to carry out activities online.

In this report, we discuss the policy landscape and legislation in place in nine jurisdictions to tackle extremism, the spread of hate speech and other illegal content online. We discuss the challenges policymakers face in tackling online extremism, as well as some of the key stakeholders involved.

Social Media, Intolerance and Extremism: Addressing the Challenges and Assessing New Developments

Canada

In late January 2017, Québécois Alexandre Bissonnette opened fire on the Islamic Cultural Centre of Quebec City, killing six and injuring five. Ensuing scrutiny of his social media presence found that Bissonnette had been active in far-right and racist online circles prior to the shooting, regularly checking the Twitter accounts of conspiracy theorists, white nationalists and alt-right online personalities such as Ben Shapiro and Alex Jones of InfoWars.⁸⁷

The Canadian government increased its attention to and investment in countering violent extremism (CVE) online following Bissonnette's attack, which demonstrated the connections between online activity and real-world harm. Canada's national CVE efforts and initiatives

⁸⁶ <https://www.nytimes.com/2006/09/30/opinion/30harris.html>.

⁸⁷ Riga, A. (17 April 2018), "Quebec Mosque Killer Confided He Wished He Had Shot More People, Court Told", *Montreal Gazette*. <https://montrealgazette.com/news/local-news/quebec-mosque-shooter-alexandre-bissonnette-trawled-trumps-twitter-feed/>. See also: Mahrouse, G. (2018), "Minimizing and denying racial violence: Insights from the Quebec Mosque shooting", *Canadian Journal of Women and the Law* vol.30 no.3: pp.471–93.

form one part of a broader, holistic CVE policy. Its response to online violent extremism, as laid out in its National Strategy on Countering Radicalization to Violence,⁸⁸ is threefold: to develop counter-narratives in conjunction with civil society, to support academic research into effective policies and to co-sponsor international initiatives with tech company involvement.

In terms of supporting CVE research, in 2019 Canada commissioned Tech Against Terrorism, an international UN-sponsored initiative that works with the global tech industry, to develop a Terrorist Content Analytics Platform (TCAP), a database that hosts verified terrorist material and content from existing datasets and open sources.⁸⁹ The platform has the ability to act as a live alert facility for smaller Internet sites that may not have the capacity or resources to comply with regulatory efforts to remove malicious and extremist content.

The TCAP uses automated content moderation mechanisms based on machine learning and natural language processing, which use data analysis to train AI to recognise the elements of harmful content online (for example, IS logos, flags, etc.) so that future videos with the same or similar elements can be identified and flagged. The TCAP, as the first unified platform for online terrorist content, will act as an invaluable source of information for developers, academics and policymakers. In providing verified terrorist content from across platforms as a historical archive, the TCAP could provide a significant technological leap forward in CVE online. The Canadian government, as co-sponsor of the platform, has demonstrated how targeted and smart investment in cross-sector initiatives can provide opportunities for academia, industry and civil society to collaborate.

The Canadian government is party to a range of international and cross-sector initiatives. Following the Christchurch mosque attacks in March 2019, Prime Minister Justin Trudeau joined the Christchurch Call to Action, a joint commitment between governments and the tech industry to 'eliminate terrorist and violent extremist content online'.⁹⁰ Alongside co-sponsoring technical developments to help track and remove extremist content, such as the Global Internet Forum to Counter Terrorism (GIFCT) hash database,⁹¹ the call also commits governments to support frameworks, capacity-building and awareness-raising activities in order to prevent the use of online services to disseminate terrorist and violent extremist content.

European Union

The EU launched the Code of Conduct on countering illegal hate speech online in May 2016, together with four technology companies: Facebook, Microsoft, Twitter and YouTube. This initiative was an effort to respond to the rise in racist and xenophobic hate speech online. The purpose of the code is to ensure that requests to remove content are dealt with swiftly. It means companies commit to

88 "National Strategy on Countering Radicalization to Violence", Public Safety Canada. <https://www.publicsafety.gc.ca/cnt/rsrscs/pblctns/ntnl-strtg-cntrng-rdclztn-vlnc/index-en.aspx#s7>.

89 The TCAP has also been covered in the Policy Landscape section of a GNET report on 'Decoding Hate: Using Experimental Text Analysis to Classify Terrorist Content.' <https://gnet-research.org/wp-content/uploads/2020/09/GNET-Report-Decoding-Hate-Using-Experimental-Text-Analysis-to-Classify-Terrorist-Content.pdf>

90 See <https://www.christchurchcall.com/>.

91 See <https://www.gifct.org/joint-tech-innovation/>.

reviewing the majority of requests within 24 hours and to removing the content if necessary – all while respecting freedom of speech.⁹² Since its launch, other social media companies have signed up to the code as well, such as SnapChat and Facebook-owned Instagram. While this is laudable, there is no quantitative metric for what ‘the majority’ of content means. A second problem is that we know that terrorist content receives the majority of its hits within the first few hours of being uploaded. While these statistics may vary for content that is not terrorist-oriented, 24 hours may still be too large a window for inappropriate content to remain online.

In addition to the code, the European Commission is currently developing a wider regulatory framework on ‘the responsibilities of information society services to address all forms of illegal content online’.⁹³ Illegal content, according to the commission, includes incitement to terrorism, but also issues such as hate speech and child sexual abuse material. As part of this regulatory package, the commission has introduced the Digital Services Act, which will seek to ‘regulate the online ecosystem across a range of areas including [...] offensive content’.⁹⁴ In 2018, the commission adopted a ‘Communication on tackling illegal content online’, which served to establish guidelines and principles for online platforms to tackle illegal content online. The communication’s purpose is to ‘facilitate and intensify the implementation of good practices for preventing, detecting, removing and disabling access to illegal content so as to ensure the effective removal of illegal content’.⁹⁵ The commission also published a recommendation on measures to tackle illegal content online effectively. This recommendation proposed a common approach to detect, remove and prevent the re-appearance of illegal content online by: having clearer ‘notice-and-action’ procedures, equating to more transparent rules; deploying more efficient tools to remove illegal content and using proactive technologies; having stronger safeguards to ensure fundamental rights; paying special attention to small companies by sharing best practices and technological solutions; ensuring closer cooperation with relevant authorities.⁹⁶

