

After Christchurch: Hate, harm and the limits of censorship

1. The terrorist attack on Christchurch mosques and the Christchurch Call

David Bromell

Working Paper 21/02



Institute for Governance
and Policy Studies
A research institute of the School of Government



INSTITUTE FOR GOVERNANCE AND
POLICY STUDIES
WORKING PAPER
21/02

MONTH/YEAR

March 2021

AUTHOR

David Bromell
Senior Associate
Institute for Governance and Policy Studies

INSTITUTE FOR GOVERNANCE AND
POLICY STUDIES

School of Government
Victoria University of Wellington
PO Box 600
Wellington 6140
New Zealand

For any queries relating to this working paper,
please contact igps@vuw.ac.nz

ACKNOWLEDGEMENT

Research on this series of working papers has
been financially supported by a fellowship at the
Center for Advanced Internet Studies (CAIS) in
Bochum, NRW, Germany (Oct 2020—Mar 2021).

DISCLAIMER

The views, opinions, findings, and conclusions or
recommendations expressed in this paper are
strictly those of the author. They do not
necessarily reflect the views of the Institute for
Governance and Policy Studies, the School of
Government, Victoria University of Wellington, or
the Center for Advanced Internet Studies (CAIS).

This is paper one in a series of seven working papers, **After Christchurch: Hate, harm and the limits of censorship**.

The series aims to stimulate debate among policy advisors, legislators and the public as New Zealand considers regulatory responses to ‘hate speech’ and terrorist and violent extremist content online following the terrorist attack on Christchurch mosques in March 2019 and the Royal Commission of Inquiry that reported in November 2020.

The seven working papers in this series are:

Title	Reference
1. The terrorist attack on Christchurch mosques and the Christchurch Call	WP 21/02
2. ‘Hate speech’: Defining the problem and some key terms	WP 21/03
3. Challenges in regulating online content	WP 21/04
4. Regulating harmful communication: Current legal frameworks	WP 21/05
5. Arguments for and against restricting freedom of expression	WP 21/06
6. Striking a fair balance when regulating harmful communication	WP 21/07
7. Counter-speech and civility as everyone’s responsibility	WP 21/08

Dr David Bromell is currently (until March 31, 2021) a research Fellow at the Center for Advanced Internet Studies (CAIS) in Bochum, North Rhine-Westphalia, Germany, which has supported his research on this series of working papers. He is a Senior Associate of the Institute for Governance and Policy Studies in the School of Government at Victoria University of Wellington, and a Senior Adjunct Fellow in the Department of Political Science and International Relations at the University of Canterbury. From 2003 to 2020 he worked in senior policy analysis and advice roles in central and local government.

He has published two monographs in Springer’s professional book series:

- *The art and craft of policy advising: A practical guide* (2017)
- *Ethical competencies for public leadership: Pluralist democratic politics in practice* (2019).

Contents

Abstract.....	5
The terrorist attack on Christchurch mosques	5
The Christchurch Call	7
Global Internet Forum to Combat Terrorism (GIFCT).....	8
Civil society reactions and responses	10
Halle shooting: first test for the Christchurch Call.....	12
The Christchurch Call: One year on.....	13
Prospects for success?	14
Conclusion: Impossibility is not an excuse to do nothing	16
Appendix: The Christchurch Call	18
Commitments by governments	18
Commitments by online service providers	18
Commitments by governments and online service providers.....	19
References	21

The terrorist attack on Christchurch mosques and the Christchurch Call

Abstract

The terrorist who attacked two Christchurch mosques on March 15, 2019 was a ‘lone actor’ but not a ‘lone wolf’. He was in contact with other far-right extremists, both online and offline. His use of social media and livestreaming ‘marks a grim new age of social media-fueled terrorism’ (Charlie Warzel).

Two months after the attack, New Zealand Prime Minister Jacinda Ardern and French President Emmanuel Macron jointly hosted a summit in Paris that brought together government and technology sector leaders to adopt the Christchurch Call ‘to eliminate terrorist and violent extremist content online’.

This paper summarises the voluntary (non-binding) commitments made by governments and online service providers that have signed up to the Christchurch Call and other initiatives, and reviews progress in implementing it since May 2019.

While the Christchurch Call is aspirational and well-intentioned, the goal of ‘eliminating’ terrorist and violent extremist content online is impossible—unless governments employ draconian measures of censorship that would severely restrict freedom of expression and commitment to a free, open and secure internet. Impossibility is not, however, an excuse for policy makers to do nothing.

The remaining six working papers in this series elaborate on defining the problem of ‘hate speech’ and other key terms (Working paper 21/03); challenges in regulating online content (Working paper 21/04); current legal frameworks (Working paper 21/05); arguments for and against restricting freedom of expression (Working paper 21/06), and the need to strike a fair balance when regulating harmful communication (Working paper 21/07); and counter-speech and civility as everyone’s responsibility (Working paper 21/08).

Tags: #ChristchurchMosque #ChristchurchAttack #ChristchurchCall #censorship #freespeech #hatespeech

The terrorist attack on Christchurch mosques

On March 15, 2019, a white nationalist terrorist shot and killed 51 people at two mosques in Christchurch, New Zealand, during Jumu’ah (Friday) prayer. Another 40 people suffered gunshot injuries; others escaped physical injury but were traumatised by witnessing the attacks (Zaki, 2020).

Brenton Tarrant, an Australian citizen, was arrested and charged with 51 murders, 40 attempted murders and engaging in a terrorist attack. He was sentenced on August 27, 2020 in the High Court at Christchurch to life imprisonment without parole after changing his plea to guilty on March 26, 2020.

Shortly before the attack, Tarrant uploaded a 74-page manifesto titled *The Great Replacement* to seven different file-sharing websites. At 1:26 pm, he updated his Facebook status with links to the

websites where he had published the manifesto. At 1:28 pm, he posted an anonymous message on 8chan that included:

I have provided links to my writings below, please do your part by spreading my message, making memes and shitposting as you usually do (cited in Royal Commission of Inquiry, 2020, p. 41).

Eight minutes before the attack, at 1:32 pm, he sent an email to 34 recipients, including New Zealand's Parliamentary Service as well as politicians, media outlets and individual journalists. He attached the manifesto to the email and provided a link within the body of the email.

The attacks began at 1:40 pm. He used a bodycam to livestream the first 17 minutes of the attack, at Masjid Al Noor, on Facebook Live. This was the first time a terrorist attack had been filmed via livestream (Macklin, 2019). It was online for nearly an hour before New Zealand Police flagged it with Facebook. This indicates that Facebook's algorithms had failed to recognise the nature of the content.¹ The livestream was viewed around 4000 times before Facebook took it down. By then, it had spread across the internet. During the next 24 hours, Facebook's updated algorithms removed 1.5 million copies of the video—1.2 million copies were blocked at upload; the other 300,000 versions of the footage were successfully uploaded and had to be removed by moderators (Rosen, 2019a; RNZ, 2019).

Copies of the livestream were re-posted on other websites and social media and file-sharing platforms, including LiveLeak, YouTube, Twitter, Reddit, 4Chan and 8Chan. At one point, YouTube reported that it was removing one copy per second (Dwoskin & Timberg, 2019). New Zealand's biggest telecommunications providers blocked access to domains hosting the material, then contacted the hosts asking them to remove the footage to have their domain unblocked (Daalder, 2019).

Initially social media algorithms may have 'recommended' the livestream to users as trending content (Mason & Errington, 2019, p. 9). Many members of the public, including children, viewed it while not fully comprehending what they were seeing (Timberg, Harwell, Shaban, Ba Tran & Fung, 2019). Families and friends of the victims were also among those who viewed the livestream as they desperately tried to discover the fate of their loved ones.

