IAN AXFORD
(NEW ZEALAND)
FELLOWSHIPS
in
PUBLIC POLICY

The Future of the Christchurch Call to Action

How to Build Multistakeholder Initiatives to Address Content Moderation Challenges

Prepared by
Rachel Wolbers

With funding from the sponsors of the
Ian Axford (New Zealand) Fellowships in Public Policy

August 2023

Established by the
New Zealand government in 1995
to facilitate public policy dialogue
between New Zealand and
the United States of America

Level 5, 26 The Terrace
PO Box 3465
Wellington 6011
New Zealand

Telephone+64 4 494 1506
E-mail fellowships@fulbright.org.nz
axfordfellowships.org.nz/

Published by Fulbright New Zealand, August 2023

ISBN 978-1-7386094-2-0 (print)
ISBN 978-1-7386094-3-7 (PDF)

IAN AXFORD FELLOWSHIP IN PUBLIC POLICY BACKGROUND

Ian Axford (New Zealand) Fellowships in Public Policy established by the New Zealand Government in 1995 to reinforce links between New Zealand and the US, Ian Axford (New Zealand) Fellowships in Public Policy provide the opportunity for outstanding mid-career professionals from the United States of America to gain firsthand knowledge of public policy in New Zealand, including economic, social and political reforms and management of the government sector.

The Ian Axford (New Zealand) Fellowships in Public Policy were named in honour of Sir Ian Axford, an eminent New Zealand astrophysicist and space scientist who served as patron of the fellowship programme until his death in March 2010.

Educated in New Zealand and England, Sir Ian held Professorships at Cornell University and the University of California, and was Vice-Chancellor of Victoria University of Wellington for three years. For many years, Sir Ian was director of the Max Planck Institute for Aeronomy in Germany, where he was involved in the planning of several space missions, including those of the Voyager planetary explorers, the Giotto space probe and the Ulysses galaxy explorer. Sir Ian was recognised as one of the great thinkers and communicators in the world of space science, and was a highly respected and influential administrator. A recipient of numerous science awards, he was knighted and named New Zealander of the Year in 1995.

Ian Axford (New Zealand) Fellowships in Public Policy have three goals:

- To reinforce United States/New Zealand links by enabling fellows of high intellectual ability and leadership potential to gain experience and build contacts internationally.
- To increase fellows' ability to bring about changes and improvements in their fields of expertise by the cross-fertilisation of ideas and experience.
- To build a network of policy experts on both sides of the Pacific that will facilitate international policy exchange and collaboration beyond the fellowship experience.

Fellows are based at a host institution and carefully partnered with a leading specialist who will act as a mentor. In addition, fellows spend a substantial part of their time in contact with relevant organisations outside their host institutions, to gain practical experience in their fields.

The fellowships are awarded to professionals active in the business, public or non-profit sectors. A binational selection committee looks for fellows who show potential as leaders and opinion formers in their chosen fields. Fellows are selected also for their ability to put the experience and professional expertise gained from their fellowship into effective use.

We acknowledge and thank the following sponsors that support the Ian Axford (New Zealand) Fellowships in Public Policy programme:

- Department of Internal Affairs
- Ministry of Defence
- Ministry of Education
- Ministry of Foreign Affairs and Trade
- Ministry of Health
- Ministry of Justice
- Ministry of Social Development New Zealand Customs Service
- Ministry of Business, Innovation and Employment
- Public Services Commission
- Victoria University of Wellington School of Government

ACKNOWLEDGEMENTS

My report and fellowship experience would not have been possible or as rewarding without the support and kindness of many people. Though by no means an exhaustive list, I would like to thank the following:

- Roy Ferguson and the rest of the Axford Fellowships Board of Directors
- Lauren Summersell, Penelope Borland, and the Fulbright New Zealand team
- The Department of Cabinet and Prime Minister Christchurch Call Unit and Ministry of Foreign Affairs and Trade including our fearless leader, Paul Ash, and the best group of Kiwi coworkers anyone could ever ask for: David Reid, Elisabeth Brown, Kristina Kirk, Aline De Vincentis, Ellen Strickland, Hisham Eldai, Hayden Glass, Aimee Jephson, Hannah Gwillim, Pip Robertson, and the Prime Minister's Special Envoy to the Christchurch Call Jacinda Ardern.
- Karl Lofgren and Ali Knott, my academic support network at the University of Victoria, Wellington.
- The Oversight Board team for allowing me to take six months off to move to New Zealand and pursue a dream of working with the Christchurch Call to Action team. A special thanks to my Communications Team colleagues who stepped up to fill in for me (Simona Sikimic-French, Tracy Manners, Matthew Sells, Francesca Scapolo, Holly Richards, Jesse Spector, Kyla-Marie Greenway, Dex Hunter-Torricke, and Jeff Gelman), the Oversight Board Executive and People Teams who granted me the distinction of being the first Oversight Board employee to take a sabbatical (Thomas Hughes, Morteza Mollanaghi, Aneka Bellamy Osakwe, David Connell) and Nino Guruli, the Oversight Board cat-sitter who truly made this trip possible for me.
- The team at NTIA who took a chance on me and inspired this journey: Fiona Alexander, Vernita Harris, David Redl, Diane Rinaldo, Susan Chalmers, Ryan Carroll, Kate O'Connor, and Jim Wasilewski.
- My wonderful mentors who wrote letter of recommendation, consistently encouraged me to chase new opportunities, taught me to believe in myself, and fueled my Section 230 obsession: John Taylor, Elizabeth Frazee, Julie Owono, Nighat Dad, Nic Suzor, Emma Llanso, Shane Tews, Kate Klonick, Daphne Keller, and Eric Goldman.
- My fellow conspirator on all things AI, Austin Carson.
- Last, but certainly not least, my family, Rich, Jens, Steph, Will and Jax, for all their love and support. And the amazing ladies who came to visit me at the bottom of the world, Lisa, Delaney, Sara, and Amy!

EXECUTIVE SUMMARY

This report explores the challenges the New Zealand Government faced after the events in Christchurch on 15 March 2019, where a violent gunman killed 51 people and live-streamed his attack on social media. The video was viewed millions of times in the days following, even as the tech companies took extraordinary efforts to reduce its virality. To find a long-term solution that ended the proliferation of this violent content while protecting human rights, the New Zealand Government decided to take a non-regulatory approach that worked alongside tech companies and civil society. The result was the creation of the Christchurch Call to Action, a multistakeholder initiative (MSI) where governments and online platforms, working with civil society, committed to 25 goals to eliminate terrorist and violent extremist content while protecting a free, open, and secure internet.

This report argues that the creation of an MSI was not only the right option for the New Zealand Government in the aftermath of Christchurch shooting, but that multistakeholderism is the best approach for addressing all issues related to the governance of user-generated content online. An MSI works with multiple actors (such as governments, corporations, civil society organisations, charitable foundations, academia, technical experts, or end-users) to find solutions to a shared problem that one group could not solve acting alone. Through an MSI, stakeholders can harness the capabilities of different actors and co-design solutions through participatory processes. The problems related to the proliferation of harmful content online cannot be solved through government regulation, and tech companies cannot, and should not, set the rules alone. Therefore, to find a solution, governments and companies must work with like-minded actors who uphold human rights principles, and meaningfully engage with civil society, technical experts, academia, and users. These solutions should be consensus-based and build in accountability mechanisms for both governments and companies. Unfortunately, the proliferation of terrorist content is only one small item on a long menu of harmful content online. However, it is an area where broader definitional consensus allows stakeholders to address thornier technical and human rights challenges. This report argues that solutions proposed addressing terrorist content could serve as a guide for other types of user-generated content where definitions remain contentious.

Part I of this report first unpacks the history of single-sided approaches to governing content online, specifically national regulatory frameworks and tech company self-regulation efforts. As this report details, neither of these approaches have successfully addressed the spread of harmful user-generated content online in a way that protects human rights, engenders consumer trust, or considers local context. To remedy these challenges, this report argues that governments, tech companies, and civil society organisations should adopt multistakeholder governance structures which have been foundational in the creation of the internet. This report provides a history of multistakeholderism and internet governance and explores recent efforts to undermine this model by authoritarian actors.

In Part II, this report distills the lessons learned from 30 years of internet governance MSIs to build a framework for MSIs seeking to address challenges related to the governance of user-generated content online. First, this part creates a taxonomy of MSIs based on their inclusion

of stakeholders and the distribution of decision-making authority. While each MSI has a unique structure, this taxonomy establishes four overarching types of MSIs: egalitarian, consultative, restricted, and curated. This report argues that MSIs addressing challenges related to user-generated content online should adopt a "curated" approach which limits participation to stakeholders who can demonstrate their commitment to human rights principles and ensures consensus-based decision-making authority. Second, this part establishes six steps for organisers to consider when running an MSI, the first four relate to its establishment, the fifth to its sustainability, and the sixth to its closure. The steps identified include deciding if an MSI is necessary, establishing the objectives and functions of the initiative, defining who is a stakeholder, setting up terms of reference, sustaining forward momentum, and deciding when the work is finished.

In Part III, this report examines the events of 15 March 2019 and the establishment of the Call. This part evaluates the progress the Call has made over the past four years towards building a multistakeholder community and eliminating terrorist and violent extremist content online while protecting a free, open, and secure internet – its two overarching objectives. Next, this part explores the future of the Call and provides suggested next steps to help the organisation create a self-sustaining MSI. Additionally, this part looks at the future of the Call and how the rise of Generative Artificial Intelligence (GenAI) may present a turning-point for the organisation. Since the start of this report in February 2023, there has been a rapid development and adoption of GenAI technologies which create both challenges and opportunities for companies seeking to combat the prevalence of terrorist and violent extremist content online. Stakeholders around the world are grappling with ways to create guardrails around GenAI without limiting its hindering innovation – precisely the type of complex issues MSIs are designed to address. Therefore, this part outlines ways the Call could expand its current work on algorithms and artificial intelligence to address issues related to the governance of GenAI and content moderation.

Table of Contents

PREFACE

This report examines one of the darkest days in New Zealand's history and the efforts made by government officials, technology company employees, and civil society organisations in its aftermath. On the afternoon of Friday, 15 March 2019, a violent extremist livestreamed on Facebook his murder of 51 people worshiping in mosques in Christchurch, New Zealand. On the other side of the world, I woke on 15 March to news of the attack a few hours before starting a new position as an internet policy specialist at the Office of International Affairs at the United States Department of Commerce's National Telecommunications and Information Administration (NTIA). This job was a continuation of my career at the intersection of the law, policy, technology, and freedom of expression issues. This new position would also place me on the policy team within the United States Government responding to the attack. From my previous work with government, tech companies, and civil society, it was clear to me that the horrific events in Christchurch would become a defining moment in the history of social media and the internet itself.

My path to New Zealand started back in 2009, when I became interested in internet policy issues after joining the Journal of Law, Technology and the Internet at Case Western Reserve University School of Law. I published my first paper on content moderation issues related to terrorists' use of the internet in 2010. In this legal article, I examined whether Wikileaks could be held liable for publishing information online that aided the Taliban and al-Qaeda in their killing of several US informants.[1] After law school, I found a job where I could put this somewhat niche policy interest to use as Legal Counsel to Representative Blake Farenthold, a self-described "tech nerd" serving on the House Judiciary Committee. While I was working on Capitol Hill, Congress was primarily concerned with copyright and surveillance issues when discussing internet law and policy. However, the discussions started to change in 2014, as the Islamic State of Iraq and Syria (ISIS) began to use American online platforms to post beheading videos and other propaganda.[2]

From 2014 to 2017, I worked at a boutique consulting firm in Washington, DC representing several internet companies grappling with how to remove terrorists and violent extremist content from their platforms. Our clients were shielded from legal liability under Section 230 of the Communications Decency Act,[3] but they were keen to find ways to work with stakeholders from government and civil society to find solutions to combat the problem of terrorist abuse of internet platforms. A passion for stakeholder engagement, and content moderation, led me to accept a job as policy director for Engine Advocacy, a non-profit organisation that worked with internet startups on a variety of policy issues. In my role at Engine Advocacy, I built coalitions with internet companies, civil society, academics, and

[1] Rachel Wolbers, *Is WikiLeaks a Hit Man Handbook?: Why WikiLeaks Cannot Claim First Amendment Immunity if the Afghan and Iraq War Logs Cause Physical Harm*, 3 CASE W. RES. J.L. TECH. & INTERNET 365 (2012), https://scholarlycommons.law.case.edu/jolti/vol3/iss2/7.
[2] Ahmad Shehabat & Teodor Mitew, *Black-boxing the Black Flag: Anonymous Sharing Platforms and ISIS Content Distribution Tactics*, 12 PERSPECTIVES ON TERRORISM 81 (Feb. 2018), https://www.jstor.org/stable/26343748.
[3] 47 U.S.C. § 230 (2021).; Section 230 will be discussed at length in Part I below.

technical experts to educate policymakers on how legal frameworks regulating speech online could foster, or hinder, technological innovation and human rights.

This brings me to 2019, when I was thrilled to join the Office of International Affairs as a career civil servant to continue working on internet policy issues. NTIA is the Executive Branch agency within the US Government with primary responsibility for advising the president on telecommunications and internet policy issues.[4] Additionally, as this report will discuss at length, NTIA has long championed multistakeholder solutions to technology policy issues. Therefore, while I had a background in internet law and a robust network of contacts within the technology industry and civil society, the job at NTIA was an opportunity for me to learn from world-leading experts on how to build successful multistakeholder initiatives. Shortly after joining the team, it was surrounded by these experts that I found myself debating whether the US government should join a new multistakeholder initiative led by New Zealand and France called the Christchurch Call to Action (the Call). While many of the goals of the Call aligned closely with existing US government policies to combat terrorist abuse of the internet, the Trump Administration ultimately decided not to support the Call. However, many American businesses and civil society organisations did. As a result, the US government monitored the work of the Call and assisted US stakeholders where their efforts aligned with US government priorities. In my role at NTIA, I was honoured to join colleagues from the White House, the State Department, the Department of Justice, and others to support the goals and priorities of the Call.

In the four years since the formation of the Call, I have left the US Government to join Meta's Oversight Board as the head of global engagement. Meta set up the Oversight Board in 2020 as an independent body that issues binding opinions on content moderation decisions taken by Facebook and Instagram and provides policy recommendations to the company. At the Board, I led outreach efforts to stakeholders from around the world to help ensure the Board's decisions and policy recommendations reflect local context and uphold international human rights principles. The Board is the first institution of its kind, and our team is continuously innovating to build multistakeholder solutions to content moderation challenges – including those relating to terrorist and violent extremist content. As part of this work, I helped create several of the Board's stakeholder engagement best practices, including a public comment process which allows stakeholders to provide feedback on cases and policy issues, monthly stakeholder roundtables that bring together experts on specific topics, and the Board's Quarterly Transparency Report. While the Oversight Board does not see itself as a multistakeholder initiative, its engagement work draws upon the lessons learned from the multistakeholder models explored in this report. As a result, I hope to bring some of these ideas back with me to the Oversight Board and put into practice what I have learned as an Axford Fellow.

I first heard about the Ian Axford Fellowship in Public Policy programme from my boss at NTIA, Fiona Alexander, an Axford alum. Having worked with the Call team throughout 2019,

[4] National Telecommunications and Information Administration, *Internet Policy*, NATIONAL TELECOMMUNICATIONS AND INFORMATION ADMINISTRATION (2023), https://ntia.gov/category/internet-policy.

I was impressed by the New Zealand Government's leadership on content moderation issues and thought the Call's multistakeholder initiative was both unique and important. With Fiona's encouragement, I applied as soon as I could in 2020. Unfortunately, the programme was put on hold during the COVID-19 pandemic. However, two years later, in a true testament to the dedication of New Zealanders, the Call was still going strong, and I was able to reapply for this amazing opportunity. In the months since I moved to New Zealand in January 2023, the content moderation landscape has rapidly changed with the advancement in generative artificial intelligence technologies. Like many stakeholders worldwide, this report, and the Call itself, has had to constantly adapt to address the enormous opportunities and challenges posed by artificial intelligence. Despite the uncertain technological future, I hope this report will be helpful to the Call team as it continues to evolve and strengthen its multistakeholder model.

Before getting to the report, I must issue a few necessary disclaimers. All the thoughts and opinions expressed here are my own and do not constitute the position of either the Department of the Prime Minister and Cabinet or the Oversight Board. All facts and information in this report are from publicly available information, and this report contains no privileged information I have learned either from working with the Christchurch Call or the Oversight Board. While this report heavily draws on the work of experts in the field, it is also reflective of the lessons I have learned over a decade of building multistakeholder coalitions to address internet governance issues. Additionally, the report reflects the many ongoing conversations I have had while living in New Zealand, including with government officials, academics, civil society, and tech companies working on the Call. Any critique given or advice offered comes from a place of tremendous respect, appreciation, and admiration for the work Call supporters are doing. The task of building a multistakeholder initiative to eliminate terrorist and violent extremist content online while upholding a free, open, and secure internet is not an easy one. If anyone is up for the job, however, it is the Christchurch Call to Action.

Introduction

On 15 March 2019, a gunman in the city of Christchurch, New Zealand turned on the GoPro video camera mounted on his helmet, linked the livestream to his Facebook account, and entered the Al Noor Mosque.[5] He proceeded to broadcast his brutal killing of 51 worshippers for 16 minutes and 55 seconds on Facebook.[6] This horrific attack was carefully planned to spread rapidly across the internet. And it did. In the first 24 hours, platforms such as YouTube, Twitter, Facebook, and Reddit removed millions of copies of the video.[7] The exploitation of social media compounded the tragedy of 15 March and New Zealanders sprang into action to eliminate this type of violence and horror online. As Prime Minister Jacinda Ardern wrote, "a terrorist attack like the one in Christchurch could happen again unless we change. New Zealand could reform its gun laws, and we did. We can tackle racism and discrimination, which we must. We can review our security and intelligence settings, and we are. But we can't fix the proliferation of violent content online by ourselves."[8]

In the weeks following, Ardern partnered with French President Emanuel Macron to bring together governments, technology companies, and civil society to adopt a set of commitments to eliminate terrorist and violent extremist content online, known as the Christchurch Call to Action (the Call). At the core of the Call, governments and tech companies agreed to make changes to prevent the posting of terrorist content online, to ensure its efficient and fast removal, and to prevent the use of livestreaming as a tool for broadcasting terrorist attacks.[9] To succeed, the group would need to work closely with civil society to ensure freedom of expression was protected and the voices of the victims and survivors heard. Emerging from this coalition was a multistakeholder initiative (MSI) designed to address a complicated problem. The answer was not one that could be solved easily through government regulation, company policies and technical measures, or civil society efforts on their own. Instead, the Call engaged a whole-of-society approach whereby stakeholders worked together to tackle the problem.[10]

Four years later, the Call remains dedicated to fulfilling the initial 25 commitments governments and companies set out on 15 May 2019.[11] Over the years, the Call has added members, partnered with similar initiatives, launched new work-streams, and adapted as technology changes. As is to be expected when addressing such complicated problems, the Call has made significant progress on some commitments and is still working on others. This report seeks to discuss why New Zealand could not stop the spread of terrorist and violent extremist

[5] Royal Commission of Inquiry into the Attack on Christchurch Mosques, *He Ara Waiora: Report of the Royal Commission of Inquiry into the Attack on Christchurch Mosques on 15 March 2019,* ROYAL COMMISSION OF INQUIRY IN NEW ZEALAND, 11 (8 Dec. 2020), https://christchurchattack.royalcommission.nz/the-report/.
[6] Jacinda Ardern, *How to Stop the Next Christchurch Massacre*, THE NEW YORK TIMES (11 May 2019), https://www.nytimes.com/2019/05/11/opinion/sunday/jacinda-ardern-social-media.html.
[7] *Id*. Facebook alone removed over 1.5 million copies of the video within the first 24 hours.
[8] *Id.*
[9] *Id.*
[10] Jacinda Ardern, *Here's the Model for Governing AI*. WASH. POST (11 June 2023), https://www.washingtonpost.com/opinions/2023/06/09/jacinda-ardern-ai-new-zealand-planning/.
[11] *The Christchurch Call to Action: To Eliminate Terrorist and Violent Extremist Content Online*, CHRISTCHURCH CALL (15 May 2019), https://www.christchurchcall.com/assets/ Documents/Christchurch-Call-full-text-English.pdf.

content online on its own – but why it may be able to meaningfully address the problem through multistakeholder solutions. This report will explain why content moderation challenges need multistakeholder solutions and how the Call can embrace this model to achieve its goals. Additionally, this report will discuss how generative artificial intelligence (GenAI) presents challenges and opportunities to eliminate terrorist and violent extremist content online and how the multistakeholder community can consider those issues.

Before discussing why New Zealand opted for a multistakeholder approach, it is important to define the problem the Call is trying to solve to prevent future attacks. As detailed by the Royal Commission of Inquiry into the Terrorist Attack on Christchurch Mosques on 15 March 2019, the Christchurch shooter[12] displayed racist views from a young age, and life experiences drove his extreme and violent behaviour towards people he considered a threat.[13] He legally purchased semi-automatic firearms and evaded police scrutiny throughout his planning process.[14] To address the problems brought to light by the attacks, the Royal Commission made 44 recommendations, including updating gun laws, building inclusive societies, making improvements in intelligence sharing and rethinking hate speech frameworks.[15] This report will only address one element of the events of 15 March – the individual's use of social media platforms to broadcast his violence and the way his terrorist and violent extremist content was able to proliferate online.

Because this report will refer to several specific issues within a much broader set of problems, it is necessary to define several terms to avoid confusion. "Extremism" is defined as a belief system held together by unwavering hostility towards a specific "out-group".[16] In line with the definition provided by the New Zealand Government, a "violent extremist" is an individual who threatens to use violence, or advocates for others to use violence, in support of their own agenda or to further their own set of beliefs.[17] As such, "terrorist and violent extremist content" (TVEC) refers to hateful or objectionable material that promotes harmful extreme views, such as articles, images, speeches, or videos that encourage violence.[18] People can (and do) debate at length on how to define TVEC.[19] However, in relation to the Call and this report, the two pieces of TVEC created by the individual – his manifesto and the video of his attack on the mosques – would fit within any reasonable definition of TVEC. The term "online service

[12] Following the precedent set by the Royal Commission, this article will not name the individual who committed the attack and will only refer to him as the "individual" or "Christchurch shooter" to ensure his name is not glorified, *see* Royal Commission of Inquiry into the Attack on Christchurch Mosques, *supra* note 5, Volume 1 at 11.
[13] *Id*. Volume 1 at 11.
[14] *Id*. Volume 1 at 11.
[15] *Id*. Volume 4 at 727, Part 10: Recommendations.
16 *See* J. M. Berger. *Extremism*. UNITED STATES: MIT PRESS, 2018 (Berger, an expert on extremist movements and terrorism, explains that extremism arises from a perception of "us versus them," intensified by the conviction that the success of "us" is inseparable from hostile acts against "them." Extremism differs from ordinary unpleasantness—run-of-the-mill hatred and racism—by its sweeping rationalization of an insistence on violence).
[17] Department of Internal Affairs, *Countering Violent Extremism: What is terrorist and violent extremist content?*, DEPARTMENT OF INTERNAL AFFAIRS OF NZ (2022), https://www.dia.govt.nz/Countering-Violent-Extremism-What-is-terrorist-and-violent-extremist-content#_ftn1.
[18] *Id*.
[19] Issie Lapowsky, *This Big Tech group tried to redefine violent extremism. It got messy*., PROTOCOL (26 June 2021), https://www.protocol.com/policy/gifct-erin-saltman (interview with Erin Saltman of the Global Internet Forum to Counter Terrorism on a months-long debate trying to define what constituted terrorist and violent extremist content online).

provider", which encompasses online platforms and social media companies, is defined as an online site or service that hosts, organises, or circulates user-generated content without producing content.[20] "Content moderation" is defined as the systems and rules online platforms use to determine how they treat user-generated content on their services.[21]

For several reasons, this report is limited to a discussion of the challenges online service providers face when moderating TVEC and does not discuss other types of harmful content online. First, this report is meant to provide policy guidance to the New Zealand government on how to improve the Call, which remains limited to TVEC.[22] Second, TVEC itself is an area in which there is general agreement that the content itself serves little to no societal value and should therefore be extremely restricted, if not entirely prohibited, from online platforms.[23] This agreement means that TVEC can be a useful test case for broader ongoing discussions around harmful content online, which often involves types of content such as hate speech, bullying, and dis/misinformation, where there is less agreement on definitions and societal value. Third, the challenges posed by TVEC online are as old as the internet itself and have been researched and discussed for decades.[24] As a result, many stakeholders, including governments, the tech industry, and civil society, have attempted to address the issue over the years, which allows for a thorough examination of what has worked – and what has not – when considering the next steps for the Call.

Next, it is important to define the harm that comes from the distribution of TVEC online and the broader societal problem for which the Call is trying to solve. First, harm occurs when viewers are traumatised because of their exposure to seeing violent content.[25] Second, the sharing of TVEC causes harm as a privacy invasion of both the surviving victims and the families of deceased victims. Third, both the Christchurch video and the manifesto are harmful because they may inspire others to commit similar acts of terrorism.[26] In fact, the Christchurch

---

[20] TARLETON GILLESPIE, CUSTODIANS OF THE INTERNET: PLATFORMS, CONTENT MODERATION, AND THE HIDDEN DECISIONS THAT SHAPE SOCIAL MEDIA (New Haven: Yale University Press, 2018); *see also* Robyn Caplan, *Content or Context Moderation? Artisanal, Community-Reliant, and Industrial Approaches*, DATA & SOCIETY, 8 (14 Nov. 2018), https://datasociety.net/library/content-or-context-moderation/.

21 Evelyn Douek, *Content Moderation as Systems Thinking*, 136 Harv. L. Rev. 526, 528 (2022) (defining "content moderation" to mean platforms' systems and rules that determine how they treat user-generated content on their services. This generally accords with Professor James Grimmelmann's definition. *See* James Grimmelmann, *The Virtues of Moderation*, 17 YALE J.L. & TECH. 42, 47 (2015) (defining "moderation" as "the governance mechanisms that structure participation in a community to facilitate cooperation and prevent abuse")).

22 *The Christchurch Call to Action: To Eliminate Terrorist and Violent Extremist Content Online, supra* note 11.

[23] There has been voluminous debate around defining both terrorism and extremism that is outside of the scope of this report. For the purposes of this report, the Christchurch shooter's 74-page manifesto and video of his attack on 15 March 2019 are both considered TVEC as they have been classified as objectionable in New Zealand.

[24] *See* Brian Fishman, *Dual-use regulation: Managing hate and terrorism online before and after Section 230 reform*, THE BROOKINGS INSTITUTION (14 March 2023), https://www.brookings.edu/articles/dual-use-regulation-managing-hate-and-terrorism-online-before-and-after-section-230-reform/.

[25] *The Christchurch Call to Action: To Eliminate Terrorist and Violent Extremist Content Online, supra* note 11.

[26] *See* Office of the New York State Attorney General Letitia James, *Investigative Report on the role of online platforms in the* tragic mass shooting in Buffalo on May 14, 2022, OFFICE OF NEW YORK STATE ATTORNEY GENERAL 17-22 (18 Oct. 2022), https://ag.ny.gov/sites/default/files/buffaloshooting-onlineplatformsreport.pdf ("But the Christchurch shooter also changed the playbook in new, deadlier ways. He was the first white supremacist to livestream his attack, and the video of the shootings went viral. He deliberately sought to create an online footprint that he hoped would be galvanizing and instructional to fellow right-wing extremists. These digital artifacts have proved to be indelible and have radicalized others, including the Buffalo shooter, who deliberately modeled his attack on the Christchurch shooter's.").

shooter credited a far-right extremist attack in Norway in 2011, which killed 77 people, for inspiring his own attack.[27] Unfortunately, over the past four years, several terrorists and violent extremists have been inspired by the Christchurch attacks to livestream their killing of minorities in a variety of places, including a supermarket in Buffalo, New York and a synagogue in Poway, California.[28] Therefore, the spread of TVEC online remains a complex and multifaceted problem.

Finally, I want to acknowledge that online platforms bring enormous societal benefits in connecting and empowering people around the world, and undue suppression of speech is a violation of human rights. As stated in the text of the Call, companies should not have to proactively scan every piece of content before it is uploaded to the internet; that would significantly restrict freedom of expression and limit the internet's ability to act as a force of good.[29] The first line of the Call is a commitment to protecting a free, open, and secure internet which is a powerful tool to promote connectivity, enhance social inclusiveness and foster economic growth.[30] Therefore, the solutions presented in this report will hopefully strike the right balance in limiting the harms caused by the spread of TVEC online while maintaining the benefits of the openness and connectivity of the internet.

This report will explore the challenges the New Zealand Government faced when trying to stop the spread of TVEC online, and why it opted for a non-regulatory solution that worked alongside tech companies and civil society. Indeed, the Call is a form of multistakeholder governance – a concept built for the 21st century and the global internet age. In the first half of the 20th century, governments increasingly relied on multilateral institutions such as the United Nations (UN) and World Trade Organization to find consensus on policies. Nation-state actors would then implement these multilateral agreements at home. However, rapid developments in technology and trade created multinational corporations, which gradually weakened the power of states to craft policies in isolation. Furthermore, fractures between democratic and non-democratic countries eroded the ability of global institutions like the UN to address nuanced global problems. Therefore, instead of turning to multilateral institutions, in certain circumstances, like-minded governments, corporations, and civil society collaborated to address various societal problems. These collaborations are frequently called multistakeholder initiatives (MSIs).

---

[27] *Id*.

[28] Mariana Olaizola Rosenblat & Paul M. Barrett, *Gaming the System: How Extremists Exploit Gaming Sites and What Can Be Done to Counter Term,* NYU STERN CENTER FOR BUSINESS AND HUMAN RIGHTS, 2 (May 2023), https://static1.squarespace.com/static/5b6df958f8370af3217d4178/t/6465b2f8be2da5102bbeb2e6/1684386554096/NYU+CBHR+Gaming_ONLINE+UPDATED+May+16.pdf ("Sure enough, copycats were quick to follow. In April 2019, a little over a month after the Christchurch tragedy, a 19-year-old male shooter opened fire at a synagogue in Poway, California, while livestreaming to his followers. One of the spectators commented during the livestream, 'get the high score' – a common phrase used among gamers. In early August of the same year, a 21-year-old man shot 23 people dead in a Walmart in El Paso, Texas. In his manifesto, he echoed the Christchurch shooter's conspiracy theory of 'white replacement,' the notion that shadowy elites are plotting to destroy white populations and culture through immigration and other policies, and mentioned a desire to live out his super soldier fantasy from the video game, Call of Duty. A month later, on Yom Kippur, another far-right militant launched a livestream on Twitch, a popular site among gamers, as he prepared to murder worshippers at a synagogue in Halle, Germany. The shooter killed two bystanders and, like those before him, left a manifesto riddled with references to far-right conspiracies couched in gaming jargon.").

[29] *The Christchurch Call to Action: To Eliminate Terrorist and Violent Extremist Content Online, supra* note 11.

[30] *Id.*

An MSI is created when two or more types of actors (such as governments, corporations, civil society, charitable foundations, academia, technical experts, or end-users) come together in a common governance enterprise to solve a problem defined by the group. The stakeholders collectively set procedural rules for decision-making and accountability. Within an MSI, governments, especially democratically elected governments, can be understood as agents of their citizens, corporations as agents of their owners or shareholders, and civil society as agents of their members. MSIs thrive because they allow a diverse group of participants to draw on multiple perspectives to produce better informed solutions to complex and interdependent problems. The diversity of possible challenges and outcomes means there is no single MSI model.[31] Instead, a wide variety of multistakeholder practices are adopted to solve unique problems. Some of the first MSIs addressed labour practices in "sweatshops", environmental degradation, the trade of "blood diamonds", standards for the vitivinicultural sector, and the distribution of development aid.[32] One area where MSIs have flourished has been relating to global internet governance challenges.

To understand why the Call chose to create an MSI in the wake of 15 March 2019, this report explores the history of how stakeholders have attempted to govern user-generated content online. Part I provides an overview of single-sided and multistakeholder governance frameworks for moderating content online. First, this part looks at single-sided frameworks created by national governments and the tech companies themselves to address the spread of TVEC online. It will examine how governments approach content online and the range of approaches taken by national regulators. This report examines a spectrum of regulation, starting with the free-speech maximalists in the US, then looking at New Zealand and the European Union (EU) as rights-respecting regimes, and finally discussing less permissive frameworks in Turkey, Russia, and China. In the absence of clear legal frameworks, tech companies have attempted to self-regulate how they moderate content to prevent TVEC online. This part also examines the history of content moderation and self-regulatory efforts. Next, because national regulation and self-regulation have not successfully addressed the problem of TVEC online, the second half of this part explores multistakeholder models. It looks at the rise of multistakeholder governance, its history in the internet governance context, and recent multilateral efforts that could undermine multistakeholderism in internet governance.

Part II distills the lessons learned from the MSIs working on internet governance issues highlighted in Part I and builds a framework for MSIs for addressing online content governance. The first section within this part proposes a taxonomy for MSIs, breaking them

[31] Bill Graham & Stephanie MacLellan, *Overview of the Challenges Posed by Internet Platforms: Who Should Address Them and How?*, CENTRE FOR INTERNATIONAL GOVERNANCE INNOVATION & STANFORD GLOBAL DIGITAL POLICY INCUBATOR, Special Report: Governance Innovation for a Connected World Protecting Free Expression, Diversity and Civic Engagement in the Global Digital Ecosystem, 12 (2018) https://www.cigionline.org/static/documents/documents/Stanford%20Special%20Report%20web.pdf ("There is no single definition to describe the multistakeholder approach. It would be counterproductive to stick to a single cookie-cutter approach; instead, the approach must be adapted to suit the nature of the problem being approached and the constellation of stakeholders to be involved in finding a solution.").