France

In France, social media companies are now obliged to remove hate speech within 24 hours of it being uploaded or risk a fine.⁹⁷ The legislation forces technology companies to take down extremist content within an hour of receiving an order from the French police or face fines of up to 4% of global revenue. The French regulator, the Superior Council of the Audiovisual (CSA) has the power to impose fines on tech companies that breach this. A key challenge for some platforms is the re-uploading of previously detected and removed content. Yet in France, the trend set by the Supreme Court appears to be that there is no obligation for platforms to prevent the reappearance of content that has been previously removed –

92 See https://ec.europa.eu/commission/presscorner/detail/en/qanda_20_1135.

93 See https://ec.europa.eu/commission/presscorner/detail/en/qanda_20_1135/

94 Stolton, S. (23 June 2020), “Platform clamp down on hate speech in run up to Digital Services Act”, *EURACTIF*, <https://www.euractiv.com/section/digital/news/platforms-clamp-down-on-hate-speech-in-run-up-to-digital-services-act/>.

95 See <https://ec.europa.eu/digital-single-market/en/news/communication-tackling-illegal-content-online-towards-enhanced-responsibility-online-platforms>.

96 See <https://ec.europa.eu/digital-single-market/en/illegal-content-online-platforms>

97 [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/655135/IPOL_STU\(2020\)655135_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/655135/IPOL_STU(2020)655135_EN.pdf), p.68.

thereby implying a limited scope of duty of care for social media companies.⁹⁸ The notice-and-action procedures are limited to illegal content, which may include extremist content but may not capture all of it.

France approved law number 2018-1202 on the ‘fight against the manipulation of information’ in 2018. It aims to ‘better protect democracy against the different ways in which fake news is deliberately spread’.⁹⁹ The focus in France lies in particular on politically sensitive times around elections. During these periods, the law stipulates that, first, there is ‘a transparency obligation for digital platforms, who need to report any sponsored content by publishing the name of the authors and the amount paid. Platforms exceeding a certain number of hits a day must have a legal representative in France and publish their algorithms.’¹⁰⁰ Second, the ‘creation of a legal injunction allowing the circulation of fake news [is] to be swiftly halted’.¹⁰¹ Outside election periods, the law assigns a ‘duty of cooperation’ to social media companies and online platforms to tackle fake news.¹⁰² The CSA has been tasked with ensuring compliance with these measures. It also has the authority to ‘prevent, suspend and stop the broadcasts of television services that are controlled by foreign states’.¹⁰³

Ghana

Ghana has had very little experience of terrorist attacks,¹⁰⁴ and therefore the Ghanaian government has not needed to develop an aggressive CVE strategy. In comparison, Ghana’s regional neighbour Nigeria has struggled with major terrorist attacks for years. Extremist groups such as Boko Haram and Islamic State West Africa Province have launched shocking attacks such as the internationally infamous kidnapping of female students in April 2014¹⁰⁵ and the January 2015 massacres.¹⁰⁶ Boko Haram utilises a range of social media platforms, such as Twitter, Facebook, and Telegram, to publicise beheadings and kidnappings and spread anti-government messaging in an effort to recruit.¹⁰⁷ In response, the Nigerian government in early 2013 intensified its anti-terror laws and governance. As well as strengthening state counter-terrorism institutions, the government can now detain and prosecute terror suspects and issue the death penalty to those found to have committed or planned to commit a terrorist act.¹⁰⁸

98 European Commission (2018), https://ec.europa.eu/commission/sites/beta-political/files/soteu2018-preventing-terrorist-content-online-regulation-640_en.pdf, p.10.

99 Government of France, ‘Against information manipulation.’ <https://www.gouvernement.fr/en/against-information-manipulation>.

100 *ibid.*

101 *ibid.*

102 *ibid.*

103 *ibid.*

104 See also: Policy Landscape section in previous GNET report, ‘Artificial Intelligence and Countering Violent Extremism: A Primer’. <https://gnet-research.org/wp-content/uploads/2020/09/GNET-Report-Artificial-Intelligence-and-Countering-Violent-Extremism-A-Primer.pdf>.

105 Mbah, F. (2019), ‘Nigeria’s Chibok schoolgirls: Five years on, 112 still missing’, Al Jazeera. <https://www.aljazeera.com/news/2019/4/14/nigerias-chibok-schoolgirls-five-years-on-112-still-missing>.

106 Amnesty International (2018), ‘Boko Haram Baga attacks: satellite images reveal destruction’. <https://www.amnesty.org.uk/nigeria-boko-haram-doron-baga-attacks-satellite-images-massacre>.

107 UN Development Programme and RAND (2018), ‘Social Media in Africa’. <https://www.africa.undp.org/content/rba/en/home/library/reports/social-media-in-africa.html>.

108 ‘Nigeria: Extremism & Counter Extremism’, Counter-Extremism Project. <https://www.counterextremism.com/countries/nigeria>.

In terms of moderating violent extremist content online, therefore, Ghana's regional neighbour has opted for a traditional, state-centric and top-down mode of governance. This form of governance centres around punitive legislative measures with less emphasis on collaborative or educational cross-sector or civil society initiatives. This mode of state-centric governance can lead to unintended dangerous outcomes: for example, government shutdowns of the Internet or governmental use of social media to suppress political dissent.¹⁰⁹ Some leaders in Africa have exploited a legacy of violent colonial laws, historically used to violate freedoms against citizenry, to 'legitimise many [...] attempts to make extra-legal demands of the private sector'.¹¹⁰ Social media platforms and Internet service providers have had to respond to extra-legal government shutdown demands, raising concerns about censorship and violating freedom of expression.¹¹¹

There is much interest by civil society groups and journalists in how Ghana will regulate social media platforms as various recent events suggest that real-world harm could emerge from social media misuse.¹¹² For instance, the Ghanaian police chief announced a possible social media shutdown ahead of the country's 2016 elections, which fortunately was not carried out.¹¹³ Moreover, generous freedom of expression laws in Ghana mean that hate speech and cyberbullying (particularly of women) have been flourishing.¹¹⁴ As calls for tighter regulation of social media platforms become louder, the Ghanaian government should engage and co-produce any future CVE strategy with civil society and cross-sector initiatives.