Selected footage from the video was used in mainstream media reporting, for example in *MailOnline*, *The Sun* and *The Daily Mirror* in the United Kingdom (Waterson, 2019), and in Sky News, channels Seven and Nine, Network Ten and SBS in Australia (Williams, 2019).² In terms of editorial decision-making, New Zealand media, with closer proximity to the event, focused largely on empathetic coverage of victims and resisted the alleged gunman's attempts to publicise his cause,

¹ Brian Fishman, Facebook's policy director for counter-terrorism, reportedly told US Congress members that its algorithm did not initially detect the Christchurch mosque attack livestream because there was 'not enough gore' (Chang, 2019).

² The Australian Communications and Media Authority investigated use of the footage by SBS and other broadcasters in Australia and concluded that finding individual contraventions of codes of practice would have little regulatory or educative benefit: 'Instead, the ACMA considers that this investigation would more usefully prompt a productive conversation with industry about whether its codes are adequately framed to deal with this type of material in the future—in particular, perpetrator-generated live streamed extreme violent material' (Australian Communications and Media Authority, 2019, p. 12). Denis Muller (2019) has criticised the ACMA's report as 'nuanced but very tame'.

while Australian media ran extended coverage of the perpetrator, along with material ruled objectionable in New Zealand (Ellis & Muller, 2020).

On March 18, 2019, New Zealand's Office of Film and Literature Classification issued a decision classifying the livestream as 'objectionable', in effect banning its possession and distribution in New Zealand (Classification Office, 2019a). On March 23, the Classification Office similarly classified the shooter's manifesto as objectionable (Classification Office, 2019b).

In the days following the attacks, and despite diplomatic protests from both New Zealand and Australia, blurred footage from the livestream video was shown during at least eight campaign rallies for Turkey's President, Recep Tayyip Erdoğan (Lowen, 2019).

More than a month after the attacks, re-edited versions of the livestream with altered digital identifiers were still being discovered on Facebook, Instagram and YouTube. One of the YouTube versions (two minutes and 17 seconds of 'highlights') had clocked 773,009 views according to a screen grab taken by New York-based hate content researcher David Feinberg on April 22, 2019 (Keall, 2019).

Between March 15 and September 30, 2019 Facebook reported taking down 4.5 million pieces of content related to the Christchurch attacks and reflected that 'the Christchurch attack was unprecedented in both the use of live streaming technology and the rapid sharing of video, image, audio and text-based content depicting the attack' (Rosen, 2019b). Versions of the video could, however, still be found online in September 2019 (Ensor, 2019).

Writing in the *New York Times*, Charlie Warzel (2019) reflected that 'the attack marks a grim new age of social media-fueled terrorism'. The manifesto and livestream inspired subsequent actual or planned attacks in the United States, Germany, Norway and Singapore. And while the Christchurch mosque attacks were carried out by a 'lone actor', referring to him as a 'lone wolf' is deceptive. He was in contact with other far-right extremists, both online and offline, and had travelled extensively (Petersmann, 2019; Knight, 2020; Royal Commission of Inquiry, 2020, pp. 109, 233–236).³ Graham Macklin (2019) reflects that 'although Tarrant acted alone, he was sustained by, and interacted with, a broader sub-cultural online environment that was electrified by his atrocity, as he knew it would be'.

The livestream, moreover, was not just a medium for the message—it was the message (Macklin, 2019). The central point of his attack was not just to kill Muslims, 'but to make a video of someone killing Muslims' (Burke, 2019). The attack on Christchurch mosques reportedly fuelled a surge in anti-Muslim hatred online: 'If 8chan was the shooter's firewood, Facebook the match, and YouTube the oxygen, Reddit is a place where the fire he set is still smoldering' (Breland, 2019).

The Christchurch Call

Two months later to the day, on May 15, 2019, New Zealand Prime Minister Jacinda Ardern and French President Emmanuel Macron jointly hosted a summit in Paris that brought together

³ On 'lone wolves connected online', see Smith, 2021.

government and technology sector leaders to ‘adopt’ the Christchurch Call by committing ‘to eliminate terrorist and violent extremist content online’.⁴

The Christchurch Call takes respect for freedom of expression as fundamental, and ‘rests on the conviction that a free, open and secure internet offers extraordinary benefits to society,’ while asserting that ‘no one has the right to create and share terrorist and violent extremist content online’ (NZ Ministry of Foreign Affairs & Trade, n.d.).

The summit noted significant steps already taken to address the issue by the European Commission and the EU Internet Forum; the G20 and the G7, including work underway during France’s G7 presidency on combating the use of the internet for terrorist and violent extremist purposes; the Global Internet Forum to Counter Terrorism (GIFCT); the Global Counterterrorism Forum; Tech Against Terrorism; and the Aqaba Process established by the Hashemite Kingdom of Jordan.⁵

The terrorist attack in Christchurch highlighted the urgent need for enhanced co-operation and action by a wide range of actors, including governments, civil society and online service providers to eliminate terrorist and violent extremist content online. The Christchurch Call seeks to do this in ways that are consistent with principles of a free, open and secure internet, and without compromising human rights and fundamental freedoms, including freedom of expression.

To this end, the Christchurch Call outlines collective, voluntary (non-binding) commitments from governments and online service providers to prevent abuse of the internet. These commitments are reproduced below in the Appendix: The Christchurch Call.

Global Internet Forum to Combat Terrorism (GIFCT)

GIFCT is an independent organisation established in July 2017 by Facebook, Microsoft, Twitter and YouTube to disrupt terrorist abuse on their platforms. Amazon, Dropbox, LinkedIn and WhatsApp joined in 2019 (GIFCT, n.d.).

In response to the terrorist attack in Christchurch and the subsequent Christchurch Call, GIFCT agreed in September 2019 to restructure as a new, independent body (still called GIFCT) led by an independent Executive Director appointed by an industry-led Operating Board. The Operating Board is guided by a multi-stakeholder Independent Advisory Committee chaired by a non-governmental representative.

GIFCT’s mission to prevent terrorists and violent extremists from exploiting digital platforms guides its work across four foundational goals.

- **Empower** a broad range of technology companies, independently and collectively, with processes and tools to prevent and respond to abuse of their platforms by terrorists and violent extremists.

⁴ The Christchurch Call summit was held alongside the Tech for Humanity meeting of G7 Digital Ministers, which France was chairing, and France’s Tech for Good summit, both held on May 15, 2019 (Ardern, 2019).

⁵ Other intergovernmental initiatives have included the Council of Europe’s 2005 Convention on the Prevention of Terrorism and 2015 Additional Protocol to the Convention on the Prevention of Terrorism and Counter-Terrorism Strategy for 2018–22 (Council of Europe, 2020), UN Security Council Resolution 2354 and the Comprehensive International Framework to Counter Terrorist Narratives (UN Security Council, 2017).

- **Enable** multi-stakeholder engagement around terrorist and violent extremist misuse of the internet and encourage stakeholders to meet key commitments consistent with the GIFCT mission.
- **Promote** civil dialogue online and empower efforts to direct positive alternatives to the messages of terrorists and violent extremists.
- **Advance** broad understanding of terrorist and violent extremist operations and their evolution, including the intersection of online and offline activities.

GIFCT’s work is being organised into three pillars.

1. **Prevent**—equip digital platforms and civil society groups with awareness, knowledge and tools, including technology, to develop sustainable programmes in their core business operations to disrupt terrorist and violent extremist activity online.
2. **Respond**—develop tools and capacity, including via regular multi-stakeholder exercises, for platforms to co-operate with one another and with other stakeholders to mitigate the impact of a terrorist or violent extremist attack.
3. **Learn**—empower researchers to study terrorism and counterterrorism, including creating and evaluating best practices for multi-stakeholder co-operation and preventing abuse of digital platforms.

Working groups engage stakeholders from government and civil society and co-ordinate multi-stakeholder funding to address, for example, interventions with respect to radicalisation, algorithmic outcomes, improving the multi-stakeholder Crisis Response Protocol and legal challenges to data sharing.