[32] Dorothée Baumann-Pauly, Justine Nolan, Auret Van Heerden & Michael Samway, *Industry-Specific Multi-Stakeholder Initiatives that Govern Corporate Human Rights Standards – Legitimacy Assessments of the Fair Labor Association and the Global Network Initiative*, UNSW LAW RESEARCH PAPER NO. 2015-12 (10 March 2015), https://ssrn.com/abstract=2576217.

down as egalitarian, consultative, restrictive, and curated. It argues that a curated MSI is the best option for the work of the Call and discusses why this format works for content governance frameworks. Next, this part sets out the necessary steps to build a curated MSI. After three decades of experimenting with MSIs, several core principles for creating a self-sustaining and effective MSI have emerged. The core steps identified include deciding if an MSI is necessary, establishing the objectives and functions of the initiative, defining who is a stakeholder, setting up terms of reference, sustaining forward momentum, and deciding when the work is finished. The purpose of outlining these best practices is to apply them to the work of the Call.

Part III examines how the New Zealand Government should look at the history of MSIs and key best practices when charting the future of the Call. First, Part III examines New Zealand's history and culture, which provide the foundations for the multistakeholder model. Next, it covers what happened on 15 March 2019 and the progress the Call has made in the four years since. Second, this part evaluates the progress the Call has made towards building a multistakeholder community and eliminating TVEC online while protecting a free, open, and secure internet – its two overarching objectives. Third, the part discusses the evolution and adoption of GenAI technologies and their impact on the moderation of TVEC online. Finally, this part applies the best practices set out in Part II to suggest steps the Call could take to create a self-sustaining MSI and how to expand the Call's work to address the impact of GenAI on content moderation.

I. GOVERNANCE FRAMEWORKS FOR CONTENT MODERATION

This part outlines how different actors have attempted to govern user-generated content online. The first section looks at single-sided governance initiatives created by national regulators and tech companies themselves. National governments have applied a spectrum of approaches to regulating content online, from free speech maximalism in the United States to the hyper-censorial regime in China. In the context of global online platforms, inconsistency between national laws often means that companies must self-regulate content moderation practices. Therefore, this section also looks at how and why tech companies have moderated user-generated content over the past 30 years. The second part of this section discusses the rise of multi-sided content governance frameworks, starting with the history of MSIs, then how multistakeholderism has evolved in the internet governance context, and concluding with recent multilateral efforts to assert government control over online internet governance.

A. Single-Sided Content Governance Frameworks

1. National Regulatory Frameworks for Content Moderation

National governments face several challenges when trying to impose legal liability on online platforms for hosting certain types of user-generated content. The first challenge arises as the technological framework that underpins the internet was designed specifically to circumvent governmental influence. The internet's origins date back to 1969, when it was a project of the US Government's Advanced Research Projects Agency (ARPA). The internet was initially used by government and academic institutions for research and communication purposes. Given the Cold War era context, ARPA designed the internet to withstand a nuclear attack by building a system that avoids single points of failure, encourages resiliency, scales effortlessly, and restricts government control.[33] This decentralisation appealed to early internet enthusiasts, who imagined a world "free of power".[34] In 1996, John Perry Barlow, a lyricist for the Grateful Dead and co-founder of the Electronic Frontier Foundation, spoke to the need for internet users to write their own rules and disparaged government control of the technology in his "Declaration of the Independence of Cyberspace".[35] Early internet protocols were heavily influenced by cyber libertarians like Barlow, who thought that the rules governing the internet should be created and enforced by online communities – not governments.[36] As a result, technologists further built the internet to interpret overt government control or censorship as

[33] Cade Metz, *Paul Baran, the Link between Nuclear War and the Internet*, WIRED (9 April 2012), www.wired.co.uk/article/h-bomb-and-the-internet.
[34] Thomas Schneider, *A vision, values, principles and mechanisms for cooperation and governance fit for purpose for the digital age*, INTERNET GOVERNANCE FORUM BERLIN, "Towards a Global Framework for Cyber Peace and Digital Cooperation: An Agenda for the 2020s," (25-29 Nov. 2019), https://www.hiig.de/wp-content/uploads/2019/11/Kleinwa%CC%88chter-Kettemann-Senges-eds.-Global-Framework-for-Cyber-Peace-2019.pdf.
[35] John Perry Barlow, *A Declaration of the Independence of Cyberspace*, ELECTRONIC FRONTIER FOUNDATION (8 Feb. 1996), https://www.eff.org/cyberspace-independence.
[36] Kate Klonick, *The New Governors: The People, Rules, and Processes Governing Online Speech*, 131 HARV. L. REV. 1598, 1616 (2018) ("In the earliest days of the internet, the regulations concerning the substance and structure of cyberspace were "built by a noncommercial sector [of] researchers and hackers, focused upon building a network… Balkin argued that the values of cyberspace are inherently democratic – bolstered by the ideals of free speech, individual liberty, and participation."); *citing* Jack M. Balkin, D*igital Speech and Democratic Culture: A Theory of Freedom of Expression for the Information Society*, 79 N.Y.U. L. REV. 1, 2 (2004) and Lawrence Lessig, CODE 2.0 (2006).

damage and route around it.[37] In effect, the early internet was a multistakeholder collaboration championed by a diverse group of actors without regard to national territorial borders or governmental controls.

By the 1990s, the internet had evolved from a communications medium owned and operated by government and academic institutions to a global platform increasingly dominated by corporations.[38] Internet adoption accelerated in the early 1990s after British computer scientist Tim Berners-Lee created the World Wide Web, which made it easier for non-technical people to access and share information online using standard protocols, thereby creating new opportunities for businesses and individuals.[39] As it grew, the internet was governed piecemeal by a variety of voluntary standard-setting bodies that empowered private companies to perform key roles as network operators and information intermediaries.[40] Throughout the 1990s, when national governments did consider regulating the internet, they largely saw the value of e-commerce and passed robust safe-harbour protections for online platforms hosting user-generated content. These legal protections led to the creation of online service providers in the early 2000s, which rapidly scaled into behemoth global companies.[41] As billions of people came online in the 2010s, many governments became wary of the free-flowing nature of the internet and started passing new regulations which threaten to undermine the decentralised internet.[42]

Starting in the early 2010s and continuing today, governments have become increasingly interested in regulating user-generated content online. However, governments have struggled to regulate online platforms for both the technical reasons described above as well as several additional reasons. First, many governments tried to fit regulation built for traditional media onto social media, which proved ineffective. The volume of content meant governments could not just hire more lawyers, police, or judges.[43] Unlike editing a newspaper, content moderation is impossible to do perfectly at scale and legal frameworks penalizing companies for every

---

[37] Nic Suzor, LAWLESS: THE SECRET RULES THAT GOVERN OUR DIGITAL LIVES, 28 (2019) ("the net interprets censorship as damage and routes around it, and when Barlow said that territorial governments had no methods of enforcement that actually worked, they were in a sense correct. The internet is fantastically hard to regulate. If your goal is to permanently remove all access to a piece of information or to prevent communications between committed, but unknown, participants, you're likely out of luck.").

[38] Internet Domain Names, Part 1: Hearing Before the Committee on Science, Subcommittee on Basic Research, 105 Cong. (25 Sep. 1997) (Statement of Jonathan B. Postel, Director, Computer Networks Division, University of Southern California), at: http://commdocs.house.gov/committees/science/hsy268140.000/hsy268140_0.HTM.

[39] *Id*.

[40] Mark Raymond & Laura DeNardis, *Multi-stakeholderism: Anatomy of an Inchoate Global Institution*, CENTRE FOR INTERNATIONAL GOVERNANCE INNOVATION AND THE ROYAL INSTITUTE OF INTERNATIONAL AFFAIRS, Research Volume 2: Global Commission on Internet Governance: Who Runs the Internet? The Global Multi-stakeholder Model of Internet Governance, 19-45 (Nov. 2016), https://www.cigionline.org/static/documents/documents/GCIG%20Volume%202%20WEB.pdf.

[41] *See* Douek at 21, at 26-27; *citing* Liat Clark, *Tim Berners-Lee: We Need to Re-Decentralise the Web*, WIRED UK (2 June 2014), https://www.wired.co.uk/article/tim-berners-lee-reclaim-the-web; Adi Robertson, *Twitter's Decentralized Social Network Project Takes a Baby Step Forward*, THE VERGE (21 Jan. 2021), https://www.theverge.com/2021/1/21/22242718/twitter-blueskydecentralized-social-media-team-project-update; Mike Masnick, *Protocols, Not Platforms: A Technological Approach to Free Speech*, KNIGHT FIRST AMENDMENT INSTITUTE (2019), https://knightcolumbia.org/content/protocols-not-platforms-a-technological-approach-tofree-speech.

[42] Adrian Shahbaz, Allie Funk & Kian Vesteinsson, *Freedom on the Net 2022: Countering an Authoritarian Overhaul of the Internet*, FREEDOM HOUSE (2023), https://freedomhouse.org/report/freedom-net/2022/countering-authoritarian-overhaul-internet#tracking-the-global-decline.

[43] Suzor, *supra* note 37 at 98.

error would be impractical to enforce.[44] Second, the speed and technological complexity of online platforms limits the states' ability to commandeer or even oversee nuanced content moderation processes.[45] Indeed, some legislative proposals have become obsolete upon enactment, as companies adopted new technology and content moderation practices. Finally, regulatory frameworks typically assume a one-size-fits all approach across a particular industry. As we will explore in the next section, there is no centralised approach to the way platforms moderate content, but rather four broad approaches: artisanal or case-by-case, community-reliant, industrial or large-scale, and no moderation whatsoever. As a result, regulators have struggled to find a legal approach for a complex industry that could be reduced to a simple one-size-fits-all checklist.[46] Moreover, legislation that divides the industry based on company size, profits, or number of users has yet to be implemented successfully.[47]

These struggles to regulate online content are found in every country, however, governments have confronted these challenges in ways that reflect their views on the freedom of expression. National regulation, therefore, falls within a wide spectrum, with the United States on one end, which allows almost all speech online, and China on the other, which closely monitors almost all speech online. As this section explores, regulation in democratic governments typically aligns with international human rights principles enshrined in Article 19 of the International Covenant on Civil and Political Rights (ICCPR).[48] Under the ICCPR, content-based restrictions on the freedom of expression are only permissible when they are clearly defined by law and are necessary and proportional to justify silencing speech – a high bar for any national law to achieve.[49] In practice, in places like the US, New Zealand and the EU where human rights are respected, laws mirroring Article 19 protections give online platforms the certainty that they can host the vast majority of user-generated content without facing legal penalties.

On the other end of the spectrum, an increasing number of governments do not adhere to Article 19 of the ICCPR when regulating content online.[50] The internet's early architecture makes it difficult to block online content based on national borders, but that has not stopped authoritarian governments from trying to force online platforms to violate human rights

[44] Mike Masnick, *Masnick's Impossibility Theorem: Content Moderation At Scale Is Impossible To Do Well*, TECHDIRT. (20 Nov. 2019), https://www.techdirt.com/articles/20191111/23032743367/masnicks-impossibilitytheorem-content-moderation-scale-is-impossible-to-do-well.shtml.
[45] Suzor, *supra* note 37 at 98; Douek, *supra* note 21 at 8 ("Even if there were not constitutional obstacles to substantive governmental regulation of content moderation, the sheer scale, speed and technological complexity of the task means state actors could not directly commandeer the operations of content moderation. This is a descriptive, not normative, observation: the state simply does not have the capacity to usurp platforms as the frontline of content moderation.").
[46] Douek, *supra* note 21 at 80 ("Content moderation, like data security, "changes too quickly and is far too dependent upon context to be reduced to a one-size-fits-all checklist."); *citing* Daniel J. Solove & Woodrow Hartzog, *The FTC and the New Common Law of Privacy*, 114 COLUM. L. REV. 583 (2014); Woodrow Hartzog & Daniel J Solove, *The Scope and Potential of FTC Data Protection*, 83 GEO. WASH. L. REV. 2230 (2015).
[47] The European Union's Digital Services Act has size-based requirements, but at the time of this writing, these measures have not gone into effect. *Proposal for a Regulation of the European Parliament and of the Council on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC*, COM (2020) 825 final (15 Dec. 2020).
[48] The International Covenant on Civil and Political Rights, *opened for signature,* 16 Dec. 1966, art. 19, S. Exec. Doc. E, 95-2, at 29 (1978), 999 U.N.T.S. 171, 178 (entered into force 23 March 1976) [hereinafter ICCPR].
[49] *Id*.
[50] Shahbaz, Funk & Vesteinsson *supra* note 42.

principles by restricting content within their country.[51] In recent years, technological advances have provided governments with solutions to break their citizens away from the global internet and control online spaces.[52] This fragmentation is sometimes referred to as the "splinternet".[53] The splinternet disrupts the previously global internet and replaces it with a system in which the internet is experienced differently by users across national jurisdictions.[54] The splinternet encompasses government restrictions on the flow of news and information, centralised state control over internet infrastructure, and barriers to cross-border transfers of user data.[55] Unfortunately, new technologies and increasing authoritarianism have led to a steady decline of global internet freedoms for the past 12 years.[56]

This next section highlights a few countries across the spectrum of national regulatory efforts: the United States, as the most-speech protective, then New Zealand, then the European Union, which has a rights-respecting framework but has passed copious amounts of legislation regulating content governance online. The section then provides examples from national regimes that subvert the protections of Article 19 of the ICCPR. There are dozens of countries that fit into this category, but this section will discuss three: Turkey, Russia, and China. Highlighting these regimes is important because China and Russia have long sought to displace the multistakeholder model of internet governance with one that promotes greater control by multilateral institutions.[57] Both countries have attempted to leverage the United Nations to endorse the right of each state to control its own "national segment of the internet".[58] As we will explore in Part II, MSIs are frequently created to fill "governance gaps," and these examples will illustrate where gaps may occur within national regulatory frameworks. In many cases, like-minded national governments will work together in multilateral or multistakeholder settings to address technological challenges. However, due to the dramatic variance of legal frameworks outlined in the next subsections, many democratic governments are unable to partner with authoritarian regimes without compromising fundamental human rights values.

1(a). The United States

The United States is undoubtedly a global outlier in its approach to free speech protections. However, understanding the US legal framework is critical when discussing internet regulations, because most large global online platforms hosting user-generated content are headquartered in the US. Overwhelmingly, global online platforms were founded by US employees who built US speech values into their content moderation systems.[59] These systems

[51] Suzor, *supra* note 37 at 28.
[52] Shahbaz, Funk & Vesteinsson *supra* note 42.
[53] Dan York, *What Is a Splinternet? And Why You Should Be Paying Attention*, INTERNET SOCIETY (23 March 2022), https://www.internetsociety.org/blog/2022/03/what-is-the-splinternet-and-why-you-should-be-paying-attention/.
[54] Suzor, *supra* note 37 at 87.
[55] Shahbaz, Funk & Vesteinsson *supra* note 42.
[56] *Id.*
[57] *Id.*
[58] *Id.*
[59] Klonick, *supra* note 36 at 1621 ("A common theme exists in all three of these platforms' histories: American lawyers trained and acculturated in American free speech norms and First Amendment law oversaw the development of company content-moderation policy. Though they might not have "directly imported First Amendment doctrine," the normative background in free speech had a direct impact on how they structured their

were further entrenched into the US system by US lawyers who used US legal principles to craft the global terms of service policies that dictate what a user can or cannot post on the online platform. Therefore, understanding the US system is critical for all other content governance analysis.

In the US, there are two foundational laws regarding the regulation of speech online: the First Amendment to the US Constitution and Section 230 of the Communications Decency Act. The First Amendment to the US Constitution states that Congress shall pass no law abridging the freedom of speech and is read to broadly protect citizens against government censorship.[60] A small number of exceptions allow the government to restrict speech, including in the cases of child sexual abuse material, fraud, obscenity, incitement to violence, speech integral to illegal conduct, speech violating intellectual property law, true threats, commercial speech, and defamation.[61] Americans are fiercely protective of their "free speech culture" and courts have strongly protected this individual right.[62] As a result, many types of speech that are restricted internationally are constitutionally protected in the US. For example, content that is published by or about terrorists or extremists would be prohibited in many jurisdictions but is protected by the First Amendment as long as the content does not imminently incite violence.[63] The First Amendment applies only to the Government's restrictions on speech, and does not pass to a company, which is under no obligation to allow all constitutionally protected speech on its platform. Indeed, the First Amendment protects private actors from government efforts to control speech, meaning that the government is not allowed to compel an online platform to restrict, remove, or promote speech.[64]

In addition to the protections under the First Amendment, online platforms also benefit from the legal framework Congress created in Section 230 of the Communications Decency Act of 1996. Congress passed the Communications Decency Act to regulate pornographic material on the internet.[65] One year after passage, the Supreme Court overturned the law for violating the First Amendment, as it was overly broad in restricting speech.[66] However, the Court upheld the safe harbour provisions for online service providers covered in Section 230. Sometimes referred to as the "26 words that created the Internet",[67] Section 230(c)(1) enables online platforms to host user-generated content without being held legally responsible for speech posted on their platforms by users.[68] Section 230(c)(2) empowers platforms to find and remove

policies … Simultaneously, there were complicated implications in trying to implement those American democratic cultural norms within a global company.").

[60] U.S. Constitution amend. I ("Congress shall pass no law … abridging the freedom of speech.").

[61] Wex Legal Dictionary*, First Amendment Overview*, CORNELL LAW SCHOOL LEGAL INFORMATION INSTITUTE, https://www.law.cornell.edu/wex/first_amendment.

[62] Genevieve Lakier, *The Non-First Amendment Law of Freedom of Speech*, 134 HARV. L. REV. 2299, 2301 (2021) ("[t]he Speech Clause of the First Amendment has for decades now served as one of the most powerful mechanisms of individual rights protection in the entire federal Constitution."); *See also* Douek, *supra* note 21 at 34 ("in content moderation, the idea of prioritizing the overall functioning of the system over individual rights is dissonant with the story American society tells itself about its free speech culture.").

[63] Eric Goldman, *The United States' Approach to 'Platform' Regulation*, SANTA CLARA UNIV. LEGAL STUDIES Research Paper No. 4404374 (2023).

[64] Fishman, *supra* note 24.

[65] Section 230 of the Communications Decency Act, 47 U.S.C. § 230 (2018).

[66] *See* Reno v. ACLU, 521 U.S. 844 (1997).

[67] JEFF KOSSEFF, THE TWENTY-SIX WORDS THAT CREATED THE INTERNET (2019).

[68] 47 U.S.C. § 230(c)(1).

material they deem objectionable content without fear of legal action from users. As such, it is sometimes referred to as the "Good Samaritan" provision of the law.[69] There are several carve-outs to Section 230 protections for internet service providers, including where the platform materially contributes to criminal behaviour, intellectual property claims, and promotions of sex trafficking and commercial sex.[70] Section 230 provides broad immunity for social media companies to host user-generated content and moderate that content as they see fit, as long as they do not significantly develop the content themselves.[71]

The protections under the First Amendment and the immunities granted by Section 230 work together to allow US online platforms to experiment with the type of content moderation that works best for their audience. In practice, if a social media company is sued for its content moderation decisions, it could assert a First Amendment defense, but Section 230 acts as a "procedural fast lane" to resolve litigation more quickly and cheaply.[72] The Section 230 "fast lane" made it possible for anyone to start a company and hosts user-generated content without being liable for what their users say or share.[73] This drove investment in the industry, particularly in Silicon Valley. Eric Goldman, a world-leading internet scholar, has called Section 230 a "globally unique solution" which has given the United States a competitive advantage when it comes to the internet.[74] As a result, the US is home to a wide diversity of online platforms that moderate user-generated content to serve different audiences including Reddit, Airbnb, Wikipedia, Yelp, and Etsy.

For over 20 years, Section 230 remained unchanged. Then, in 2018, Congress added a new carve-out to the law with the passage of the Allow States and Victims to Fight Online Sex Trafficking Act and the Stop Enabling Sex Trafficking Act, known as FOSTA-SESTA or just FOSTA.[75] Leading up to the passage of FOSTA, Backpage.com was sued by victims of sex trafficking who claimed the website had helped facilitate the criminal activity they experienced.[76] These lawsuits were dismissed by the courts, which convinced the trafficking victims to lobby Congress for an amendment to Section 230 related to promotion of sex trafficking and commercial sex.[77] FOSTA's passage was a turning point for Section 230, as it catapulted the relatively unknown and uncontroversial limited liability provisions for websites into the centre of a national debate around the power of "big tech" companies. Five years later,

[69] *See* Zeran v. Am. Online, Inc., 129 F.3d 327, 330 (4th Cir. 1997) (noting that the purposes of intermediary immunity in § 230 were not only to incentivize platforms to remove indecent content but also to protect the free speech of platform users).
[70] 47 U.S.C. § 230(e).
[71] *See* Fair Housing Council of San Fernando Valley v. Roommates.com, LLC, 521 F.3d 1157 (9th Cir. 2008).
[72] Eric Goldman, *Why Section 230 Is Better Than the First Amendment*, 95 NOTRE DAME L. REV. REFLECTION 34 (2019).
[73] *Id*.
[74] Caplan, *supra* note 20 at 27; *quoting* comments made by Eric Goldman at the Content Moderation at Scale Conference in Washington, D.C., on May 7, 2018.
[75] 47 U.S.C. § 230(e); Allow States and Victims to Fight Online Sex Trafficking Act of 2017, Pub. L. No. 1115-164, 132 Stat. 1253 (codified as amended in scattered sections of 18 and 47 U.S.C.) (2018).
[76] Eric Goldman, *The Complicated Story of FOSTA and Section 230*, 17 FIRST AMENDMENT L. REV. 279 (2019).
[77] *Id*.; *E.g*., Backpage.com, LLC v. Dart, 807 F.3d 229 (7th Cir. 2015); Jane Doe No. 1 v. Backpage.com, LLC, 817 F.3d 12 (1st Cir. 2016); Backpage.com, LLC v. Cooper, 939 F. Supp. 2d 805 (M.D. Tenn. 2013); Backpage.com, LLC v. Hoffman, 2013 WL 4502097 (D.N.J. Aug. 20, 2013); Backpage.com, LLC v. McKenna, 881 F. Supp. 2d 1262 (W.D. Wash. 2012); M.A. ex rel. P.K. v. Vill. Voice Media Holdings, LLC, 809 F. Supp. 2d 1041 (E.D. Mo. 2011).

this debate continues, without any political consensus on how to resolve it. While Democrats are pushing companies to restrict speech that is hateful or bullying, Republican states, including Florida and Texas, have passed laws requiring social media companies to leave up all constitutionally protected speech.[78] These laws are embroiled in litigation and likely to end up before the Supreme Court.[79] Until then, divisive partisanship has entrenched a legislative stalemate and no federal laws related to Section 230 have passed since 2018.

Potentially due to this legislative stalemate, in April of 2021, Supreme Court Justice Clarence Thomas wrote a concurrence on the dismissal of a case relating to internet policies attacking Section 230 and the powers of the First Amendment.[80] As part of Thomas's concurrence he invited lawyers to bring cases challenging Section 230 to court.[81] A year later, the Supreme Court granted certiorari to two cases relating to the culpability of social media companies for a deadly Islamic State attack, which the perpetrators discussed on their platforms.[82] The family members of victims who died in an ISIS attack in Europe presented their case to the Court in February of 2023, arguing that Twitter, Facebook, and YouTube should be held liable because of ISIS's general presence on their platforms.[83] In May 2023, the Court dismissed the cases, stating that the social media companies did not provide knowing or substantial assistance to ISIS necessary to find them culpable under the Anti-Terrorism Act."[84] However, the Court expressly declined to rule on the Section 230 issues, including on whether the law applies to algorithmic promotion of content, leaving in place the broad scope of Section 230.[85]

1(b). New Zealand

The next legal framework on our spectrum is that of New Zealand, which has enshrined legal provisions aligning with Article 19 of the ICCPR in the New Zealand Bill of Rights Act 1990 and the Human Rights Act 1993. These laws guarantee the right to freedom of thought, conscience, and religion, including the freedom to hold opinions without interference and to seek, receive, and impart information and ideas of all kinds. However, in New Zealand, freedom of expression is not absolute. There are certain limitations and restrictions, including on speech that incites violence, hatred, or discrimination; defamation; harassment; and copyright infringement.[86] This right is also limited under the Summary Offences Act 1981, which prohibits threatening or violent speech.[87] Unlike the United States and many other democratically governed countries, New Zealand does not have a legal regime that specifically provides safe harbour protections for online intermediaries hosting user-generated content.

[78] Texas House Bill 20 (H.B. 20, enacted September 9, 2021); Florida S.B. 7072, 2021 Leg. (Fla. 2021).
[79] NetChoice, LLC, v. Paxton, No. 21-51178 (5th Cir.), No. 22-555 (Supreme Court).
[80] Mark MacCarthy, *Justice Thomas sends a message on social media regulation*, THE BROOKINGS INSTITUTION (9 April 2021), https://www.brookings.edu/blog/techtank/2021/04/09/justice-thomas-sends-a-message-on-social-media-regulation/.
[81] Bobby Allyn, *Justice Clarence Thomas Takes Aim At Tech And Its Power 'To Cut Off Speech,"* NATIONAL PUBLIC RADIO (5 April 2021), https://www.npr.org/2021/04/05/984440891/justice-clarence-thomas-takes-aims-at-tech-and-its-power-to-cut-off-speech.
[82] Twitter, Inc. v. Taamneh, 598 U.S. __ (2023).
[83] *Id*.
[84] *Id*.
[85] *Id*.
[86] Human Rights Act 1993 (NZ).
[87] Royal Commission of Inquiry into the Attack on Christchurch Mosques, *supra* note 5, at Part 9, chapter 4.

Instead, it has a patchwork of laws governing content moderation, hate speech, and the distribution of TVEC online. Four primary statutes impose liability on social media companies hosting objectionable speech: the Summary Offences Act 1981, the Harmful Digital Communications Act 2015, the Broadcasting Act 1989, and the Films, Videos, and Publications Classification Act 1993.[88]

In the context of assessing user-generated content posted online, New Zealand has two statutes regulating content. First, the Broadcasting Act 1989 sets standards for traditional media 'broadcasters', but some standards apply online.[89] Second, the Films, Videos, and Publications Classifications Act creates a consumer advisory system for age suitability and warnings for content in "films". It also specifies what "publications" are illegal (or "objectionable") for distribution across mediums in New Zealand.[90] It was under this Act that the Christchurch shooter's video and manifesto were deemed objectionable in the days immediately following the attack. New Zealand's Chief Censor "called in" the livestream video and manifesto for classification, and the office decided to ban the materials on 20 and 23 March, respectively.[91] This designation made it illegal to hold or distribute the video or manifesto. To comply with this legal restriction, many online platforms now work with third-party hash-sharing systems to automatically detect and remove this content, as discussed in the next section.[92]

Next, New Zealand's Harmful Digital Communications Act of 2015 regulates issues such as cyberbullying, harassment, and other forms of harmful online behaviour.[93] It defines harmful digital communications as those that are threatening, intimidating, or otherwise harmful to an individual, and that are made using a digital communication device, such as a computer, smartphone, or social media platform.[94] To help enforce these rules and settle disputes with companies, the Act has an "approved agency" receive and investigate complaints about harmful digital communications.[95] The current approved agency is Netsafe, a non-profit entity that receives funding from the Ministries of Justice and Education and assists victims exposed to harmful digital content.[96] Netsafe works closely with technology companies to resolve these complaints, and with the Police and the Department of Internal Affairs, which have set up separate processes.

In May 2021, New Zealand's Government initiated the Content Regulation Systems Review to align some of the statutory obligations of internet media companies with those of their traditional media counterparts.[97] This review is unlikely to be finalised before the October 2023

[88] *Id*.
[89] Broadcasting Act 1989 (NZ).
[90] Films, Videos, and Publications Classification Act 1993 (NZ).
[91] David Shanks, *Classification Office response to the March 2019 Christchurch terrorist attack,* CLASSIFICATIONS OFFICE (9 Dec. 2020), https://www.classificationoffice.govt.nz/news/news-items/response-to-the-march-2019-christchurch-terrorist-attack/.
[92] Tech Against Terrorism (@TechvsTerrorism), Twitter (18 March 2023), https://twitter.com/techvsterrorism/status/1636761830558167043.
[93] Harmful Digital Communications Act of 2015 (NZ).
[94] *Id*.
[95] *Id*.
[96] Netsafe, *About Netsafe*, NETSAFE (2023), at https://netsafe.org.nz/aboutnetsafe/partners/.
[97] New Zealand Department of Internal Affairs, *Media and online content regulation* (1 June 2023), at https://www.dia.govt.nz/media-and-online-content-regulation.

election. However, while it is under way, large online service providers in New Zealand, including Meta and Google, have worked with the New Zealand Tech Alliance to create the Aotearoa New Zealand Code of Practice for Online Safety and Harms, which provides guidance to companies on how to enhance safety and mitigate harm online.[98] Launched in July 2022, this self-regulatory Code of Practice requires companies to make "best efforts" towards a set of commitments that will reduce harmful content, increase transparency, and empower users.[99] When a company signs onto the code's framework, it identifies which of its products the code will apply to, and can further choose to opt out of any measures it feels are not relevant to the company's products.[100] Critics have argued that this is an attempt by the companies to pre-empt regulation and that the effort lacks legitimacy and community accountability.[101] However, due to the lack of explicit legal provisions regulating social media in New Zealand, many companies have experimented with this type of self-regulatory mechanism.

In June 2023, the Department of Internal Affairs put forward a discussion document on their proposal to regulate online platforms.[102] The document acknowledges that it can be difficult for citizens to navigate the five industry complaint bodies they can approach if they feel content is unsafe or breaches the company's terms of service.[103] The regulation proposed by the Department of Internal Affairs would create codes of practice which set out specific safety obligations for larger or riskier platforms and would be enforceable by an independent regulator.[104] This new independent industry regulator would provide a clear "home" for consumer safety on online platforms, and industry groups would develop new codes with input from, and approval by, the regulator.[105] The Department is accepting feedback on its policy proposals until 31 July 2023.[106]

1(c). The European Union

Next on the spectrum of national regulatory frameworks governing online platform liability of user-generated content is the European Union. The EU has several rights-based restrictions on speech and legal liability frameworks set out in national laws and EU-level regulations and directives. Freedom of expression is codified in Article 10 of the European Convention on Human Rights, which has been incorporated into EU law through the Charter of Fundamental Rights of the European Union.[107] This framework mirrors the ICCPR, mentioned above.

---

[98] Curtis Barnes, Tom Barraclough, & Allyn Robins, *Platforms Are Testing Self-Regulation in New Zealand. It Needs a Lot of Work*, LAWFARE (2 Sep. 2022), https://www.lawfaremedia.org/article/platforms-are-testing-self-regulation-new-zealand-it-needs-lot-work.
[99] *Id*.
[100] *Id*.
[101] *Id*.; *see also* Tom Pullar-Strecker, *Social media firms advance NZ's controversial 'world first' code of conduct*, STUFF (1 April 2023), https://www.stuff.co.nz/business/131613136/social-media-firms-advance-nzs-controversial-world-first-code-of-conduct.
[102] New Zealand Department of Internal Affairs, *Discussion Document, Safer Online Services and Media Platforms*, DEPARTMENT OF INTERNAL AFFAIRS NZ (June 2023), https://www.dia.govt.nz/diawebsite.nsf/Files/online-content-regulation/$file/Safer-Online-Services-and-Media-Platforms-Discussion-Document-June-2023.pdf.
[103] *Id*.
[104] *Id*.
[105] *Id*.
[106] *Id*. This report will be published on 1 August 2023 and will therefore not detail the outcome of the proposal.
[107] European Convention for the Protection of Human Rights and Fundamental Freedoms, *opened for signature* Nov. 4, 1950, art. 10, Euro. T. S. No. 5, 213 U.N.T.S. 221 (entered into force 3 Sep. 1953).

However, until 2000, EU member states took different approaches to regulating content online. In most cases, online speech was subject to the same legal framework that applied to traditional media such as newspapers, television, and radio – with a great deal of variation between member states.[108] This patchwork approach created legal uncertainty for online platforms and threatened the growing e-commerce industry. As such, in 2000, the EU passed the e-Commerce Directive, which created a safe harbour for online intermediaries like the one found in Section 230, adding in a caveat that illegal content be removed "expeditiously".[109]

While the EU and US frameworks mirror each other in form and function, the definitions of "illegal" speech vary greatly. In the US, "illegal" speech exists only under the limited carve-outs of the First Amendment, and Section 230 immunity ensures that if illegal content is posted online, the user who posted the content is liable and not the platform itself. This is not the case in the EU, where member states have passed several regulations increasing liability for online intermediaries. Under the EU framework, national regulators are able to define broad categories of speech as "illegal" because there is less of a presumption against speech restrictions.[110] As a result, over the past 20 years, the EU has enacted a wide range of rules making types of speech illegal, ranging from the right to be forgotten found in the General Data Protection Regulation to hate speech laws in Germany under the Netzwerkdurchsetzungsgesetz (commonly known as NetzDG) and the restrictions on harmful speech passed recently in the Digital Services Act (DSA). Scholarly analysis of these laws will fill hundreds of textbooks; this section will only detail the regulations surrounding TVEC online. The TVEC regulations not only provide a helpful insight into the rulemaking process for content moderation more broadly in the EU; they are also relevant to the work of the Christchurch Call and this report.