Japan

Japan has two different approaches to CVE online, hinging around whether the terror threat is perceived to be foreign or domestic. Cold War-era activities to combat the ostensible communist threat have left an imprint on the way in which Japan approaches domestic threats. Responses are largely coordinated by law enforcement agencies. Prefectural police (overseen by the National Police Agency) and the Public Security Intelligence Agency, Japan's national intelligence agency, lead intelligence gathering and counter-terrorism efforts on Japanese soil.¹¹⁵

Traditional policing and security architectures, therefore, are mobilised in response to domestic terror activity online. Japan leads the world for innovative technological developments, which is reflected in its security strategy. The Japanese government has invested heavily in

109 Ilori, T. (2020), "Content Moderation Is Particularly Hard in African Countries", Information Society Project at Yale Law School. <https://law.yale.edu/isp/initiatives/wikimedia-initiative-intermediaries-and-information/wiii-blog/moderate-globally-impact-locally-content-moderation-particularly-hard-african-countries>.

110 Ilori, T. (2020), "Stemming digital colonialism through reform of cybercrime laws in Africa", Information Society Project at Yale Law School. <https://law.yale.edu/isp/initiatives/wikimedia-initiative-intermediaries-and-information/wiii-blog/stemming-digital-colonialism-through-reform-cybercrime-laws-africa>.

111 Ranking Digital Rights, "2019 RDR Corporate Accountability Index". <https://rankingdigitalrights.org/index2019/assets/static/download/RDRIndex2019report.pdf>.

112 Majama, K. (2019), "Africa in urgent need of a homegrown online rights strategy", Association for Progressive Communications. <https://www.apc.org/en/news/africa-urgent-need-homegrown-online-rights-strategy>.

113 Olukotun, D. (16 August 2019), "President of Ghana says no to internet shutdowns during coming elections", *AccessNow*. <https://www.accessnow.org/president-ghana-says-no-internet-shutdown-elections-social-media/>.

114 Endert, J. (2018), "Digital backlash threatens media freedom in Ghana", *DW Akademie*. <https://www.dw.com/en/digital-backlash-threatens-media-freedom-in-ghana/a-46602904>.

115 Kotani, K. (2013), 'A Reconstruction of Japanese Intelligence: Issues and Prospects', in Philip H. J. Davies & Kristian C. Gustafson (edd.), *Intelligence Elsewhere: Spies and Espionage Outside the Anglosphere* (Washington D.C.: Georgetown University Press): pp.181–99.

AI-led solutions, including large-scale facial recognition, biometric authentication and behaviour detection systems.¹¹⁶ These solutions suggest a governance model centred around early detection and prevention, operationalised through traditional police and security tactics.

Supporting these efforts, the former prime minister Shinzo Abe forced a 'brutal'¹¹⁷ anti-terror bill through parliament in mid-2017.¹¹⁸ The legislation criminalises planning to commit any of over 270 'serious crimes', which includes sit-in protests and music copyright infringements; its enforcement extends to social media.¹¹⁹ Civil rights activists have voiced concerned about the law, given its scale and the power it grants law enforcement in Japan to surveil.¹²⁰

With regard to counter-terrorism activities internationally, Japan's approach diverges dramatically from its domestic emphasis on criminalisation. Its overseas counter-terrorism efforts are regional, capacity-building and cooperative. The Association of Southeast Asian Nations (ASEAN) is the forum through which many of Japan's overseas counter-terrorism efforts are funnelled.¹²¹ ASEAN's declarations commit the signatories to 'prevent, disrupt and combat international terrorism through information exchange, intelligence sharing and capacity building', establishing a precedent for regional cooperation to counter violent extremism and terrorism.¹²² Japan has twice hosted the annual ASEAN-Japan Counter Terrorism Dialogue and has engaged in bilateral talks with a range of global actors.¹²³ For example, in late 2019, Japan and the UK held discussions on 'the current situation of international terrorism, domestic measures to counter terrorism, and also on current counter-terrorism capacity building cooperation particularly in third [sic] countries.'¹²⁴

It is not clear the extent to which Japan's regional efforts are performative, giving the government licence to pursue an approach based on security, policing and surveillance on its own soil. Any meaningful CVE strategy needs to take heed of civil society groups and respect citizens' rights to freedom and privacy, rather than camouflage efforts to erode these via a criminalisation approach.

-
- 116 The Government of Japan, "All is Ready for a Safe and Secure Tokyo Games", Autumn/Winter 2019. <https://www.japan.go.jp/tomodachi/2019/autumn-winter2019/tokyo2020.html>; NEC Corporation (2015), "NEC Becomes a Gold Partner for the Tokyo 2020 Olympic and Paralympic Games". https://www.nec.com/en/press/201502/global_20150219_01.html; Kyodo News (29 January 2018), "Kanagawa police eye AI-assisted predictive policing before Olympics". <https://english.kyodonews.net/news/2018/01/5890d824baaf-kanagawa-police-eye-ai-assisted-predictive-policing-before-olympics.html>.
- 117 Allen-Ebrahimian, B. (16 June 2017), "Japan Just Passed a 'Brutal,' 'Defective' Anti-Terror Law", *Foreign Affairs*. <https://foreignpolicy.com/2017/06/16/japan-just-passed-a-brutal-defective-anti-terror-law/>.
- 118 The Bill passed via 'the unusual step of skipping a vote in the Upper House Committee on Judicial Affairs'. Japan Federation of Bar Associations (15 June 2017), "Statement on the Enactment of the Bill to Revise the Act on Punishment of Organized Crimes and Control of Crime Proceeds, including the Criminalization of Conspiracy". <https://www.nichibenren.or.jp/en/document/statements/170615.html>.
- 119 McCurry, J. (15 June 2017), "Japan passes 'brutal' counter-terror law despite fears over civil liberties", *The Guardian*. <https://www.theguardian.com/world/2017/jun/15/japan-passes-brutal-new-terror-law-which-opponents-fear-will-quash-freedoms>; Adelstein, J. (15 June 2017), "Japan's Terrible Anti-Terror Law Just Made 'The Minority Report' Reality", *The Daily Beast*. <http://www.thedailybeast.com/japans-terrible-anti-terror-law-just-made-the-minority-report-reality>.
- 120 Japan Federation of Bar Associations (15 June 2017), "Statement on the Enactment of the Bill to Revise the Act on Punishment of Organized Crimes and Control of Crime Proceeds, including the Criminalization of Conspiracy". <https://www.nichibenren.or.jp/en/document/statements/170615.html>.
- 121 "Japan: Extremism & Counter Extremism", Counter-Extremism Project. <https://www.counterextremism.com/countries/japan>.
- 122 "ASEAN-Japan Joint Declaration for Cooperation to Combat International Terrorism", ASEAN. https://asean.org/?static_post=asean-japan-joint-declaration-for-cooperation-to-combat-international-terrorism-2.
- 123 "Japan: Extremism & Counter Extremism", Counter-Extremism Project. <https://www.counterextremism.com/countries/japan>.
- 124 Ministry of Foreign Affairs of Japan (4 December 2019), "The 4th Japan-the UK Counter-Terrorism Dialogue". https://www.mofa.go.jp/fp/is_sc/page1e_000297.html.