In its annual transparency report for 2020, GIFCT (2020c) reported on implementation of its new structure and joint tech innovations. The largest cross-platform technical tool supported by GIFCT is the Hash Sharing Consortium, which shares ‘hashes’ (unique digital fingerprints) of known terrorist images and videos.⁶ There are now over 300,000 hashes in the database.

Because of differing definitions of ‘terrorism’ and ‘terrorist content’, the hash-sharing database is limited to content related to organisations on the United Nations Security Council’s consolidated sanctions list, and to declared Content Incident Protocols. Using its agreed taxonomy, the transparency report shows the breakdown of content in the database as at July 2020:

- Imminent credible threat (0.1%);
- Graphic violence against defenseless people (16.9%);
- Glorification of terrorist acts (72%);
- Radicalisation, recruitment, instruction (2.1%); and
- Christchurch, New Zealand, attack and Content Incident Protocols:
 - Christchurch, New Zealand perpetrator content (6.8%)
 - Halle, Germany, perpetrator content (2%)
 - Glendale, Arizona, US, perpetrator content (0.1%).

⁶ GIFCT (2020c) explains: ‘The image or video is “hashed” in its raw form and is not linked to any source original platform or user data. Hashes appear as a numerical representation of the original content, which means it cannot be easily reverse engineered to create the image and/or video.’

Since January 2019, GIFCT has also developed a URL-sharing programme. When terrorist content is shared on one platform, linked to content hosted on another platform, companies only have jurisdiction to remove content hosted on their own services. This means they can remove a post, but the source link and hosted content remain intact on the third-party platform. Conversely, while the primary site on which content was shared might remove it and disallow re-posting, this does not ensure its removal from other platforms to which it may have been migrated. In the case of the Halle shooting, for example, even though the livestream was removed from Twitch, edited versions of it continued to surface on the encrypted platform, Telegram (Robertson, 2019).

In the URL-sharing programme, a GIFCT company has a safe mechanism to share a URL link with the industry partner to whom the content belongs. The notified platform can then review the link to decide whether the content violates its terms of service. GIFCT has adapted the programme through a 12-month pilot project with SITE Intelligence, a firm that provides subscription-based monitoring and analysis of terrorist content and other online harms. This gave GIFCT members access to SITE's SourceFeed, with extra context around a given URL, including the organisational affiliation of terrorist content and translation of content into English. Through this programme, GIFCT has shared around 24,000 URLs since its launch (GIFCT, 2021).

In partnership with Tech Against Terrorism, the GIFCT has held 13 workshops around the world, bringing together tech platforms with policy makers, law enforcement, civil society, academic experts and practitioners to share knowledge, best practice and models for cross-sector collaboration. The GIFCT has also supported the Global Research Network on Terrorism and Technology, aimed at developing research and recommending policy to prevent terrorist exploitation of technology.

Civil society reactions and responses

A Christchurch Call Advisory Network has been built with interested civil society groups that represent a range of perspectives, including human rights, freedom of expression, digital rights, counter-radicalisation, victim support and public policy (NZ Ministry of Foreign Affairs & Trade, n.d.).

Several New Zealand-based non-governmental organisations and think tanks have published commentary and reports on the Christchurch Call (Thompson, 2019, p. 89). InternetNZ facilitated a 'civil society' statement informed by the Voices for Action meeting with Prime Minister Jacinda Ardern in Paris the day before the Christchurch Call summit (InternetNZ, 2019). The statement documents include civil society, academic and technical community perspectives on terrorist and violent extremist content online.

Discussions broadly supported some key values in the Call, specifically the affirmation of human rights and fundamental freedoms, including freedom of expression, to which Voices for Action proposed adding freedom of association, religion and assembly; and of a free, open and secure internet, with the proposed addition of 'globally connected' and 'inter-operable'.⁷

⁷ Internet interoperability allows information exchange across platforms; e.g., web pages being readable on any device or browser. It is critical for the future of the Internet of Things (IoT).

Issues with the most shared concern were:

- The definition of ‘terrorism and violent extremism’ and the risk in leaving interpretation of this to states individually, where this could result in disproportionate suppression of political critics, human rights defenders, journalists, ethnic or religious minorities, refugees, asylum seekers and migrants;
- The need for broader societal (rather than closed-door) discussions about the respective roles of tech companies and governments;
- The need for an evidence-based, systemic approach to the issues;
- The need to address issues of terrorism and violent extremism in our societies, not just online;
- The process and timeline of the Call, including concerns about the siloed approach to negotiations and the exclusion of civil society, academic experts, journalists and news media representatives and the technical community from the summit;
- The need for governments and tech companies to engage in dialogue with each other, as well as with civil society, academics, journalists and news media, and the technical community, as well as survivors and families of victims; and
- The need for technical solutions to reflect commitments to human rights and a free, open and secure internet, particularly:
 - differentiating between user-generated content platforms and internet infrastructure; and
 - excluding upload filters as inconsistent with human rights and because they prevent collection of evidence of human rights abuses (InternetNZ, 2019).

The statement is explicit that governments should not outsource speech regulation or governance to technology companies:

Governments should meet their own responsibilities under human rights instruments by relying upon democratically-enacted and judicially reviewable law—not proprietary terms of service. Outsourcing speech regulation to unaccountable private actors is no more acceptable, and no more permitted under human rights law, than delegating other core functions of government (InternetNZ, 2019, p. 3).

This is a key concern among civil rights advocates. Where tech companies are required to police content, high fines could incentivise them to over-block potentially unlawful or harmful material that may be legitimate expression of opinion. This also gives too much power to social media companies to determine the limits of free speech, which should remain subject to judicial determination and review and due process (York, 2019; Mason & Errington, 2019, p. 12; Etzioni, 2019).⁸

Differentiation between content platforms and internet infrastructure is also important. Jordan Carter, chief executive of InternetNZ, and Konstantinos Komaitis, senior director of global policy

⁸ This issue came to the fore on January 6, 2021, when Twitter and then Facebook locked the accounts of US President Donald Trump after he published a string of inaccurate and inflammatory messages on a day of violence on Capitol Hill (Conger, Isaac, & Frenkel, 2021). For private companies to censor an incumbent, democratically elected President is as extraordinary as Trump’s inflammatory pronouncements. On the de-platforming of President Trump, see Working paper 21/04, **Challenges in regulating online content**.

development and strategy at the Internet Society, wrote in an opinion piece published on the day of the Christchurch Call summit:

It is governments' job to protect us, and this involves difficult trade-offs. But if new laws and regulations are not carefully targeted at the appropriate layer of the internet, they will damage its infrastructure and shut down possibilities for future innovation. Interfering with the infrastructure layer, even unintentionally, to fix problems at the content layer creates unintended consequences that hurts everyone's ability to communicate legitimately and use the internet safely and securely (Carter & Komaitis, 2019).

A significant objection to upload filters and rapid erasing of content is that it can be counter-productive to understanding and taking action against terrorism and violent extremism. The International Criminal Court and prosecutors in the European Union have relied on such evidence to conduct investigations and bring charges. Ullmann and Tomalin (2019) have argued for quarantining rather than deletion of online hate speech, to support a more justifiable balance between freedom of expression and appropriate censorship. When platforms immediately delete content as it appears without reaching out to investigators, opportunities to address the root causes of extremism are also deleted.⁹ 'Counter speech',¹⁰ journalistic reporting and other protected speech is vital to countering extremism (InternetNZ, 2019, pp. 5–6; Human Rights Watch, 2020).

Halle shooting: first test for the Christchurch Call

The GIFCT activated its Content Incident Protocol for the first time on October 9, 2019, following terrorist shootings in Halle, Germany. Communication between GIFCT founding members began at around 5:00 am and companies began platform-specific monitoring and enforcement operations. Early reports by eyewitnesses indicated the attacker wore a video camera. This was confirmed at around 9:15 am when GIFCT became aware of the shooter's video and that copies of the livestream were circulating on non-GIFCT member platforms. By this time, the archived video on Twitch had been removed and the GIFCT Operational Board formally activated the Content Incident Protocol at 10:09 am (GIFCT, 2020a).