As related to TVEC, the safe-harbour provisions for online platforms found in the 2000 e-Commerce Directive started to erode in 2008 after laws implementing the EU's counter-terrorism agenda were updated to criminalise the incitement to terrorism online.[111] These updates included requirements for internet platforms to cooperate with law enforcement to receive safe harbour protections.[112] The EU first explored multistakeholder options to assist online platforms with this work, including the creation of the Radicalisation Awareness Network, in 2011, which provides guidance to policymakers from civil society organisations

[108] Alexandre De Streel & Martin Husovec, *The e-commerce Directive as the cornerstone of the Internal Market, Assessment and options for reform,* EUROPEAN PARLIAMENT, Study requested by the IMCO committee (May 2020), https://www.europarl.europa.eu/RegData/etudes/STUD/2020/648797/IPOL_STU(2020)648797_EN.pdf.
[109] Pablo Baistrocchi, *Liability of Intermediary Service Providers in the EU Directive on Electronic Commerce*, 19 SANTA CLARA HIGH TECH. L.J. 111 (2002).
[110] Danielle Keats Citron, *Extremist Speech, Compelled Conformity, and Censorship Creep*, 93 NOTRE DAME L. REV. 1035, 1038 (2018).
[111] Santina Musolino, *EU policies for preventing violent extremism: a new paradigm for action?*, REVISTA CIDOB D'AFERS INTERNACIONALS, ISSUE 128, 39-57 (September 2021), doi. org/10.24241/rcai.2021.128.2.39/en ("The amendment of the Framework Decision 2002/475/JHA53 in 2008 added several more activities to the list of those already criminalised and shifted the focus on criminalising preparatory acts and incitement to terrorism. Moreover, it stressed the importance of reconsidering the potentialities of a preventive action. The adoption of the EU Internal Security Strategy in Action in 2010 and the creation, in 2011, of the EU Radicalisation Action Network outlined the importance of creating a network connecting first-line experts from various EU member states.").
[112] *Id.*

working to prevent and counter radicalisation.[113] After a spate of deadly terror attacks and hate crimes in 2015, European regulators began to place more responsibility on social media companies for the violence.[114] In December 2015, the EU created the EU Internet Forum to bring together tech platforms, law enforcement authorities, and civil society to reduce the availability of terrorist material online through programmes like the EU Internet Referral Unit.[115] As part of this work, in 2016, EU regulators worked with tech companies to create a Voluntary Codes of Conduct to remove illegal hate speech – including terrorist content.[116] Under the Code of Conduct, companies agreed to voluntarily comply with any requests from the EU Internet Referral Unit and remove content within 24 hours.[117] This new framework faced significant backlash from civil society, which had been excluded from the conversation and viewed the arrangement as both overreaching and censorial because it required companies to remove speech without questioning the validity of the government's request.[118]

The EU issued its first assessment of the Voluntary Code of Conduct in December 2016, which criticised the online platform's "success rate" at actioning removal requests.[119] Moreover, EU lawmakers deemed self-regulation attempts by the online platforms to be insufficient and decided to impose legal measures to combat terrorist radicalisation online.[120] In 2018, as part of an update to the Audio Visual Media Services Directive, the EU compelled member states to pass laws that would prevent the upload and dissemination of harmful material, including terrorist content.[121] Despite these changes, the EU again updated its laws again in the Regulation on Preventing the Dissemination of Terrorist Content Online (TCO) in 2021.[122] The TCO requires online platforms to remove terrorist content within one hour of receiving a removal order from a competent authority in an EU member state or face a fine of up to

---

[113] *Id.*; *See also About RAN Practitioners*, European Commission, Migration and Home Affairs (2023), https://home-affairs.ec.europa.eu/networks/radicalisation-awareness-network-ran/about-ran_en.
[114] Citron, *supra* note 110 at 1040; *See also* Lizzie Plaugic, *France Wants to Make Google and Facebook Accountable for Hate Speech*, THE VERGE (27 Jan. 2015), https://www.theverge.com/2015/1/27/7921463/google-facebookaccountable-for-hate-speech-france.
[115] European Commission, Migration and Home Affairs, *European Union Internet Forum (EUIF)*, EUROPEAN COMMISSION (2023), https://home-affairs.ec.europa.eu/policies/internal-security/counter-terrorism-and-radicalisation/prevention-radicalisation/terrorist-content-online_en.
[116] European Commission Press Release IP/15/6243, *EU Internet Forum: Bringing Together Governments*, EUROPOL AND TECHNOLOGY COMPANIES TO COUNTER TERRORIST CONTENT AND HATE SPEECH ONLINE (3 Dec. 2015), http://europa.eu/rapid/press-release_IP-15- 6243_en.htm.
[117] *Id*.
[118] Citron, *supra* note 110 at 1041 ("Although civil society organizations participated in early meetings held by the European Internet Forum, they were excluded from the negotiations that resulted in the Code. *EDRi and Access Now Withdraw from the EU Commission IT Forum Discussions*, EDRi (May 31, 2016), https://edri.org/edri-access-now-withdraw-eu-commissionforum-discussions. As the civil society group European Digital Rights (EDRi) explained, the European Commission refused to give the groups access to the negotiations and drafts of the agreement. Maryant Fernandez Perez, *New Documents Reveal the Truth Behind the Hate Speech Code*, EDRi (7 Sep. 2016), https://edri.org/new-documents-reveal-truth-behindhate-speech-code; Jennifer Baker, *Europol's Online Censorship Unit Is Haphazard and Unaccountable Says NGO,* ARS TECHNICA (4 July 2016), https://arstechnica.com/tech-policy/2016/07/europol -iru-extremist-content-censorship-policing/.").
[119] Citron, *supra* note 110 at 1042; *See also* European Comm'n, *Code Of Conduct On Countering Illegal Hate Speech Online: First Results On Implementation* 1 (Dec. 2016), http://ec.europa.eu/informa tion_society/newsroom/image/document/2016-50/factsheet-code-conduct-8_40573.pdf.
[120] Citron, *supra* note 110 at 1042.
[121] Tech Against Terrorism, *The Online Regulation Series – European Union (update)* (Nov. 2021), https://www.techagainstterrorism.org/2021/12/10/the-online-regulation-series-european-union-update/.
[122] Jan Penfrat, *Digital Services Act, The EDRi guide to 2297 amendment proposals*, European Digital Rights (EDRi) (Oct. 2021), https://edri.org/wp-content/uploads/2021/10/EDRi-policy-paper-Digital-Services-Act-Nov-2021.pdf.

4 per cent of their total revenue.[123] The TCO received significant pushback from civil society organisations for three reasons. First, civil society worried that overclassification by law enforcement, combined with the tight timeline, would stifle freedom of expression.[124] Second, civil society actors noted that the TCO seems to be in conflict with the EU's ePrivacy Directive, which limits the ability of platforms to scan more private surfaces for terrorist material.[125] Finally, civil society argued that the regulation grants national governments too much power to order the removal of speech with only minimal judicial oversight.[126] The TCO went into effect in July 2022, and there has yet to be much reporting from national authorities as to how they are implementing the regulatory tools.[127]

The other notable piece of EU legislation regarding content moderation more broadly, is the newly enacted DSA. The DSA is a sweeping legislative effort to "create safer digital space in which the fundamental rights of all users of digital services are protected".[128] While the DSA does not replace the TCO, the new rules in the DSA cover detection, flagging, and removal of "illegal content" as defined by either the member states or the EU itself.[129] For online platforms, compliance with the DSA will be extraordinarily challenging as new measures include: updating user safeguards, creating transparency and oversight processes, bans on advertising, and additional liability regimes.[130] Given the complexity of the DSA, its broader impact on content moderation and the future of MSIs will be hard to assess for years to come as pieces of the DSA are implemented both at the member state and the EU level.[131] With the DSA, the EU has increased the liability of online platforms in ways that might make them less likely to try new voluntary initiatives.

---

[123] Clothilde Goujard, *Online platforms now have an hour to remove terrorist content in the EU*, POLITICO (7 June 2022) https://www.politico.eu/article/online-platforms-to-take-down-terrorist-content-under-an-hour-in-the-eu/; *see also* Regulation (EU) 2021/784 of the European Parliament and of the Council of 29 April 2021 on addressing the dissemination of terrorist content online, EUROPEAN PARLIAMENT (29 April 2021), https://home-affairs.ec.europa.eu/policies/internal-security/counter-terrorism-and-radicalisation/prevention-radicalisation/terrorist-content-online_en.

[124] EDRi, *Terrorist Content Regulation: Document Pool*, EDRi (21 Jan. 2019) https://edri.org/our-work/terrorist-content-regulation-document-pool/ ("A major concern for the functioning and freedom of the internet is the extension of the upload filter regime the EU is currently about to introduce for copyright to terrorist content. Requiring internet companies to monitor everything we say on the web does not only have grave implications for the freedom of speech, but it also follows a dangerous path of outsourcing and privatising law enforcement.").

[125] Fishman, *supra* note 24 ("The Terrorism Content Online regulation focuses on removing public material supporting terrorism, while the ePrivacy Directive limits the ability of platforms to scan more private surfaces for terrorist material.").

[126] *Id*. ("Although companies do have the ability to appeal such orders, only a few companies are likely to have the legal capacity to file such appeals at scale and they will take months, if not years, to adjudicate. The regulation effectively grants national governments extraordinary latitude to order the removal of speech with minimal judicial oversight.").

[127] *See* Regulation (EU) 2021/784 of the European Parliament and of the Council of 29 April 2021 on addressing the dissemination of terrorist content online, *supra* note 123.

[128] European Commission, *The Digital Services Act Package, Shaping Europe's digital future*, EUROPEAN COMMISSION (6 June 2023) https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package.

[129] European Commission, *Questions and Answers: Digital Services Act*, EUROPEAN COMMISSION (25 April 2023), https://ec.europa.eu/commission/presscorner/detail/en/qanda_20_2348 ("What constitutes illegal content is defined in other laws either at EU level or at national level – for example terrorist content or child sexual abuse material or illegal hate speech is defined at EU level. Where a content is illegal only in a given Member State, as a general rule it should only be removed in the territory where it is illegal.").

[130] Daphne Keller, *The EU's new Digital Services Act and the Rest of the World*, VERFASSUNGSBOLG (7 Nov. 2022), https://verfassungsblog.de/dsa-rest-of-world/.

[131] *Id.*

1(d). Turkey

Turkey is next on the spectrum of national regulations governing user-generated content, as it is increasingly less free. Turkey is a democratic regime that has become more authoritarian in recent years by passing restrictive speech laws, heavily monitoring speech online, and increasingly threatening online platforms like Wikimedia and Twitter.[132] Starting in 2016, the country has implemented several laws that allow for the censorship of online content, including the 2016 Law on the Regulation of Publications on the Internet and Suppression of Crimes Committed by Means of Such Publication.[133] This law grants authorities the power to block websites and social media accounts that are deemed to be harmful to national security or public order. Online criticism of the government or the president can result in prosecution, and many journalists and social media users have been arrested for their online activities. Additionally, the government has required social media companies to establish local offices in Turkey and to comply with government requests to remove content.[134] In 2022, lawmakers went a step further ahead of upcoming elections, enacting new amendments which gave the government power to enact severe penalties against tech companies for failure to comply with take-down requests, ensuring companies will be complicit in censorship.[135] In the days before the 2023 election, Twitter restricted access to content in Turkey at the request of the government.[136] Regarding the restriction of TVEC online, the Turkish state has adopted a very broad definition of terrorism that increasingly covers peaceful acts of dissidence.[137] Unfortunately, Turkey's legal framework no longer complies with many of the provisions in Article 19 of the ICCPR.

1(d). Russia

The second most restrictive national regulatory system on our spectrum is Russia. However, most commentators agree that Russia would be just as restrictive as China if it had the technological capabilities to enact those restrictions. [138] In Russia, the government has the power to block websites if the state deems the content as extremist or harmful to the country's security or sovereignty.[139] Many global online service providers proactively left the Russian market in 2019, when Russia introduced a law that required all online communications to be

---

[132] Freedom House, *Freedom in the World 2023: Turkey*, FREEDOM HOUSE (9 March 2023), https://freedomhouse.org/country/turkey/freedom-world/2023.
[133] Tech Against Terrorism, *The Online Regulation Series – Turkey*, TECH AGAINST TERRORISM (2023), https://www.techagainstterrorism.org/2020/10/23/the-online-regulation-series-turkey ("The Regulation of Publications on the Internet and Suppression of Crimes Committed by means of Such Publication, 2007, widely known as the "Internet Law 5651" or "Law No. 5651." This regulates prohibited content, such as child abuse images and obscenity, on the Internet and enables the blocking of websites.").
[134] *See Freedom in the World 2023: Turkey*, *supra* note 132.
[135] Human Rights Watch, *Turkey: Dangerous, Dystopian New Legal Amendments*, HUMAN RIGHTS WATCH (14 Oct. 2022), https://www.hrw.org/news/2022/10/14/turkey-dangerous-dystopian-new-legal-amendments.
[136] Ashley Belanger, *Musk defends enabling Turkish censorship on Twitter, calling it his "choice",* ARS TECHNICA (May 16, 2023), https://arstechnica.com/tech-policy/2023/05/musk-defends-enabling-turkish-censorship-on-twitter-calling-it-his-choice/.
[137] Nazli Ozekici, *Turkey's broad definition of terrorism does nothing to halt radicalisation*, OPENDEMOCRACY (20 Jan. 2022), https://www.opendemocracy.net/en/north-africa-west-asia/turkeys-broad-definition-of-terrorism-does-nothing-to-halt-radicalisation/.
[138] *Russia is trying to build its own great firewall*, THE ECONOMIST (19 Feb. 2022), https://www.economist.com/business/russia-is-trying-to-build-its-own-great-firewall/21807706.
[139] Freedom House, *Freedom in the World 2023: Russia*, FREEDOM HOUSE (9 March 2023), https://freedomhouse.org/country/russia/freedom-world/2023.

stored for six months and made accessible to the government upon request.[140] In 2022, the government used this law to issue massive fines on platforms that refused to remove content and localise user data.[141] After Russia's invasion of Ukraine, authorities passed more restrictive legislation that granted more powers to state bodies tasked with regulation of the internet, expanded the grounds for what content could be deemed illegal, and required media outlets to refer to the war as a "special military operation."[142] While Russia is still connected to the broader global internet, the Russian government has hastened its progress toward infrastructural isolation. Regarding the moderation of TVEC online, in 2022, the Russian government blocked prominent social media platforms, including Facebook, Instagram, and Twitter, and labelled the companies as "extremist organizations".[143] Time and again, Russia has shown little regard for human rights principles when it comes to protecting the freedom of expression.

1(e). China

On the furthest end of the regulatory spectrum is China which has demonstrated little interest in protecting human rights. China is home to one of the world's most restrictive media environments and its most sophisticated system of censorship, which started in the late 1990s with the banning of pornography and media sites.[144] The country has a comprehensive censorship system known as the "great firewall", which blocks access to foreign websites and restricts content that is deemed politically sensitive or harmful to the country's social stability. As a result, almost no foreign global platforms are allowed to operate in the country, and domestic international platforms are tightly regulated. The government actively monitors online activities and requires online service providers to store user data within the country's borders, making it easier to monitor and censor content.[145] Additionally, the government has introduced laws that hold internet companies accountable for the content shared on their platforms, resulting in self-censorship by these companies to avoid legal repercussions.[146] China has been rated as the world's worst environment for internet freedom for eight straight years.[147] Regarding the regulation of TVEC online, Chinese officials and state media label a

[140] Masha Borak, *How Russia killed its tech industry*, MIT TECHNOLOGY REVIEW: TECH POLICY (4 April 2023), https://www.technologyreview.com/2023/04/04/1070352/ukraine-war-russia-tech-industry-yandex-skolkovo/.
[141] *Freedom in the World 2023: Russia, supra* note 138.
[142] David Ignatius, *Russia hasn't stopped maneuvering for a role in internet oversight*, WASH. POST (6 July 2023), https://www.washingtonpost.com/opinions/2023/07/06/russia-internet-governance-united-nations/; *see also* Shahbaz, Funk & Vesteinsson, *supra* note 42 ("Internet freedom in Russia reached an all-time low following the government's brutal invasion of Ukraine.").
[143] *Freedom in the World 2023: Russia, supra* note 138.
[144] *Internet Information Service Management Measures*, DigiChina, STANFORD UNIVERSITY (25 Sep. 2000), https://digichina.stanford.edu/work/internet-information-service-management-rules/ ("In September 2000, State Council Order No. 292 created the first set of content restrictions for Internet content providers. China-based websites cannot link to overseas news websites or distribute news from overseas media without separate approval.").
[145] *Freedom in the World 2023: China*, FREEDOM HOUSE (9 March 2023), https://freedomhouse.org/country/china/freedom-world/2023.
[146] *China to tighten grip on social media comments, requiring sites to employ sufficient content moderators*, SOUTH CHINA MORNING POST (18 June 2022) https://finance.yahoo.com/news/china-tighten-grip-social-media-093000585.html.
[147] Shahbaz, Funk & Vesteinsson, *supra* note 42 ("In China, the government has been fairly successful in pairing systematic censorship of foreign services with robust investment in domestic platforms that are beholden to the ruling party.").

wide range of activity as terrorism or violent extremism, including protests in Hong Kong, uprisings in Xinjiang and Tibet, and even a tennis star's accusation of a high-ranking Chinese Communist Party official of sexual assault.[148]

2. Self-Regulation by Social Media Companies

This section provides a brief history of company self-regulation of content moderation practices which took place in phases: early efforts before 2009, the rise of industrial content moderation in 2009–2017, and improved technology alongside increasing legal requirements beginning in 2017, through to today. As the analysis of national laws did, this section will also specifically look at how platforms moderate TVEC. Many platforms look to the UN Human Rights Guiding Principles on Business and Human Rights and Article 19 of the ICCPR to guide their governance practices.[149] However, all online platforms moderate user-generated content slightly differently, and self-regulation efforts have varied between companies over the years. Generally, online platforms have moderated content through four broad approaches: case-by-case, community-reliant, industrial, or large-scale, and no moderation whatsoever.[150]

Self-regulation by online platforms hosting user-generated content started in the earliest days of internet bulletin board services, when companies like CompuServe and Prodigy set rules for their subscribers to follow when posting content.[151] Indeed, it was specifically to protect the content moderation practices of these early internet companies that Congress passed Section 230 in 1996.[152] When modern-day online platforms launched in the early 2000s, content moderation was largely ad hoc, and most companies presented themselves as neutral intermediaries to avoid being held responsible for what their users said and did.[153] However, even in the early days, all commercially viable platforms moderated some content, to ensure their services were not overrun with spam, nudity, or other toxic content.[154] As Charlotte Willner, one of Meta's first content moderators, noted, the ethos of the pre-2008 moderation guidelines was, "if it makes you feel bad in your gut, then go ahead and take it down."[155] This

---

[148] Murray Scot Tanner & James Bellacqua, *China's Response to Terrorism*, CAN (Jun. 2016), https://www.uscc.gov/sites/default/files/Research/Chinas%20Response%20to%20Terrorism_CNA061616.pdf.

[149] T.G. Thorley & E. Saltman, *GIFCT Tech Trials: Combining Behavioural Signals to Surface Terrorist and Violent Extremist Content Online, Studies in Conflict & Terrorism*, TAYLOR & FRANCIS ONLINE, https://www.tandfonline.com/doi/full/10.1080/1057610X.2023.2222901 ("Focused on companies' applications of policies, the UNHR's Guiding Principles on Business and Human Rights is a bedrock for tech companies. Guiding principles for tech companies in moderation practices and data collection dictate that policies should dictate actions deemed as necessary, lawful, legitimate, and proportionate, and that the right to restriction should be tied to a defined and defendable threat.").

[150] Caplan, *supra* note 20 at 16 ("Three major categories of platform companies according to their size, organization, and content moderation practices: (1) the artisanal approach, where case-by-case governance is normally performed by between 5 and 200 workers; (2) community-reliant approaches, which typically combine formal policy made at the company level with volunteer moderators; and (3) the industrial approach, where tens of thousands of workers are employed to enforce rules made by a separate policy team."). Caplan's analysis misses an emergent set of companies that claim to do no content moderation whatsoever, including platforms like 4chan, 8kun and Gab.

[151] *Section 230: Legislative History*, ELECTRONIC FRONTIER FOUNDATION, https://www.eff.org/issues/cda230/legislative-history.

[152] Kosseff, *supra* note 67 at 75-76.

[153] Suzor, *supra* note 37 at 15; *see also* Tarleton Gillespie, *The Politics of Platforms*, NEW MEDIA & SOCIETY 12, NO. 3 (1 May 2010), 347–364, https://doi.org/10.1177/1461444809342738.

[154] James Grimmelmann, *The Virtues of Moderation*, YALE JOURNAL OF LAW AND TECHNOLOGY 17, 42 (2015).

[155] Klonick, *supra* note 36 at 1631; *citing*, Telephone Interview with Dave Willner, Former Head of Content Policy, Facebook & Charlotte Willner, Former Safety Manager, User Operations, Facebook (23 March 2016).

ethos, still found in the artisanal approach to content moderation, shifted to become more industrial as online platforms expanded internationally and companies sought to make their products attractive to global users.[156]

Starting in 2009, and continuing through to 2016, companies began to craft global platform rules. This led to a more industrial process of content moderation where companies enforced rules globally on millions of pieces of content.[157] The rules are sometimes referred to as the "terms of service" or "community standards" which a user agrees to follow when signing up for a platform.[158] As described by Kate Klonick, an academic focusing on internet policies, in her seminal article on content moderation, "The New Governors", social media companies developed global standardised content rules to manage: "(1) the increase in both users and volume of content; (2) the globalization and diversity of the online community; and (3) the increased reliance on teams of human moderators with diverse backgrounds."[159] Klonick argues that online platforms self-regulated because they were economically motivated to create a hospitable environment to incentivise engagement.[160] She goes on to say that companies try to keep up as much speech as possible while upholding their ideals of corporate responsibility.[161] As processes developed, many self-regulatory models adopted a "common-law" approach, to maintain consistency in the decision-making process.[162] Even with "common-law" precedent, platforms were constantly updating their policies to adapt to global norms in response to: "(1) government request, (2) media coverage, (3) third-party civil society groups, and (4) individual users' use of the moderation process."[163] Throughout the 2000s and 2010s, companies were largely free to write their own rules, because Section 230 and other safe harbour regimes did not draw clear lines around acceptable or unacceptable content.[164]

As the large social media companies came to dominate the global landscape and to draw increased scrutiny, a societal shift was taking place. Starting in 2017, events such as Russian

---

[156] Gillespie, *supra* note 20 at 4; *see also* Caplan, *supra* note 20 at 23.
[157] Klonick, *supra* note 36; *see also* Douek, *supra* note 21 at 11 ("Once those rules are written, it's simply a matter of applying them over and over … and over again—the standard picture conceives of content moderation as simply the aggregation of millions of daily paradigm cases. The scale is hard to comprehend: in Q3 2021, Facebook took down 933,426,800 pieces of content, YouTube took down 4,806,042 channels and 6,229,882 videos, and in Q2 2021 TikTok removed 81,518,334 videos. These figures do not include every time these platforms decided to leave up content flagged for review (which would greatly exceed decisions to remove content) or appeals.").
[158] Citron, *supra* note 110 at 1037 ("From the start, tech companies' commitment to free expression admitted some exceptions. Terms of service and community guidelines banned child pornography, spam, phishing, fraud, impersonation, and copyright violations. Threats, cyber stalking, nonconsensual pornography, and hate speech were prohibited after extended discussions with advocacy groups. The goal was to strike an appropriate balance between free expression and abuse prevention while preserving platforms' market share.").
[159] Klonick, *supra* note 36 at 1635.
[160] *Id*. at 1618 ("[companies] are private, self-regulating entities that are economically and normatively motivated to reflect the democratic culture and free speech expectations of their users.").
[161] *Id*.
[162] Caplan, *supra* note 20 at 18 ("One legal counsel compared the model they took to a "common-law system" based on precedent, while others described a process similar to a grounded theory approach, a methodology used in the social sciences to inductively build up categories, through the aggregation of individual cases or data points."); *citing* Interview with Alex Feerst, head of Legal at Medium.
[163] Klonick, *supra* note 36 at 1649.
[164] Caplan, *supra* note 20 at 28 ("Within the United States, Section 230 of the Communications Decency Act provides platforms like those discussed above with the freedom to organize their content moderation teams as they see fit, as long as they are taking care to remove copyright protected and illegal content. As platforms deploy the other right given to them by Section 230 and the "Good Samaritan" provision, platforms told us they are finding it difficult to draw lines in ways that make sense both ethically and organizationally.").

meddling in the 2016 election, the genocide in Myanmar, the Cambridge Analytica scandal and, subsequently, the Christchurch attack, raised the public awareness of the potential harms of "big tech". As a result, social media companies adopted a defensive posture, and many platforms looked to self-regulatory solutions as a low-cost way to repair reputational damage and stave off government regulation.[165] Self-regulatory efforts were frequently championed by tech company employees who wanted to create change from the inside, and sky-high profits meant the companies had cash to spend on these experiments.[166] In this vein, tech companies used their money and soft power to work with civil society, journalists, and academics to institutionalise self-regulation practices through organisations like the Global Internet Forum to Counter Terrorism (GIFCT), Meta's independent Oversight Board,[167] and Alphabet's Jigsaw project, which researched how to curb extremism and misinformation across products.[168]

Another shift began in 2017 when new technologies transformed the content moderation industry. Online platforms started to deploy automated tools to detect and filter harmful content alongside predictive models that relied on AI to learn and recognise patterns.[169] Many companies incorporated automation into their content moderation systems, with the encouragement of policymakers, who were increasingly calling on them to restrict content they deemed harmful.[170] In addition to automation, moderation itself became more nuanced, as companies thought beyond the binary decision of keeping up or taking down content. As Evelyn Douek, an internet law academic, notes, platforms adopted a variety of tools, including "sticking labels on posts; partnerships with fact-checkers; greater platform and government collaboration; adding friction to how users share content; giving users affordances to control their own online experience; looking beyond the content of posts to how users behave online to determine what should be removed; and tinkering with the underlying dynamics of the very platforms themselves."[171] Indeed, over the past few years, the evolving work of ensuring the safety and security of online platforms has become so sophisticated that it has created an entire industry of "trust and safety" professionals.[172]

While these new tools affected a wide range of content, preventing the spread of TVEC online was one area of content moderation in which companies invested significantly in self-regulation after years of government pressure.[173] Terrorist and violent extremist use of the

[165] Kate Klonick, *The End of the Golden Age of Tech Accountability*, THE KLONICKLES (4 March 2023), https://klonick.substack.com/p/the-end-of-the-golden-age-of-tech.
[166] *Id.*
[167] Kate Klonick, *Inside the Making of Facebook's Supreme Court*, THE NEW YORKER (12 Feb. 2021), https://www.newyorker.com/tech/annals-of-technology/inside-the-making-of-facebooks-supreme-court.
[168] Jane Wakefield, *TED 2018: Alphabet firm's tools to combat extremism*, BBC (13 April 2018), https://www.bbc.com/news/technology-43760213.
[169] Carey Shenkman, Dhanaraj Thakur & Emma Llanso, *Do You See What I See?*, CENTER FOR DEMOCRACY AND TECHNOLOGY (May 2021), https://cdt.org/wp-content/uploads/2021/05/2021-05-18-Do-You-See-What-I-See-Capabilities-Limits-of-Automated-Multimedia-Content-Analysis-Full-Report-2033-FINAL.pdf.
[170] *Id*. at 10 ("Policymakers worldwide are increasingly calling on social media companies to identify and restrict text, photos, and videos that involve illegal, harmful, or false information. Many services are voluntarily incorporating automation into their content moderation systems, and government agencies are also exploring the use of automated content analysis."), *see also* Robert Gorwa et al., *Algorithmic Content Moderation: Technical and Political Challenges in the Automation of Platform Governance*, 7 BIG DATA & SOC. 1 (2020).
[171] Douek, *supra* note 21 at 5.
[172] Fishman, *supra* note 24; *See also* TS 2023 Research Collaborative, *Trust & Safety Curriculum*, TRUST & SAFETY PROFESSIONALS ASSOCIATION (2023), https://www.tspa.org/curriculum/ts-curriculum/.
[173] Fishman, *supra* note 24.

internet is not a new phenomenon. Indeed, al-Qaeda was operating online by the mid-1990s,[174] and the prevalence of white supremacism online was so great by 1996 that the Anti-Defamation League started tracking it.[175] Like the early cyber libertarians, early terrorists and violent extremists saw the internet as a great place to find like-minded individuals and discuss ideas free from government censorship.[176] By the early 2000s, terrorists and violent extremists were drawn to social media for the same reasons as everyone else: social media platforms are a simple and reliable way to share ideas and connect with a vast network of people.[177] But terrorist use of social media did not go unnoticed. In 2008, during a Senate hearing, US Senator Lieberman demanded that YouTube remove Al-Qaeda training videos.[178] In a response many would now find shocking, the company's representative defended the terrorist organisation's right to express unpopular viewpoints on their platform.[179] Indeed, the companies did not seriously try to self-regulate until 2015, when the so-called Islamic State began using social media to recruit and inspire violence in Europe, leading lawmakers to threaten regulation.[180]

In the years following 2015, social media companies attempted to self-regulate TVEC on their platforms by establishing robust internal processes and industry collaboration. First, social media companies cleaned up their platforms individually in several ways, including: writing rules defining what constitutes a terrorist organisation and TVEC; identifying and removing policy violations; providing data on TVEC in transparency reports; and limiting access to product features to decrease the virality of TVEC.[181] Additionally, platforms started to proactively work with governments and law enforcement officials to remove content from entities who were designated as terrorist organisations.[182] As companies implemented these measures, they were quick to share results with lawmakers in an attempt to stave off regulation. For example, Twitter, a company who branded itself the "free-speech wing of the free-speech party" since its founding, reported that it had suspended over 125,000 ISIS-related accounts in 2016, and Meta announced it had hired 3,000 more people to stop the spread of terrorist propaganda.[183]

Second, tech companies started to work together as an industry to self-regulate through several projects. The most notable TVEC-related self-regulatory initiative was the Global Internet

---

[174] Gabriel Weimann, *Terror on the Internet: The New Arena, the New Challenges*, U.S. INSTITUTE OF PEACE (2010), https://www.usip.org/publications/2010/05/terror-internet.
[175] Fishman, *supra* note 24; *citing* David H. Strassler, et al., *The Web of Hate: Extremists Exploit the Internet*, ANTI-DEFAMATION LEAGUE (1996), https://www.adl.org/sites/default/files/documents/assets/pdf/combating-hate/ADL-Report-1996-Web-of-Hate-Extremists-exploit-the-Internet.pdf.
[176] Fishman, *supra* note 24.
[177] *Id*.
[178] Timothy B. Lee, *YouTube Rebuffs Senator's Demands to Remove Islamist Videos*, ARS TECHNICA (20 May 2008), https://arstechnica.com/tech-policy/2008/05/youtube-rebuffssenatorss-demands-for-removal-of-islamist-videos/.
[179] *Id.*
[180] Liat Clark, *Facebook and Twitter Must Tackle Hate Speech or Face New Laws*, WIRED UK (5 Dec. 2016), http://www.wired.co.uk/article/us-tech-giants-must-tackle-hate-speech-orface-legal-action.
[181] Fishman, *supra* note 24.
[182] Klonick, *supra* note 36 at 1638; *See also* Natalie Andrews & Deepa Seetharaman, *Facebook Steps Up Efforts Against Terrorism*, WALL STREET JOURNAL (11 Feb. 2016), http://on.wsj.com/1T; Joseph Menn & Dustin Volz, *Google, Facebook Quietly Move Toward Automatic Blocking of Extremist Videos*, REUTERS (24 June 2016), https://www.reuters.com/article/us-internet-extremism-video-exclusive/exclusive-google-facebookquietly-move-toward-automatic-blocking-of-extremist-videos-idUSKCN0ZB00M.
[183] *Id*. at Andrews & Seetharaman.

Forum to Counter Terrorism (GIFCT). In 2016, the idea was floated that companies should create a shared database of banned TVEC, which would operate like PhotoDNA, a tool developed to remove child sexual abuse material.[184] At first, online platforms and civil society organisations were wary of the idea of a TVEC database, as there was no agreed-upon definition for what constituted "terrorist content".[185] However, the tech companies reversed course in December 2016, the day before the European Commission released a damning report condemning their efforts to remove TVEC.[186] In 2017, Facebook, Microsoft, Twitter, and YouTube launched the GIFCT as an industry initiative to apply technology, share knowledge, and support research on terrorists' abuse of the platforms.[187] This new project included the creation of a database to which companies could upload terrorist content found on their platforms and "hash" the images and videos. These "hashes", frequently called "digital fingerprints", were entered into the database, and the technology prevented the upload of hashed images on any of the cooperating platforms.[188] By 2019, this database included over 200,000 pieces of content.[189] Despite claims of success by tech companies, the hash-sharing database was frequently criticised by civil society for not being more transparent in regard to the content in the database and by governments that wanted to ensure the images they perceived as TVEC were included.[190] As Part III will explore, after the Christchurch attack in 2019, reforming the GIFCT from an industry-led project into an MSI became a top priority.