New Zealand

Similar to Canada (above), New Zealand's overarching counter-terrorism strategy shows that governance of CVE online involves the coordination of several agencies and bodies. These range from the Cabinet External Relations and Security Committee; to police, intelligence and security communications agencies; to foreign affairs, trade, defence, transport, innovation and development agencies.

New Zealand is looked upon favourably in the international sphere for its leadership in transnational and cross-sector initiatives. The Christchurch Call to Eliminate Terrorist and Violence Extremist Content Online, launched in the aftermath of the shocking Christchurch mosque shootings in March 2019, brings together a partnership of heads of state and social media and technology companies.¹²⁵ Protecting freedom of expression and privacy concerns are at the core of the call, along with robust legislative efforts to counter violent extremism online and frameworks, capacity-building and awareness-raising activities.

Since the shooter's rationale for the attack on the Christchurch mosque was so heavily formed by consuming radical content online, the Christchurch Call also holds companies – including Amazon, Facebook, Google, Twitter and YouTube – to greater industry standards of accountability and transparency. For instance, platforms must enforce their community standards and terms of service by prioritising content moderation and removal actions, as well as identifying content in real-time for review and assessment. Collectively, the partnership is developing efforts with civil society to promote community-led activities in order to intervene in the processes of online radicalisation.

The Christchurch Call also acted as the vehicle through which the GIFCT was overhauled. As part of the overhaul, GIFCT's remit expanded to include a suite of preventative, response and educational activities in the effort to counter violent extremism online.¹²⁶

New Zealand's efforts to co-sponsor a range of cross-sector global initiatives demonstrate a more horizontal approach to governing extremists' use of tech platforms. The approach encompasses conventional security and intelligence structures as well as initiatives that bring together practitioners, academia, policymakers and tech leaders to formulate responses to emerging violent extremist threats online.

¹²⁵ See <https://www.christchurchcall.com/>.

¹²⁶ Global Internet Forum to Counter Terrorism (23 September 2019), "Next Steps for GIFCT". <https://gifct.org/press/next-steps-gifct/>.

United Kingdom

The United Kingdom's approach to combatting extremist use of online platforms follows a traditional mode of governance that centres state institutions. The main agency responsible for counter-terrorism legislation is the Home Office, which also coordinates with the Government Communications Headquarters, the country's primary security and intelligence body. The Home Office has also created specific initiatives and bodies, such as the UK Council for Internet Safety, the National Counter Terrorism Security Office and the Commission on Countering Extremism, in partnership with other government institutions (most often the Department for Digital, Culture, Media, and Sport) and Parliament.¹²⁷

Along with Japan (above), the UK follows a twofold approach in its CVE strategy. The first approach is focused on conventional policing, security and intelligence institutions, bolstered by counter-terrorism legislation that enjoys strong public support. In Spring 2020, the British government introduced new proposed counter-terrorism legislation that targets suspects of terrorist activities. Under the new legislation, suspects 'who have not been convicted of any offense could potentially face expanded and increased surveillance measures',¹²⁸ which would no longer be subject to a two-year cap. Additionally, terrorism prevention and investigation measures (known as Tpims) – including mandatory relocation, electronic monitoring tagging, exclusion from specific places and limits on travel, association, financial services and use of communications – will now be easier to impose under the proposed reduction of burden of proof.¹²⁹

Regulating social media and technology platforms is the UK's second approach. In April 2019, the government published its Online Harms White Paper, which set out a robust case for national regulation of social media.¹³⁰ Under this new framework, social media and technology companies will bear a new statutory duty of care to their users, enforceable via Ofcom, the UK's regulatory body for communications. Ofcom will subject platforms to financial and technical penalties – websites could be blocked at ISP level and fined up to 4% of their global turnover – for non-compliance with the framework and violations of the statutory duty of care.¹³¹ At the time of writing, the Online Harms Bill, the legislative operationalisation of the White Paper, has been delayed for several years.¹³²

127 Gov.uk, UK Council for Internet Safety. <https://www.gov.uk/government/organisations/uk-council-for-internet-safety>; Gov.uk, Commission for Countering Extremism. <https://www.gov.uk/government/organisations/commission-for-countering-extremism>; Gov.uk, National Counter Terrorism Security Office. <https://www.gov.uk/government/organisations/national-counter-terrorism-security-office>.

128 "United Kingdom: Extremism & Counter Extremism", Counter-Extremism Project. <https://www.counterextremism.com/countries/unitedkingdom>.

129 Grierson, J. (20 May 2020), "Unconvicted terrorism suspects face indefinite controls under UK bill", *The Guardian*. <https://www.theguardian.com/politics/2020/may/20/unconvicted-terrorism-suspects-face-indefinite-controls-under-uk-bill>.

130 HM Government (April 2019), "Online Harms White Paper". https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/793360/Online_Harms_White_Paper.pdf.

131 Crawford, A. (29 June 2020), "Online Harms bill: Warning over 'unacceptable' delay", *BBC*. <https://www.bbc.co.uk/news/technology-53222665>.