GIFCT (2020a) reports that overall, content from the Halle shooting had significantly less impact online than content relating to the attack on Christchurch mosques. During the Halle Content Incident Protocol, GIFCT members shared hashes (unique digital fingerprints) related to 36 visually-distinct videos via its collective database—far fewer variations than they saw during a similar period following the Christchurch attack.

Similarly, following the shooting in Glendale, Arizona, on May 20, 2020, GIFCT activated the Content Incident Protocol. 198 visually distinct videos (just under 1200 hashes) were added to the hash-sharing database. Following a noticeable reduction in attempted uploads of violating content on

⁹ The state media authority for North Rhine-Westphalia (Landesanstalt für Medien NRW) has a campaign, *Verfolgen statt nur Löschen* (Pursue instead of just deleting). The initiative brings together representatives from media oversight agencies, law enforcement authorities and media companies to identify and sanction illegal hate speech on the internet and facilitate criminal prosecution. Notifications informed action by police in nine federal states in November 2019. See further Working paper 21/08, **Counter-speech and civility as everyone's responsibility**.

¹⁰ On counter-speech, see Working paper 21/08.

member platforms, GIFCT formally concluded the Content Incident Protocol at 9:34 pm on May 21, 2020. Any additional copies of the video surfaced by GIFCT member platforms were hashed and uploaded to the shared industry database.

The Christchurch Call: One year on

On May 15, 2020, New Zealand Prime Minister Jacinda Ardern and French President Emmanuel Macron re-affirmed their commitment to the Christchurch Call a year after its launch. In a joint statement, they reported that:

48 countries, the European Commission, two international organisations, and eight tech companies are now actively contributing to the implementation of the Call commitments, together with an Advisory Network of 44 international civil society representatives, which include non-governmental organisations, academia, think-tanks and others who focus on human rights (including freedom of expression and digital rights), counter-radicalization, victim support and public policy (Ardern, 2020).

The statement reported progress in three areas.

- Multi-stakeholder collaboration means co-ordination is now possible across governments and tech companies to respond to attempts to use the internet as a terrorist weapon.
- Tech companies have reformed the GIFCT into an independent organisation with dedicated resources, a 24/7 crisis management function and an enhanced governance structure.
- A Content Incident Protocol has been adopted and tested in both controlled environments and in response to real-life events, so GIFCT member companies can become aware of, quickly assess and act on potential content circulating online resulting from a real-world terrorism or violent extremist event.

In October 2019, as part of a European Union response to the Christchurch Call, the EU Internet Forum committed to a Crisis Protocol, to enable rapid response to contain the viral spread of terrorist and violent extremist content online (European Commission, 2019).

Noting that there is still much to do, Ardern and Macron stated that New Zealand and France will continue their work to broaden commitment to the Call as members of the GIFCT Independent Advisory Committee, and seek to engage the tech industry more broadly, including small and medium-sized companies.

While much has been achieved in a relatively short time, the Christchurch Call raises wider questions about regulation of social media companies and other digital intermediaries (Thompson, 2019, pp. 91–92), including:

- Limitations of industry self-regulation, given the non-binding nature of the Christchurch Call, limited transparency and public oversight, and commercial competition, convergence and network dominance stemming from digital intermediaries' control over the architectures and algorithms of content discovery and promotion;
- The need for regulation within domestic jurisdictions as well as multi-lateral, multi-stakeholder global agreements and initiatives, balanced by the risk of state 'over-reach' that impinges on freedom of expression and protection of user privacy;

- Engagement with civic actors (and not only as a marginal or token gesture), particularly given the tangled web of interests between governments and tech companies, including reliance of governments and political parties on social media for their own public communications and political campaigning; and
- Broadening the dialogue from a narrow focus on terrorist and violent extremist content to address wider concerns about online hate speech and extremism while protecting freedom of expression and encouraging digital competence and civility.

Prospects for success?

While good progress appears to have been made by GIFCT since the Christchurch Call was adopted in May 2019, questions remain about social media and digital intermediaries' selective adoption of modes of content monitoring, especially where this may compromise their commercial interests (Thompson, 2019, p. 84).

A 2017 UK government report had criticised the big social media companies for taking insufficient action to address illegal and hateful content:

It is shocking that Google failed to perform basic due diligence regarding advertising on YouTube paid for by reputable companies and organisations which appeared alongside videos containing inappropriate and unacceptable content, some of which were created by terrorist organisations. We believe it to be a reflection of the laissez-faire approach that many social media companies have taken to moderating extremist content on their platforms. We note that Google can act quickly to remove videos from YouTube when they are found to infringe copyright rules, but that the same prompt action is not taken when the material involves hateful or illegal content ... The biggest and richest social media companies are shamefully far from taking sufficient action to tackle illegal and dangerous content, to implement proper community standards or to keep their users safe (House of Commons Home Affairs Committee, 2017, p. 21).

Thompson (2019, p. 84) notes that although Facebook is only one of many social media companies, proponents of stronger state-based regulation of the tech sector often target Facebook for particular attention because of its poor track record in blocking extremist content. In 2018, for example, Channel 4's *Dispatches* uncovered Facebook's 'shielded review' moderation system, which permitted high-traffic, right-wing hate speech posted by groups such as the English Defence League and Britain First to remain online despite numerous complaints. Only after the Christchurch mosque attacks did Facebook acquiesce and block the content in April 2019.

Through an analysis of internal moderation training documents, a Vice Motherboard investigation in 2018 uncovered details of Facebook's response to the Unite the Right rally in Charlottesville, Virginia, on August 11–12, 2017. The Facebook policy was to block praise, support and representation of white supremacy as an ideology, while allowing praise, support and representation of white nationalism and white separatism as ideologies. After backlash from civil rights groups and historians, Facebook's Content Standards Forum also agreed in March 2019 to block white nationalist and white separatist content on Facebook and Instagram, and to re-direct users who search for or try to post white nationalism, white separatism, or white supremacist content to the website for Life After Hate (<https://www.lifeafterhate.org/>), a non-profit founded by

ex-white supremacists that is dedicated to getting people to leave hate groups (Cox, 2018; Cox & Koebler, 2019a).

Implicit and coded white nationalism and white separatism are, however, difficult to detect and enforce. The alt-right commonly cloaks ideology in generalities and uses coded and covert signals, irony, sarcasm, images, memes and in-jokes to shroud borderline content and extreme views in plausible deniability (Binder, Ueberwasser, & Stark, 2020, p. 60; Chen, 2020, pp. 157–159; Fielitz & Schwarz, 2020, pp. 45–48; Gilbert & Elley, 2020; Miller-Idriss, 2020, pp. 65–66; Schmitt, Harles, & Rieger, 2020; Owen, 2019).¹¹ A two-year internal audit of Facebook’s performance on vital social issues found that the company has not implemented the auditors’ recommendation to prohibit veiled and not just explicit references to white nationalist or white separatist ideology (Ortutay, 2020).

And despite Facebook’s new ruling, an investigation by *The Independent* and confirmed by a Counter Extremism Project showed that even after the Christchurch mosque attacks, Facebook allowed neo-Nazi groups to remain on Facebook. These included Combat 18, the Misanthropic Division, chapters of Be Active Front USA (a racist skinhead group), the neo-Nazi British Movement, Green Line Front and the US National Alliance Reform and Restoration Group, which calls for supporters to mount a ‘white revolution ... and take back our homeland from the invaders’ (Dearden, 2019).