B. Multi-Sided Content Governance Frameworks

Single-sided efforts by national lawmakers and online platforms were successful to some extent in reducing the proliferation of TVEC online. However, many argued that a new framework was necessary because democratic countries were limited in their ability to regulate content, and self-regulation was falling short.[191] One potential solution was the creation of MSIs, which would bring together governments, companies, civil society, and outside experts to identify a solution and implement it across sectors. Through an MSI, stakeholders can harness the

[184] Kaveh Waddell, *A Tool to Delete Beheading Videos Before They Even Appear Online*, THE ATLANTIC (22 June 2016), https://www.theatlantic.com/technology/archive/2016/06/a-tool-to-delete-beheading-videos-before-they-even-appear-online/488105/.
[185] Citron, *supra* note 110 at 1044; *noting*, lawmakers were uninterested in hearing reasons why TVEC was a fundamentally different problem to child sexual abuse material which was universally considered to be illegal and abhorrent.
[186] *Partnering to Help Curb Spread of Online Terrorist Content*, META NEWSROOM (5 Dec. 2016), https://about.fb.com/news/2016/12/partnering-to-help-curb-spread-of-online-terrorist-content/.
[187] Christchurch Call News & Updates, *Significant progress made on eliminating terrorist content online*, CHRISTCHURCH CALL TO ACTION (24 Sep. 2019), https://www.christchurchcall.com/media-and-resources/news-and-updates/new-news-article-page-8/.
[188] Global Internet Forum to Counter Terrorism, *Who we are: Story, 2017 Year in Review*, GLOBAL INTERNET FORUM TO COUNTER TERRORISM (2023), https://gifct.org/about/story/#2017-year-in-review.
[189] Global Internet Forum to Counter Terrorism, *Who we are: Story, May 2019, the Christchurch Call to Action*, GLOBAL INTERNET FORUM TO COUNTER TERRORISM (2023), https://gifct.org/about/story/#may-2019---christchurch-call-to-action.
[190] Courtney Radsch, *GIFCT: Possibly the Most Important Acronym You've Heard Of*, JUST SECURITY (30 Sep. 2020), https://www.justsecurity.org/72603/gifct-possibly-the-most-important-acronym-youve-never-heard-of/.
[191] Douek, *supra* note 21 at 78 ("An underlying theme and motivation of this Article has been that the limits of direct governmental regulation of online speech are significant, making it necessary to find an approach that leverages and legitimates platform self-regulation. Governmental oversight of platforms should aim to maximize the private sector's resources, expertise and dynamism in finding innovative and effective methods for tackling content moderation challenges while requiring platforms to explain, justify and verify those methods. By allowing platforms to experiment, government oversight would avoid locking in the status quo at the major platforms.").

capabilities of different actors and co-design solutions through participatory processes. This section will first explore the rise of multistakeholder frameworks and their suitability to addressing global challenges. It will then explore the successes of multistakeholderism within the internet governance space as a template for the content moderation governance problem.[192] Finally, it looks at emerging multilateral initiatives and their effort to frame themselves as multistakeholder without truly being MSIs.

1. The Transition from Multilateral to Multistakeholder

Unlike our concept of multistakeholderism, modern-day concepts of multilateralism can be traced back to 1648 and the signing of the Peace of Westphalia, which recognised the sovereignty of individual states and promoted the idea of non-interference in the affairs of other states.[193] That treaty created a world order based on interaction, negotiation, and cooperation among sovereign states.[194] John Gerald Ruggie, in his seminal article on multilateralism, defines multilateralism as "the practice of coordinating national policies in groups of three or more states, through ad hoc arrangements or by means of institutions."[195] In this system, government representatives act on behalf of their citizens and implement the terms of any agreement within their borders to resolve international issues.[196] Following two devastating World Wars, nations strengthened multilateral institutions and developed new ones, notably the United Nations, to prevent further violent conflict and set human rights standards.[197] However, by the 1980s, multilateral frameworks were failing to address many global issues as governments lacked the internal capacity to implement policies due to a gradual erosion of trust and (in many cases) extensive corruption.[198]

Rapid globalisation in the 1980s and 1990s compounded geopolitical tensions and exposed many of the underlying problems with multilateral frameworks. During this timeframe, national governments found their monopoly on public policy making increasingly contested, with the emergence of three powerful groups: transnational corporations, civil society, and an

[192] This report uses the definition from Raymond & DeNardis for "internet governance" to broadly describe six technical functions for the internet including: critical internet resources such as domain names and IP addresses, internet standards for interoperability, interconnection between networks, cyber-security, information intermediation, and intellectual property rights enforcement. Content moderation happens primarily at the information intermediation layer of the internet stack, so while many of these MSIs include some coordination on content, they typically address a wider range of internet functions; *see* Raymond & DeNardis, *supra* note 40, 19-45.
[193] Brad Smith and Carol Ann Browne, TOOLS AND WEAPONS, 330 (2021).
[194] John Gerald Ruggie, *Multilateralism: the Anatomy of an Institution*, THE MIT PRESS VOL. 46, NO. 3, 565 (Summer, 1992).
[195] *Id.*
[196] Harris Gleckman, *Multistakeholderism: a corporate push for a new form of global governance*, TRANSNATIONAL INSTITUTE (19 Jan. 2016), https://www.tni.org/en/publication/multi-stakeholderism-a-corporate-push-for-a-new-form-of-global-governance.
[197] *Id*.
[198] Christopher Ansell & Jacob Torfing, HANDBOOK ON THEORIES OF GOVERNANCE, 7 (24 June 2016); explaining the rise of multistakeholder governance ("in the fields of public administration, public law and public policy, this question arose out of the attempt to address challenges posed by administrative complexity, poor policy implementation and fiscal austerity. In the field of development studies, it developed in response to the frustration of achieving development goals in partnership with weak or corrupt developing states. In the field of international relations, economics and environmental studies, the question grew out of the need to address collective action problems and the management of common pool resources.").

independent media.[199] First, transnational corporations grew so large that their economic power and cultural authority sometimes exceeded that of many states. Second, growing international links between civil society organisations connected disparate movements, which provided a larger platform for human rights advocacy. As such, civil society and non-government organisations came to be viewed as legitimate actors in the formulation, implementation, and evaluation of public policy.[200] Third, as literacy rates and access to information increased, citizens became more skeptical of state-run media organisations, lending credibility to independent journalists. The newly empowered media was quick to expose governmental inability to hold corporations responsible for their wrongdoing, which increased public pressure on corporations to respect human rights.[201]

Increased pressure on transnational corporations and the governments that failed to hold them accountable forced parties to consider collaborative approaches with new groups of stakeholders.[202] For corporations, stakeholder collaboration was valuable in jurisdictions where governments could not or would not uphold basic human rights, leaving governance gaps for unregulated business practices.[203] Additionally, these discussions could provide corporations with local knowledge and new insights into diverse problems, which sometimes yielded better return on investment.[204] Next, democratic governmental actors were quick to participate in stakeholder discussions, as results from this type of governance could demonstrate an impact to constituents without the need to pass legislation.[205] Finally, civil society found this collaboration beneficial when it provided new opportunities to demonstrate their soft power.[206] Over time, this consultation with stakeholders provided a helpful form of checking and balancing.[207] As a result, collaboration between stakeholders led to the co-production of public solutions which increased their legitimacy.[208] While often not legally binding, if executed

---

[199] *Id.*
[200] *Id*.
[201] Baumann-Pauly et al., *supra* note 32 at 10 ("media interest focused on headline grabbing issues, such as the use of sweatshops by well-known brands like Nike, Disney and Levi Strauss.").
[202] Ansell & Torfing, *supra* note 198 at 7.
[203] Gleckman, *supra* note 196.
[204] Ariel Babcok, Nathan Barrymore, Christopher Bruno, Allen He, et al., *Walking the Talk: Valuing a Multi-Stakeholder Strategy*, FCLT GLOBAL AND WHARTON UNIVERSITY OF PENNSYLVANIA, 5-6, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4023510 ("while there is often a natural gravitational pull to prioritize one set of stakeholders over another (shareholders in many cases), prioritizing one group continuously is not a winning long-term strategy … Future-fit, long-term companies need more durable performance to succeed – and that requires attention to a broader group of stakeholders.").
[205] Nick Buxton, *Multistakeholderism: a critical look*, TRANSNATIONAL INSTITUTE (19 Jan. 2016), https://www.tni.org/en/publication/multistakeholderism-a-critical-look; *see also* Lawrence E. Strickling & Jonah Force Hill, *Multi-stakeholder Governance Innovations to Protect Free Expression, Diversity and Civility*, CENTRE FOR INTERNATIONAL GOVERNANCE INNOVATION & STANFORD GLOBAL DIGITAL POLICY INCUBATOR, Special Report: Governance Innovation for a Connected World Protecting Free Expression, Diversity and Civic Engagement in the Global Digital Ecosystem, 45 (2018) ("Multistakeholder processes can be resource intensive, but they are still generally less financially burdensome than traditional regulatory proceedings or litigation. Reaching multi-stakeholder consensus can be difficult and time-consuming, but compare the time it takes to achieve consensus to the time it takes the US Congress to enact legislation. New entrants may have a strategic disadvantage in multi-stakeholder settings, but they at least have a seat at the table and a say in the outcome. Traditional government and multilateral rulemaking settings afford them no such right.").
[206] Raymond & DeNardis, *supra* note 40.
[207] Admin. Conference of the U.S., Recommendation 2018-7, Public Engagement in Rulemaking, 84 Fed. Reg. 2139, 2146 (6 Feb. 2019) ("Robust public participation is vital to the rulemaking process. By providing opportunities for public input and dialogue, agencies can obtain more comprehensive information, enhance the legitimacy and accountability of their decisions, and increase public support for their rules.").
[208] Ansell & Torfing, *supra* note 198 at 7.

properly, this type of collaboration establishes and reinforces standards that one party could not achieve by acting on its own. These successes contributed to the formalisation of multistakeholderism and declining reliance on multilateralism.[209]

Multistakeholderism is defined as two or more classes of actors engaged in a common governance enterprise to solve a wider problem, where decision-making authority is distributed between actors based on procedural rules.[210] An MSI is created when two or more types of actors come together in a structured organisation to solve a problem defined by the group. Different types of actors with a potential stake in an MSI include businesses, civil society, governments, universities, academics, technical experts, investors, and consumers.[211] In recent decades, multistakeholderism has emerged as an alternative to, and sometimes a direct competitor with, traditional multilateral approaches, for several key reasons. First, MSIs are frequently created when an industry or government finds itself facing a significant amount of public pressure to fix a problem that it cannot solve on its own.[212] Sometimes this occurs shortly following a tragic event. For example, after rampant human rights violations in the diamond trade were made public, the Kimberley Process created an MSI that urged governments to pass regulation, companies to certify the source of the diamonds, and civil society to oversee the process.[213] Second, MSIs are created to help fill governance gaps in regulatory frameworks. In this situation, MSIs establish guidelines or best practices for stakeholder behaviour where local or national regulators cannot or do not uphold human rights principles. For example, many MSIs were created in the 1990s to address the use of "sweatshops" in countries where governments did not enforce fair labour practices. [214] Finally, MSIs are frequently created to address technological advances where development of the technology requires broader societal input to protect human rights. The need to address technological advancement is one of the primary drivers of the proliferation of MSIs in the internet governance space.

---

[209] Gleckman, *Multistakeholderism: a new way for corporations and their new partners to try to govern the* world, CIVICUS (Oct. 2018), https://www.civicus.org/index.php/re-imagining-democracy/overviews/3377-multistakeholderism-a-new-way-for-corporations-and-their-new-partners-to-try-to-govern-the-world ("Even for the proponents of multistakeholderism, the transition from the nation-state as the actor in international affairs to 'stakeholders' as global governors has been an uneven process. One major element of the transition for these new claimants as global leaders is learning to work with a heterogeneous group of organisations, some or which were, or still are, institutional opponents. The differences in types of power external to an MSG group create a fundamental asymmetry of power within the group.").

[210] Raymond & DeNardis, *supra* note 40 at 20 ("Multi-stakeholderism is defined here as two or more classes of actors engaged in a common governance enterprise concerning issues they regard as public in nature, and characterized by polyarchic authority relations constituted by procedural rules.").

[211] Baumann-Pauly et al., *supra* note 32.

[212] John Gerald Ruggie, *Business and Human Rights: The Evolving International Agenda*, THE AMERICAN JOURNAL OF INTERNATIONAL LAW VOL. 101, No. 4, 819-840 (Oct. 2007) ("[d]riven by social pressure, [and]… seek to close regulatory gaps that contribute to human rights abuses. But they do so in specific operational contexts, not in any overarching manner. Moreover, recognizing that some business and human rights challenges require multi-stakeholder responses, they allocate shared responsibilities and establish mutual accountability mechanisms within complex collaborative networks. These can include any combination of host and home States, corporations, civil society actors, industry associations, international institutions and investors groups.").

[213] *What is the Kimberley Process?,* KIMBERLEY PROCESS (2023), https://www.kimberleyprocess.com/en/what-kp.

[214] Baumann-Pauly et al., *supra* note 32 at 2 ("The absence of state regulation presents major business challenges for corporations. Clothing retailers like Walmart and H&M face unsafe factory conditions in Bangladesh in the wake of the Rana Plaza tragedy. Internet service providers like Facebook and Google wrestle with their users' expectations to guarantee freedom of expression in China and other non-democratic regimes. Oil and mining companies like Shell and Newmont operating in conflict zones from the Congo to Iraq struggle to provide security for their people and facilities in these inherently dangerous places. In these contexts, multistakeholder initiatives (MSIs) have become the default response for addressing so-called "governance gaps".).

2. Multistakeholderism in Internet Governance

The 1990s were a pivotal decade for both the development of MSIs and the growth of the internet. It is therefore not surprising that the two rose to prominence together. As noted above, early internet adopters identified with many of the core principles of multistakeholderism, including the need to be collaborative, interconnected, and global.[215] The internet of today is a byproduct of multistakeholder collaboration between engineers, individuals, government agencies, and businesses. Internet standards were created by an international group of stakeholders who shared a common goal to create a decentralised network.[216] Over the years, internet governance MSIs have encompassed a wide range of approaches, procedures, formats, and outcomes.[217]

One early example of an MSI for internet governance was the Internet Engineering Task Force (IETF). Initially started in 1987 as a quarterly meeting at which a dozen US researchers could exchange ideas, by 1992 over 750 stakeholders from government, civil society, and industry were attending to set internet standards.[218] Around the same time, in 1991, Vint Cerf, Bob Kahn, and other internet entrepreneurs created the multistakeholder Internet Society to "promote the open development, evolution and use of the internet for the benefit of all people throughout the world."[219] In 1992, the two organisations merged, with the Internet Society providing a legal umbrella for the IETF to help manage its growth and maintain independence from the US government.[220] Both the IETF and Internet Society are still relevant for our discussion of successful MSIs in internet governance, as they demonstrate an early focus on multistakeholderism in the community. At the IETF, stakeholders set standards through a bottom-up process whereby decisions are based on what has been called "rough consensus and running code".[221]

Another important, albeit controversial, internet governing body created during this period was the Internet Corporation for Assigned Names and Numbers (ICANN). The roots of ICANN date back to 1969, when researchers began sending electronic messages to each other through the Arpanet.[222] To make it easier to track and send messages via the network, Jon Postel, a

---

[215] Internet Society, *Internet Governance: Why the Multistakeholder Approach Works*, INTERNET SOCIETY (26 April 2016), https://www.internetsociety.org/wp-content/uploads/2016/04/IG-MultiStakeholderApproach.pdf.
[216] Konstantinos Komaitis, *Global Digital Compact – Additional submission,* UNITED NATIONS TECH ENVOY FILES (6 April 2023), https://www.un.org/techenvoy/sites/www.un.org.techenvoy/files/GDC-submission_Konstantinos-Komaitis.pdf ("The Internet and multistakeholder governance are tightly interwoven. The Internet is a byproduct of a pure collaborative process between engineers, individuals, government agencies and businesses. It emerged because this different set of people shared a common goal despite their often diverse and distinct viewpoints; that goal was to create a network that would be decentralized and could respond to any type of failure.").
[217] Strickling & Hill, *supra* note 205 at 45 ("There is no one single concept of what is appropriately viewed to be a multi-stakeholder approach. There are, instead, numerous models currently in use today, each with its own unique contours. Few, if any, of the models currently in use are static; rather, they are constantly evolving to meet new and yet uncharted governance challenges.").
[218] Scott Bradner, *The Internet Engineering Task Force*, OPEN SOURCES: VOICES FROM THE OPEN SOURCE REVOLUTION, 1ST EDITION (Jan. 1999), https://www.oreilly.com/openbook/opensources/book/ietf.html.
[219] Barry M. Leiner, Vinton G. Cerf, David D. Clark, Robert E. Kahn, et al., *Brief History of the Internet*, INTERNET SOCIETY (1997), https://www.internetsociety.org/internet/history-internet/brief-history-internet/.
[220] Bradner, *supra* note 218.
[221] Raymond & DeNardis, *supra* note 40 at 32.
[222] *ICANN History Project*, INTERNET CORPORATION OF ASSIGNED NAMES AND NUMBERS (Oct. 2016), https://www.icann.org/history.

researcher in southern California, created a registry to manage the coordination of messages, which ultimately became the Domain Name System.[223] In practice, the Domain Name System became the "phonebook" of the internet – allowing people to easily look up other people. In the beginning, it was possible for Postel to maintain this function on his own, but the burden of providing for the technical management of the Domain Name System increased rapidly. As Postel testified to Congress, in 1993 there were 30,000 domain names; by 1997 there were 1.6 million globally.[224]

As the scale and complexity of the Domain Name System grew, the US Government sought to relinquish its historic control over technical internet functions by creating ICANN, a non-profit entity dedicated to the task. After lengthy stakeholder engagement, in 1998, the Department of Commerce signed a memorandum of understanding with ICANN which outlined how ICANN would manage key functions, including by allocating IP number blocks, overseeing the root server system, and coordinating technical parameters.[225] Most critically, under this agreement, ICANN managed the Internet Assigned Numbers Authority (IANA), which administers functions of the Domain Name System. It is important to note that while the IANA function serves the global internet, at that time its funding came from the US Government, and it was considered a government asset.[226] As a result, ICANN was beholden to the US Government's reporting requirements, with the understanding that the organisation would eventually become fully independent.[227] While the US Government's role was largely procedural, there was mounting resentment from other nations over perceived "American control of the internet."[228] This issue threatened to divide the global internet space.

Tensions surrounding ICANN's structure escalated in the early 2000s. Many governments wanted to see the UN manage ICANN's responsibilities through the multilateral International Telecommunication Union (ITU). The ITU is a body within the UN that regulates radio spectrum, satellite orbits and certain worldwide technical standards.[229] Standard-setting at the ITU is top-down and bureaucratic: a method preferred by governments that were seeking to control their citizens' access to information and tax the burgeoning internet economy. This structure was anathema to the bottom-up, decentralised, and interoperable internet governance system that had developed since 1969. As a result, many stakeholders from civil society, industry, and democratic governments saw the possibility of ITU control over the IANA function as undermining both the functionality and freedoms of the global internet. These

[223] *Id.*
[224] Internet Domain Names, Part 1: Hearing Before the Committee on Science, Subcommittee on Basic Research, *supra* note 38.
[225] Joe Sims & J. Beckwith Burr, *Memorandum of Understanding Between the U.S. Department of Commerce and Internet Corporation for Assigned Names and Numbers*, INTERNET CORPORATION OF ASSIGNED NAMES AND NUMBERS RESOURCES (31 Dec. 1999), https://www.icann.org/resources/unthemed-pages/icann-mou-1998-11-25-en.
[226] Steve Crocker*, On Creating Internet Governance Organizations: A Comment on the ICANN Experience*, INTERNET GOVERNANCE FORUM BERLIN, "Towards a Global Framework for Cyber Peace and Digital Cooperation: An Agenda for the 2020s," 148 (25-29 Nov. 2019), https://www.hiig.de/wp-content/uploads/2019/11/Kleinwa%CC%88chter-Kettemann-Senges-eds.-Global-Framework-for-Cyber-Peace-2019.pdf.
[227] Sims & Burr, *supra* note 225.
[228] Raymond & DeNardis, *supra* note 40 at 27.
[229] United Nations Secretariat of the Internet Governance Forum (IGF), *About the IGF*, INTERNET GOVERNANCE FORUM (2023), https://www.intgovforum.org/en/about.

tensions came to a head in 2003, when the battle over the future of ICANN and the Domain Name System root zone management was brought up at the ITU's World Summit on the Information Society (WSIS). The entrenchment of both factions meant that no agreement was reached. However, two years later, at the second phase of WSIS in Tunis, UN members agreed on a compromise that would forestall giving ICANN oversight to the ITU by creating the Internet Governance Forum (IGF). The IGF is an MSI still under the oversight of the UN that identifies and defines the public policy issues that are relevant to internet governance.[230]

Following the directive set for the IGF in Tunis in 2005, the UN held two rounds of consultations to establish the objectives and format of the IGF.[231] The first meeting of the IGF was in 2006 in Athens. Over 1,200 participants attended from government, the private sector, civil society, academia, and technical communities.[232] In the years following, the IGF created processes to be more inclusive, including the creation of a dedicated Multistakeholder Advisory Group to help with planning and participation, starting an open consultation process to allow the public to submit suggestions regarding the program for the IGF, and instituting a host country selection process whereby countries could bid to host the event.[233] Over the years, the IGF expanded its stakeholder engagement and helped develop a broader sense of multistakeholderism throughout the internet governance community.

In 2013, events threatened to undo the successes of multistakeholderism. Edward Snowden's leak of thousands of documents revealed an extensive spying programme the US National Security Agency conducted over internet infrastructure. Many world leaders turned to the UN in hopes of finding a multilateral solution to government surveillance.[234] One voice calling for multilateral intervention was then-President of Brazil Dilma Rousseff, who had her personal cell phone targeted for the content of calls, emails, and messages by the National Security Agency.[235] In the days following the leaks, she urged the UN and government actors to get involved to enforce rules governing the internet. However, shortly after her speech to the UN, Brazil instead decided to organise the Global Multistakeholder Meeting on the Future of Internet Governance, which came to be known as NETmundial.[236] The NETmundial conference took place in April 2014, bringing together over 1,400 people from all over the world.[237] Stakeholders collaborated in small working groups over several days to create an outcome document which outlined principles for internet governance and a roadmap for the future of the internet governance ecosystem.[238] By all measures, this was a significant

[230] *Id.*
[231] *Id.*
[232] United Nations Secretariat, *Internet Governance Forum to Hold Inaugural Session in Athens from 30 October to 2 November, Press ReleasePI/1747*, UNITED NATIONS: MEETINGS COVERAGE AND PRESS RELEASES (25 Oct. 2006), https://press.un.org/en/2006/pi1747.doc.htm.
[233] United Nations Secretariat of the Internet Governance Forum (IGF), *supra* note 229.
[234] Deborah Brown & Anriette Esterhuysen, *Extracting lessons from NETmundial: Achieving bottom-up and multistakeholder outcomes from global internet governance policy discussions*, ASSOCIATION FOR PROGRESSIVE COMMUNICATIONS (2016), https://www.apc.org/sites/default/files/ExtractingLessonsFromNETmundial.pdf.
[235] *Id.* at 6.
[236] *Id.*
[237] cgi.br (Brazilian Internet Steering Committee) & /1net, *NETmundial Multistakeholder Statement*, NETMUNDIAL GLOBAL MEETING ON THE FUTURE OF INTERNET GOVERNANCE (24 April 2014), https://netmundial.br/about/.
[238] *Id.*

achievement for the multistakeholder model, which had not traditionally produced consensus-driven outcomes.

NETmundial's successes fostered goodwill between stakeholders in the internet governance ecosystem. Seeking to maintain momentum, just a few months later, the conference organisers teamed up with ICANN and the World Economic Forum to start the NETmundial Initiative (NMI).[239] The NMI was meant to "carry forward the cooperative spirit of São Paulo and work together to apply the NETmundial Principles."[240] However, it ran into trouble almost immediately when it was revealed that the three lead organisers had awarded themselves "permanent seats" on its 25-member council, isolating key stakeholders and directly undermining the NMI's claims to be a bottom-up MSI.[241] The NMI was further undermined by a lack of transparency, accountability, and inclusivity – all values called for in the outcome documents from NETmundial.[242] Finally, it was hard to justify the need for a separate initiative when the reforms outlined in the outcomes document had been enacted by the IGF and ICANN.[243] As a result, NMI's "mandate" to ICANN and the World Economic Forum expired in 2016, and the initiative was shut down.

One issue that received a lot of attention at NETmundial was the ongoing debate related to the US Government's oversight of ICANN, and the IANA functions. Stakeholders contended that internet governance could never be multistakeholder as long as the US Government still maintained oversight.[244] However, six weeks before NETmundial, the US Government announced its intent to transition its stewardship role of the IANA function to the global multistakeholder community.[245] In June 2014, ICANN started a multistakeholder process to transition the IANA function oversight away from the US Government.[246] Over the next two years, participants held more than 600 meetings to finalise the details of the transition and on

---

[239] World Economic Forum and Internet Corporation for Assigned Names and Numbers, *NETmundial Initiative for Internet Governance Cooperation & Development*, WORLD ECONOMIC FORUM (28 Aug. 2014), https://www3.weforum.org/docs/WEF_1NetmundialInitiativeBrief.pdf.

[240] *Id.*

[241] Internet Society, *Internet Society Statement on the NETmundial Initiative, Press Release,* INTERNET SOCIETY (17 Nov. 2014), https://www.internetsociety.org/news/press-releases/2014/internet-society-statement-on-the-netmundial-initiative/ ("Based on the information that we have to date, the Internet Society cannot agree to participate in or endorse the Coordination Council for the NETmundial Initiative. We are concerned that the way in which the NETmundial Initiative is being formed does not appear to be consistent with the Internet Society's longstanding principles, including: Bottom-up orientation, Decentralized, Open, Transparent, Accountable, Multi-stakeholder.").

[242] Larry Strickling, *Remarks of Assistant Secretary Strickling on the Self-Governing Internet at Georgia Institute of Technology*, NATIONAL TELECOMMUNICATIONS AND INFORMATION ADMINISTRATION (26 Oct. 2016), https://ntia.gov/speechtestimony/remarks-assistant-secretary-strickling-self-governing-internet-georgia-institute ("Yet despite support from the United States government and others, the NetMundial Initiative never got off the ground. Why? Because it lacked the support and participation of all the relevant stakeholders, most notably the business community and the Internet Society. It was developed in a top-down way, without bottom-up support and input from the community. In the eyes of many key stakeholders, the initiative lacked the legitimacy it needed to succeed.").

[243] *Id.; see also* cgi.br & /1net, *supra* note 237 for outcomes document text.

[244] cgi.br & /1net, *supra* note 237.

[245] ICANN's Major Agreements and Related Reports *Transition of NTIA's Stewardship of the IANA Functions*, ICANN, https://www.icann.org/resources/pages/process-next-steps-2014-06-06-en#:~:text=On%2014%20March%202014%20the,to%20the%20global%20multistakeholder%20community.

[246] Fiona Alexander, *Global Digital Cooperation: Conditions for Success*, INTERNET GOVERNANCE FORUM BERLIN, "Towards a Global Framework for Cyber Peace and Digital Cooperation: An Agenda for the 2020s," 70-73 (25-29 Nov. 2019), https://www.hiig.de/wp-content/uploads/2019/11/Kleinwa%CC%88chter-Kettemann-Senges-eds.-Global-Framework-for-Cyber-Peace-2019.pdf.

1 October 2016, the process was completed.[247] As part of these negotiations, ICANN added another layer of governance, called the Empowered Community, which promoted multistakeholderism in its processes supporting its internet governance activities.[248]

Another issue debated at NETmundial was the future of the IGF, which was set to be reviewed in 2015 by the UN General Assembly in a process called "WSIS+10". In December 2015, much of the advice provided in the NETmundial outcomes document was incorporated into the 10-year renewal of the IGF.[249] Part of this renewed mandate included a commitment to the UN's Sustainable Development Goals, an ambitious blueprint for global peace and prosperity established in 2015.[250] Goal 17 recognises multistakeholder partnerships as important vehicles for mobilising and sharing knowledge, expertise, technologies, and financial resources to support the Sustainable Development Goals in all countries.[251] With this mandate, IGF continued its role as a global convener for multistakeholderism in internet governance, with no additional powers to bind stakeholders to standards or rules. One stakeholder, Nnenna Nwakanma, spoke to the consensus: "IGF is not what we want it to be. But we do not have a better option. We all wish to be happy, but since we cannot all be happy in our own ways, we settle for collective dissatisfaction."[252] The IGF remained a worthwhile initiative for many stakeholders.

Despite the criticism, the IGF helped entrench an ethos of multistakeholderism in the internet governance space, for three key reasons. First, global multistakeholder attendance at the conference brought together people from around the world and across sectors who shared common goals and beliefs. These connections were invaluable to the internet governance ecosystem, which requires a high degree of trust between disparate groups. Second, the language of multistakeholderism was so pervasive that many organisations sought to adopt similar messaging to increase the credibility of their policy solutions. Third, the IGF was purposefully created to avoid regulatory approaches. Instead, it encouraged bottom-up, collaborative solutions. This allowed new organisations to fill the policy vacuum and start new MSIs that could create policy between smaller groups of stakeholders. As a result, in the past 18 years, hundreds of internet governance MSIs have been created to address global issues. There are too many to name and analyse here, but it is worth mentioning a few of the pivotal MSIs that formed following the creation of the IGF in 2006 that are still relevant today.

---

[247] Larry Strickling, *Remarks of Assistant Secretary Strickling at The Internet Governance Forum USA*, NATIONAL TELECOMMUNICATIONS AND INFORMATION ADMINISTRATION (14 July 2016), https://ntia.gov/speechtestimony/remarks-assistant-secretary-strickling-internet-governance-forum-usa-1.
[248] Crocker, *supra* note 226 at 152.
[249] Internet Society, *Understanding the WSIS+10 Review Process*, INTERNET SOCIETY (May 2015), https://www.internetsociety.org/wp-content/uploads/2017/08/WSISplus10-Overview.pdf.
[250] United Nations Department of Economic and Social Affairs, *The 17 Goals*, UNITED NATIONS (2023), https://sdgs.un.org/goals.
[251] United Nations Department of Economic and Social Affairs *Sustainable Development, Multi-stakeholder partnerships*, UNITED NATIONS (2023), https://sdgs.un.org/topics/multi-stakeholder-partnerships.
[252] Nnenna Nwakanma, *Because I am involved!,* INTERNET GOVERNANCE FORUM BERLIN, "Towards a Global Framework for Cyber Peace and Digital Cooperation: An Agenda for the 2020s," 198-200 (25-29 Nov. 2019), https://www.hiig.de/wp-content/uploads/2019/11/Kleinwa%CC%88chter-Kettemann-Senges-eds.-Global-Framework-for-Cyber-Peace-2019.pdf.

One topic frequently discussed at the IGF is government censorship and privacy violations. In 2008, there was a series of high-profile instances where technology companies had legally complied with the Chinese Government's requests for access to data. This data was then used to jail journalists and activists. Finding themselves in a no-win situation – either they violated a legally issued government order or they undermined human rights – certain global companies teamed up with civil society, investors, and academia to create the Global Network Initiative (GNI).[253] GNI established the Principles of Free Expression and Privacy to create a baseline of human rights commitments that participating stakeholders agreed to uphold globally.[254] As an MSI, GNI collaborates to find solutions to the challenges of protecting digital rights globally by drawing on the perspectives, leverage, credibility, and expertise of many different stakeholders.[255] One unique aspect of GNI is its independent assessment process, through which participating companies undergo a third-party review of their efforts to implement the GNI Principles and their more detailed Implementation Guidelines. These assessments focus on internal company systems and emblematic case studies, providing insights to non-company GNI members on sensitive, non-public information and scenarios. GNI's Board is then charged with determining whether each company has implemented the Principles and Implementation Guidelines "in good faith, with improvement over time." Over the past 15 years, GNI has continued to be a leading MSI on internet governance issues, fostering multistakeholder collaboration to push back on government censorship, enhance shared learning, and provide tools to support responsible decision making by tech companies.

Another topic frequently discussed at the IGF was how global technology companies should operate if national regulations conflict. In 2011, Internet & Jurisdiction Policy Network (I&J) was formed to address the idea that governments, internet companies, civil society, and academics should come together to advance legal interoperability online.[256] I&J focused on specific issues-based problems, believing that cooperation in the internet governance space needed to be addressed with joint agenda setting and policy development by all relevant stakeholders to foster the mutual trust needed for implementation.[257] After four years of meetings and stakeholder consultations, I&J has launched three workstreams: Data & Jurisdiction, Content & Jurisdiction, and Domains & Jurisdiction. These workstreams eventually led to policy options papers and toolkits for governments, tech companies, and civil society that continue to be relevant and useful for stakeholders across the internet governance sector.

Overall, each of the MSIs discussed above (IETF, Internet Society, ICANN, IGF, NETmundial, GNI, and I&J) have succeeded in bringing stakeholders together to address challenging internet governance problems that could not be solved through national laws and tech industry self-

[253] Baumann-Pauly et al., *supra* note 32.
[254] Global Network Initiative, *The GNI Principles*, GLOBAL NETWORK INITIATIVE (June 2011), https://globalnetworkinitiative.org/gni-principles/; *see also* Baumann-Pauly et al., *supra* note 32.
[255] Global Network Initiative, *About GNI*, GLOBAL NETWORK INITIATIVE https://globalnetworkinitiative.org/about-gni/.
[256] Internet & Jurisdiction Policy Network, *History*, INTERNET & JURISDICTION POLICY NETWORK (2020), https://www.internetjurisdiction.net/about/history.
[257] Bertrand de la Chapelle, *Towards a Governance Protocol for the Social Hypergraph*, INTERNET GOVERNANCE FORUM BERLIN, "Towards a Global Framework for Cyber Peace and Digital Cooperation: An Agenda for the 2020s," 106-109 (25-29 Nov. 2019), https://www.hiig.de/wp-content/uploads/2019/11/Kleinwa%CC%88chter-Kettemann-Senges-eds.-Global-Framework-for-Cyber-Peace-2019.pdf.

regulation on their own. The past 30 years of MSIs have produced a rich, normative framework of stakeholder collaboration to ensure internet governance is a highly interdependent process.[258] However, resurgent top-down multilateral efforts in the internet governance space threaten to undermine this progress.