132 *ibid.*

The attacks at Fishmongers' Hall in the City of London in November 2019 and Streatham High Road in February 2020¹³³ meant that public support for stricter counter-terrorism measures, like those described above, was boosted.¹³⁴ In this context, moderating intolerance and violent extremism may increasingly move away from regulatory efforts and come to rely purely on a criminalisation approach. Under the proposed bill, the burden of proof for subjecting a citizen to Tpims will be reduced to 'reasonable grounds'.¹³⁵ This generous legislation may mean that a greater number of UK citizens are at risk of being subject to counter-terrorism measures for accessing or spreading extremist content.

United Nations Counter-Terrorism Executive Directorate

In 2018, the UN Secretary General launched the UN Strategy and Plan of Action on Hate Speech to respond to a worldwide increase in hate speech. Its purpose was to address hate speech through a coordinated response that tackles the root causes and drivers of hate speech and highlights its impact on victims and societies.¹³⁶ It put together a 13-point action plan, called the Key Commitments, ranging from monitoring and analysing to using technology and education to engaging in advocacy and developing guidance, which together form a strategy to combat intolerance.¹³⁷

In addition to this, the UN also initiated action plans worldwide based in civil society to prevent violent extremism. The UN Alliance of Civilisations launched a #SpreadNoHate initiative to 'engage global media in a dialogue on hate speech and the sharing of best practices to promote counter narratives in media'.¹³⁸ The UN Development Programme has also released an action plan to address radicalisation and violent extremism, which includes policy dialogues, programme support, learning of best practices and a mechanism to prevent violent extremism grants.¹³⁹

133 Department of Justice, (20 May 2020), "Press release: 14-year minimum jail terms for most dangerous terror offenders". <https://www.gov.uk/government/news/14-year-minimum-jail-terms-for-most-dangerous-terror-offenders>.

134 In a September 2017 report that included polling on attitudes towards extremist content online, nearly three quarters of respondents said they would support new legislation criminalising the possession and consumption of extremist content online. See: Frampton, M. (2017), "The New Netwar: Countering Extremism Online", *Policy Exchange*. <https://policyexchange.org.uk/wp-content/uploads/2017/09/The-New-Netwar-1.pdf>.

135 Amnesty International UK (June 2020), "Counter-Terrorism and Sentencing Bill 2019–21: Submission to the Public Bill Committee". <https://publications.parliament.uk/pa/cm5801/cmpublic/CounterTerrorism/memo/CTSB07.pdf>.

136 [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/655135/IPOL_STU\(2020\)655135_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/655135/IPOL_STU(2020)655135_EN.pdf), p.28.

137 [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/655135/IPOL_STU\(2020\)655135_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/655135/IPOL_STU(2020)655135_EN.pdf), p.28; see https://www.un.org/en/genocideprevention/documents/UN%20Strategy%20and%20PoA%20on%20Hate%20Speech_Guidance%20on%20Addressing%20in%20field.pdf.

138 [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/655135/IPOL_STU\(2020\)655135_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/655135/IPOL_STU(2020)655135_EN.pdf), p.29.

139 <https://www.undp.org/content/dam/norway/undp-ogc/documents/Discussion%20Paper%20-%20Preventing%20Violent%20Extremism%20by%20Promoting%20Inclusive%20%20Development.pdf>, p.37.

United States

Prior to the Trump administration, the USA's national policy approach to CVE was uneven, irregular and deferential to protecting social media platforms' interests, but generally in good faith. The Department of Homeland Security, the Department of Justice, the Federal Bureau of Investigation, the National Counter Terrorism Center, the National Security Council and Congress, amongst others, are the agencies primarily responsible for the strategy.¹⁴⁰ It has included a range of activities, including 'counter messaging, awareness briefings, partnerships, and legislation'.¹⁴¹ The USA has co-sponsored various cross-national initiatives, such as Tech Against Terrorism and the Global Counterterrorism Forum, but involvement has been limited in deference to freedom of expression rights.

Given the sheer scale of the task, activities were often uncoordinated and ineffective. In 2011, the Obama administration established the Countering Violent Extremism Task Force in order to 'unify the domestic CVE effort'.¹⁴² The force is intended to bring together practitioners from the bodies listed above in order to coordinate engagement with civil society, develop intervention models, promote investments in research and cultivate communications and digital strategies.¹⁴³

However, the Trump administration has damaged the USA's CVE efforts and reputation both on home soil and overseas. In early 2017, Trump considered restructuring the task force to remove white supremacist terrorism from its remit and renaming the programme 'Countering Radical Islamic Extremism'.¹⁴⁴ Funding for activities engaging with communities and civil society, such as Life After Hate, an initiative that works with individuals to leave white supremacist and neo-Nazi groups,¹⁴⁵ was halted. In Spring 2017, the Trump administration cut all funding to CVE programmes.¹⁴⁶ By late October 2018, the task force had been shut down entirely.¹⁴⁷ These actions demonstrate the Trump administration's unwillingness to counter violent extremism online and its implicit support for white supremacist and racist terrorist actions.

The incoming Biden administration must establish a robust strategy to counter dangerous online content. As we have seen with the 6 January 2021 insurrection at the US Capitol, if violent extremist content can spread relatively unchecked online, the consequences will continue to be terrifying and tragic.

140 Alexander, A. (2019), "A Plan for Preventing and Countering Terrorist and Violent Extremist Exploitation of Information and Communications Technology in America", *George Washington University Program on Extremism*: p.5. <https://extremism.gwu.edu/sites/g/files/zaxdzs2191/f/A%20Plan%20for%20Preventing%20and%20Countering%20Terrorist%20and%20Violent%20Extremist.pdf>.

141 *ibid.*

142 Department of Homeland Security, "Countering Violent Extremism Task Force". <https://www.dhs.gov/cve/task-force>.

143 *ibid.*

144 Ainsley, J. et al. (3 February 2017), "Exclusive: Trump to focus counter-extremism program solely on Islam – sources", *Reuters*. https://www.reuters.com/article/idUSKBN15G5VO?feedType=RSS&feedName=topNews&utm_source=twitter&utm_medium=Social.