Twitter and YouTube did not follow Facebook’s lead in blocking white nationalism (Cox, 2019) and both platforms have tended to lag behind Facebook in controlling extremist content and in doing so consistently (Thompson, 2019, p. 84). A 2019 Motherboard report found, for example, that Twitter’s robust policing of Islamic State content has not been matched by its policing of alt-right content. While society at large unequivocally demanded that Twitter take action against ISIS when beheading videos spread on the platform, the same eradicate-everything approach, applied to white nationalism is much more controversial (Cox & Koebler, 2019b). Algorithms to block or remove it could, for example, capture content from democratically elected politicians. Content moderation issues are not, therefore, merely technical and algorithmic—they are also moral and political.

Thompson (2019, p. 85) notes that a week before the Christchurch mosque attacks, Facebook’s Mark Zuckerberg outlined a privacy-oriented model for social networking to give more control to private individuals through, for example, end-to-end encryption and data storage and deletion options for all Facebook apps. This will allow people to communicate in the digital equivalent of the ‘living room’, and not only in the digital equivalent of a ‘town square’. He acknowledged the tension between protecting legitimate dissidents and ‘the privacy of people doing bad things’, noting ‘a growing concern among some that technology may be centralizing power in the hands of governments and companies like ours’. His objective is ‘working towards a world where people can speak privately and live freely knowing that their information will only be seen by who they want to see it and won’t all stick around forever’ (Zuckerberg, 2019a).

¹¹ The Royal Commission of Inquiry (2020, p. 11) cites a style guide for far-right website the Daily Stormer leaked in December 2017 to illustrate this. See further Working paper 21/04, **Challenges in regulating online content**.

Two weeks after the Christchurch mosque attacks, however, Zuckerberg (2019b) published an op-ed in the *Washington Post* inviting ‘a more active role for governments and regulators’ in four areas: harmful content, election integrity, privacy and data portability:

Internet companies should be accountable for enforcing standards on harmful content. It’s impossible to remove all harmful content from the Internet, but when people use dozens of different sharing services—all with their own policies and processes—we need a more standardized approach.

One idea is for third-party bodies to set standards governing the distribution of harmful content and to measure companies against those standards. Regulation could set baselines for what is prohibited and require companies to build systems for keeping harmful content to a bare minimum.

Thompson (2019) notes that Facebook is seeking to redress reputational damage from its facilitation of livestreamed terrorism and abuses of user data, including the Cambridge Analytica scandal (Cadwalladr, 2017), and that beyond clearer regulatory definitions of ‘baselines’, Zuckerberg continues to promote industry actors as the principal agents of regulatory intervention. Thompson (p. 85) comments:

Indeed, statutory regulation potentially offers strategic advantages for social media and digital intermediaries, especially if they are involved in its design and implementation ... Formal regulation (such as defining what content is deemed harmful) would allow digital intermediaries and social media to maintain a default defence of compliance with applicable laws.

Further, ‘by offloading decisions about harmful content, privacy rules, and elections onto third-parties, Facebook may not have to take as much of the heat when mistakes are made’ (McMaken, 2019; cf. Isaac, 2019).

Tech companies have taken steps, particularly through GIFCT, to clean up their act since the Christchurch mosque attacks, but as private companies, their primary imperative is to generate profit and their primary accountability is to their shareholders, not to the public. This issue is compounded by the global dominance of the so-called FAANG (Facebook, Apple, Amazon, Netflix, Google) or FAMGA (Facebook, Apple, Microsoft, Google, Amazon) digital media and communications platform companies and a tendency towards monopoly or oligopoly in the sector (Flew, Martin & Suzor, 2019, p. 34).

Conclusion: Impossibility is not an excuse to do nothing

While the Christchurch Call is aspirational and well-intentioned, its goal of ‘eliminating’ terrorist and violent extremist content online is impossible. This is particularly so when the net is cast wider, to capture ‘hate speech’ and not only terrorist and violent extremist content.

There are two main reasons for this. First, a wide range of actors (governments, private companies, the media, civil society and individual citizens) with divergent, competing and conflicting interests would need to work together in co-ordinated ways to eliminate harmful communication online. What is required is some combination of governmental and inter-governmental regulation including internationally aligned antitrust and competition regulation, industry self-regulation, multi-lateral, multi-stakeholder agreements and initiatives, technology innovation, and market pressure by

advertisers, consumers and service users. While this is exactly the commitment invited by the Christchurch Call, there is little evidence to date that it will or can occur to an extent that would 'eliminate' harmful use of the internet.

Secondly, this in turn suggests that the goal of eliminating harmful digital communication could be achieved only by governments exercising draconian measures of censorship that would severely restrict freedom of expression and commitment to a free, open and secure internet. Those who call for more laws should be careful what they wish for, and passing laws does not in any case solve complex social problems.

Impossibility is not, however, an excuse for policy makers to do nothing. Governments need to exercise their expressive powers, not only their regulatory powers, and encourage and support counter-speech strategies as alternatives or complements to prohibition and censorship.

Options include investment in public education programmes in civics, human rights, conflict resolution and digital literacy; in building stronger partnerships with communities, civil society groups, public sector institutions and industry; reducing inequalities and marginalisation on all fronts, and outreach, early intervention and rehabilitation to prevent extremism from taking root; and in well-funded public broadcasting that provides access to authoritative information and diverse ideas. All of this costs money. Governments can also withhold non-profit status and tax privileges to groups that fail to respect the values of freedom and equality that underpin any democratic society.

But there are limits to what the state can and should be expected to achieve. Recovery of civility is everyone's responsibility. That includes pulling back from angry, 'woke' virtue signaling and 'cancel culture', and 'calling in' instead of 'calling out' those with whom we disagree (Bennett, 2020). As President Joe Biden said in his inauguration speech on January 20, 2021, it's time to 'stop the shouting, and lower the temperature' (Biden, 2021).

The remaining six working papers in this series elaborate on defining the problem of 'hate speech' and other key terms (Working paper 21/03), challenges in regulating online content (Working paper 21/04), current legal frameworks (Working paper 21/05), arguments for and against restricting freedom of expression (Working paper 21/06), striking a fair balance when regulating harmful communication (Working paper 21/07), counter-speech as an alternative or complement to regulation, and civility as everyone's responsibility (Working paper 21/08).

Appendix: The Christchurch Call

Commitments by governments

Governments that are parties to the Christchurch Call¹² entered into the following commitments.

- **Counter the drivers of terrorism and violent extremism** by strengthening the resilience and inclusiveness of our societies to enable them to resist terrorist and violent extremist ideologies, including through education, building media literacy to help counter distorted terrorist and violent extremist narratives, and the fight against inequality.
- **Ensure effective enforcement of applicable laws** that prohibit the production or dissemination of terrorist and violent extremist content, in a manner consistent with the rule of law and international human rights law, including freedom of expression.
- **Encourage media outlets to apply ethical standards** when depicting terrorist events online, to avoid amplifying terrorist and violent extremist content.
- **Support frameworks, such as industry standards**, to ensure that reporting on terrorist attacks does not amplify terrorist and violent extremist content, without prejudice to responsible coverage of terrorism and violent extremism.
- **Consider appropriate action** to prevent the use of online services to disseminate terrorist and violent extremist content, including through collaborative actions, such as:
 - awareness-raising and capacity-building activities aimed at smaller online service providers;
 - development of industry standards or voluntary frameworks; and
 - regulatory or policy measures consistent with a free, open and secure internet and international human rights law.

Commitments by online service providers

Online service providers¹³ entered into the following commitments.

- **Take transparent, specific measures seeking to prevent the upload of terrorist and violent extremist content and to prevent its dissemination on social media and similar content-sharing services, including its immediate and permanent removal, without prejudice to law enforcement and user appeals requirements, in a manner consistent with human rights and fundamental freedoms.** Co-operative measures to achieve these outcomes may include technology development, the expansion and use of shared databases of hashes and URLs, and effective notice and takedown procedures.