3. Recent Multilateral Efforts in Internet Governance

As the internet is increasingly intertwined with other global issues, the UN has tried to move internet governance away from multistakeholderism and back into a multilateral framework. To do this it has launched multi-year initiatives that will culminate in the multilateral negotiations of the Global Digital Compact in 2024.[259] The timeline and development of this work is troubling to the broader multistakeholder internet governance community. First, in July 2018, the UN Secretary-General convened a High-Level Panel on Digital Cooperation to advance proposals to strengthen cooperation in the digital space.[260] This kicked off two years of debate around the UN's role in internet governance and culminated in June 2020 with the "Roadmap for Digital Cooperation", which the Secretary-General's Office of the Envoy on Technology was set to implement.[261] Following the Roadmap for Digital Cooperation, in 2021, the UN Secretary-General's Envoy on Technology put forward "Our Common Agenda", which proposed a "Global Digital Compact – an Open, Free and Secure Digital Future for All".[262] In 2023, the UN sought input from all stakeholders on the Global Digital Compact, which it plans to integrate into a policy brief to help aid future negotiations on internet governance policies.[263] This negotiation will culminate in 2024, when the UN will host the "Summit of the Future", at which the member states will agree on multilateral solutions to "strengthen" global internet governance.[264] Part of this work will include rules for internet governance for "ensuring the protection of human rights in the digital era."[265] Therefore, while the Global Digital Compact will have multistakeholder input, the final agreement will be multilateral in nature. This is troubling because it creates top-down rules that give countries like Russia and China the ability to weaken the strong human rights protections put in place by a multistakeholder framework.

Unfortunately, the Global Digital Compact is only one slice of the work the UN has launched related to internet governance in the past few years. Additionally, the United Nations Educational, Scientific and Cultural Organization (UNESCO), which promotes freedom of expression, access to information, and digital transformation, has also taken an interest in internet governance issues.[266] UNESCO's "Internet for Trust" is developing "guidelines for

[258] Komaitis, *supra* note 216.
[259] United Nations Office of the Secretary General's Envoy on Technology, *Report of the Secretary-General Roadmap for Digital Cooperation*, THE UNITED NATIONS TECH ENVOY (May 2020), www.un.org/techenvoy/sites/www.un.org.techenvoy/files/general/Roadmap_for_Digital_Cooperation_9June.pdf.
[260] *Id.*
[261] *Id.*
[262] United Nations Office of the Secretary General's Envoy on Technology, *Our Common Agenda Policy Brief 5, A Global Digital Compact – an Open, Free and Secure Digital Future for All*, THE UNITED NATIONS TECH ENVOY (May 2023), https://www.un.org/techenvoy/global-digital-compact.
[263] *Id.*
[264] *Id.*
[265] *Id.*
[266] United Nations Educational, Scientific and Cultural Organization, *Safeguarding freedom of expression and access to information: guidelines for a multistakeholder approach in the context of regulating digital platforms*

regulating digital platforms: a multistakeholder approach to safeguarding freedom of expression and access to information".[267] UNESCO is trying to build upon the work it has done in the domain of broadcast regulation, which established principles for internet universality known as the ROAM principles: rights, openness, accessibility to all, and multistakeholder participation.[268] This MSI has been criticised as unnecessary as it is unclear how these guidelines will work with other UN initiatives, including the work of the Envoy on Technology and the IGF.[269] Additionally, the proposed guidelines for regulation are consistent with Article 19 of the ICCPR, which leaves stakeholders to wonder why UNESCO is trying to rewrite settled principles.[270] One theory would be that government actors who do not agree or abide by Article 19 principles and are seeking to weaken existing protections.

The underlying problem with both initiatives is that they are top-down, bureaucratic, and inherently multilateral.[271] As Konstantinos Komaitis, an internet scholar and practitioner, describes it, this model "is based purely on state actors making all decisions at the exclusion of other stakeholders who can make valuable and informed contributions … It will not advance the Internet; on the contrary, it will break it into small pieces. It will fragment it."[272] Multilateral negotiations on internet governance are particularly troubling when non-democratic nations like Russia and China are given a seat at the table. Democratic governments know that the autocrats will not uphold human rights commitments relating to freedom of expression and privacy, so any negotiation will likely weaken the legal commitments already in place. The UN understands the mistrust of civil society and democratic governments and, therefore, cloaks

---

Draft 3.0, CI-FEJ/FOEO/3 Rev, UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION DIGITAL LIBRARY (27 April 2023), https://unesdoc.unesco.org/ark:/48223/pf0000384031.locale=en.

[267] *Id.*

[268] *Id.; see also* United Nations Educational, Scientific and Cultural Organization, *Internet Universality Indicators: Background*, UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION (2019), https://www.unesco.org/en/internet-universality-indicators/background.

[269] Global Network Initiative, *Global Network Initiative Comments on UNESCO's "Guidance for Regulating Digital Platforms: A Multistakeholder Approach",* GLOBAL NETWORK INITIATIVE (April 2023), https://globalnetworkinitiative.org/wp-content/uploads/2023/01/GNI-Comments-on-UNESCO-draft-Guidance-FINAL.docx-1-1.pdf ("The shortcomings of the consultation process are underscored by the lack of any clear and compelling justification for why this process is being expedited, and the resulting lack of risk-benefit analysis or impact assessment. Perhaps due to this rushed process, the Guidance makes only passing mention of the "Our Common Agenda" report, the "Global Digital Compact," the "UN Summit of the Future," and the WSIS+20 process, and does not offer any clear articulation of how this initiative has been or will be coordinated with other relevant UN offices and initiatives, including the Tech Envoy's office, UN Human Rights (OHCHR), and the Internet Governance Forum.").

[270] David Kaye, *UNESCO Guidelines for Regulating Digital Platforms: A Rough Critique*, UCI LAW INTERNATIONAL JUSTICE CLINIC (21 Feb. 2023), https://ijclinic.law.uci.edu/2023/02/21/unesco-guidelines-for-regulating-digital-platforms-a-rough-critique/ ("Second, the draft provides limited if any guidance as to the definition of the problem it is meant to address. From a legality perspective ("provided by law"), this is deeply concerning. Early on, the draft emphasizes "content that is illegal under international human rights law and content that risks significant harm to democracy and the enjoyment of human rights." …I blanched when I saw that [definition], since generally speaking (with two exceptions) international law does not make content illegal; it provides a framework of guaranteed individual rights (Article 19: seek, receive and impart information and ideas of all kinds, regardless of frontiers) along with a set of narrow limitations as to when the state may restrict those rights. It is true that Article 20 of the International Covenant on Civil and Political Rights (ICCPR) obligates states to prohibit "propaganda for war" and "advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence". But if that's what the draft means to address, why not say that directly? This may seem like an editing issue, but the lack of clarity opens the door to state arguments that categories of content many want to limit (e.g., defamation of religion, lèse-majesté, false information, extremism, and so on) are not merely subject to restriction but illegal under international law. This could amount to a major win for governments not, shall we say, entirely enamored of Article 19 of the ICCPR.").

[271] Komaitis, *supra* note 216.

[272] *Id.*

these initiatives as a "multistakeholder effort" by involving the private sector, civil society, and other stakeholders in consultations.[273] In this framework, the UN envisions itself as a convener for multistakeholder policy dialogues, but calls upon member states to develop and implement regulatory frameworks.[274] Therefore, both initiatives embrace the term "multistakeholder" without actually being multistakeholder. As a result, the rights-respecting internet governance community is increasingly uncomfortable with the efforts by the UN to set standards.

[273] See United Nations Office of the Secretary General's Envoy on Technology, *Report of the Secretary-General Roadmap for Digital Cooperation*, *supra* note 259 at 22.
[274] *Id.* at 24.

. II. CREATING A MULTISTAKEHOLDER INITIATIVE FOR CONTENT GOVERNANCE

The New Zealand government set up the Call as an MSI to address online user-generated content governance frameworks, understanding that a patchwork of national laws and self-regulation has not been sufficient to eliminate TVEC online while protecting a free, open, and secure internet. The Call drew upon the history of the internet, which is a by-product of multistakeholder collaboration between engineers, governments, tech companies, and civil society.[275] However, despite the deep history of multistakeholderism in internet governance over the past 30 years, many MSIs have only recently started to consider multistakeholder solutions for governance of user-generated content online.[276] While the early MSIs provide a good guide, current models can be slightly adjusted to better address content governance issues and new technologies.[277] This part first outlines the different types of MSIs found in the internet governance space and then proposes six best practices for creating a self-sustaining MSI. This framework will help guide the Call as it considers the future of its work in the content governance space and the development of new technologies.

Overall, this part argues that stakeholders should embrace MSIs to effectively address problems with content governance, for three important reasons. First, the content online crosses borders and cannot be effectively legislated by national governments, leaving governance gaps. This is particularly important because not all governments are willing to govern content in a way that respects human rights. To solve this problem, MSIs can exclude bad actors without compromising protections. Second, online platforms will continue to struggle to create their own standards without more input from governments, civil society, and technologists. These inputs can help balance national security interests with freedom of expression and provide local context and accountability. Finally, multistakeholderism is already built into the internet's foundation, and it can therefore be easily imported into new initiatives. Today, the effectiveness of multistakeholderism appears to be taken at face value; almost all internet policymaking initiatives have adopted the model.[278] As the Internet Governance Project, a non-profit organisation affiliated with the Georgia Institute of Technology, has argued, the embrace of multistakeholderism is generally a positive development, as it ensures that civil society, governments, and tech companies all have a seat at the table.[279] But not all MSIs are created equal, and this framing can mean that the term "multistakeholder" is sometimes applied unequally, and therefore critical analysis is required.[280]

---

[275] Komaitis, *supra* note 216.

[276] *See* de la Chapelle, *supra* note 257 at 106 ("A distributed institutional ecosystem was progressively developed for governance OF the internet58. It efficiently enabled this unique creation of mankind to now serve more than half the world's population. However, equivalent efforts were not devoted to developing the necessary policy-making tools for governance ON the internet, i.e. to organize its uses and mitigate in respect of human rights abuses it can allow. As a result, we witness a legal arms race.").

[277] Strickling & Hill, *supra* note 205.

[278] Jyoti Panday, Milton Mueller, Farzaneh Badiei, *Multistakeholderism & Platform Content Governance: An Assessment Framework with Applications*, UNIVERSITY OF GEORGIA TECH SCHOOL OF PUBLIC POLICY INTERNET GOVERNANCE PROJECT (20 Jan. 2022), https://www.internetgovernance.org/wp-content/uploads/MS-Content.docx-1.pdf ("The term "multistakeholder" (MS) is now claimed as a legitimizing feature of various international, Internet-related policy development entities. Civil society in particular tends to demand multistakeholder governance in order to gain entry into decision-making processes otherwise controlled by business or government. While in many ways the advance of MS governance is a good thing, it also means that the term can be applied loosely or even deceptively. We need to ask what multistakeholderism really means in a particular policy environment, and we need to assess critically how these organizations are being set up.").

[279] *Id.*

[280] *Id.*

A. Typology of MSIs

The variation of challenges, actors, and structures can make it difficult to have one definition of "MSI". As the former administrator of NTIA Larry Strickling notes, multistakeholder models have their own unique contours, but, "few, if any, of the models currently in use are static; rather, they are constantly evolving to meet new and yet uncharted governance challenges."[281] Part I defines an MSI as two or more classes of actors engaged in a common governance enterprise to solve a wider problem, where decision-making authority is distributed between actors based on procedural rules.[282] Therefore, two core elements that define an initiative as multistakeholder are: the inclusion of multiple types of actors and the distribution of decision-making authority based on procedural rules.[283]

These two core elements frequently take two forms within any given MSI, creating four overarching types of MSIs. In terms of the first element, the inclusion of stakeholders, there are two principal systems: either anyone who is interested can participate, or the MSI only allows stakeholders who meet certain criteria to join. The second element, regarding how the MSI distributes decision-making authority based on procedural rules, is slightly more complex. One option is for decisions to be made by the consensus of all stakeholders; in this case, the governance itself takes a multistakeholder form.[284] Consensus-based institutions are considered more multistakeholder in nature and can increase the possibility that the solution presented by the MSI is adopted in the long-run.[285] The second option is for stakeholders to serve a purely consultative purpose; in this case, decision-making happens unilaterally by the designated authority.[286] This can also be considered "ancillary" multistakeholder governance, because it involves the multistakeholder body acting as an appendage to a decision-making body.[287] While neither of these decisions are straightforward or strictly binary, it is helpful to make distinctions to examine what types of MSIs are best suited for each unique situation.

There are thus four types of MSIs:

- egalitarian: any stakeholder, consensus decision-making
- consultative: any stakeholder, unilateral decision-making
- restricted: limited stakeholders, unilateral decision-making

[281] Strickling & Hill, *supra* note 205 at 40.
[282] Raymond & DeNardis, *supra* note 40.
[283] Panday, Mueller & Badiei, also includes "funding" as a distinctive category. This report addresses funding considerations as part of the terms of reference rather than its own category, *see* Panday, Mueller & Badiei *supra* note 278.
[284] Jan Aart Scholte, *Multistakeholderism: Filling the Global Governance Gap?*, STOCKHOLM: GLOBAL CHALLENGES FOUNDATION, 4 (6 April 2020), https://globalchallenges.org/multistakeholderism-filling-the-global-governance-gap/.
[285] Strickling & Hill, *supra* note 205 at 49 ("Also, to maximize the possibility of success, participants must be the ones who make the final decision on a particular issue, not the convening body. This feature is one of the fundamental differences between a multi-stakeholder process and consultation. If participants are not empowered to make a final decision, then a process is merely consultative. By contrast, multi-stakeholder processes that place responsibility for final decision making on the participants themselves are generally viewed as more legitimate. They also tend to be more successful because the prospect of fashioning policy, and not just offering commentary, frequently induces the participants to put in the extra effort needed to reach a consensus. Further, entrusting the participants with the power to make decisions also reduces the possibility of non-participants mounting a collateral challenge of the outcome by appealing to others who did not choose to participate.").
[286] *See* Panday, Mueller & Badiei *supra* note 278; the authors have a third category where MSIs have a consultative body that is advisory in status, but its formal advice triggers some kind of procedure and cannot be ignored. However, for purposes of this report, this middle ground will be included in the consultative function.
[287] Scholte, *supra* note 284 at 4.

- curated: limited stakeholders, consensus decision-making.

1. Egalitarian MSIs: Any Stakeholder, Consensus Decision-making

This type of MSI was the vision of early internet adopters – people like John Perry Barlow thought that the rules for the internet would emerge through community engagement and consensus.[288] As such, we saw examples of this type of MSI in the early days of the internet.[289] Egalitarian MSIs look like Athenian democracy, where all stakeholders must participate directly in the decision-making. While this may sound aspirational, there are many MSIs that operate in this manner. The IETF is an example. It does not have an official or defined membership; rather, it allows anyone to participate.[290] Many of the stakeholders come from industry, government, civil society, and the technical community, but everyone participates in their personal capacity. The IETF has no formal voting process but makes decisions based on what has been called "rough consensus and running code".[291] The IETF's process of standard-setting is a significant investment in time and energy by stakeholders, but in the end, the community is able to progress with the greatest amount of input and consensus possible.

The IETF has sustained this type of MSI for decades, but not all egalitarian MSIs have succeeded. For example, in 2009, Facebook (now Meta) experimented with its own egalitarian MSI after changes to its privacy policy were significantly criticised.[292] Facebook announced that it would develop the site's terms of service through consensus building, by asking users to weigh in on company policies. In 2012, Facebook tested this approach by putting forward two different privacy policies; it asked users to vote, committing that if more than 30 per cent of all active registered users participated, their decision would be binding.[293] When it came time to vote, only 665,654 people voted – about 0.3 per cent of Facebook's 200 million users at the time.[294] Facebook followed the majority opinion of the lackluster showing, but since most people voted for the proposed changes, the decision was criticised for being a cover for the company to do a thing it already wanted to do.[295] In the end, Meta scrapped the initiative, which the *Los Angeles Times* called "a homework assignment no one did."[296] Overall, an egalitarian MSI approach works best where relevant stakeholders are deeply invested in the outcome and highly motivated to find consensus.

4. Consultative MSIs: Any Stakeholder, Unilateral Decision-making

A consultative MSI is created when one stakeholder has unilateral decision-making authority but seeks input from all interested stakeholders. The stakeholder input is considered but not dispositive to the final decision. This type of multistakeholder governance is found throughout many democratic institutions, in places like the US Administrative Procedures Act, which requires a "notice and comment

[288] Barlow, *supra* note 35.
[289] Komaitis, *supra* note 216.
[290] P. Resnick, *On Consensus and Humming in the IETF*, INTERNET ENGINEERING TASK FORCE (June 2014), https://datatracker.ietf.org/doc/html/rfc7282; *see also* Niels ten Oever, *Plus Hum Now: Decision Making at the IETF*, HACK_CURIO (March 2018), https://hackcur.io/please-hum-now/.
[291] Raymond & DeNardis, *supra* note 40 at 32.
[292] Suzor, *supra* note 37 at 20; Mark Zuckerberg, *Update on Terms*, FACEBOOK (17 Feb. 2009), www.facebook.com/ notes/facebook/update-on-terms/54746167130.
[293] Adi Robertson, *Mark Zuckerberg wants to democratize Facebook — here's what happened when he tried*, THE VERGE (6 April 2018), https://www.theverge.com/2018/4/5/17176834/mark-zuckerberg-facebook-democracy-governance-vote-failure.
[294] Id.
[295] Id.
[296] David Sarno, *Facebook governance vote is a homework assignment no one did*, LOS ANGELES TIMES (23 April 2009), https://www.latimes.com/archives/blogs/technology-blog/story/2009-04-23/facebook-governance-vote-is-a-homework-assignment-no-one-did.

period" before a regulatory agency can issue a final rule.[297] In these MSIs, the multistakeholder community acts as a sounding board or advisor to the decision maker.[298] One example of this type of consultative MSI is Meta's Oversight Board, which hears appeals from users regarding content moderation decisions taken by Meta on Facebook and Instagram and issues binding decisions to the company.[299] As part of the Board's work when deciding cases, it opens up a "public comment process", which allows any stakeholder to submit their thoughts on how the company should have moderated a piece of content or crafted its policies.[300] The Board seeks advice from all stakeholders to gain local context and subject matter expertise to improve the quality of their decisions.[301]

Other consultative MSIs previously discussed include UNESCO's Internet for Trust and its development of "guidelines for regulating digital platforms: a multistakeholder approach to safeguarding freedom of expression and access to information" and the UN's Roadmap for Digital Cooperation.[302] In these cases, the UNESCO Secretariat or the UN Secretary-General is developing the new guidelines and, as part of the process, will conduct multistakeholder consultations.[303] In these examples, a multistakeholder community is consulted, but the authority for deciding ultimately sits with one stakeholder. The advantage of a consultative MSI is that it allows for a wide range of voices to participate in a process without relying on consensus to reach a final decision. This expedites the process and provides all stakeholders an opportunity to weigh in. However, this type of MSI can be deceptive if the term "multistakeholder" is used to legitimise a process without clearly explaining stakeholders' lack of decision-making authority.[304]

5. Restricted MSIs: Limited Stakeholders, Unilateral decision-making

A restricted MSIs allows only qualified stakeholders to participate in the initiative and decision-making to happen unilaterally. In some respects, the US Supreme Court's amicus curiae process is a restricted MSI, because it allows for the consideration of stakeholders' views on current cases but restricts those stakeholders to attorneys admitted to practice before the court.[305] Given the restrictions, many would not consider this to be a multistakeholder process. Another decision-making institution that embraces

---

[297] Justia, *The Notice and Comment Process Legally Provided for Agency Rulemaking*, JUSTIA (May 2023), https://www.justia.com/administrative-law/rulemaking-writing-agency-regulations/notice-and-comment/.; *see also* Strickling & Hill, *supra* note 205 at 49 ("Notwithstanding the desire of government officials to allow a group of stakeholders to reach a consensus decision, the laws of the government, such as the Administrative Procedures Act in the United States, may prohibit giving the decisionmaking power to a group of stakeholders and require the agency to conduct subsequent notice and comment on the rule-making processes, thus diminishing the incentive of stakeholders to work together to reach consensus in the multi-stakeholder discussions.").

[298] Panday, Mueller & Badiei *supra* note 278 at 2 ("the additional stakeholders serve a purely advisory or consultative function; they act as a sounding board or advisor to the decision maker.").

[299] The Oversight Board, *Governance: Rulebook*, THE OVERSIGHT BOARD (Oct. 2022), https://oversightboard.com/; *noting* the Board opens up public comment processes for both case decisions and Policy Advisory Opinions.

[300] *Id.*

[301] *Id.; see also* The Oversight Board 2022 Annual Report, 13 (May 2023), https://oversightboard.com/news/560960906211177-2022-annual-report-oversight-board-reviews-meta-s-changes-to-bring-fairness-and-transparency-to-its-platforms/ ("As a Board, our achievements so far have been made possible by listening to and collaborating with researchers, civil society groups and others who have worked for many years on the issues we are dealing with. To find practical solutions to our strategic priorities, and the enormously challenging issues they raise, the subject-matter expertise and local knowledge of these stakeholders is essential.").

[302] United Nations Educational, Scientific and Cultural Organization, *Safeguarding freedom of expression and access to information: guidelines for a multistakeholder approach in the context of regulating digital platforms Draft 3.0, supra* note 266.

[303] *Id.*

[304] Panday, Mueller & Badiei *supra* note 278.

[305] Scott S. Harris, *Memorandum to Those Intending to File an Amicus Curiae Brief in The Supreme Court Of The United States*, Supreme Court Of The United States Office Of The Clerk (Jan. 2023), https://www.supremecourt.gov/casehand/AmicusGuide2023.pdf.

the restricted MSI model is the ITU, which allows non-state actors to join the multistakeholder processes through restricted participation. A non-state stakeholder is referred to as a "sector member" and must apply to join and be sponsored by a member state.[306] Sector members can participate in day-to-day standards-setting work within the working groups of the ITU, but any recommendations made by the working groups are ultimately approved exclusively by member states.[307]

One example of a restricted MSI that is working to address content moderation challenges surrounding TVEC online is the Organisation for Economic Co-operation and Development (OECD), which developed a Voluntary Transparency Reporting Framework for TVEC online.[308] Created after the Christchurch attack in 2019, the OECD created the Framework "in collaboration with member countries, business, civil society and academia, to develop a multi-stakeholder, consensus-driven framework" for voluntary transparency reporting by companies.[309] The OECD limited participation to its member states and invited guests, and the OECD had final decision-making authority on what criteria would be included in the Framework. In the past three years, the OECD has published an annual report that takes stock of the "current policies and procedures related to TVEC of the world's leading online platforms and other online content sharing services."[310] This MSI has provided insight into global efforts to reduce TVEC online. However, in the cases of both the OECD and the ITU, stakeholder participation is limited and decisions on the outcome are made by one actor.

6. Curated MSIs: Limited Stakeholders, Consensus Decision-making

Of the internet governance MSIs mentioned in Part I, curated MSIs are the most common type; this is the approach taken by the Christchurch Call. Stakeholders in these MSIs uphold the ideals of consensus in the same way the early internet adopters envisioned through shared decision-making authority. However, stakeholder participation is limited to those who meet a defined set of criteria. One important reason for limiting participation in internet governance MSIs stems from the divide between democratic and authoritarian regimes and their vastly different approaches to upholding the human rights principles as they relate to content moderation practices. Therefore, in many instances, a demonstrated commitment to upholding human rights is a baseline for participation in the MSI to safeguard the consensus-driven outcomes.

GNI and I&J are two examples of internet governance curated MSIs. To participate in GNI, stakeholders must support the organisation's established Principles of Free Expression and Privacy and undergo a due diligence check.[311] GNI develops governance structures for companies to implement through a consensus-based multistakeholder process.[312] Likewise, I&J works with relevant stakeholders committed to preserving "the cross-border nature of the internet, [protecting] human rights, [fighting] abuses, and [enabling] the global digital economy."[313] I&J working groups develop consensus-based

[306] *Membership Terms & Conditions*, INTERNATIONAL TELECOMMUNICATIONS UNION (2023), https://www.itu.int/hub/membership/become-a-member/member-terms-conditions/.
[307] Raymond & DeNardis, *supra* note 40 at 33.
[308] Organisation for Economic Co-operation and Development, *Transparency reporting on terrorist and violent extremist content online 2022*, OECD DIGITAL ECONOMY PAPERS, No. 334 (25 Oct. 2022), https://www.oecd.org/digital/vtrf/.
[309] Maddie Cannon, *A Review of International Multi-Stakeholder Frameworks for Countering Terrorism and Violent Extremism Online*, GLOBAL NETWORK ON EXTREMISM & TECHNOLOGY (16 March 2022), https://gnet-research.org/2022/03/16/a-review-of-international-multi-stakeholder-frameworks-for-countering-terrorism-and-violent-extremism-online/.
[310] *See* Organisation for Economic Co-operation and Development, *supra* note 308 at 16.
[311] Global Network Initiative, *The GNI Principles*, *supra* note 254.
[312] *Id.*; *see also* Baumann-Pauly et al., *supra* note 32.
[313] *Internet & Jurisdiction Policy Network Progress Report 2021*, INTERNET & JURISDICTION (2021), https://www.internetjurisdiction.net/uploads/pdfs/IJPN-Progress-Report-2021.pdf.

toolkits that strive to fill an institutional gap in internet governance.[314] Additionally, many curated MSIs address content moderation challenges surrounding TVEC online, including the EU Internet Forum and the GIFCT. As a curated MSI, the EU Internet Forum provides a collaborative environment in which partners discuss and address the challenges posed by malicious and illegal content online – including TVEC.[315] The Forum is chaired by the European Commission, which invites stakeholders from the internet industry and civil society actors to participate.[316] The GIFCT has a multistakeholder Independent Advisory Committee that advises the Operating Board through regular meetings and makes decisions by consensus.[317] The Independent Advisory Committee appoints representatives from civil society organisations and academia as well as representatives from governments who are members of the Freedom Online Coalition, a group of governments dedicated to human rights.[318]

B. Building a Curated MSI

The Call seeks to build a curated MSI, which ensures participating stakeholders are committed to protecting a free, open, and secure internet. Through the Call's leader-led format, all decisions are reached by consensus. This section will outline best practices for creating a self-sustaining a curated MSI. The uniqueness of each MSI precludes a one-size-fits-all approach. However, it is possible to distill a step-by-step guide that entails deciding if an MSI is necessary, establishing the objectives and functions of the initiative, deciding who is a stakeholder, setting up terms of reference, sustaining forward momentum, and deciding when the work is finished.

1. Step 1: Deciding if an MSI is Necessary

1(a). Define the Problem

MSIs are frequently created in three circumstances: to allow stakeholders to fill governance gaps in regulatory frameworks, to respond to a high-profile industry or government failure (sometimes in the wake of a tragic event), or when technological developments pose new challenges. An MSI is necessary when one party acting on its own cannot solve the problem.[319] The organisers should start by defining the problem caused by the change in circumstances and why input from multiple stakeholders is necessary to address the situation.

1(b). Map the Landscape

Next, the organisers should conduct a mapping exercise to see if other groups are already addressing the defined problem. Organisers should then consider if they are in the best position to offer solutions to the defined problem. For example, would litigation be a better instrument to remedy the harm caused?

---

[314] *Id.*
[315] European Commission, Migration and Home Affairs, *European Union Internet Forum (EUIF), Membership*, EUROPEAN UNION (2023), https://home-affairs.ec.europa.eu/networks/european-union-internet-forum-euif_en#membership.
[316] *Id.*
[317] Global Internet Forum to Counter Terrorism, *Governance, Independent Advisory Committee*, GLOBAL INTERNET FORUM TO COUNTER TERRORISM (2023), https://gifct.org/governance/#iac-governance.
[318] Global Internet Forum to Counter Terrorism Independent Advisory Committee, *Terms of Reference,* GLOBAL INTERNET FORUM TO COUNTER TERRORISM (2023), https://gifct.org/wp-content/uploads/2021/09/GIFCT-IAC-Terms-of-Reference.pdf.
[319] *See* Part I(B)(1); *see also* Baumann-Pauly et al., *supra* note 32 at 2 ("We understand "multistakeholder initiative" or "MSI" to mean an entity that works with multiple stakeholders (usually business and civil society, along with others, including governments, universities, and/or investors) to solve a business and human rights problem that no actor can solve alone.").

Can the problem only be properly addressed through government regulation? Is the problem already being solved by another MSI that the organisers could join instead of setting up a new MSI? Stakeholders frequently turn to MSIs to address complex and intractable problems, but sometimes additional MSIs can undermine an organiser's objective, by fracturing stakeholders or interfering with ongoing processes. Understanding the full landscape before proceeding will be helpful to defining the scope of a new MSI. Organisers should ensure they are bringing a new solution to the defined problem.

1(c). Question whether an MSI is the Best Approach

After defining the issue and mapping the landscape, organisers should decide whether a new MSI is the best approach to solve the problem. A multistakeholder approach is optimal "where:

- decisions impact a wide and distributed range of people and interests
- there are overlapping rights and responsibilities across sectors and borders
- different forms of expertise are needed, such as technical expertise
- the legitimacy and acceptance of decisions directly impact implementation."[320]

2. Step 2: Establishing the Objectives and Functions of the MSI

2(a). Set Objectives

Once organisers have decided that a new MSI is the best solution to the problem, they should establish primary objectives and set goals for the new initiative. Clear objectives for the group will ensure everyone is working towards a common purpose and align stakeholders when the group is inevitably pulled in different directions.[321] The established objective can become the mission statement of the group, which will also help clarify which stakeholders can or should be involved in the MSI.

2(b). Determine the Function

MSIs can perform a wide variety of functions, which may evolve over time, depending on strengths and opportunities. Each MSI has a different approach, encompassing a range of procedures, formats, resolution mechanisms, and outcomes, and MSIs are constantly adapting to meet new challenges.[322]

---

[320] Internet Society, *Internet Governance: Why the Multistakeholder Approach Works, supra* note 215 at 2 ("The multistakeholder approach allows us to protect and further develop the complex systems we rely on while allowing those systems to go on working.").

[321] Internet Society*, Internet Governance: Why the Multistakeholder Approach Works, supra* note 215 at 5; *noting* in setting up a multistakeholder approach, organisers should ask ("Before the substantive discussions begin, does everyone agree on shared goals to guide the process and ensure the core questions are not debated multiple times? Is it clear from the outset – when shared goals are defined – that an outcome can feasibly be implemented by all relevant stakeholders? Is there a common understanding across stakeholder groups about how decisions will be made? Has everything been done to ensure that those who operate the infrastructure or are most affected by this decision have been part of making it? Has the process been sufficiently inclusive and transparent to maximise the ease of implementation?).

[322] Strickling & Hill, *supra* note 205 at 46 ("Multi-stakeholder approaches are just that, approaches. They encompass a range of procedures, formats, resolution mechanisms and outcomes. In the same way that democratic governments may follow a parliamentary or a presidential system of governance, so too do multi-stakeholder approaches vary and adapt to fit the particular governance question at hand. Some models lead to decisions while others are merely consultative. Some have established membership rules and criteria, while others allow anyone to participate. Some models are intended to last decades, while others are one-off processes designed to address a specific challenge of the day.").

Some organisations may serve multiple purposes, and functions may change over time. However, generally, there are three overarching approaches that most MSIs take:

- a standard-setting initiative, to create industry standards or offer accreditation, certification, or verification
- a policy-oriented initiative, to collaborate on best practices, define problems, or create a mutual understanding between parties
- a project-oriented initiative, set up to accomplish a goal, raise awareness, distribute funding, implement capacity building, or bring parties together.[323]

3. Step 3: Deciding Who is a Stakeholder

In an egalitarian or consultative MSI, anyone with a "stake" in the outcome proposed by an MSI could be included in the MSI. For a curated MSI, organisers should try to include a diverse set of stakeholders that are necessary to garner support and increase buy-in of solutions. However, curated MSIs organisers do not need to create an overly expansive group at the cost of reaching meaningful consensus on solutions.[324] To increase the success of the MSI and build legitimacy, a curated MSI should establish criteria to join, a process to include all relevant stakeholders, and mechanisms to ensure meaningful representation.