145 Life After Hate, 'About Us'. <https://www.lifeafterhate.org/about-us-page>.

146 Ainsley, J. (23 May 2017), "White House budget slashes 'countering violent extremism' grants", *Reuters*. <https://www.reuters.com/article/us-usa-budget-extremism-idUSKBN18J2HJ>.

147 Beinart, P. (29 October 2018), "Trump Shut Programs to Counter Violent Extremism", *The Atlantic*. <https://www.theatlantic.com/ideas/archive/2018/10/trump-shut-countering-violent-extremism-program/574237/>.

Deepening Fault Lines: Investigating Violent Extremism and Underlying Social Issues

In their report above, Jordan Newton, Yasmira Moner, Kyaw Nyi Nyi and Hari Prasad (hereafter Newton et al.) show how violent extremism online follows the contours of ‘underlying ethnic, religious and class divisions’. Taking India as an example, Hindutva (Hindu nationalist) narratives that have been mainstreamed into the state and enforced by violent paramilitary organisations have caused untold suffering and death among non-Hindu, particularly Muslim Indians.¹⁴⁸ State sponsorship of the demonisation, dehumanisation and marginalisation of minority populations is perhaps the greatest determinant of violent extremism online – as well as real-world harm offline.

In this section, we answer Newton et al.’s call to look at violent extremism online ‘as an extension of the human experience’ rather than as a technical issue to solve. We take the case of Brenton Tarrant, the Australian-born perpetrator of the Christchurch mosque shootings in March 2019, to examine some of the underlying racial and social divisions in Australia and how they contributed to the attack.

European imperial expansion to Asia, Africa and the Middle East, beginning in the late 15th and early 16th centuries, forced many indigenous peoples into contact with European colonisers and voyagers. After an initial Dutch encounter in 1605, English Lieutenant James Cook was charged by the monarchy to search for “a Continent or Land of great extent” in the south-west of the Pacific and “with the Consent of the Natives to take possession of Convenient situations in the Country in the name of the King”.¹⁴⁹ In 1770, Cook laid claim to an island in the Torres Strait Islands in Northern Australia, naming it Possession Island.¹⁴⁹

By this time, the idea of racial hierarchy had already entered the European intelligentsia and discourse. A so-called ‘Great Chain of Being’, which ranked humans, animals, plants and minerals in a hierarchical chain leading to God and angelic beings, formed the basis of modern taxonomy. Carl Linnaeus’ *Systema Naturæ*, published in 1737, subdivided the human species into four based on continent and skin colour, placing white Europeans at the top of the racial hierarchy.¹⁵⁰

These notions of property rights and racial supremacy shaped the British colonial encounter with Australia in 1788. When a British penal colony was established in Botany Bay, New South Wales, continental Australia was declared to be *terra nullius* (land owned by no-one) and therefore claimed by the British Crown.¹⁵¹ The settlers devastated the indigenous population, raping or slaughtering women, children and elderly people, as well as kidnapping young

148 Chandra, K. (13 June 2019), “How Hindu Nationalism Went Mainstream”, *Foreign Policy*, <https://foreignpolicy.com/2019/06/13/how-hindu-nationalism-went-mainstream/>.

149 Broome, J. (2010), *Aboriginal Australians: A History Since 1788* (Allen & Unwin); pp.15–18.

150 Müller-Wille S. (2014), “Linnaeus and the Four Corners of the World”, in Coles, K. A., Bauer, R., Nunes, Z. and Peterson C. L. (edd.), *The Cultural Politics of Blood, 1500–1900* (London: Palgrave Macmillan). See also: Hudson, N. (1996), “From ‘Nation’ to ‘Race’: The Origin of Racial Classification in Eighteenth-Century Thought”, *Eighteenth-Century Studies*, vol.29, no.3: pp.247–64.

151 This is in contrast to its colonies in North America, southern Africa, India and New Zealand whereby the British made treaties and agreements with indigenous populations. Broome: p.19.

children for labour.¹⁵² Diseases introduced by settler colonists, such as smallpox, influenza, tuberculosis and venereal diseases, also killed thousands of indigenous peoples.¹⁵³ Frontier massacres between colonists and the indigenous populations continued into the early 20th century.¹⁵⁴

When gold was discovered in Australia, the total population tripled due to immigration from China, Europe and the United States. Immigration from China, following earlier links with Macau and Canton as trading ports with the Australian colony, blossomed in the 1850s as thousands arrived to labour in the goldmines in Victoria. Settler Australians 'looked to the state to exclude the inferior races by law',¹⁵⁵ particularly Chinese immigrants, onto whom the British 'projected an array of racial, economic, cultural and strategic anxieties'.¹⁵⁶ Legislators in the state of Victoria implemented a head tax on Chinese people entering Australia and imposed strict limits on the number of Chinese immigrants that could legally be carried by ship. Similar provisions were implemented in other states in the 1850s and 1860s, culminating in what became known from the 1880s as the White Australia policy.

Once a federal, independent Australia was formed in 1901, the new government immediately passed the Immigration Restriction Act. Although the act could not explicitly legislate against Chinese, Indian and Japanese immigrants, it effectively did so by adopting a mandatory dictation passage. If an immigrant failed to transcribe a passage in English, they were excluded from entering the country.¹⁵⁷ The White Australia policy succeeded in its aims: 'by 1947 the non-European population, other than Aborigines, was measured by the Census as 0.25 percent of the total. Australia had become one of the "whitest" countries in the world outside northwestern Europe.'¹⁵⁸

In 1958, the Australian government replaced the 1901 act with the Migration Act 1958, which abolished the dictation test, but the White Australia policy continued implicitly. Immigration officers judged 'the degree of ["coloured"] blood in the veins of applicants for settlement', demonstrating a continued obsession with racial purity and with tightly controlling who was permitted to enter Australia.¹⁵⁹

The Migration Act 1958 has left profound legacies on Australia's immigration policy and racial politics. The act was formally updated in 1989 to introduce a points-based immigration system. Points-based systems score applicants on a number of criteria including age, career history, educational qualifications, language and prior work experience in the country of application. The explicit rationale for competitive immigration systems is economic: governments can adjust the selection criteria based on a demand for skills lacking in the current population and strive to achieve

152 Jalata, A. (August 2013), "The Impacts of English Colonial Terrorism and Genocide on Indigenous/Black Australians", *SAGE Open*: p.8. <https://journals.sagepub.com/doi/full/10.1177/2158244013499143>.