¹² Founding governments of the Christchurch Call were New Zealand and France. Founding supporter governments announced in Paris on May 15, 2019 were Australia, Canada, the European Commission, Germany, Indonesia, India, Ireland, Italy, Japan, Jordan, The Netherlands, Norway, Senegal, Spain, Sweden and the United Kingdom. Additional supporters announced in New York on September 23, 2019 were Argentina, Austria, Belgium, Bulgaria, Chile, Colombia, Costa Rica, Cyprus, Denmark, Finland, Georgia, Ghana, Greece, Hungary, Iceland, Ivory Coast, Kenya, Latvia, Lithuania, Luxembourg, Maldives, Malta, Mexico, Mongolia, Poland, Portugal, Romania, South Korea, Slovenia, Sri Lanka, Switzerland, UNESCO and the Council of Europe (NZ Ministry of Foreign Affairs & Trade, n.d.).

¹³ Online service providers that support the Christchurch Call are Amazon, Daily Motion, Facebook, Google, Microsoft, Qwant, Twitter, YouTube and LINE (NZ Ministry of Foreign Affairs & Trade, n.d.).

- **Provide greater transparency in the setting of community standards or terms of service**, including by:
 - outlining and publishing the consequences of sharing terrorist and violent extremist content; and
 - describing policies and putting in place procedures for detecting and removing terrorist and violent extremist content.
- **Enforce those community standards or terms of service** in a manner consistent with human rights and fundamental freedoms, including by:
 - prioritising moderation of terrorist and violent extremist content, however identified;
 - closing accounts where appropriate; and
 - providing an efficient complaints and appeals process for those wishing to contest the removal of their content or a decision to decline the upload of their content.
- **Implement immediate, effective measures to mitigate the specific risk that terrorist and violent extremist content is disseminated through livestreaming**, including identification of content for real-time review.
- **Implement regular and transparent public reporting**, in a way that is measurable and supported by clear methodology, on the quantity and nature of terrorist and violent extremist content being detected and removed.
- **Review the operation of algorithms and other processes that may drive users towards and/or amplify terrorist and violent extremist content** to better understand possible intervention points and to implement changes where this occurs. This may include using algorithms and other processes to redirect users from such content or the promotion of credible, positive alternatives or counter-narratives. This may include building appropriate mechanisms for reporting, designed in a multi-stakeholder process and without compromising trade secrets or the effectiveness of service providers' practices through unnecessary disclosure.
- **Work together to ensure cross-industry efforts are co-ordinated and robust**, for instance by investing in and expanding the GIFCT, and by sharing knowledge and expertise.

Commitments by governments and online service providers

Governments and online service providers further agreed to work together, and with others.

- **Work with civil society to promote community-led efforts** to counter violent extremism in all its forms, including through the development and promotion of positive alternatives and counter-messaging.
- **Develop effective interventions**, based on trusted information sharing about the effects of algorithmic and other processes, to redirect users from terrorist and violent extremist content.
- **Accelerate research into and development of technical solutions** to prevent the upload of and to detect and immediately remove terrorist and violent extremist content online, and share these solutions through open channels, drawing on expertise from academia, researchers and civil society.
- **Support research and academic efforts to better understand, prevent and counter terrorist and violent extremist content online**, including both the offline and online impacts of this activity.

- **Ensure appropriate co-operation with and among law enforcement agencies** for the purposes of investigating and prosecuting illegal online activity in regard to detected and/or removed terrorist and violent extremist content, in a manner consistent with rule of law and human rights protections.
- **Support smaller platforms as they build capacity to remove terrorist and violent extremist content**, including through sharing technical solutions and relevant databases of hashes or other relevant material, such as the GIFCT shared database.
- **Collaborate, and support partner countries, in the development and implementation of best practice** in preventing the dissemination of terrorist and violent extremist content online, including through operational co-ordination and trusted information exchanges in accordance with relevant data protection and privacy rules.
- **Develop processes allowing governments and online service providers** to respond rapidly, effectively and in a co-ordinated manner to the dissemination of terrorist or violent extremist content following a terrorist event. This may require the development of a shared crisis protocol and information-sharing processes, in a manner consistent with human rights protections.
- **Respect, and for Governments protect, human rights**, including by avoiding directly or indirectly contributing to adverse human rights impacts through business activities and addressing such impacts where they occur.
- **Recognise the important role of civil society** in supporting work on the issues and commitments in the Call, including through:
 - offering expert advice on implementing the commitments in this Call in a manner consistent with a free, open and secure internet and with international human rights law;
 - working, including with governments and online service providers, to increase transparency; and
 - where necessary, working to support users through company appeals and complaints processes.
- **Affirm our willingness to continue to work together**, in existing fora and relevant organisations, institutions, mechanisms and processes to assist one another and to build momentum and widen support for the Call.
- **Develop and support a range of practical, non-duplicative initiatives to ensure that this pledge is delivered.**
- **Acknowledge that governments, online service providers and civil society may wish to take further co-operative action to address a broader range of harmful online content, such as the actions that will be discussed further during the G7 Biarritz Summit, in the G20, the Aqaba Process, the Five Country Ministerial, and a range of other fora.**

References

- Ardern, J. (2019). NZ and France seek to end use of social media for acts of terrorism. Media release, April 24, 2019. Accessed February 8, 2021, from <https://www.beehive.govt.nz/release/nz-and-france-seek-end-use-social-media-acts-terrorism>
- Ardern, J. (2020). Christchurch Call: One year anniversary. Joint statement with President Emmanuel Macron, May 15, 2020. Accessed October 6, 2020, from <https://www.beehive.govt.nz/release/christchurch-call-one-year-anniversary>
- Australian Communications and Media Authority. (2019). ACMA investigation into coverage by Australian television broadcasters of the Christchurch terrorist attack, July 2019. Accessed October 8, 2020, from <https://www.acma.gov.au/node/576>
- Bennett, J. (2020). What if instead of calling people out, we called them in? *New York Times*, November 19, 2020. Accessed November 24, 2020, from <https://nyti.ms/32UOqNN>
- Biden, J. (2021). Inaugural Address by President Joseph R. Biden, Jr. The White House, January 20, 2021. Accessed January 21, 2021, from <https://www.whitehouse.gov/briefing-room/speeches-remarks/2021/01/20/inaugural-address-by-president-joseph-r-biden-jr/>
- Binder, L., Ueberwasser, S., & Stark, E. (2020). Gendered hate speech in Swiss WhatsApp messages. In G. Giusti & G. Iannàccaro (Eds.), *Language gender and hate speech: A multidisciplinary approach* (pp. 59–74). Venice: Edizioni Ca' Foscari. <https://doi.org/10.30687/978-88-6969-478-3>
- Breland, A. (2019). Anti-Muslim hate has been rampant on Reddit since the New Zealand shooting. *Mother Jones*, March 21, 2019. Accessed December 9, 2020, from <https://www.motherjones.com/politics/2019/03/reddit-new-zealand-shooting-islamophobia/>
- Burke, J. (2019). Technology is terrorism's most effective ally. It delivers a global audience. *The Guardian*, March 17, 2019. Accessed December 9, 2020, from <https://www.theguardian.com/commentisfree/2019/mar/17/technology-is-terrorisms-most-effective-ally-it-delivers-a-global-audience>
- Cadwalladr, C. (2017). The great British Brexit robbery: how our democracy was hijacked. *The Guardian*, May 7, 2017. Accessed October 12, 2020, from <https://www.theguardian.com/technology/2017/may/07/the-great-british-brexite-robbery-hijacked-democracy>
- Carter, J., & Komaitis, K. (2019). How to regulate the internet without shackling its creativity. *Stuff*, May 15, 2019. Accessed October 15, 2020, from <https://www.stuff.co.nz/national/christchurch-shooting/112704241/how-to-regulate-the-internet-without-shackling-its-creativity>
- Chang, D. (2019). Jacinda Ardern says Facebook should have removed gunman's livestream immediately. *NZ Herald*, April 11, 2019. Accessed October 7, 2020, from https://www.nzherald.co.nz/nz/news/article.cfm?c_id=1&objectid=12221523
- Chen, S. (2020). The spread of online fascism. In A. Chen (Ed.), *Shouting zeros and ones: Digital technology, ethics and policy in New Zealand* (pp. 151–223). Wellington: Bridget Williams Books.
- Classification Office. (2019a). Christchurch mosque attack livestream. Notice of decision under Section 38(1). Reference 1900148.000, March 18, 2019. Wellington: Office of Film and Literature Classification. Accessed August 23, 2020, from <https://register.classificationoffice.govt.nz>
- Classification Office. (2019b). The Great Replacement. Notice of decision under Section 38(1). Reference 1900149.000, March 23, 2019. Wellington: Office of Film and Literature Classification. Accessed August 23, 2020, from <https://register.classificationoffice.govt.nz>
- Conger, K., Isaac, M., & Frenkel, S. (2021). Twitter and Facebook lock Trump's accounts after violence on Capitol Hill. *New York Times*, January 6, 2021. Accessed January 7, 2021, from <https://nyti.ms/3q8Lpmz>