3(a). Criteria to Join

The organiser of a curated MSI must set parameters and act as a gatekeeper to ensure that stakeholders are willing to abide by group decisions and contribute to the organisation in a meaningful way.[325] To build legitimacy for the MSI, it is important that the organisation clearly defines who can and cannot join. This may require a different set of criteria for different types of stakeholders. For example, a business might need to have a threshold market share in the industry, while a civil society organisation might be required to provide transparency on their funding source. The rules need to be clearly stated

[323] Gleckman, *supra* note 209 ("Multistakeholderism as a form of global governance has a number of different structures. MSG groups tend to fall into three different categories depending on the governance gaps they are designed to address and their self-defined scope of activity… Policy-oriented multistakeholder governance groups, Product and process-oriented standard-setting multistakeholder groups, Project-oriented multistakeholder groups); *see also* Baumann-Pauly et al., *supra* note 32 at 3 ("While the current landscape of MSIs is at this point not comprehensively tracked or documented… MSIs differ on the functions they perform: Some MSIs mainly focus on fostering dialogue among stakeholders (e.g., the Ethical Trading Initiative) while others create standards and come up with mechanisms to enforce them. Some MSIs have certification schemes (e.g., Rainforest Alliance); others accredit the management systems of participants and verify remedial efforts at suppliers (e.g., Fair Labor Association."); *see also* Mariette van Huijstee, *Multi-stakeholder initiatives: a strategic guide for civil society organizations*, CENTRE FOR RESEARCH ON MULTINATIONAL CORPORATIONS, 17 (March 2012), https://www.somo.nl/wp-content/uploads/2012/03/Multi-stakeholder-initiatives.pdf.

[324] The Stanley Foundation, *Policy Dialogue Brief: Multistakeholder Coalitions Innovating or Complicating Global Governance?*, THE STANLEY FOUNDATION, NEW AMERICA, THE STIMSON CENTER (May 2016), https://stanleycenter.org/publications/pdb/MultistakeholderPDB816.pdf ("Broad and diverse membership helps to garner widespread support and global buy-in as well as engage those who have the power to affect change at any number of levels. Yet a diversity of actors can also mean conflicting interests or perspectives, and the resulting bureaucracy and decision-making processes are often slow and drawn out.").

[325] Gleckman, *supra* note 209 ("In a multilateral world, the role of gatekeeper and guard of legitimacy in global affairs is performed by the nation-state, working with other nation-states directly or collectively through the UN system. In a multistakeholder world, the convener of an MSG group takes on this gatekeeping function. The convener provides the leadership in choosing the appropriate collection of multistakeholder categories that are required for a new group.").

and upheld, to maintain cohesion and protect the process. These criteria will also be necessary in case the curated MSI decides to remove a stakeholder in the future.

3(b). Ensuring Inclusivity

Once an organiser has set parameters to participate, it should ensure that as many relevant stakeholders as possible are included in the process.[326] This can be achieved through a variety of outreach strategies, such as holding public meetings, creating online forums through which stakeholders may sign up, or hiring recruiters. However, the best method will be the time-intensive process of individual outreach by the organisers to key stakeholders. This individual outreach is likely to pay off in the long term, as it will build trust in a process that is often fraught with mistrust. The organisers should be aware of some of the common tensions that could lead to mistrust. First, companies within industries can be fiercely competitive with each other, and may not want to collaborate with their rivals. Second, civil society may mistrust industries or governments, who they frequently seek to hold accountable. Third, civil society may see joining an MSI as compromising their values. Finally, governments may not want to join an MSI with other governments, to preserve negotiating power in other diplomatic areas. Organisers must overcome these preconceived notions and foster an inclusive process to build legitimacy for the MSI.[327] One organisers can build a culture of inclusivity is to provide training to stakeholders on how to effectively participate in an MSI, which could increase the likelihood of building consensus.[328]

3(c). Meaningful Representation

Once organisers have identified key stakeholders, it is important to ensure there is meaningful and diverse representation. Unbalanced representation undermines the work of the MSI and will ultimately decrease the legitimacy of the work accomplished. For example, over the years, many MSIs have been set up in the Global North to tackle problems in the Global South without much representation or buy-in from the stakeholders in those areas.[329] Without key stakeholder buy-in, MSIs can exacerbate the

---

[326] Strickling & Hill, *supra* note 205 at 48 ("More than a well-regarded convening agency, more even than a history of sound practices, the legitimacy of any process derives from its openness to any participant, its conscious inclusion of a diversity of stakeholders and its commitment to reaching decisions by consensus.").
[327] Internet Society, *Internet Governance: Why the Multistakeholder Approach Works, supra* note 215 at 5 ("Inclusiveness is the basis of legitimacy in collaborative decision-making. Those significantly affected by a decision should have the chance to be involved in making it. Inclusiveness is not just an admirable goal, but an essential part of an effective process. The less inclusive a process is, the less likely it is to engender the trust and support of those outside of the process.").
[328] Lawrence E. Strickling & Grace M. Abuhamad, *The Feasibility of Expanding the Use of Multistakeholder Approaches for Internet Governance: Final Report to the Internet Society*, INTERNET SOCIETY, 7 (26 Oct. 2017), https://www.internetsociety.org/wp-content/uploads/2018/01/Feasibility-Study-Final-Report-Oct-2017.pdf ("Providing training on how to be an effective participant in multistakeholder processes serves two important functions. First, the presence of trained stakeholders improves the efficiency of multistakeholder convenings and increases the likelihood that the convening will result in a successful consensus outcome. Second, providing training around the globe will help educate stakeholders about the multistakeholder approach and will give them the skills and confidence to organize their own multistakeholder discussions in their locality or region. Given these benefits, our interviewees enthusiastically supported this activity for the initiative.").
[329] van Huijstee, *supra* note 323 at 41 ("MSIs mainly originate from the Northern hemisphere, even though they are almost exclusively targeted at solving issues in the South. Southern CSOs and Southern businesses often have more distrust towards each other than their Northern counterparts, Southern CSOs are often more directly confronted with the negative impacts of corporate conduct, and may even have experienced corporate corruption and repression. This may make these organisations wary and sceptical of collaboration with business through MSIs in the first place… . It is advisable for Northern CSOs to consult with Southern partners when designing and implementing an MSI."); *see also* Gleckman, *supra* note 209 ("Without a UN system-style rulebook or a governing convention, there is no mechanism to ensure that the convener's selection of stakeholder members reflects geographic balance, gender balance, worker inclusion, the communities most likely to be affected by the actions of the group, or the participation of organisations and individuals at the bottom end of the global social structure.

problems. Organisers must ensure a meaningful mix of participants who can speak to the problem firsthand. In many cases this means that organisers need to provide stakeholders with limited resources the funding they require to engage with the work of the MSI.[330] Additionally, it is important to consider how to best integrate new entrants from civil society, small companies, and governments of developing countries, who may struggle with time commitments and expertise, but have important perspectives to share.[331]

4. Step 4: Setting up Terms of Reference

Once the problem has been defined and the group of stakeholders brought together, the organisers must set up the foundations of how the curated MSI will operate: the terms of reference.[332] The terms of reference will serve as the roadmap for how the MSI implements its proposed solutions. While terms of reference can be legally binding, they can also be a simple agreed-upon set of principles for how the MSI will operate which is enforceable only as a matter of practice. As no two MSIs are the same, there is wide variety in terms of reference, ranging from just a few sentences to hundreds of pages of legalese. Neither approach is wrong, but there are five key decisions that need to be codified in the terms of reference to ensure success: Who sets the agenda? How are decisions made? What does transparency look like? What accountability mechanisms are in place? What funding is required?

4(a). Leadership and Agenda Setting

To ensure longevity and sustainability, an MSI needs to appoint someone to keep the work of the MSI on track. The terms of reference must designate the person, or group of people, who will make those calls. Once the leadership is determined, the terms of reference must address how the group will set the agenda for meetings, assign tasks to members, and divide up the workload.[333] Many MSIs will appoint an executive director or a board of directors to oversee the work and ensure commitments are upheld. However, the nature of multistakeholder work lends itself to a wide range of potential structures,

Designating a person to participate in an MSG group from one of the under-represented communities does not solve the challenge. Having one token woman or one member of the global southern community among 12 men or an equal number of international experts can easily mean that, while there is a minimum presence, that voice is dwarfed by the unbalanced weight of an entire MSG group.").

[330] Strickling & Hill, *supra* note 205 at 50 ("Multi-stakeholder processes are generally quite resource-intensive, both in terms of time and money. A single initiative focusing on a specific policy issue can take months from start to completion. Many multistakeholder organizations hold multiple meetings a year, often in far-flung places across the globe. For stakeholders with limited resources, in-person attendance can be prohibitively expensive. While most venues try to provide remote participation opportunities for stakeholders who are unable to travel, there is a sense that stakeholders who participate in person can have more impact on the group decision than those who engage remotely.").

[331] Strickling & Abuhamad, *supra* note 328 at 10-11 ("New entrants often lack these competences, and as a result, their views are less likely to be incorporated into the group's decision-making. This handicap, combined with resource constraints, is one of the primary reasons why stakeholders from the developing world are so often frustrated by the approach. They wish to contribute, and they have important interests to advance, but may find that they lack the technical and institutional knowledge, and/or longstanding personal and trust relationships needed to have one's views incorporated into the group's consensus. We believe, and our interviewees confirmed, that the inclusion of underrepresented groups is critical to the success of multistakeholder governance, and this initiative, if it goes forward, will need to address this disparity between well and poorly-resourced stakeholders.").

[332] Gleckman, *supra* note 209 ("Questions on internal governance: Are the terms of reference for the group clear and acceptable? How should decision-making operate within the group? What should be the dispute resolution system for the group? Where are the resources coming from to fund the group? Where are the resources coming from to finance the programmes and recommendations of the group?").

[333] Gleckman, *supra* note 209 ("The convenor will have a lead role in designating the individuals and organisations assigned to 'represent' categories of stakeholders, and are more likely to designate individuals and organisations that are supportive of the convener's preference regarding the expected outcome of the group.").

depending on the goals of the group. Some MSIs will be more stakeholder-driven (involving the entire group in decisions) and others will appoint a designated leader to set the agenda and next steps.[334]

4(b). Decision-making Authority

Consensus-based decision-making can help build trust, promote cooperation, and neutralise power differences between stakeholders.[335] However, consensus-based decision-making requires dedication from both the organisers and stakeholders to reach a compromise, often reached after lengthy and contentious negotiations.[336] For a curated MSI, the terms of reference must address the process by which consensus-based decisions are reached. Some curated MSIs have a voting structure; decisions are made by a simple majority vote. Others have an informal process where a decision is reached when the solution benefits the greatest number of stakeholders. For a curated MSI, the terms of reference should ensure that all stakeholders are heard and clearly define the decision-making process.

4(c). Transparency

The MSI's terms of reference should address how the organisation will provide transparency internally, between the organisation's stakeholders, and externally to the public.[337] Transparency will ultimately build accountability and contribute to broader acceptance of the solutions the MSI proposes.[338]

- Internal transparency is necessary to build trust and cohesion among stakeholders. Some secrecy can be necessary to make progress towards a solution, but this can undermine broader goals if it is not disclosed. Some stakeholders may be reluctant to provide external transparency, but maintaining internal transparency will create a collaborative environment.[339] Terms of reference should foster internal transparency by ensuring:

---

[334] Strickling & Hill, *supra* note 205 at 51; *noting* that one opinion on what is an MSI relates to how the agenda is set ("It is proposed that an "authentic" multi-stakeholder process display the following attributes: stakeholder-driven: stakeholders determine the process and decisions, from agenda setting to workflow, rather than simply fulfilling an advisory role.").

[335] Internet Society, *Internet Governance: Why the Multistakeholder Approach Works, supra* note 215 at 1.

[336] Strickling & Hill, *supra* note 205 at 49 ("Consensus decision making requires parties to persuade one another of the merits of their position. In consensus decision making, participants must compromise if they are to accomplish anything; they must ultimately either persuade, or be persuaded by, the other participants, at least insofar as it is necessary to achieve the required consensus. What is consensus? A standard of unanimity is nearly impossible to achieve. If the standard is not unanimity, how should it be defined and who sets the standard? There is no one standard that works for every situation, but many convenings have found that a standard of "can you live with it?" works well. Perhaps the best solution is to leave the definition of consensus to the participating stakeholders in the process.").

[337] van Huijstee, *supra* note 323 at 57-58 ("An important condition for the smooth functioning of an MSI is the formalisation of how the parties will interact with each other… One element that needs to be addressed in this regard is access to information. A problem that is often experienced by CSOs that participate in MSIs is that they are not so well informed as their business counterparts. In some cases, this may be caused by lack of capacity, but in many cases, the information that is crucial for judging a situation is 'owned' by the company (eg. information on their supplier base). For CSOs to represent their beneficiaries effectively, it is crucial to make clear arrangements with regard to access to information.").

[338] The Stanley Center, *supra* note 324 at 6 ("Because the legitimacy of multistakeholderism hinges on accountability, which is often internally enforced, external transparency is critical. Only through transparent financial and administrative operations can multistakeholder coalitions prove their internal accountability and thereby solidify their legitimacy in the eyes of the external community. In some instances, transparency and accountability can be achieved through third party monitors, but this tool will not be an appropriate approach in all cases of multistakeholderism.").

[339] Baumann-Pauly et al., *supra* note 32 at 22 ("Conflict in those early negotiations arose particularly over the level of transparency that the initiatives would require. Accountability required the review of corporate conduct against a substantive standard and reasonable levels of disclosure, but this was not an easy sell, particularly for the companies involved... With strengthened levels of trust in the organization, the participating stakeholders realized that greater transparency could lend their efforts greater credibility.").

- access to all working documents stakeholders have created
- notifications and/or invitations to all stakeholders when substantive meetings are happening
- opportunities for stakeholders to contribute to the work product before finalising the decision-making process.

- External transparency builds MSIs' broader legitimacy and accountability. However, the organisation does not need to disclose every contribution of every stakeholder.[340] Ensuring some privacy for stakeholders can encourage difficult conversations and prevent grand standing by individual stakeholders. Ideally, frank internal discussions will allow the group to reach consensus and to speak to external parties with one voice. The work products an MSI should externally disclose include:
  - organisational and decision-making processes
  - summaries of stakeholder inputs to demonstrate all voices were heard
  - the outcomes and/or next steps determined by stakeholders.

4(d). Accountability

The terms of reference must establish mechanisms by which to hold stakeholders accountable. Some MSIs may have dozens of outputs, and others may just have one meeting – either way, it is important to define how, and to whom, the MSI will be held accountable to produce results. Formal accountability structures can help assess each stakeholder's commitment or dedication to the MSI, which will help guide the direction of the organisation.[341] Finally, without accountability mechanisms, MSIs are susceptible to mission creep, which can stretch an organisation so thin that it can no longer pursue its goals.[342] Defining accountability mechanisms in the terms of reference from the beginning will build trust and legitimacy in the long term.

4(e). Funding

All MSIs will require resources, whether in the form of money, in-kind contributions, or participant time. The terms of reference should set out how the organisation will receive, distribute, and disclose funding sources. Some MSIs are set up as a purely voluntary organisation, where all stakeholders cover their own expenses and donate their time. However, the majority of curated MSIs will need some form

[340] *Id.* at 23 ("Transparency requirements have a history of reluctant acceptance by corporations, but when corporations assume a role of both regulator and 'regulatee', the transparency of standards, monitoring results and governance structure are key in helping to boost the legitimacy quota of the regulatory initiative. However transparency, like the MSIs themselves, is most effective when targeted.").
[341] The Stanley Center, *supra* note 324 at 3 ("the legitimacy of cooperative multistakeholder action is directly tied to accountability. Considering that multistakeholder coalitions frequently fill governance gaps, their own accountability is most often internally enforced. Answers to questions such as "Who is in charge?" and "How are decisions made?" can be unclear or even unknown. Without formal monitoring structures in place to inform stakeholders and the external community of how any given multistakeholder coalition operates, room remains for corruption, nepotism, and manipulation. This lack of accountability can hamper a multistakeholder coalition's internal operations and external legitimacy.").
[342] Kim Jonker & William F. Meehan III, *Mission Matters Most*, STANFORD SOCIAL INNOVATION REVIEW (19 Feb. 2014), https://ssir.org/articles/entry/mission_matters_most, *noting* that accountability stems from a ("strong mission statement that reflects an organization's true mission is the first, best tool to ensure that an organization will resist mission creep.").

of funding to cover expenses such as travel and events costs, translation services, or salaried staff. Who funds the MSI is often tied to legitimacy and efficacy; therefore, funding decisions should be considered in the terms of reference.[343] Ideally, the MSI would diversify its funding sources to avoid reliance on one stakeholder or group of stakeholders.[344] The primary ways MSIs are funded include:

- *Charitable donations*. Many MSIs are funded by charities who believe the mission of the MSI aligns with their own charitable goals and objectives. Charitable donations to an MSI can be used to fund research, support advocacy efforts, or build new partnerships.

- *Corporate partnerships.* For industry based MSIs, corporate sponsorship can signify that the businesses have a stake in the outcome and are "bought in" on the solution. However, in this case the MSI should work to diversify contributors within the industry and not receive all funding from a small group of companies. The perception that an MSI is beholden to a small subset of companies can undermine its legitimacy or cause it to be labelled a "shill" for one group at the expense of another.

- *Government grants*. Typically, MSIs are created because governments are unable to provide a regulatory solution; however, there are a lot of instances in which governments wish to support public-private partnerships and MSIs. If government funding is provided, the MSI should consider if this funding is tied to political outcomes and ensure that it is not over-exposed to political shifts.

- *Membership-based or event-based funding*. Membership or event funding can spread the cost of an MSI between stakeholders and can provide something of value to participating stakeholders (e.g., networking opportunities or professional credential). MSIs with this type of funding source should consider whether the costs are a barrier to entry for stakeholders who may be necessary to implement the proposed solution but unable to afford the up-front cost. When possible, the MSI should consider scholarships, grants, or membership costs based on resource constraints.

5. Step 5: Sustaining Forward Momentum

As discussed in Part I(B), many curated MSIs are formed in the aftermath of a devastating event or to address rapid technological developments. When this type of MSIs starts out, there is a shared sense of purpose and an urgency to "get things done". Governments will want to show their constituents they are addressing the problem, civil society will want to justify their work to their funders, and businesses will expect an outcome they can report back to shareholders in the form of a public relations boost or

[343] The Stanley Center, *supra* note 324 at 5 ("funding, which is often tied to limited political will, is an issue that cooperative multistakeholder action also grapples with, since funding is a crucial determinant of these coalitions' sustainability. Overall, it is still unclear what roles governments, businesses, foundations, and civil society organizations ought to play in forming and maintaining these multistakeholder coalitions and which issue types are more appropriate for each body. Determining who should be the lead or primary funder is a question that is often tied to legitimacy and efficacy; government funding can serve as the seal of approval for a coalition, but if a particular initiative targets business behavior, it is more likely to be successful if it has buy-in from the business community. Some entities tend to prefer to fund more-salient issues, while others are more open to obscure ones.").

[344] van Huijstee, *supra* note 323 at 31 ("it is advisable to have a certain degree of differentiated funding and not be completely dependent on participating companies. This is because, once the organisation is completely dependent on corporate participant funding, it gets an interest in maintaining this funding base, making it harder to force companies out of the initiative when they do not deliver. Some share of public funding thus seems a good alternative, although complete dependence on such funding may lead to other problems, most importantly lack of ownership by the parties it wishes to involve.").

cost-saving mechanism. A sense of urgency can indeed be extraordinarily helpful to pressuring stakeholders to find consensus before too many parties lose interest in the process.[345] Ultimately, the MSI must prove that it can get results, or stakeholders will start to leave. To sustain forward momentum, MSIs need to prioritise maintaining trust, documenting results, staying relevant, and fighting stakeholder burnout.

5(a). Maintaining Trust between Stakeholders[346]

For a curated MSI, selecting the right group of stakeholders and backing up the process with strong terms of reference can help build trust early on.[347] To maintain the trust, it is crucial for organisers to ensure all voices are heard, represented fairly, and dissenting opinions are acknowledged respectfully.[348] While it is natural for factions to form among stakeholders, within a curated MSI, it is important to balance competing priorities and acknowledge underlying motivations to maintain trust.[349] One way to ensure open communication is to create a small amount of bureaucracy to document and share communications with all stakeholders. This bureaucracy will ensure that decisions are made transparently and not overturned later.[350]

5(b). Delivering and Documenting Results

Curated MSIs are frequently set up to address complex problems, which means it is unlikely that they will quickly solve the underlying issue. Therefore, MSIs should demonstrate that they are delivering incremental results, to build momentum. In many cases, the MSI will need to break down the steps they are taking to accomplish the stated goal and continuously report on progress. Measuring the impact of "norm creation" or "policy reform" is not as straightforward as measuring economic or even social impact. Therefore, MSIs should systematically catalog any incremental results, to better articulate the impact of their achievements.[351]

---

[345] Strickling & Abuhamad, *supra* note 328 at 5 ("On this issue of timeliness, interviewees acknowledged that for the discussions to be successful, they need to take up issues for which stakeholders feel an urgency to find consensus solutions now, not later. Absent a sense of urgency, stakeholders may lose interest in the process and fail to put in the sustained effort necessary to reach consensus outcomes.").

[346] George Schultz, *The 10 most important things I've learned about trust over my 100 years*, WASH. POST (Dec. 2020) ("I'm struck that there is one lesson I learned early and then relearned over and over: Trust is the coin of the realm. When trust was in the room, whatever room that was – the family room, the schoolroom, the locker room, the office room, the government room or the military room – good things happened. When trust was not in the room, good things did not happen. Everything else is details.").

[347] The Stanley Center, *supra* note 324 at 6 ("There is a period of trust building that prepares the ground for ensuing conversations and negotiations between these actors with competing interests. As a natural consequence of competing interests, stakeholders tend not to be in agreement at first. They have to break down stereotyping barriers and be open to listening to ideas that they might not agree with. After some level of mutual trust is established, a productive and safe environment where disagreement can take place is built.").

[348] *Id*. at 5 ("Successful multistakeholder coalitions create room for dissent and disagreement.").

[349] *Id.* ("The fostering of stakeholders' understanding of each others' roles is a crucial component of success. Clear delimitations of actors' goals ensures that even when certain actors appear resistant or inflexible, all interlocutors are treated as legitimate ones. The development of interpersonal and intercultural communication skills works best if it is framed as an opportunity for members to pick up new skills rather than as a punitive exercise.").

[350] *See* van Huijstee *supra* note 323 at 56 for a detailed checklist on processes and procedures to build trust between stakeholders.

[351] The Stanley Center, *supra* note 324 at 5 ("Monitoring and evaluating cooperative multistakeholder action poses a challenge. Measuring the impact of "norm creation" or "policy reform" is not as straightforward as measuring economic or even social impact at the local level. As one participant stated, "It is difficult to measure the dog that doesn't bark." Multistakeholder coalitions should systematically catalog changes in policy and practice, such as changes in military manuals or city bylaws, in order to better articulate the impact of their achievements. …Multistakeholder coalitions should systematically catalog changes in policy and practice in order to better articulate and track the impact of their achievements.").

Once the results have been measured or cataloged, the MSI should share results with all relevant stakeholders. This may seem obvious, but many curated MSIs make significant progress without communicating their success to the public. Documenting accomplishments on a website, in newsletters, and to the media is as crucial as delivering the results themselves. Creating awareness of an MSI's accomplishments contributes to the "Tinkerbell effect" – where a belief or idea becomes real or gains power because people believe in it or act as if it is real.[352] The Tinkerbell effect suggests that if individuals or groups believe in the power and effectiveness of a coalition, it is more likely to succeed. When people have confidence in the collective efforts of a coalition, they are more likely to contribute their resources, time, and support to achieve the coalition's goals. This is particularly true for a curated MSI, where the impact of the results can be arbitrary and hard to measure.

5(c). Staying Relevant

As an MSI matures and becomes part of a broader conversation or solution, it is important that it maintains its relevance. To do this, the MSI must identify opportunities to showcase the work the group is doing and stay part of the ongoing conversation in the targeted industry. This can mean building new partnerships, showcasing work at industry conferences, or inserting the MSI's work into an ongoing media narrative. To build legitimacy for the MSI in the eyes of external actors, organisers need to advocate for their solution with a broader audience. However, in the pursuit of relevance, an MSI should not neglect its core work. Finally, if the MSI is going to take on new work with stakeholders or partners to stay relevant, all stakeholders should have a say in that new direction. In a curated MSI, without a broader consensus, organisers risk losing trust between stakeholders.

5(d). Motivating Stakeholders and Fighting Burnout

Fatigue is inevitable in all MSIs, but it can be particularly acute when a curated MSI is created following a tragic event, because stakeholders work so hard to address a problem. It is important for MSIs to find ways to keep stakeholders motivated and fight burnout. An MSI should look to motivate different stakeholders in the ways that fit their organisations best.[353] For government stakeholders, the MSI should prioritise the importance of its work to key policy goals for elected officials. This will ensure stakeholders are not pulled into other projects as media attention shifts or election cycles reshape priorities. For company stakeholders, the MSI should prioritise how its work returns values to shareholders. MSIs are typically viewed as a cost-centre; therefore, organisers should highlight public relations benefits, added efficiencies from multistakeholderism, and the likely expense of the regulatory process by comparison. For civil society, the MSI should find ways to meaningfully give stakeholders a seat at the decision-making table. Civil society resources are always scarce, so it is helpful for organisers to show how the work of stakeholders meaningfully drives change.

The best way to fight burnout and ensure longevity is to consistently bring new stakeholders into the process. A fresh set of eyes and new ideas will drive success and lessen the burden for the original

[352] *Id*. ("If a multistakeholder coalition is proving effective in advocating its agenda or governing its target area, this builds its legitimacy in the eyes of external actors by default. Ironically, while proven effectiveness can often boost coalition legitimacy, most collective multistakeholder action requires a certain amount of external legitimacy in order to be effective. This paradox demonstrates the delicate and complex balance multistakeholders must strike.").

[353] van Huijstee, *supra* note 323 at 24; *noting* stakeholder motivations for remaining in an MSI ("It is crucial to be aware of the power dynamics and motivations and interests in the value chain to be able to make the judgement about whether and how an MSI can be useful, and how it should be organised. As a general rule, one should keep in mind that an incentive needs to be created for [stakeholders] to participate in the MSI and make it an attractive alternative compared to 'business as usual'.").

stakeholders.[354] This means an MSI must have an ongoing recruitment and on-boarding process, to ensure new stakeholders are integrated into the organisation's work seamlessly. Creating a recruitment committee will likely pay off in the long run, as stakeholder turnover is inevitable.

6. Step 6: Deciding When the Work is Done

It is important for the leadership of a curated MSI to have a clear idea of when and how the project will end. Ideally, all MSIs would conclude when their solution is implemented. In some cases, this is easily accomplished. However, in many MSIs, there is no clear end-date, because the organisation is addressing an intractable problem. Where this is the case, MSIs have a variety of options, including to shut down, to join forces with another MSI, and to pivot.

6(a). Shut Down

An MSI typically shuts down when the majority of stakeholders can no longer justify the financial cost or ideologically support the work of the initiative. In the best-case scenario, an MSI shuts down because stakeholders agree they have accomplished everything they could do to address the problem, and it dissolves amicably.[355] In the worst-case scenario, the stakeholders withdraw their support for the initiative after trust erodes, forcing a shut-down. The NETmundial Initiative is an example of a MSI that shut down because stakeholders were no longer able to ideologically support the initiative due to a lack of trust. [356] The NMI was created following the NETmundial conference when organisers tried to implement the work of the outcomes document produced. However, the NMI leaders gave themselves permanent positions within the organisation causing key stakeholders to withdraw their support from an MSI they saw as untrustworthy and lacking transparency.[357] No matter what the reason is for shutting down, MSIs should communicate publicly why the MSI was shuttered as it could help stakeholders carry forward solutions to the problem the MSI was trying to address.

6(b). Join Forces

In many cases, an MSI is not the only organisation working towards solving an identified problem. For example, in the internet governance space, the Internet Society and IETF joined forces in 1992 when the Internet Society provided a corporate home for the IETF.[358] The combination of these MSIs was necessary to streamline funding sources and coordinate activities to promote the overall goals of both

---

[354] Strickling & Hill, *supra* note 205 at 49-50.
[355] United Nations Environmental Development Forum, *Designing a Multi-stakeholder Process: Issues & Questions to be Addressed*, UNED FORUM (2000), http://www.earthsummit2002.org/msp/report/chapter9-1.html ("MS[I]s need clear goals, cut-off points and concrete, identifiable outcomes. Participants need to develop a sense of ownership not only of the process but also of an output that they feel comfortable promoting – a concrete set of suggestions, toolkits or subsequent agreed actions. Once the group agrees that this point is reached, an MS[I] should be closed.").
[356] *See* full discuss in Part I(B)(2) above.
[357] *See also* Brown & Esterhuysen, *supra* note 234; Internet Society, *Internet Society Statement on the NETmundial Initiative, Press Release, supra* note 241; Strickling, *supra* note 242; Jeremy Malcolm, *Internet Governance and the NETmundial Initiative: A Flawed Attempt at Turning Words into Action,* ELECTRONIC FRONTIER FOUNDATION (28 Aug. 2014), https://www.eff.org/deeplinks/2014/08/internet-governance-and-netmundial-initiative-flawed-attempt-turning-words-action.
[358] Internet Society, *About the Internet Engineering Task Force (IEFT),* INTERNET SOCIETY (2023), https://www.internetsociety.org/about-the-ietf/ ("The Internet Society and the IETF have long been aligned on the advocacy of multistakeholder processes and open standards. Since 1992, we have supported the IETF by providing a corporate home for the IETF LLC—the administrative entity of the IETF—and through annual financial contributions.").

organisations.[359] Joining forces can benefit many stakeholders where duplication of efforts was straining resources or creating conflicting outcomes.

6(c). Pivot

If an MSI realises their solution no longer adequately addresses the problem, it can decide to pivot the work to arrive at a better solution. A pivot could include reframing the discussion, repurposing the group to address a new problem, or shrinking the scope of the MSI.[360] For a curated MSI, a pivot should happen only when there is a clear consensus among stakeholders which is clearly communicated to all participants. Failure to bring all stakeholders along with the proposed changes can undermine the trust and support of the stakeholder community writ large. A pivot may also require additional resources and staffing, along with new terms of reference. If a curated MSI is considering a pivot, it should go back to Step 1 and follow each step. Many internet governance MSIs see pivoting as an essential way to live up to the Silicon Valley culture of "innovate or die".[361]

[359] Vint Cerf, *IETF and the Internet Society*, INTERNET SOCIETY (18 July 1995), https://www.internetsociety.org/internet/history-of-the-internet/ietf-internet-society/ ("In contemplation of the need for a mechanism for aggregating funding from many sources, it was proposed to form an Internet Society and to use its resources, in part, to provide funds in support of IETF. The plan was for the Society to engage in a variety of activities including conferences, workshops, and raise funds from industry and other institutional sources. It does so on an international basis, and acts as a neutral and internationally recognized body, devoted to the support of Internet administrative infrastructure, including, for example, IAB, IETF, IRTF and IANA.").
[360] The Stanley Center, *supra* note 324 at 4 ("Multistakeholder coalitions are created, renewed, reframed, merged, enlarged, or shrunk throughout the course of their lives, and these transitions often require different key stakeholders and activities. The types of actors and activities involved shift significantly from the agenda- and norm-setting phases to the implementation phase, and again into the monitoring and evaluation phases. Multistakeholder coalitions tackle each of these phases as they come, often reinventing the wheel and sacrificing stakeholder and donor momentum in order to do so.").
[361] Gary Shapiro, NINJA FUTURE: SECRETS TO SUCCESS IN THE NEW WORLD OF INNVOATION (31 Dec. 2018).

III. THE FUTURE OF THE CHRISTCHURCH CALL TO ACTION

This part will apply the lessons learned from multistakeholder governance to the Christchurch Call to Action. The first section examines New Zealand's unique history, global position, and political framework, which provide the foundation for a multistakeholder solution. It then discusses what happened on 15 March 2019, with an emphasis on how the Christchurch shooter exploited social media to amplify his terrorist attack. Next, this section catalogs the creation of the Call and provides an overview of the work it has done over the past four years. The second section provides an analysis of the progress of the Call, including its achievements and where the Call is still working to fulfill its commitments. The third section explores the future of the Call as it relates to GenAI. Finally, the last section applies the framework of how to create a curated MSI to the work of the Call to help it to achieve long-term sustainability and pivot towards new and emerging technologies – including GenAI.

A. History of the Christchurch Call to Action

1. New Zealand's History, Culture, and Place in the World

The foundations of the Call can be found in the unique cultural history, global position, and political framework of New Zealand. Some of the same reasons the New Zealand Government decided to pursue a multistakeholder solution – fairness, humility, inclusiveness, tolerance of diversity, and collective understandings – were reasons the attacker, an Australian, chose New Zealand as his target for the attack.

From its beginning, New Zealand understood its place in the world as a country distant from others, isolated by oceans, and yet strongly connected to global events. New Zealand was discovered by Polynesian navigators between 1200 and 1300 AD, Europeans did not settle on the island until the early 1800s.[362] In 1840, New Zealand became a British colony when a treaty was signed by representatives of Queen Victoria of England and more than 500 Māori chiefs, called Te Tiriti o Waitangi or the Treaty of Waitangi.[363] Unlike previous agreements made between colonial empires and indigenous populations, the Treaty of Waitangi was based on partnership and equal treatment, and required the Crown to act with the utmost good faith, fairly, and impartially.[364] As a Commonwealth country, New Zealand was English-speaking and "western", closely aligned with the United Kingdom and other British colonies. As a Pacific Island, New Zealand was closely aligned with other Polynesian countries, and has taken a leading role in many Asia-Pacific partnerships. In later years, New Zealand joined the United States, United Kingdom, Australia, and Canada as a Five Eyes partner.[365] However, more than shared cultural heritage or treaty, New Zealand aligns itself closely with countries who support global rules-based systems such as the Universal Declaration on Human Rights or the World Trade

[362] Te Ahukaramū Charles Royal, *'Māori', Te Ara - the Encyclopedia of New Zealand*, TE ARA GOVERNMENT OF NEW ZEALAND (8 Feb. 2005). https://teara.govt.nz/en/maori (New Zealand was the last large and livable place in the world settled by humans).
[363] *Id.* at chapter 3.
[364] Waitangi Tribunal, *Te Ropu Whakamana i te Triti o Waitangi,* WAITANGI TRIBUNAL (2020), https://www.waitangitribunal.govt.nz/publications-and-resources/waitangi-tribunal-reports/ngatiwai-mandate-inquiry/chapter-3/#.
[365] The National Counterintelligence and Security Center, *Five Eyes Intelligence Oversight and Review Council*, OFFICE OF THE DIRECTOR OF NATIONAL INTELLIGENCE (2023), https://www.dni.gov/index.php/ncsc-how-we-work/217-about/organization/icig-pages/2660-icig-fiorc.