153 *ibid.*, p.2.

154 See "Colonial Frontier Massacres in Australia, 1788–1930", University of Newcastle. <https://c21ch.newcastle.edu.au/colonialmassacres/>.

155 Jupp, J. (2010) *From White Australia to Woomera: The Story of Australian Immigration*, 2nd edition (Cambridge: Cambridge University Press): p.9.

156 <https://docs.lib.purdue.edu/cgi/viewcontent.cgi?article=1004&context=histpubs>.

157 *ibid.*, pp.7–8.

158 Jupp, p.10.

159 *ibid.*

'immigration surplus', whereby the additional tax income from immigrants outstrips the 'net fiscal burden imposed on taxpayers'.¹⁶⁰ Australia also operates a 'character test', which excludes anyone with a criminal record or who could 'incite discord in the Australian community'.¹⁶¹

Australia's points-based system, itself a legacy from the racialised and exclusionary White Australia policy, acts as the 'legitimation of a distinction ... between "good" immigrants, who are skilled and deserve to be welcomed because they have acquired education to help themselves, and "bad" or "undeserving" immigrants, who are unskilled, and a drain on society and cost to the economy.'¹⁶² Education, then, is utilised as 'a tool of exclusion and division',¹⁶³ as well as a tool of racial and gender discrimination. This is because immigration systems based on education and skill discriminate against applicants from poorer nations with under-developed educational systems, as well as people from particular socioeconomic backgrounds, racial and ethnic minority groups or genders that lack equal access to education and skilled employment.¹⁶⁴ More importantly, these immigration controls function to create an 'implicit national profile, a cultural preference for the cloned or same and an elevation of those who fit the streamlining profile over those who don't'.¹⁶⁵ Although the points-based system is heralded as a transition away from Australia's racist immigration policies of the past, it continues to uphold underlying notions of racial supremacy and exclusion, continuing *de facto* racism into the modern era.

Furthermore, Australia operates one of the world's most punitive detention-based immigration systems. In 1992, the country adopted a mandatory detention policy in which anyone who is not an Australian citizen and does not have a valid visa is to be detained until their visa is granted or they are deported. At the same time, the law removed a detention cap of 273 days, so that under Australian law, anyone can be detained indefinitely in a detention camp. Many of these camps are offshore, located on the islands Nauru, Manus Island and Christmas Island, under a policy known, until March 2019, as Australia's 'Pacific Solution'.¹⁶⁶

Civil society groups have launched legal challenges to the mandatory indefinite detention policy on human rights grounds, but the High Court of Australia ruled it constitutional in 2004.¹⁶⁷ Australia's detention facilities have been likened to concentration camps.¹⁶⁸ The human impact of Australian immigration detention has been devastating: Ghader, an Iranian man who protested state-endorsed oppression and fled certain death was held in detention on Christmas Island for two years. Once he was granted a bridging visa, he married and had a child. Before his child was

160 Murray, A. (1 January 2011), "Britain's points based migration system", Centre Forum: p.15. <https://www.yumpu.com/en/document/read/30522810/britains-points-based-migration-system-centreforum>.

161 <https://www.bbc.co.uk/news/uk-politics-29594642>.

162 <https://doi.org/10.1111/j.1467-8330.2010.00864.x>, p.1333. For discussion of the neoliberalisation of race, see: Goldberg, D. T. (2011) *The Threat of Race: Reflections on Racial Neoliberalism*, chapter 8.

163 Tannock, p.1332.

164 Tannock, p.1336.

165 Goldberg, p.181.

166 https://www.aph.gov.au/About_Parliament/Parliamentary_Departments/Parliamentary_Library/pubs/BN/2012-2013/PacificSolution#_Toc334509636.

167 <https://www.refugeecouncil.org.au/detention-policies/2/>.

168 <https://www.smh.com.au/opinion/our-detention-centres-are-intentionally-cruel-and-must-be-closed-20160504-golr04.html>.

born, immigration officials cancelled his visa and Ghader was ordered either to return to Iran or remain in indefinite detention in Australia. At the time of writing, Ghader will have been in detention on Christmas Island, unable to see his family, for over a decade.¹⁶⁹

The cumulative effect of Australia's history of settler colonialism and immigration control is to perpetuate implicitly an idea of a natural 'White Australia'. We can see an explicit expression of this in the 2005 Cronulla riots, in which white youths attacked Australians of Middle Eastern background with machetes, home-made bombs, knives and guns.¹⁷⁰ One rioter said, 'You came in chains, we came in planes',¹⁷¹ an unintentionally ironic¹⁷² reflection of the racial supremacist ideas underpinning Australia's immigration policy.

More insidiously, indigenous Australians are suffering from the effects of decades of institutional racism and generational trauma.¹⁷³ Rates of life expectancy, child mortality, early childhood education, literacy and numeracy, school attendance, incarceration, housing and employment outcomes all demonstrate the disadvantaged position of indigenous Australians in contemporary society.¹⁷⁴ Generational trauma and cumulative disempowerment have also led to an elevated prevalence of mental illness, specifically substance abuse,¹⁷⁵ a pattern also found among indigenous peoples in Canada, New Zealand and elsewhere.¹⁷⁶

Contemporary racism in Australian society, fuelled by historical white supremacy, enjoys a degree of institutional support. In August 2018, Senator Fraser Anning made a speech in parliament to call for a return of the White Australia policy as 'the final solution' to the 'immigration problem'. Anning contended that:

"The record of Muslims who have already come to this country in terms of rates of crime, welfare dependency and terrorism are the worst of any migrant and vastly exceed another immigrant group. The majority of Muslims in Australia of working age do not work and exist on welfare. Muslims in New South Wales and Victoria are three times more likely than other groups to be convicted of crimes. We have black African Muslim gangs terrorising Melbourne, we have Isis-sympathising Muslims trying to go overseas to try and fight for Isis and while all Muslims are not terrorists, certainly all terrorists these days are Muslims. So why would anyone want to bring more of them here?"