- Council of Europe. (2020). Details of Treaty No.196, Council of Europe Convention on the Prevention of Terrorism. Accessed October 7, 2020, from <https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/196>
- Cox, J. (2018). These are Facebook's policies for moderating white supremacy and hate. *Vice Motherboard*, May 29, 2018. Accessed October 8, 2020, from https://www.vice.com/en_us/article/mbk7ky/leaked-facebook-neo-nazi-policies-white-supremacy-nationalism-separatism
- Cox, J. (2019). Twitter and YouTube won't commit to can white nationalism after Facebook makes policy switch. *Vice Motherboard*, April 3, 2019. Accessed October 8, 2020, from https://www.vice.com/en_us/article/mbzz8x/twitter-youtube-wont-ban-white-nationalism-facebook
- Cox, J. and Koebler, J. (2019a). Facebook bans white nationalism and white separatism. *Vice Motherboard*, March 28, 2019. Accessed October 8, 2020, from https://www.vice.com/en_us/article/nexpbx/facebook-bans-white-nationalism-and-white-separatism
- Cox, J. and Koebler, J. (2019b) Why won't Twitter treat white supremacy like Isis? Because it would mean banning some republican politicians too. *Vice Motherboard*, April 26, 2019. Accessed October 8, 2020, from https://www.vice.com/en_us/article/a3xgq5/why-wont-twitter-treat-white-supremacy-like-isis-because-it-would-mean-banning-some-republican-politicians-too
- Daalder, M. (2020). The ISP scramble to block the Christchurch video. *Newsroom*, November 20, 2019. Accessed January 4, 2021, from <https://www.newsroom.co.nz/how-isps-scrambled-after-christchurch>
- Dearden, L. (2019). Neo-Nazi groups allowed to stay on Facebook because they 'do not violate community standards'. *The Independent*, March 24, 2019. Accessed October 8, 2020, from <https://www.independent.co.uk/news/uk/home-news/facebook-new-zealand-neo-nazis-white-supremacists-a8837886.html>
- Dwoskin, E., & Timberg, C. Christchurch mosque shootings: Inside YouTube's struggles to shut down video—and the humans who outsmarted its systems. *Washington Post*, March 19, 2019. Accessed December 9, 2020, from <https://www.nzherald.co.nz/business/christchurch-mosque-shootings-inside-youtubes-struggles-to-shut-down-video-and-the-humans-who-outsmarted-its-systems/XUD63DODCSGK3R2RC73SI7KWBO/>
- Ellis, G., & Muller, D. (2020). The proximity filter: The effect of distance on media coverage of the Christchurch mosque attacks. *Kōtuitui: New Zealand Journal of Social Sciences Online*, 15(2), 332–348. <https://doi.org/10.1080/1177083X.2019.1705358>
- Ensor, J. (2019). Christchurch shooting videos remain on social media—expert. *Newshub*, September 10, 2019. Accessed October 8, 2020, from <https://www.newshub.co.nz/home/new-zealand/2019/09/christchurch-shooting-videos-remain-on-social-media-expert.html>
- Etzioni, A. (2019). My experience with social media restrictions on free speech. *The National Interest*, December 15, 2019. Accessed October 16, 2020, from <https://nationalinterest.org/feature/my-experience-social-media-restrictions-free-speech-105127>
- European Commission. (2019). Fighting Terrorism Online: EU Internet Forum committed to an EU-wide Crisis Protocol. October 7, 2019. Accessed October 7, 2020, from https://ec.europa.eu/commission/presscorner/detail/en/ip_19_6009
- Fielitz, M., & Schwarz, K. (2020). *Hate not found?! Deplatforming the far right and its consequences*. Institut für Demokratie und Zivilgesellschaft / Amadeu-Antonio-Stiftung. Accessed December 10, 2020, from <https://www.idz-jena.de/forschung/hate-not-found-das-deplatforming-der-extremen-rechten/>
- Flew, T., Martin, F., & Suzor, N. (2019). Internet regulation as media policy: Rethinking the question of digital communication platform governance. *Journal of Digital Media & Policy*, 10(1), 33–50. https://doi.org/10.1386/jdmp.10.1.33_1
- GIFCT. (n.d.). *Global Internet Forum to Counter Terrorism*. Accessed October 5, 2020, from <https://gifct.org/>

- GIFCT. (2020a). Update to GIFCT statement on Halle shooting. October 17, 2019. Accessed February 16, 2021, from <https://gifct.org/2019/10/17/update-to-gifct-statement-on-halle-statement/>
- GIFCT. (2020b). Update on Glendale, AZ shooting, May 22, 2020. Accessed February 23, 2021, from <https://gifct.org/2020/05/22/update-on-glendale-az-shooting/>
- GIFCT. (2020c). GIFCT Transparency report, July 2020. Accessed October 6, 2020, from <https://gifct.org/transparency/>
- GIFCT. (2021). Joint tech innovation. Accessed February 22, 2021, from <https://gifct.org/joint-tech-innovation/>
- Gilbert, J. & Elley, B. (2020). Shaved heads and Sonnenrads: Comparing white supremacist skinheads and the alt-right in New Zealand. *Kōtuitui: New Zealand Journal of Social Sciences Online*, 15(2), 280-294. <https://doi.org/10.1080/1177083X.2020.1730415>
- House of Commons Home Affairs Committee. (2017). Hate crime: abuse, hate and extremism online. Fourteenth report of session 2016–17, April 25, 2017. Accessed October 12, 2020, from <https://publications.parliament.uk/pa/cm201617/cmselect/cmhaff/609/609.pdf>
- Human Rights Watch. (2020). 'Video unavailable': Social media platforms remove evidence of war crimes, September 10, 2020. Accessed October 16, 2020, from <https://www.hrw.org/report/2020/09/10/video-unavailable/social-media-platforms-remove-evidence-war-crimes>
- InternetNZ. (2019). Civil society positions on Christchurch Call pledge. May 14, 2020. Accessed October 15, 2020, from <https://docs.google.com/document/d/10RadyVQUNu1H5D7x6IJVKbqmaeDeXre0Mk-FFNklVxs/edit>
- Keall, C. (2019). As Jacinda Ardern prepares social media crackdown, alleged gunman's video reappears—and gets huge traffic. *NZ Herald*, April 23, 2019. Accessed October 7, 2020, from https://www.nzherald.co.nz/business/news/article.cfm?c_id=3&objectid=12224469
- Isaac, M. (2019). Mark Zuckerberg's call to regulate Facebook, explained. *New York Times*, March 30, 2019. Accessed October 12, 2020, from <https://nyti.ms/2UI4gOK>
- Keall, C. (2019). As Jacinda Ardern prepares social media crackdown, alleged gunman's video reappears—and gets huge traffic. *NZ Herald*, April 23, 2019. Accessed October 7, 2020, from https://www.nzherald.co.nz/business/news/article.cfm?c_id=3&objectid=12224469
- Knight, B. (2020). Neo-Nazi attack survivors create tool to track racist extremists. *Deutsche Welle*, December 16, 2020. Accessed December 16, 2020, from <https://www.dw.com/en/neo-nazi-attack-survivors-create-tool-to-track-racist-extremists/a-55959989>
- Lowen, M. (2019). Christchurch shootings: Why Turkey's Erdogan uses attack video. *BBC News*, March 20, 2019. Accessed October 8, 2020, from <https://www.bbc.com/news/world-europe-47638936>
- Macklin, G. (2019). The Christchurch attacks: Livestream terror in the viral video age. *Combating Terrorism Centre*, 12(6), July 2019. Accessed December 9, 2020, from <https://ctc.usma.edu/christchurch-attacks-livestream-terror-viral-video-age/>
- McMaken, R. (2019). Three reasons why Facebook's Zuckerberg wants more government regulation. *Mises Institute Wire*, April 1, 2019. Accessed October 12, 2020, from <https://mises.org/wire/3-reasons-why-facebooks-zuckerberg-wants-more-government-regulation>
- Mason, C., & Errington, K. (2019). *Anti-social media: Reducing the spread of harmful content on social media networks*. Auckland: Helen Clark Foundation. Accessed October 15, 2020, from <https://helenc Clark.foundation/reports/anti-social-media/>
- Miller-Idriss, C. (2020). *Hate in the homeland: The new global far right*. Princeton, NJ: Princeton University Press.