Organisation.[366] This stems from fundamental ideas of fairness embraced by Kiwis as well as a realistic understanding that, for a small country, binding larger countries to rules-based diplomacy is the best way to ensure stability in the modern era.

Starting in the early 20th century, New Zealand gained a reputation as the "social laboratory" of the world for its ability to experiment with progressive policy ideas.[367] The New Zealand Government has been able to quickly enact bold policy ideas in part because it has a centralised unicameral government. Additionally, the small population of New Zealand is relatively cohesive, and trust in public officials ranks highest in the world alongside Denmark and Finland.[368] For example, these qualities were tested by New Zealand's response to fighting the COVID-19 pandemic where elected officials enforced strict lock-downs, but the country was able to contain the virus better than almost any country in the world.[369] The same qualities were also crucial in the response to the Christchurch attack; less than one month after the tragedy, the New Zealand Parliament voted 119-1 in favor of a gun reform bill that banned military-style, semi-automatic weapons.[370] In many policy areas, New Zealand Government officials are able to find creative solutions to large challenges because they have built a culture of trust and their government is much smaller than other democracies. Indeed, deciding to build a new MSI to address the problems with TVEC online is closely aligned with New Zealand's ability to experiment with policy ideas, bring together a diverse group of stakeholders, and create a rules-based system.

7. 15 March 2019

On the Friday morning of 15 March 2019, the Christchurch shooter drove from his home in Dunedin, New Zealand to Christchurch, a city with a small but growing Muslim population.[371] At 1.18 pm, the individual emailed his 74-page manifesto with details of the attack plans to dozens of government officials and media organisations.[372] Eight minutes later, at 1.26 pm, he updated his Facebook status with links to seven different file-sharing websites that contained copies of a manifesto he had written explaining his motivation.[373] He then posted to 8chan, an online message board frequently used by white supremacists, a link to his Facebook account with the message, "well lads, it's time to stop shitposting and time to make a real life effort post. I will carry out an attack on the invaders, and will even live

[366] David Hackett Fischer, FAIRNESS AND FREEDOM: A HISTORY OF TWO OPEN SOCIETIES: NEW ZEALAND AND THE UNITED STATES, 354 (10 Feb. 2012); *see also* Department of the Prime Minister and Cabinet, *Aotearoa's National Security Strategy* (12 Aug. 2022) (""New Zealand is a liberal democracy based on a bicultural relationship established in te Tiriti o Waitangi/Treaty of Waitangi. New Zealand upholds human rights and civil liberties, and is committed to being a good global citizen, supporting the rules-based system, and working with our international partners.").
[367] Hackett, *supra* note 366 at 340 ("In the period 1890–1920 New Zealand was at times regarded by foreign observers as a "social laboratory" in which new policy initiatives were being set in train.").
[368] Te Kawa Mataaho, Public Service Commission, *New Zealand Public Service among most trusted in the world*, NEW ZEALAND PUBLIC SERVICE COMMISSION (26 Jan. 2022),https://www.publicservice.govt.nz/news/new-zealand-public-service-among-most-trusted-in-the-world/ ("The survey found 81 percent of New Zealanders trust public services based on their personal experience, while trust in the Public Service brand sits at 62 percent – compared to 49 percent in the private sector.").
[369] Michael Baker & Nick Wilson, *New Zealand's Covid strategy was one of the world's most successful – what can we learn from it?,* THE GUARDIAN (5 April 2022), https://www.theguardian.com/world/commentisfree/2022/apr/05/new-zealands-covid-strategy-was-one-of-the-worlds-most-successful-what-can-it-learn-from-it.
[370] Graham Macklin, *The Christchurch Attacks: Livestream Terror in the Viral Video Age*, COMBATING TERRORISM CENTER AT WEST POINT VOLUME 12, ISSUE 6 (July 2019), https://ctc.westpoint.edu/christchurch-attacks-livestream-terror-viral-video-age/.
[371] Royal Commission of Inquiry into the Attack on Christchurch Mosques, *supra* note 5 at Volume 1 at 40.
[372] *Id.*
[373] *Id*. Volume 1 at 41.

stream the attack via facebook … I have provided links to my writings below, please do your part by spreading my message, making memes and shitposting as you usually do."[374]

At 1.33 pm, he linked the feed of the GoPro camera on his helmet to his mobile phone and started a Facebook livestream of the footage.[375] At 1.40 pm, he entered the Masjid an-Nur mosque and opened fire on the worshippers gathered for Friday prayers.[376] After completing the first attack, he went back to his car and drove to a second nearby location, the Linwood Islamic Centre, arriving there at 1.52 pm. There, he opened fire on worshippers again.[377] The individual got back in his car to attack a third location, the Al-Nur Early Childhood Education and Care Centre, but was arrested after two New Zealand police officers rammed his vehicle with their car.[378] In total, 51 people died and 40 people suffered gunshot injuries.[379] On 27 August 2020, the individual was sentenced to life imprisonment without parole for the murder of 51 individuals and designated as a terrorist entity under the Terrorism Suppression Act 2020.[380]

On Facebook, the live feed continued throughout the attack. It remained on the individual's page for another 12 minutes before Facebook was notified by police and removed the content.[381] The video of the attack was viewed 4,000 times before it was taken down by Facebook.[382] Responding to the call to action from the individual's message on 8chan, like-minded extremists copied and shared the footage across the internet on platforms such as Twitter, YouTube, and Reddit. In the first 24 hours after the attack, Facebook removed or blocked over 1.5 million uploads of the video.[383] As quickly as social media platforms could take down the content, it was re-uploaded, "sometimes spliced into new video clips, making it impossible to detect quickly."[384] The video was widely viewed across New Zealand as it reappeared on social media – sometimes promoted by the company's algorithms, which amplified trending content. While it is impossible to know how many New Zealanders saw the video, in the first week after the attacks 8,000 people who saw it called mental health support lines.[385]

As Kevin Roose of the *New York Times* noted, the Christchurch massacre:

> … felt like a first – an internet-native mass shooting, conceived and produced entirely within the irony-soaked discourse of modern extremism. The attack was teased on Twitter, announced on the online message board 8chan and broadcast live on Facebook. The footage was then replayed endlessly on YouTube, Twitter and Reddit, as the platforms scrambled to take down the clips nearly as fast as new copies popped up to replace them.[386]

Clips made their way to the mainstream platforms after spreading across smaller social media sites like 8chan, 4chan, Discord, and Gab. The individual, who was steeped in internet subcultures, carefully planned his attack to go viral on these sites by providing followers with many "in-joke" opportunities

[374] *Id*.
[375] *Id*. Volume 1 at 42.
[376] *Id*. Volume 1 at 43.
[377] *Id*.
[378] *Id*. Volume 1 at 45.
[379] *Id*. Volume 1 at 46.
[380] *Id*. Volume 1 at 47.
[381] Chris Sonderby, *Update on New Zealand*, FACEBOOK NEWSROOM (18 March 2019), https://about.fb.com/news/2019/03/update-on-new-zealand/.
[382] *Id*.
[383] *Id*.
[384] Royal Commission of Inquiry into the Attack on Christchurch Mosques, *supra* note 5, Volume 1 at 46.
[385] Ardern, *supra* note 6.
[386] Kevin Roose, *A Mass Murder of, and for, the Internet,* THE NEW YORK TIMES (15 March 2019), https://www.nytimes.com/2019/03/15/technology/facebook-youtube-christchurch-shooting.html.

to create memes. He had stated in his manifesto, "memes have done more for the ethnonationalist movement than any manifesto."[387] Without any meaningful content moderation or cooperation, the Christchurch manifesto and video content circulated freely on the smaller unmoderated platforms whose users quickly produced memes and spread the individual's message.

As the video and manifesto migrated from these smaller sites, the large platforms faced several challenges in the first 24 hours. First, companies frequently rely on "hash" technology to remove objectively awful content such as child sexual abuse material and ISIS beheading videos.[388] "Hashing" works by creating a digital fingerprint of the unique pixels of the image.[389] In the case of the Christchurch attack, extremists sympathetic to the shooter were slightly altering the video before uploading it, or creating memes, which evaded detection by the hash technology.[390] Second, graphic content is primarily removed from platforms using AI. However, in 2019, the online platform's AI tools could not identify first-person shooting videos as graphic content because no large dataset of videos existed to train the algorithm.[391] Third, removal efforts were made more difficult as clips of the video were included in reporting on mainstream media outlets. The media clips spliced the footage from the shooter making it impossible to effectively filter the video that included the more graphic scenes of the massacre. Eventually, YouTube stopped trying to differentiate between media footage and the massacre video and blocked all videos using the footage.[392]

After addressing the initial technical problems, in the days following the attack, tech industry leaders issued their well-rehearsed *mea culpa* about the proliferation of harmful content online. The tech executives acknowledged that they needed to improve structures to stop this type of event from happening again. Meta's chief operating officer, Sheryl Sandberg, said in a letter responding to the attacks, "many of you have also rightly questioned how online platforms such as Facebook were used to circulate horrific videos of the attack … We have heard feedback that we must do more – and we agree."[393] Microsoft's president, Brad Smith, stated, "it's clear that we need to learn from and take new action based on what happened in Christchurch."[394] YouTube's chief product officer, Neal Mohan, said, "this incident has shown that, especially in the case of more viral videos like this one, there's more work to be done."[395] While livestreaming a terrorist attack of this magnitude was unprecedented, the act itself was unfortunately all too common.[396]

---

[387] Macklin, *supra* note 370.
[388] Kate Klonick, *Inside the Team at Facebook that Dealt with the Christchurch Shooting*, THE NEW YORKER (25 April 2019), https://www.newyorker.com/news/news-desk/inside-the-team-at-facebook-that-dealt-with-the-christchurch-shooting.
[389] *See* Part I, Section 2 for more background on how hash-sharing technology works.
[390] Klonick, *supra* note 388.
[391] Sonderby, *supra* note 381; *see also* Evelyn Douek, *Australia's 'Abhorrent Violent Material' Law: Shouting 'Nerd Harder' and Drowning Out Speech*, 94 AUSTRALIAN LAW JOURNAL 41 (26 Aug. 2019).
[392] Alex Hern, *Facebook and YouTube Defend Response to Christchurch Videos*, THE GUARDIAN (19 March 2019), https://www.theguardian.com/world/2019/mar/19/facebook-and-youtube-defend-response-to-christchurch-videos.
[393] Julia Carrie Wong, *Facebook finally responds to New Zealand on Christchurch attack*, THE GUARDIAN (29 March 2019), https://www.theguardian.com/us-news/2019/mar/29/facebook-new-zealand-christchurch-attack-response.
[394] Brad Smith, *A tragedy that calls for more than words: The need for the tech sector to learn and act after events in New Zealand*, MICROSOFT (24 March 2019), https://blogs.microsoft.com/on-the-issues/2019/03/24/a-tragedy-that-calls-for-more-than-words-the-need-for-the-tech-sector-to-learn-and-act-after-events-in-new-zealand/.
[395] Elizabeth Dwoskin & Craig Timberg, *Inside YouTube's struggles to shut down video of the New Zealand shooting — and the humans who outsmarted its systems*, THE WASH. POST (18 March 2019), https://www.washingtonpost.com/technology/2019/03/18/inside-youtubes-struggles-shut-down-video-new-zealand-shooting-humans-who-outsmarted-its-systems/.
[396] A shooter in Cleveland had livestreamed his violent attack years earlier, *see* Jane Morice, *Facebook killer chooses victim at random, laughs about killing in videos*, CLEVELAND.COM (17 April 2017),

After years of promises from tech companies to clean up their platforms, these statements fell short of convincing lawmakers that they could solve the issue alone. The UK Home Secretary, Sajid Javid, said, "online platforms have a responsibility not to do the terrorists' work for them. This terrorist filmed his shooting with the intention of spreading his ideology. Tech companies must do more to stop his messages being broadcast on their platforms."[397] In the US, the Chairman of the House Homeland Security Committee, Bennie Thompson, told tech company executives during a congressional hearing weeks later, "your companies must prioritise responding to these toxic and violent ideologies with resources and attention. If you are unwilling to do so, Congress must consider policies to ensure that terrorist content is not distributed on your platforms."[398] In perhaps the most extreme response, just weeks after the attack, the Australian Government passed the Criminal Code Amendment (Sharing of Abhorrent Violent Material).[399] Without much debate or input from stakeholders, the Australian legislation created criminal and civil penalties for tech companies if users post abhorrent violent material, including the Christchurch video and manifesto.[400]

Like others, the New Zealand Government considered its options for how to move forward. Prime Minister Jacinda Ardern said in a speech to Parliament in the days after the attack, "we cannot simply sit back and accept that these platforms just exist and that what is said on them is not the responsibility of the place where they are published … They are the publisher, not just the postman. It cannot be a case of all profit, no responsibility."[401] In the aftermath of the event, Ardern would go on to say:

> I don't think anyone wants platitudes. We didn't want just a response to that individual act. If anything, we wanted to make sure that the pain and the horror of not just the act itself, but the fact that it was then broadcast, didn't occur elsewhere. Governments will not be able to regulate their way out of this problem. Tech companies, perhaps, if they continue to work on their own may not find solutions, but through collaboration together, I do believe we can make progress.[402]

8. The Creation of the Christchurch Call to Action

In the days following 15 March 2019, Prime Minister Ardern explored options to ensure this type of attack never happened again. Ardern and her team understood that the world's outrage would eventually dissipate, so they needed to use their moral authority to build an initiative that could have a lasting impact.[403] With this in mind, two weeks after the attacks, Ardern and her team met Microsoft President

https://www.cleveland.com/metro/2017/04/accused_facebook_live_killer_c.html.

[397] *See* Hern, *supra* note 392.

[398] Lauren Feiner, *House Homeland Security chair calls on Facebook, YouTube, Twitter and Microsoft to explain the spread of mosque shooting video*, CNBC (19 March 2019), https://www.cnbc.com/2019/03/19/rep-bennie-thompson-asks-tech-to-explain-mosque-shooting-video-spread.html.

[399] Criminal Code Amendment (Sharing of Abhorrent Violent Material) Act 2019 (Cth) s 474.31, AU; *see also* Douek, *supra* note 391.

[400] Douek, *supra* note 391 at 42 ("The Australian Parliament showed little regard for these complexities when it passed the Criminal Code Amendment (Sharing of Abhorrent Violent Material) Act 2019 (Cth) (AVM Act) on 4 April 2019. The Act made its way through both houses of Parliament in less than two days and came into force two days later. The government did not consult with experts, civil society or industry. Proudly touting the legislation as a "world first", the government did not stop to ask itself why that was.").

[401] Matt Novak, *New Zealand's Prime Minister Says Social Media Can't Be 'All Profit, No Responsibility'*, GIZMODO (20 March 2019), https://www.gizmodo.com.au/2019/03/new-zealands-prime-minister-says-social-media-cant-be-all-profit-no-responsibility/.

[402] Jacinda Ardern & Brad Smith, *Prime Minister Jacinda Ardern: Can we work together to end violent extremism online?*, TOOLS AND WEAPONS PODCAST WITH BRAD SMITH (6 July 2022), https://tools-and-weapons-with-brad-smith.simplecast.com/episodes/jacinda-ardern-can-tech-and-government-end-extremism-online/transcript.

[403] Smith and Browne, *supra* note 193 at 153.

Brad Smith, who happened to be in New Zealand for a visit, planned long before the attack.[404] After discussing multiple options, the teams sketched out the idea of a "Christchurch Call to Action" as an MSI. The solution was akin to the Paris Call for Trust and Security in Cyberspace – an MSI designed to protect international norms against cyber-attacks that had launched six months prior.[405] Due in part to France's successful leadership of the Paris Call, the New Zealand Government found a willing partner in President Macron to set up a similar initiative to address the problem of TVEC online. The timing was fortuitous, as France was set to host the Tech for Good Summit alongside the G7 Digital Ministers Meeting just a few weeks later. Organisers of the Call hoped to formally announce it at the Summit, and sign on other government leaders.[406]

To pull off a launch just six weeks later, the New Zealand Government and Microsoft teams worked around the clock to secure additional partners in the major platforms, including Google (and its subsidiary YouTube), Facebook, Twitter, and Amazon as well as two French companies, Dailymotion and Qwant.[407] The eight companies had very different platforms, business models, engineering capabilities, and experiences with TVEC online, but they were able to find commonality in wanting to prevent another Christchurch-type attack.[408] As a result, the companies worked with New Zealand and France to come up with nine steps they could take to address TVEC online. Five of these steps would be for individual companies to take: tighten their terms of service, better manage live videos, respond to user reports of abuse, improve technology controls, and publish transparency reports. Four of the steps were industry-wide: launch a crisis response protocol, develop open source-based technology, improve user education, and support additional research to prevent TVEC online.[409]

Throughout the initial creation of the Christchurch Call, civil society was skeptical that the initiative could produce any meaningful outcomes or commit to a human rights-respecting framework.[410] Behind this skepticism was a general frustration with the EU, which was in the process of passing a regulation on "Preventing the Dissemination of Terrorist Content Online". As discussed above, these regulations were incredibly controversial with many civil society organisations viewing them as threatening to free expression and human rights.[411] Additionally, the regulations came on the heels of the EU passing several laws regulating content moderation that were thought to be technologically unworkable, restrictive of human rights, ambiguously drafted and massively overreaching.[412] As a result, there was little trust between civil society organisations and European regulators, which meant that the inclusion of France as a co-lead for the Call raised concerns.

The hostility towards European regulators hung over the room as several civil society organisations met with the New Zealand Prime Minister the day before the Call was to be launched in Paris. At that meeting, civil society organisations presented a letter with input from dozens of signatories detailing

---

[404] *Id*.
[405] *Id*.
[406] The Christchurch Call to Action, *New Zealand and France Seek to End Use of Social Media for Acts of Terrorism*, CHRISTCHURCH CALL (24 April 2019), https://www.christchurchcall.com/media-and-resources/news-and-updates/nz-and-france-seek-to-end-use-of-social-media-for-acts-of-terrorism/.
[407] Smith and Browne, *supra* note 193 at 154.
[408] *Id*.
[409] *Id*.
[410] *See* Courtney Radsch, *Taking down terrorism online while preserving free expression*, MEDIUM (16 May 2019), https://medium.com/@old-cradsch/taking-down-terrorism-online-while-preserve-our-free-expression-678ab1100a67; *See also* Jillian C. York, *The Christchurch Call: The Good, the Not-So-Good, and the Ugly*, ELECTRONIC FRONTIER FOUNDATION (16 May 2019), https://www.eff.org/deeplinks/2019/05/christchurch-call-good-not-so-good-and-ugly.
[411] *See* Part I(A)(1)(c) related to the EU's TCO and discussion with civil society organizations happening in 2019. *See also* Citron, *supra* note 110 at 1038.
[412] Penfrat, *supra* note 122.

their concerns.[413] Included in the letter were complaints civil society had raised with European regulators in previous conversations, including the lack of clear definitions of "terrorism" and "violent extremism", the need to differentiate between social media companies and internet infrastructure providers, and the importance of governmental transparency around take-down requests.[414] Additionally, the letter expressed concern that civil society had been left out of the early stages of negotiations and a perceived lack of desire for meaningful input from civil society by governments.[415] The New Zealand and French teams worked closely with the group to resolve some of these issues, and won over a number of civil society representatives. In the end, all stakeholders pledged to work together to better incorporate civil society views into the text of the Call commitments themselves.

Civil society was not the only recalcitrant stakeholder in May 2019. Despite being the corporate home of most of the major tech platforms, the US Government declined to join the Call, stating it was "not in a position to join the endorsement" because of issues regarding the First Amendment.[416] However, the White House said it "stands with the international community in condemning terrorist and violent extremist content online" and supported the Call's goals.[417] In 2019, the Trump Administration had been focused on "political censorship" of speech by social media companies. As such, many conservatives saw the Call as a threat to free speech.[418] Behind the scenes, US Government officials stayed in touch with their New Zealand and French counterparts and did what they could to support the effort, consistent with US policy.[419] However, on the day the Call was signed, the White House announced the creation of a "tool" Americans could use to report if their speech was removed by social media companies due to "political bias".[420] Two years later, under the Biden Administration, the US Government joined the Call, noting that they would not take any action to "undermine the First Amendment".[421]

On 15 May 2019, just two months after the attacks, New Zealand and France formally announced the creation of the Christchurch Call to Action, a set of commitments by governments and online service providers to eliminate TVEC online while protecting the free, open, and secure internet. To stand up the organisation, the Governments of New Zealand and France formed the Call Secretariat, which would be staffed by government officials. The original Call text of 25 commitments was supported by 17

---

[413] Jillian C. York, *et al., Civil Society Positions on Christchurch Call Pledge*, ELECTRONIC FRONTIER FOUNDATION (May 14, 2019), https://www.eff.org/files/2019/05/16/community_input_on_christchurch_call.pdf, *noting*, letter includes signatures from 20 contributors to the document stating ("The document was created with input from dozens of members of civil society, including some who were in attendance on 14 May and some who were not. A non-exhaustive list of those individuals is at the bottom of this document.").

[414] *Id.*

[415] *Id.*

[416] Adam Satariano, *Trump Administration Balks at Global Pact to Crack Down on Online Extremism*, THE NEW YORK TIMES (15 May 2019), https://www.nytimes.com/2019/05/15/technology/christchurch-call-trump.

[417] Tony Romm, *Christchurch Call: White House won't sign deal to stop online extremism*, WASH POST (15 May 2019), https://www.washingtonpost.com/technology/2019/05/15/white-house-will-not-sign-christchurch-pact-stamp-out-online-extremism-amid-free-speech-concerns/.

[418] Charlie Warzel, *The World Wants to Fight Online Hate. Why Doesn't President Trump?,* THE NEW YORK TIMES (16 May 2019), https://www.nytimes.com/2019/05/16/opinion/christchurch-online-extremism-trump.html.

[419] As mentioned in the preface, I was one of these US government employees helping behind the scenes. This report does not contain any confidential information from my work with NTIA.

[420] *See* Warzel, *supra* note 418.

[421] Jen Psaki, *Statement by Press Secretary Jen Psaki on the Occasion of the United States Joining the Christchurch Call to Action to Eliminate Terrorist and Violent Extremist Content Online*, THE WHITE HOUSE (7 May 2021), https://www.whitehouse.gov/briefing-room/statements-releases/2021/05/07/statement-by-press-secretary-jen-psaki-on-the-occasion-of-the-united-states-joining-the-christchurch-call-to-action-to-eliminate-terrorist-and-violent-extremist-content-online/.

countries, the EU, and eight tech companies.[422] Of these commitments, five apply to only the governments, seven apply only to the tech companies, and the other 13 apply to both.[423] The commitments include developing tools to prevent the upload of TVEC, countering the drivers of violent extremism through education, increasing transparency around the removal and detection of content, and reviewing how companies' algorithms direct users to violent extremist content.[424] The commitments are careful to balance freedom of expression with the need for governments and companies to do more to counter extremism – both online and offline.

As a curated MSI, the text of the commitments is important for a few key reasons. First, the Call acknowledges that there are already several other forums addressing the issue of TVEC online, including multilateral efforts at the G7 and G20 and tech industry efforts such as the GIFCT and Tech Against Terrorism (TAT).[425] The drafters understood that this was not a new idea, but that it would be the first of its kind to bring a broader group of stakeholders together to address TVEC online – breaking down the silos of many of the other initiatives. Second, while the Call commitments include a provision to consider regulation, there is no commitment to impose new regulations on tech companies or law enforcement. This multistakeholder approach and stood in contrast to the discussions happening in some multilateral forums at the time. Third, civil society is not formally committed to the Call; instead, several of the commitments within the Call require governments and online service providers to work with civil society to promote community-led efforts. As such, the supporters commit to recognising the important role of civil society in offering advice and increasing transparency.

9. Overview of the Work of the Christchurch Call to Action

After the launch in May 2019, work on the Call steadily increased over the next few months, culminating in September that year, when leaders reconvened at the United Nations General Assembly in New York. At this meeting, the leaders acknowledged the progress that had been made towards fulfilling the Call commitments and welcomed 31 new countries and two international organisations as partners.[426] Among the Call's accomplishments was the establishment of a Christchurch Call Advisory Network (CCAN), to advise on the implementation of the Call.[427] CCAN was initially a group of 40 organisations, including representatives from civil society, human rights defenders, technical experts, and free speech advocates. In 2019, this group was formally recognised to provide expertise to the Call's government and company supporters on how they can fulfill the commitments in the Call.[428]

Other accomplishments coming out of the Call's 2019 Leaders' Summit were in relation to the GIFCT. By 2019, the GIFCT database included over 200,000 pieces of content, but there was still a strong focus on ISIS propaganda and beheading videos.[429] Despite the growth of the hash-sharing database, and the

[422] Christchurch Call to Action: News and Updates, *The Christchurch Call to Action: To Eliminate Terrorist and Violent Extremist Content Online*, CHRISTCHURCH CALL (15 May 2019), https://www.christchurchcall.com/media-and-resources/news-and-updates/christchurch-call-adopted/. The original supporters of the Call were: France, New Zealand, Canada, Indonesia, Ireland, Jordan, Norway, Senegal, the UK, and the European Commission as well as Amazon, Facebook, Dailymotion, Google, Microsoft, Qwant, Twitter, and YouTube, Australia, Germany, India, Italy, Japan, the Netherlands, Spain and Sweden.
[423] The Christchurch Call to Action: Full English Text*, supra* note 11.
[424] The Christchurch Call to Action: News and Updates, *supra* note 422.
[425] The Christchurch Call to Action: Full English Text*, supra* note 11.
[426] Christchurch Call to Action: News and Updates, *supra* note 422.
[427] *Id.*
[428] Christchurch Call Advisory Network, *About Us – History*, CHRISTCHURCH CALL ADVISORY NETWORK (2023), https://christchurchcall.network/about-us/history/.
[429] Global Internet Forum to Counter Terrorism, *Transparency Report – July 2019*, GLOBAL INTERNET FORUM TO COUNTER TERRORISM (July 2019), https://gifct.org/wp-content/uploads/2020/10/GIFCT-Transparency-Report-July-2019-Final.pdf ("At the end of 2018 the GIFCT gave itself the goal of reaching 200k hashes by the end of 2019.

inclusion of new social media companies, the GIFCT was not a standalone organisation. Instead, the founding member companies (Microsoft, Facebook, YouTube, and Twitter) rotated leadership each year, meaning processes were updated and staffed ad hoc by each company.[430] This had proved challenging on 15 March 2019, when the companies tried to quickly stop the spread of the Christchurch massacre video and manifesto. The GIFCT reported that it hashed more than 800 visually distinct versions of the video in the first 48 hours.[431] The attack highlighted the overall importance of this tool to the safety of billions of users around the world, and the Call's company supporters agreed to GIFCT reforms.

As part of their Call commitments, companies outlined five steps they would take as individual companies and four they would take as an industry.[432] The four industry commitments would largely be enacted through changes to the GIFCT.[433] At the Leaders' Summit in September 2019, the GIFCT announced the creation of a standalone organisation with a dedicated structure and staff, as well as the creation of working groups focused on research, algorithms, and information sharing.[434] Another important announcement at the Summit was the creation of a multistakeholder Independent Advisory Committee (IAC), which would include representatives from governments, civil society, and academia to guide the GIFCT Operating Board on organisational priorities.[435] Finally, the GIFCT and governments worked together to establish a "Content Incident Protocol" to provide a more systematic way of addressing terrorist content in the wake of an attack.[436] These changes were remarkable achievements for a multistakeholder institution to accomplish in just four months.

Going into 2020, the Call made progress on several other commitments, but COVID-19 slowed its momentum, as governments and tech companies needed to prioritise their responses to the pandemic. Therefore, the Call Secretariat set out to conduct a stock-taking exercise in 2020 with input from the Call community and publish the results at the Leaders' Summit in 2021. Given the wide range of efforts happening worldwide to reduce TVEC online, this was also an attempt to understand the landscape and assess where the Call could add value. On 14 April 2021, the Call published its first Christchurch Call Community Consultation Report.[437] The Call Secretariat sent out a questionnaire to all the signatories of the Call, as well as civil society organisations affiliated with CCAN. In total, there were 99 parties contacted, and 39 participated in the study, including 24 countries, six companies, and nine civil society organisations.[438] The overarching goal of the survey was to establish a baseline of progress to inform the future direction of the Call.[439]

We are pleased to say that the Hash Sharing Consortium has reached over 200k unique pieces of terrorist content. Companies often have slightly different definitions on "terrorism" and "terrorist content". For the purposes of the hash sharing database, and to find an agreed upon common ground, founding companies in 2017 decided to define terrorist content based on content relating to organizations on the UN Terrorist Sanctions lists.").

[430] Global Internet Forum to Counter Terrorism, *Who we are: March 2019, Cross-Industry Collaboration*, GLOBAL INTERNET FORUM TO COUNTER TERRORISM (2023), https://gifct.org/about/story/#march-2019---cross-industry-collaboration.

[431] *Id*. ("In response to the Christchurch mosque shootings in New Zealand, members of GIFCT utilized channels of communication that GIFCT had developed as well made use of the hash-sharing database to share more than 800 visually-distinct videos related to the attack.").

[432] Smith and Browne, *supra* note 193 at 154.

[433] *Id.*

[434] Christchurch Call to Action: News and Updates, *supra* note 422.

[435] Global Internet Forum to Counter Terrorism, *Governance, Independent Advisory Committee*, *supra* note 317.

[436] Christchurch Call to Action: News and Updates, *supra* note 422.

[437] Christchurch Call to Action, *Christchurch Call Community Consultation, Final Report*, CHRISTCHURCH CALL TO ACTION (May 2021), https://www.christchurchcall.com/assets/Documents/Chch-Call-Community-Consulation-Report-2021.pdf.

[438] *Id.* at 4 ("The consultation was open to submissions from 21 September to 30 October 2020. In total, members of the Call community submitted 39 responses.").

[439] *Id.*

The stock-taking report found that the Call community had undertaken dozens of new initiatives in their home jurisdictions, and companies had created new policies to fulfill the commitments of the Call. When asked what the most important accomplishment the Call had achieved, 50 per cent of the respondents answered it was the creation of a multistakeholder approach to preventing the abuse of the internet by terrorist and violent extremists.[440] Another 26 per cent believed it was raising awareness of the issue of TVEC online.[441] The remaining 24 per cent believed it was reforming the GIFCT and creating Crisis Incident Protocols.[442] The response to where respondents wanted to go next were mixed, but the majority supported increasing collaboration on a multistakeholder approach and recruitment efforts to increase the number of Call supporters.[443]

With the mandate to strengthen a multistakeholder model, the Call community reunited (virtually) in May 2021 for the two-year anniversary of the Call. The 2021 Leaders' Summit produced a concrete work plan for what the Call would accomplish in the next three years.[444] 2021 priorities included developing an advisory function for CCAN, enhanced information sharing, increased tech company membership, sharing best practices, and strengthening the links between the Call and the GIFCT.[445] Additionally, the Call put together four work plans for community building, crisis response, tech and government transparency, and algorithms and positive interventions.[446] Each work plan detailed what the Call had accomplished since 2019 and what the working groups would do in the next six months, one year and three years to fulfill the Call commitments.