Anning was not removed from parliament for his remarks. One of Australia's most well-known pundits, Andrew Bolt, published a full-page opinion piece in a major national newspaper with the headline 'The foreign invasion', which warned against a 'tidal wave

169 <https://www.aljazeera.com/news/2020/1/8/refugee-diaries-10-years-in-immigration-detention-in-australia>.
170 Goldberg, pp.184–5.

171 <https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKewjTzayX-8ruAhU9ShUIHbUtDm0QFjABegQIAxAC&url=https%3A%2F%2Fwww.abc.net.au%2Fnews%2F2015-12-11%2Ftaha-how-the-riots-shaped-who-i-am-today%2F7019998&usg=AOvVaw1cHs3FyWN8JunBQLIT-0yf>.

172 The first settlers in Australia were British convicts.

173 Perhaps the most abhorrent example of institutional racism and generational trauma is the cultural genocide of indigenous people in Australia known as 'Stolen Generations'. State legislation, starting in 1869, authorised the forcible removal of mixed-race children – the offspring of white and indigenous Australian parents – into missions, government institutions and children's homes described as 'concentration camps'. Children in placements were often subject to physical and sexual abuse. <https://humanrights.gov.au/our-work/bringing-them-home-report-1997>.

174 See baseline data at <https://www.closingthegap.gov.au/data>.

175 <https://ojs.stanford.edu/ojs/index.php/surj/article/view/1346>.

176 https://www.who.int/substance_abuse/activities/indigenous/en/.

of immigrants sweep[ing] away our national identity' in August 2018, demonstrating a tacit level of approval for racist and anti-immigration notions in Australian society.¹⁷⁷

In a political context such as this, the connections between underlying and historical racial divisions in Australia and the violent attack carried out by Brenton Tarrant become clear.¹⁷⁸ The events of 15 March 2019 in Christchurch mark New Zealand's worst massacre in modern history. Tarrant, a white male, opened fire in two mosques, killing fifty-one and injuring fifty more. He posted a poisonous racist manifesto to 8chan and broadcast the massacre to a Facebook live feed. The manifesto and video liberally refer to memes and in-jokes from an Internet subculture, intended to misdirect mainstream audiences and impress fellow 8chan users.¹⁷⁹

Prior to his shooting spree, Tarrant published a manifesto on Twitter and 8chan. It was entitled 'The Great Replacement', a reference to the 'Great Replacement' conspiracy theory. This theory posits that European populations will be replaced, due to both the immigration of non-European (non-white) populations and the decline of European (white) birth rates. Tarrant contends that this represents an 'assault on the European people that, if not combatted, will ultimately result in the complete racial and cultural replacement of the European people', which he called 'white genocide'.¹⁸⁰ Tarrant's manifesto, which 'barely mentions Australia and instead trains his focus on Europe' echoes Anning's comments that distinguished 'European' and 'non-European' migration,¹⁸¹ itself symptomatic of a wider impulse in Australian politics to be considered part of the Anglosphere.¹⁸²

Tarrant's attack can be seen as a continuation of the preoccupation in Australian society, since the earliest colonial encounters with indigenous peoples, with racial purity and the defilement of white bloodlines. These white supremacist ideas have been reframed for the modern era and are part of the accepted mainstream discourse. The idea that 'Europe is being swamped by Third World migrants, and especially by Muslims', is mainstream discourse in the West.¹⁸³ This 'war' narrative 'intellectually equips those with catastrophised subjectivities to take their proclaimed state of emergency as a green light for desperate measures'; in other words, the mainstream discourse on race and immigration in Australia provides a rationale for individuals to take up arms.¹⁸⁴

Newton et al. contend that 'No amount of technical interventions can resolve the underlying issue of othering and demonisation against out-groups' and that 'Rather than looking at malign activities in the online space as a problem, it is more prudent to look at them as an extension of human experience.' In placing human experience in a sweeping historical context, we can become more innovative about the ways in which we think about moderating content online.

177 <https://www.theguardian.com/media/2018/aug/02/andrew-bolts-tidal-wave-of-immigrants-article-prompts-press-council-complaint>.

178 Australian mainstream conservative politicians were also approvingly mentioned in the manifesto authored by Anders Breivik, perpetrator of the July 2011 Norway attacks.

179 <https://www.bellingcat.com/news/rest-of-world/2019/03/15/shitposting-inspirational-terrorism-and-the-christchurch-mosque-massacre/>.

180 <https://www.tandfonline.com/doi/full/10.1080/14623528.2019.1599493?scroll=top&needAccess=true>.

181 <https://www.theguardian.com/australia-news/2018/oct/25/australian-senator-who-called-for-final-solution-to-immigration-expelled-from-party>.

182 <https://theconversation.com/friday-essay-australias-dangerous-obsession-with-the-anglosphere-97443>.

183 <https://www.tandfonline.com/doi/full/10.1080/14623528.2019.1599493?scroll=top&needAccess=true>.

184 *ibid.*

As Newton et al. suggest, we can perhaps begin to conceptualise violent extremism online as a symptom – rather than a problem in and of itself – of underlying and historical harms. When we do so, there is an opportunity to radically reframe our approach to CVE policy. For instance, placing CVE within the scope of programmes intended to redress historical abuses could transform CVE, moving it away from a technical exercise towards a meaningful educational programme on the historical basis of underlying social divisions. In so doing, there is an opportunity to intervene in harmful mainstream discourses that tip over into real-world harm, as well as centre the historical and contemporary experiences of victims and survivors of violent extremism. Rather than perpetuate the same focus on perpetrators of these attacks, CVE policy has the potential to amplify alternative voices in order to address deep-rooted social issues.



CONTACT DETAILS

For questions, queries and additional copies of this report, please contact:

ICSR
King's College London
Strand
London WC2R 2LS
United Kingdom

T. **+44 20 7848 2098**
E. **mail@gnet-research.org**

Twitter: **[@GNET_research](https://twitter.com/GNET_research)**

Like all other GNET publications, this report can be downloaded free of charge from the GNET website at www.gnet-research.org.

© GNET