- Muller, D. (2019). Media watchdog's report into Christchurch shootings goes soft on showing violent footage. *The Conversation*, July 24, 2019. Accessed October 8, 2020, from <https://theconversation.com/media-watchdogs-report-into-christchurch-shootings-goes-soft-on-showing-violent-footage-120577>
- NZ Ministry of Foreign Affairs & Trade. (n.d.) The Christchurch Call. Accessed October 5, 2020, from <https://www.christchurchcall.com/>
- Ortutay, B. (2020). Facebook has 'devastating' civil rights record, its own audit says. *Stuff*, July 10, 2020. Accessed October 9, 2020, from <https://www.stuff.co.nz/technology/social-networking/300053939/facebook-has-devastating-civil-rights-record-its-own-audit-says>
- Owen, T. (2019). Decoding the racist memes the alleged New Zealand shooter used to communicate. *Vice*, March 15, 2019. Accessed December 14, 2020, from <https://www.vice.com/en/article/vbwn9a/decoding-the-racist-memes-the-new-zealand-shooter-used-to-communicate>
- Petersmann, S. (2019). Christchurch: The myth of the lone wolf attacker. *Deutsche Welle*, March 18, 2019. Accessed October 12, 2020, from <https://p.dw.com/p/3FHBr>
- RNZ. (2019). 4000 people watched Facebook video before anyone reported it. Radio New Zealand, March 19, 2020. Accessed October 7, 2020, from <https://www.rnz.co.nz/news/national/385127/4000-people-watched-facebook-video-before-anyone-reported-it>
- Robertson, A. (2019). An anti-Semitic shooting in Germany was live-streamed on Twitch. *The Verge*, October 9, 2019. Accessed October 7, 2020, from <https://www.theverge.com/2019/10/9/20906637/halle-germany-synagogue-shooting-attack-live-stream-twitch-amazon-video-unavailable>
- Rosen, G. (2019a). A further update on New Zealand terrorist attack. Facebook, March 20, 2019. Accessed October 15, 2020, from <https://about.fb.com/news/2019/03/technical-update-on-new-zealand/>
- Rosen, G. (2019b). An update on content removals following the Christchurch attack. Facebook Australia & New Zealand Policy, November 13, 2019. Accessed October 5, 2020, from <https://www.facebook.com/notes/facebook-australia-new-zealand-policy/an-update-on-content-removals-following-the-christchurch-attack/2472694583057800/>
- Royal Commission of Inquiry. (2020). *Ko tō tātou kāinga tēnei. [This is our home.] Report: Royal Commission of Inquiry into the terrorist attack on Christchurch masjidain on 15 March 2019*. November 26, 2020. Accessed December 8, 2020, from <https://christchurchattack.royalcommission.nz/>
- Schmitt, J., Harles, D., & Rieger, D. (2020). Themen, Motive und Mainstreaming in rechtsextremen Online-Memes. *Medien & Kommunikationswissenschaft*, 68(1–2), 73–93. <https://doi.org/10.5771/1615-634X-2020-1-2-73>
- Smith, L. (2021). Lone wolves connected online: A history of modern white supremacy. *New York Times*, January 26, 2021. Accessed January 28, 2021, from <https://nyti.ms/3a8fiN4>
- Thompson, P. (2019). Beware of geeks bearing gifts: Assessing the regulatory response to the Christchurch Call. *Political Economy of Communication*, 7(1), 83–104. Accessed October 7, 2020, from <https://iamcr.org/swg/section/poe/journal-7-1>
- Timberg, C, Harwell, D., Shaban, H., Ba Tran, A., & Fung, B. (2019). The New Zealand shooting shows how YouTube and Facebook spread hate and violent images — yet again. *Washington Post*, March 15, 2019. Accessed February 6, 2021, from <https://www.washingtonpost.com/technology/2019/03/15/facebook-youtube-twitter-amplified-video-christchurch-mosque-shooting/>
- Ullmann, S., & Tomalin, M. (2020). Quarantining online hate speech: technical and ethical perspectives. *Ethics and Information Technology*, 22, 69–80. <https://doi.org/10.1007/s10676-019-09516-z>
- UN Security Council. (2017). Unanimously Adopting Resolution 2354 (2017), Security Council Urges Member States to Follow New Guidelines on Countering Terrorist Narratives. May 24, 2017. Accessed October 7, 2020, from <https://www.un.org/press/en/2017/sc12839.doc.htm>

- Warzel, C. (2019). The New Zealand massacre was made to go viral. *New York Times*, March 15, 2019. Accessed October 5, 2020, from <https://nyti.ms/2Y2MhMj>
- Waterson, J. (2019). Facebook removed 1.5m videos of New Zealand terror attack in first 24 hours. *The Guardian*, March 17, 2019. Accessed October 8, 2020, from <https://www.theguardian.com/world/2019/mar/17/facebook-removed-15m-videos-new-zealand-terror-attack>
- Williams, D. (2019). Terror on TV news: Lessons from Australia. *Newsroom*, August 1, 2019. Accessed October 8, 2020, from <https://www.newsroom.co.nz/2019/08/01/710147/terror-on-tv-news-lessons-from-australia#>
- York, J. (2019). The EU's Proposal to curb the 'dissemination of terrorist content' will have a chilling effect on speech. Electronic Frontier Foundation, February 5, 2019. Accessed October 15, 2020, from <https://www.eff.org/deeplinks/2019/02/eus-proposal-curb-dissemination-terrorist-content-will-have-chilling-effect-speech>
- Zaki, A. (2020). Christchurch mosque attacks: Victims call for more support. *NZ Herald*, December 21, 2020. Accessed December 30, 2020, from <https://www.nzherald.co.nz/nz/christchurch-mosque-attacks-victims-call-for-more-support/6P32CVA2XAJKGMP7ARFTVYOMY/>
- Zuckerberg, M. (2019a). A privacy-focused vision for social networking. *Facebook Comment*, March 6, 2019). Accessed October 12, 2020, from <https://www.facebook.com/notes/mark-zuckerberg/a-privacy-focused-vision-for-social-networking/10156700570096634/>
- Zuckerberg, M. (2019b). The Internet needs new rules. Let's start in these four areas. *Washington Post*, March 30, 2019. Accessed October 12, 2020, from https://www.washingtonpost.com/opinions/mark-zuckerberg-the-internet-needs-new-rules-lets-start-in-these-four-areas/2019/03/29/9e6f0504-521a-11e9-a3f7-78b7525a8d5f_story.html