First, the Community Work Plan outlined the work of the Call to foster multistakeholderism and give all stakeholders a seat at the table. To maintain this momentum, one of the most pressing tasks was to fund a Secretariat to assist CCAN rather than expecting CCAN supporters to volunteer for administrative tasks.[447] This was crucial, as most of the civil society organisations involved are run on very tight budgets and were juggling dozens of similarly related initiatives on TVEC. In the short term, the New Zealand and French Governments provided the funding to hire the Secretariat. Next, the Call wanted to further develop the advisory function of CCAN by increasing its membership, creating a technological solution to enable intersessional dialogue between Call community members, and

[440] *Id.* at 66 ("50% of responses referenced, in some capacity, the unique multistakeholder approach embraced in the development and implementation of the Call. For the first time, governments, major tech companies, and civil society representatives have created an innovative, flexible coalition, working cooperatively to stop and prevent attacks like Christchurch being broadcast and spread online.").
[441] *Id.*
[442] *Id.* at 70, *see* figure 3.
[443] *Id.*
[444] Christchurch Call to Action, *Christchurch Call Community Work Stream Work Plan*, CHRISTCHURCH CALL TO ACTION (May 2021), https://www.christchurchcall.com/assets/Documents/Community-Work-Stream-Work-Plan.pdf.
[445] *Id.* at 1 ("As we move into the third year of the Call, it is incumbent on the Call community to ensure all members are equipped to participate fully in the work of the Call. The work plan developed by the Call Community work stream seeks to facilitate a flourishing community where every stakeholder has a seat at the table as envisaged by the text of the Call. The plan also seeks to foster trusted relationships between all stakeholders. Without this, no other work stream will reach its full potential. This plan identifies areas where more work is needed to achieve this ambition. It seeks to build trust across the multi-stakeholder community through improved information sharing and increased channels of communication, including through the use of technology. Recognising the value of increased industry participation in the Call, it promotes involvement of the entire Call community in the on-boarding of new supporters, to ensure the continued integrity of the Call principles. The Call commitments are voluntary, therefore any mechanisms for understanding how supporters are carrying out the commitments in the Call must be grounded in trust-based dialogue between members. In order to achieve these objectives, resourcing will be required. It is our hope that the community will rise to this challenge, bringing their different capacities and capabilities to bear on the project.").
[446] *See all*, Christchurch Call to Action, *Christchurch Call Reports and Publications*, CHRISTCHURCH CALL TO ACTION (May 2021), https://www.christchurchcall.com/media-and-resources/reports-and-publications/.
[447] *See* Christchurch Call to Action, *Christchurch Call Community Work Stream Work Plan, supra* note 444.

developing more accountability mechanisms.[448] Finally, the Community Work Plan envisioned a closer link between CCAN and the GIFCT, as the GIFCT is the Call's "primary partner for delivery against Call commitments through its multistakeholder working groups".[449]

Second, the Crisis Response Work Plan set out key objectives for improving processes for crisis response under the Call.[450] While coordination between governments and tech companies had improved remarkably since 2019, terrorists and violent extremists were still turning to social media to broadcast their attacks and promote radicalisation efforts. Additionally, the Call was looking to civil society to help improve crisis response tools to reflect due process and human rights considerations.[451] Therefore, the Call set out to conduct a review of the Call's Crisis Response Protocol, along with a comprehensive mapping exercise of all content incident protocols to identify where there were overlaps or gaps.[452] Finally, as the Community Work Plan had done, the Crisis Response Work Plan called for broadening the Call's membership and involving civil society and academia in the discussion.[453]

Third, the Transparency and Reporting Work Plan discussed how increased transparency could build trust among stakeholders, help prevent and reduce harm from TVEC online, and protect human rights and fundamental freedoms.[454] Many of the objectives of this work plan focused on the need to raise awareness of, and guide stakeholders to, the ongoing transparency reporting-related work happening at the GIFCT, TAT, and OECD.[455] However, one key initiative that the Call was undertaking that was not happening in other fora was related to how governments can be more transparent about when they ask companies to remove TVEC.[456] As one of the only MSIs in this space with governments, civil society and tech companies at the table, the Call was unique in asking government leaders to examine their practices and provide guidance on how they could improve processes in line with human rights principles.

Finally, the Algorithms and Positive Interventions Work Plan looked at ways to better understand user journeys and the role algorithms play in the radicalisation process.[457] In 2021, there were several MSIs working on issues related to this topic. Two of these MSIs involved government stakeholders: the GIFCT's Content-Sharing Algorithms, Processes, and Positive Interventions Working Group[458] and the Global Partnership on Artificial Intelligence, created by Canada and France at the G7 Digital Ministerial

---

[448] *Id.* at 2, *noting* "medium term objectives achievable within 6-12 months" include "develop civil society advisory function of the Call, through addressing gaps in diversity and mapping and utilization of diverse expertise within the network.".
[449] *Id.* at 1.
[450] Christchurch Call to Action, *Christchurch Call Crisis Response Work Plan*, CHRISTCHURCH CALL TO ACTION (May 2021), https://www.christchurchcall.com/assets/Documents/Christchurch-Call-Crisis-Reponse-Workplan.pdf.
[451] *Id.*
[452] *Id.* at 1 ("Since its launch in May 2019, the Call has developed a dedicated Crisis Response Protocol (Christchurch Call Crisis Response Protocol). Other protocols at an international, domestic and organisational level have also been developed. Some of these protocols are geographically specific, whilst others are more global in nature and seek to coordinate a swift response.").
[453] *Id.*
[454] Christchurch Call to Action, *Christchurch Call Transparency Work Plan*, CHRISTCHURCH CALL TO ACTION (May 2021), https://www.christchurchcall.com/assets/Documents/Christchurch-Call-Transparency-Work-Plan.pdf.
[455] *Id.*
[456] *Id.*
[457] Christchurch Call to Action, *Christchurch Call Algorithms & Positive Interventions Work Plan*, CHRISTCHURCH CALL TO ACTION (May 2021), https://www.christchurchcall.com/assets/Documents/Algorithms-and-Positive-Interventions-WorkPlan.pdf.
[458] Global Internet Forum to Counter Terrorism, *Content-Sharing Algorithms, Processes, and Positive Interventions Working Group*, GLOBAL INTERNET FORUM TO COUNTER TERRORISM (July 2021), https://gifct.org/wp-content/uploads/2021/07/GIFCT-CAPI2-2021.pdf.

Meeting in 2020 and hosted by the OECD.[459] Therefore, the Work Plan set out action items in this area to avoid duplicating efforts. As a result, this Work Plan included building understanding of recommender algorithms and user journeys, empowering community-driven online interventions, and mechanisms for TVEC removal including transparency and redress.[460] In line with the literature in 2021, the emphasis was on positive intervention measures to redirect a person away from extremist or terrorist content.

To carry out each of these work plans, community members met frequently throughout 2021 and 2022. Given that much of the work was happening at the working-group level, the Call Secretariat conducted a survey of community members to understand how stakeholders felt the work plans had progressed and published the results in August 2022. Additionally, the Secretariat hosted two community-wide meetings to discuss the work plans and evaluate resourcing.[461] The 2022 Community Update contains feedback from these meetings and survey results.[462] Overall, community members thought the greatest achievements of the Call since 2021 were: the creation of a new stakeholder on-boarding process for Call supporters, a review and update of the Crisis Response Protocol, better ties with the GIFCT, increased awareness of the Call's work, and improved communication through monthly calls with CCAN and additional stakeholders.[463]

Alongside the Community Update, Call community supporters made several statements ahead of the 2022 Leaders' Summit detailing their progress in fulfilling the Call commitments.[464] Five governments and the European Commission outlined actions they had taken to address TVEC online, including Australia's passage of the Online Safety Act 2021, the EU's regulation on "preventing the spread of extremist content online", Japan's efforts to improve the capacity of Association of Southeast Asian Nations countries to prevent TVEC online, and India's media standards framework.[465] Other organisations gave updates on their work, including the GIFCT, which stated it had responded to over 270 attacks since creating the Crisis Incident Protocol; the Global Partnership on Artificial Intelligence, which summarised its work on recommender algorithms; and Inclusive Aotearoa Collective Tāhono, which detailed its work in New Zealand to build more inclusive communities.[466] CCAN provided a response to the community statements document expressing the desire for civil society to play a more pronounced role in policy development and urging supporters to engage with them more frequently. [467]

---

[459] The Global Partnership on Artificial Intelligence, About GPAI – Our Mission, THE GLOBAL PARTNERSHIP ON ARTIFICIAL INTELLIGENCE (2021), https://www.gpai.ai/about/.
[460] Christchurch Call to Action, *Christchurch Call Algorithms & Positive Interventions Work Plan*, *supra* note 457.
[461] Christchurch Call to Action, *Christchurch Call 2022 Community Update*, CHRISTCHURCH CALL TO ACTION (May 2022), https://www.christchurchcall.com/assets/Documents/Christchurch-Call-2022-Community-Update.pdf ("This report reflects an overview by the Secretariat of Call Community efforts and progress under each of the work plans. It draws on input from responses to our 2022 Community Survey, which asked Community members to share their thoughts on the allocation of effort and progress made under the work plans, and their assessment of risks, opportunities, and priority areas as this work continues. In addition to the Community Survey, the Community came together over the course of two meetings to collectively reflect on progress on the work plans, and opportunities and priorities for the future.").
[462] *Id.*
[463] *Id.*
[464] Christchurch Call to Action, *Christchurch Call Community Statements 2022*, CHRISTCHURCH CALL TO ACTION (May 2022), https://www.christchurchcall.com/assets/Documents/Christchurch-Call-Community-Statements-2022.pdf.
[465] *Id.*
[466] *Id.*
[467] Christchurch Call Advisory Network, *Christchurch Call Advisory Network (CCAN) position statement, Christchurch Call Summit, 2022*, CHRISTCHURCH CALL ADVISORY NETWORK (Sep. 2022), https://christchurchcall.network/wp-content/uploads/Summit-Sept-22-CCAN-Statement.pdf ("Finally, we believe civil society should have a more pronounced role in policy development. Just as we advocate for online service providers to include civil society earlier in the design process, so too should governments in creating their policies. We urge the supporter companies and states to consult with CCAN to ensure that the Call values are

In addition to their message in the community statements document, CCAN announced a separate initiative to evaluate the work of the Call.[468] The CCAN evaluation document, published in September 2022, expressed frustrations with the work of Call, including lack of transparency on commitments, lack of concrete evidence that human rights due diligence processes were in place, a failure on the part of government and company leaders to meaningfully engage civil society, and the creation of silos around the Crisis Response Protocols.[469] To remedy these problems, CCAN recommended regular reporting to CCAN from governments and companies of their actions, which could be done via publicly accessible repositories.[470] To start this work, CCAN decided to undertake an evaluation process of governments' and companies' efforts to fulfill the Call commitments.[471] This evaluation would survey six governments and four companies and cover overarching themes related to transparency, human rights due diligence, civil society engagement, and cross-Call collaboration.[472] On 30 June 2022, the survey was sent out to identified participants and CCAN members volunteered to do additional research to supplement responses.[473] As of the time of writing, this evaluation is still under way.

The CCAN evaluation followed a similar approach to another report released by CCAN in September 2022, which analysed anti-dehumanisation policies.[474] In March 2022, CCAN distributed a request for information to government and company supporters of the Call to map current approaches combating dehumanising speech.[475] This research was important to the work of the Call, as dehumanisation is a common feature of terrorist and violent extremist propaganda. Dehumanising speech is separate from hate speech; it aims to lower an audience's moral reflexes towards a particular group, which can lead to offline violence, as seen in the Christchurch shooter's manifesto.[476] The evaluation found that, of the companies and governments surveyed, only Twitter had specific policies regarding dehumanising speech. However, companies and governments alike had rules and laws that could cover dehumanising speech if applied correctly.[477] Therefore, the report suggested that the Call members could work together on strategies to counter the production and dissemination of dehumanising speech, including through

incorporated and that the commitments enumerated in the Call to Action are undertaken in a manner that is consistent with the rule of law and international human rights law, and in a way that meets the needs of people and communities most impacted by TVEC.").

[468] Christchurch Call Advisory Network, *Evaluating the Impact of Government and Company Commitments Under the Christchurch Call to Action A Pilot Project of the Christchurch Call Advisory Network*, CHRISTCHURCH CALL ADVISORY NETWORK (Sep. 2022), https://www.christchurchcall.com/assets/Documents/Christchurch-Call-2022-CCAN-Evaluation-Project.pdf.

[469] *Id.* at 1 ("In contrast, it was much harder to find evidence that supporters had implemented their commitments under the Call beyond declarations of intent to do so. If work was undertaken in response to the Call, it was rarely identified as such, making measurement of the Call's impact difficult. This raises questions about the consistency of the Call's impact across its many government and company supporters.").

[470] *Id.*

[471] *Id.*

[472] *Id.* at 3 ("We also selected a small sample of the supporting governments and companies to include in this first evaluation. We chose six governments—New Zealand, France, Australia, Canada, United Kingdom, and India—and four companies—Microsoft, Meta, Twitter and Google. We chose these signatories based on their role as leaders of the Call (in the case of New Zealand and France), the longevity of their support for the Call, and our internal capacity to conduct this analysis, such as familiarity with language, legal systems, and access to resources.").

[473] *Id.*

[474] Christchurch Call Advisory Network, *CCAN Report on Anti-Dehumanization Policy*, CHRISTCHURCH CALL ADVISORY NETWORK (Sep. 2022), https://christchurchcall.network/wp-content/uploads/CCAN-Report-on-Anti-Dehumanization-Policy.pdf.

[475] *Id.*

[476] *Id.* at 4 ("Dehumanization is a distinct concept from hate speech and Terrorist and Violent Extremist content (TVEC), although it often features in both. Dehumanizing language or speech (e.g., referring to a race of people as a disease) is a type of hate speech, broadly defined, and can create a heightened environment for violence.").

[477] *Id.* at 4 ("Except for Twitter, there were no existing laws, rules or policies distinctly on dehumanising speech or language. However, there were laws, rules or policies that conceivably could cover dehumanising speech or language.").

frameworks related to hate speech, disinformation, harmful digital communications, and tort law.[478] The Call community welcomed this thoughtful feedback. The report is an outstanding example of multistakeholderism advancing policy changes.

Before turning to the discussions at the 2022 Leaders' Summit, it is important to acknowledge one event that brought renewed attention and urgency to the work of the Call – the mass shooting in Buffalo, New York on 14 May 2022. In Buffalo, a white 18-year-old male killed 10 people in a supermarket in a predominantly black neighborhood.[479] The shooter wore a GoPro camera and attempted to livestream his attack on Twitch, a gaming platform, but the company disabled the livestream within two minutes.[480] An investigative report into the incident by the Office of the New York State Attorney General details the Buffalo shooter's radicalisation online and his use of social media platforms, including Reddit, Discord, 4chan, 8kun, and others to connect with violent extremists.[481] Notably, in the Buffalo shooter's manifesto, he stated that the Christchurch attack was a "catalyst" and inspired him towards ethno-nationalist beliefs.[482] Unfortunately, the Buffalo attack has not been the only Christchurch-inspired attack; there have been others in Poway, El Paso, Dayton, Halle, Glendale, Nakhon Ratchasima, Nice, and Vienna.[483] While the companies had improved their capabilities to stop the spread of the video and manifesto, it was clear more work needed to be done.

The Buffalo attack was top-of-mind at the 2022 Leaders' Summit on the sidelines of the UN General Assembly in New York in September. The meeting was an opportunity to welcome new industry supporters and partner organisations, including Roblox, Zoom, Mega, Clubhouse, the Global Community Engagement and Resilience Fund, and TAT.[484] Additionally, the meeting sought to provide a strategic direction for the upcoming year, prioritising three areas: improving incident response, understanding how algorithms and social drivers can lead to radicalisation, and future-proofing the Call.[485] With regard to the latter, leaders added two new workstreams – one exploring new technologies and the other exploring the drivers of violent extremism, including gender-based hate.[486] The Summit's joint statement also mentioned how the Call's multistakeholder model could help similar MSIs combatting disinformation, harassment, hatred online, and issues affecting youth, including Tech for Democracy, the Summit for Democracy, the Global Partnership for Action on Gender Based Online Harassment and Abuse, and the Global Partnership on Artificial Intelligence.[487]

Another important announcement in September 2022 was the launch of the Christchurch Call Initiative on Algorithmic Outcomes (CCIAO). The CCIAO is a project funded by Microsoft, Twitter, and the Governments of New Zealand and the US to develop new technologies to understand the impacts of

[478] *Id.*
[479] *See* Office of the New York State Attorney General Letitia James, *supra* note 26.
[480] *Id.* at 10 ("The shooter began livestreaming using the online platform Twitch at approximately 2:08 p.m., using a GoPro video camera attached to his helmet.").
[481] *Id.* at 6-9.
[482] *Id.* at 17.
[483] Christchurch Call to Action, *Christchurch Call Second Anniversary Summit Co-Chair Statement 2021*, CHRISTCHURCH CALL TO ACTION (May 2021), https://www.christchurchcall.com/assets/Documents/Christchurch-Call-2nd-Anniversary-Summit-Co-chair-Statement-2021.pdf ("Despite our achievements so far, the many attacks since Christchurch - in Colombo; El Paso; Dayton; Halle; Glendale; Nakhon Ratchasima; Conflans Sainte-Honorine; Nice; and Vienna among others – bear witness to the challenge we still face").
[484] Jacinda Ardern and Emmanuel Macron, *Co-Chair Statement Christchurch Call Leaders' Summit*, CHRISTCHURCH CALL TO ACTION (Sep. 2022), https://www.christchurchcall.com/assets/Documents/Christchurch-Call-Joint-Statement-2022-English-version.pdf.
[485] Christchurch Call to Action, *Our Work: Leaders' Summits*, CHRISTCHURCH CALL TO ACTION (2023), https://www.christchurchcall.com/about/leaders-summits/.
[486] Ardern & Macron, *supra* note 484.
[487] *Id.*

algorithms on people's online experiences.[488] Working with OpenMined, an open-source non-profit organisation, the CCIAO is developing tools to provide access to researchers to study how individuals are radicalised across platforms. In the past, it has been difficult to carry out extensive research on TVEC because quality research requires access to sensitive information across platforms. The CCIAO is developing software through a privacy-enhancing technology that will enable data scientists to study algorithms across multiple online platforms. This technology provides cross-platform analysis which will give researchers a better understanding of how people are radicalised online and how to more effectively intervene to protect people, both online and offline.[489] If proven successful in the Call context, this could open up a new field of algorithmic research for a much wider application.[490] Work on the CCIAO is under way; researchers are beginning to access data from Twitter, DailyMotion, and LinkedIn through the privacy-enhancing technology to explore the ways in which AI and humans interact online.[491]

On 19 January 2023, Prime Minister Ardern announced she was resigning from office and would not seek re-election.[492] However, Ardern was committed to staying involved with the work of the Call. On 4 April 2023, Prime Minister Chris Hipkins announced he was appointing Ardern as special envoy for the Christchurch Call.[493] During a virtual gathering on the fourth anniversary of the creation of the Call, Special Envoy Ardern stated the Secretariat's intention to host a 2023 Leaders' Summit in September.[494] Ardern and other speakers during the virtual meeting called for more attention in several areas, including understanding the impact of algorithmic systems on radicalisation, confronting the reality of gender-based hatred and abuse as a factor in radicalisation and violence, and considering emergent technologies including generative AI.[495]

B. Evaluation of the Christchurch Call to Action

Evaluating the work of the Call is not a mere check-box exercise, as the initiative is a bottom-up, large-scale collaboration between various stakeholders who all have their own motivations and reasons for implementing the commitments of the Call.[496] Ardern summarised the accomplishments of the Call in a June 2023 op-ed in the *Washington Post*, saying:

---

[488] Christchurch Call to Action, *Christchurch Call Initiative on Algorithmic Outcomes*, CHRISTCHURCH CALL TO ACTION (Sep. 2022), https://www.christchurchcall.com/media-and-resources/news-and-updates/christchurch-call-initiative-on-algorithmic-outcomes/.

[489] *Id.* ("That system will help us to answer questions such as: "What are the distinct features of a user journey for someone that engages with TVEC?" "What is the before/after impact of positive interventions, or changes to ranking systems or other platform features designed to reduce toxicity or risk of harm?" "What do user journeys for 'at risk' user types look like between and across platforms?" "How effective and fair are the automated systems that identify and remove TVEC?").

[490] Ardern, *supra* note 10 ("We're also taking on some of the more intransigent problems. The Christchurch Call Initiative on Algorithmic Outcomes, a partnership with companies and researchers, was intended to provide better access to the kind of data needed to design online safety measures to prevent radicalization to violence. In practice, it has much wider ramifications, enabling us to reveal more about the ways in which AI and humans interact.").

[491] *Id.*

[492] Beehive Press Release, *Prime Minister Jacinda Ardern announces resignation*, NEW ZEALAND GOVERNMENT (19 Jan. 2023), https://www.beehive.govt.nz/release/prime-minister-jacinda-ardern-announces-resignation.

[493] Christchurch Call to Action, *New Zealand Special Envoy for the Christchurch Call announced*, CHRISTCHURCH CALL TO ACTION (4 April 2023),https://www.christchurchcall.com/media-and-resources/news-and-updates/new-zealand-special-envoy-for-the-christchurch-call-announced/.

[494] Christchurch Call to Action, *Four years of the Christchurch Call*, CHRISTCHURCH CALL TO ACTION (15 May 2023), https://www.christchurchcall.com/media-and-resources/news-and-updates/four-years-of-the-christchurch-call/.

[495] *Id.*

[496] Ardern, *supra* note 10.

> … we have developed new policies and ways of working that holistically address the complexities of terrorist and violent extremist content. We have established new crisis protocols to respond effectively and in a coordinated manner to attacks with an online component. We worked as a community to establish the Global Internet Forum to Counter Terrorism as an independent NGO. This created the opportunity for the GIFCT to become a more fully multistakeholder construct, develop integrated solutions, and share information and expertise, should it choose to. I know we still have work to do to fulfill this vision. We now better understand the online ecosystem and the experiences of affected communities, having led collaborative research across our community. And we have built a strong and diverse multistakeholder community.[497]

Rather than go through each of the original 25 commitments in the Call, this section examines overarching themes of the Call's work in two areas: building a multistakeholder community to address the drivers of TVEC and taking steps to eliminate TVEC online while protecting a free, open, and secure internet. This section will discuss the work of both the Call and some of the other MSIs that collaborate with the Call community.

Building a Multistakeholder Community

Among the goals of the Call is to counter the drivers of TVEC through a whole-of-society approach to addressing the problem, via a multistakeholder framework. Call supporters agreed to work collectively on 12 commitments which fall into three broader buckets of work: raising awareness to widen support for the Call, working with civil society to address the drivers of TVEC, and accelerating research.[501] This section evaluates how the Call has accomplished these three overarching goals.

1(a). Raising Awareness

The events of 15 March 2019 were a harsh wake-up call to governments and tech companies alike. In their aftermath, the New Zealand Government received an out-pouring of support and had the moral authority to lead an MSI to tackle the issue.[498] Since 2019, the Call has done a remarkable job of keeping the Christchurch shooting front and centre in global content moderation discussions. As part of this effort, the Call community has partnered with dozens of other MSIs to confront the challenges of TVEC online, including the IGF, GNI, I&J, the Summit for Democracy, and the EU Internet Forum (to name a few MSIs we have already examined in this report).[499] Additionally, the New Zealand Government has partnered with governments and civil society in their efforts to eliminate TVEC online, including the Jakarta Centre for Law Enforcement Cooperation, the Pacific Working Group on Counter Terrorism and Transnational Organised Crime, the Global Community Engagement and Resilience Fund, the UN Office on Drugs and Crime, and the Aqaba Process (to name a few global non-MSI forums). The Call's supporters, often working alongside CCAN, have also attended a wide range of conferences to build awareness for the Call, including RightsCon, the Paris Peace Forum, and the Trust and Safety Professional Association's "TrustCon." This effort has brought together new stakeholders who may not

[497] *Id.*
[498] Smith and Browne, *supra* note 193.
[499] The Christchurch Call to Action: Full English Text, *supra* note 11 ("Tech for Democracy, the Summit for Democracy, the Freedom Online Coalition, the Declaration for the Future of the Internet, the Aqaba Process, the Global Partnership for Action on Gender Based Online Harassment and Abuse, the Global Partnership on Artificial Intelligence, and the International Call to Stand up for Children's Rights Online, and where there is multistakeholder interest in new work programmes separate to the Call.").

have been impacted by the events of 15 March 2019, but are now coming together to share ideas on how to combat TVEC online.

One way the Call has ensured attention on its work has been through annual Leaders' Summits, where supporters meet to confirm priorities and identify areas of focus.[500] Ahead of these summits, the Call Secretariat convenes working groups to undertake multistakeholder efforts throughout the year and encourage stakeholders to act independently in their commitments.[501] These summits bring together stakeholder "leaders" – meaning heads of governments, CEOs, and top leaders from civil society or academia. One goal for having these conversations at the "leader-level" is to ensure the issue remains a top priority. This framing ensures the top officials are aware of the ongoing work, but it can present challenges to the overall inclusiveness of the event. Many heads of state and CEOs have incredibly busy schedules, which can conflict with the timing of the meeting, resulting in key supporters being left out of the discussion. In the long run, this can negatively impact implementation efforts as these leaders do not typically do the day-to-day work of implementing commitments. Therefore, this framing risks disenfranchising supporters who may feel less bought-in on the process. Making these summits more inclusive and accessible could be one way to improve raising awareness around the work of the Call.

1(b). Working with Civil Society

The Call's supporters – governments and tech companies – work with civil society primarily through CCAN. CCAN represents a diverse group of civil society actors, including victims of the Christchurch attack, human rights organisations, technical experts, and free speech advocates.[502] CCAN has worked closely alongside Call supporters over the years to provide expert advice in a manner consistent with a free, open, and secure internet and international human rights principles. In many ways, CCAN is a separate curated MSI that sits alongside the Call itself. It has its own website, terms of reference, and leadership structure.[503] Additionally, CCAN has its own recruitment and approval process, which has changed over the years and was most recently updated in a 2022 terms of reference.[504] Despite the growing number of civil society organisations working on content moderation problems, CCAN has not grown much in four years: from 40 members to 46.[505] One reason for this may be the terms of reference, which limits the amount of funding an organisation can receive from governments or companies to less than 25 per cent.[506] Unfortunately, this threshold has the unintended consequence of limiting many of the most relevant non-profit organisations researching and developing solutions for

[500] Ardern, *supra* note 10.
[501] *Id.*
[502] Christchurch Call Advisory Network, *About Us – History, supra,* note 428.
[503] *Id.; see also* Christchurch Call Advisory Network, *Terms of Reference*, CHRISTCHURCH CALL ADVISORY NETWORK (Sep. 2022), https://christchurchcall.network/governance/.
[504] *Id.*
[505] Christchurch Call Advisory Network, *About Us – Members*, CHRISTCHURCH CALL ADVISORY NETWORK (Sep. 2022), https://christchurchcall.network/about-us/members/.
[506] *See* Christchurch Call Advisory Network, *Terms of Reference*, *supra*, note 503 ("Members must be independent of governments and companies. To qualify for membership, they should, if applicable: 1. Establish that they have organizational and accountability structures in place, such as being registered as a non-governmental organization in their country or providing visibility of their operations through a published statement of purpose and meeting minutes; 2. Include in their application a statement that their work is not directed or strongly influenced by a government or private sector company. 3. To the extent applicants receive significant funding (more than 25% of their operating budget) from governments or corporations, the application should disclose the total percentage of their operating revenue that comes from these sources, the specific governments or companies that provide funding, and what measures they take to ensure/maintain independence from those funders. 4. Organizations need not disclose how much funding they receive from any particular source, and financial information will not be shared with anyone not directly involved in determining membership eligibility for the applicant; such disclosure could pose serious legal, reputational, or security risks to the applicant or its partners.").

addressing TVEC online. Therefore, one way the Call has worked to ensure broader inclusion of new stakeholders has been to create a "partners" category, which includes organisations such as TAT, UNESCO, and the Council of Europe.[507] This solution helps ensure a wider variety of perspectives.

Tension between civil society, governments, and industry on policy direction is common within any MSI, because those groups tend to view their roles very differently. In many cases, civil society often see themselves as individual advocates instead of implementation partners. Indeed, over the years, CCAN has requested a "more pronounced role in policy development" from the government and company supporters of the Call.[508] Unfortunately, these tensions have been amplified by the fact that CCAN members are not formal supporters of the Call commitments, but serve an advisory role.[509] As the Call looks towards new projects and initiatives, finding ways to more directly incorporate CCAN into the structure of the Call could help address some of these tensions and increase interest among potential new supporters. One way the Call has addressed these tensions is by directly incorporating CCAN members into the Call's working groups. Additionally, the Call supporters continue to build trust between stakeholders through summits, on-going conversations, and internal transparency processes. All this work will hopefully contribute to the appeal of joining the Call and increase the diversity of supporters.

1(c). Accelerating Research

Over the years, Call supporters have invested heavily in research initiatives addressing the problems of TVEC online. Companies, through their contributions to the GIFCT, support the Global Network on Extremism and Technology, which is the GIFCT's academic research arm, exploring the nexus between online behaviours and offline harms.[510] The GIFCT also commissions research about the evolving tactics, capabilities, and identities of violent extremist groups and shares them with the Call community more broadly. Additionally, the GIFCT, and its member companies, work with social scientists and extremism experts in various regions to help them develop the skills to identify and counter extremism. TAT receives both government and company funding to support third-party researchers on projects.[511] Individual companies have their own initiatives such as Google's Jigsaw project, which explores threats to online discourse, and Meta's Oversight Board, which is currently exploring how the company moderates content related to dangerous individuals and organisations.[512] Governments have also supported research; two examples include Canada's Centre for Community Engagement and Prevention of Violence, which seeks to counter radicalisation to violence,[513] and New Zealand's He Whenua Taurikura, the National Centre of Research Excellence for Preventing and Countering Violent Extremism.[514]

---

[507] Christchurch Call to Action, *Our Community – Partners*, CHRISTCHURCH CALL TO ACTION (4 April 2023). https://www.christchurchcall.com/our-community/partners/.
[508] Christchurch Call Advisory Network, *Christchurch Call Advisory Network (CCAN) position statement, Christchurch Call Summit, 2022, supra* note 467.
[509] *Id.*
[510] Global Internet Forum to Counter Terrorism, *Research*, GLOBAL INTERNET FORUM TO COUNTER TERRORISM (2023), https://gifct.org/research/.
[511] Tech Against Terrorism, *Research and Publications*, TECH AGAINST TERRORISM (2023), https://www.techagainstterrorism.org/research/.
[512] The Oversight Board, *Oversight Board announces a review of Meta's approach to the term "shaheed"*, THE OVERSIGHT BOARD (March 2023), https://www.oversightboard.com/news/1299903163922108-oversight-board-announces-a-review-of-meta-s-approach-to-the-term-shaheed/.
[513] Public Safety Canada, *Canada Centre for Community Engagement and Prevention of Violence*, GOVERNMENT OF CANADA (5 Dec. 2022), https://www.publicsafety.gc.ca/cnt/bt/cc/index-en.aspx.
[514] Department of the Prime Minister and Cabinet, *He Whenua Taurikura*, DEPARTMENT OF THE PRIME MINISTER AND CABINET (22 Sep. 2021), https://www.dpmc.govt.nz/our-programmes/national-security/counter-terrorism/he-whenua-taurikura.

In addition to the individual research-supporting efforts, the Call launched its own research initiative, the CCIAO, mentioned above. The CCIAO is funded by Microsoft, Twitter and the governments of New Zealand and US to create new technology to understand the impacts of algorithms on people's online experiences.[520] This cross-industry and government project was necessary because studying the impact of algorithmic outcomes, and the way they impact a user's journey to radicalization, is incredibly difficult to do in a way that allows researchers access to highly sensitive datasets using privacy-respecting technologies. While many online platforms claim they have made progress in improving algorithmic recommendation systems, without independent study it is impossible to measure the impacts of these changes. The CCIAO seeks to address these challenges by providing researchers access to anonymized datasets to test how people are radicalized online.[521] The technology, if proven successful in the Call context, could open up a new field of algorithmic research with a much wider application.[522] Work on the CCIAO is now underway, with researchers beginning to access the platform to explore the ways in which AI and humans interact online.[523] This research project is set to be a cornerstone of the Call's future work on AI and automation.

Eliminating TVEC Online

It would be impossible to calculate the percentage of content online that qualifies as TVEC, and whether that number has increased or decreased since 2019. Even without this data, we know that the world remains a long way from "eliminating" TVEC online. However, the Call has been an important catalyst for efforts to achieve this goal, by coming up with a plan and getting stakeholders to agree to it. In 2019, the mere fact that companies and governments could agree to work together to solve a broader societal challenge was novel. The Call deserves credit for bringing together stakeholders to work collectively to address the issue and publicly commit to a plan.[515] This planning itself represents progress, as it generated proactive thinking on solutions and highlighted key blind spots in both company and governmental actions to eliminate TVEC online.[516] Another positive outcome from the Call's creation was that the commitments put public pressure on companies to invest in policy and technical solutions.[517] Finally, over the past four years, the Call has served as a rallying point for greater cross-industry reporting, which has helped improve compliance standards and create best practices.[518] Through convening stakeholders and publicly committing to a plan of action, the Call has helped companies find new ways to eliminate TVEC online – both individually and as an industry.

2(a). Individual Company Solutions

As part of their commitments to the Call, companies outlined steps they would take to address TVEC on their own platforms, undertaking to "tighten their terms of service, better manage live videos, respond to user reports of abuse, improve technology controls, and public transparency reports."[519] There is no doubt many companies have implemented changes in all five of these areas. However, as Evelyn Douek, a scholar who fastidiously tracks changes to social media companies' policies, notes, it

[515] Douek, *supra* note 21 at 71 ("First, requiring planning forces platforms to think proactively and methodically about potential operational risks. The process of having to articulate a plan itself engenders proactivity and highlights blind spots. Platforms are known for failure to anticipate key risks, so "making [platforms] think" is meaningful, and a useful counterweight to the "Move Fast and Break Things" culture of Silicon Valley.").
[516] *Id*.
[517] *Id*. at 72 ("transparent plans facilitate broader policy learning for regulators and across industry. Comparative information would show industry best (or worst) practices"); *citing* Margot E. Kaminski, *Binary Governance: Lessons from the GDPR's Approach to Algorithmic Accountability*, 92 S. CAL. L. REV. 1529, 1533–34 (2019).
[518] *Id*. Public planning efforts create some accountability on the companies and governments to improve their practices related to content moderation.
[519] Smith and Browne, *supra* note 193 at 154.

can be difficult to know exactly what changes companies implemented specifically as a commitment to the Call and what changes they made because it happened to align with other company priorities.[520] Unfortunately, this is part of a broader accountability problem for internet companies who rarely publicly explain how they enforce their own rules and the systems they have in place.[521] However, several companies have stated that they changed their policies to fulfill Call commitments. For example, a representative from Twitter testified to Congress that the Christchurch Call made the company realise they needed a real-time communications strategy in a crisis.[522] Additionally, Meta testified in that same hearing that the company introduced reforms in line with their Call commitments to limit access to certain features – notably live streaming – for users that had violated its Dangerous Organizations policy. [523] CCAN is currently undertaking an evaluation process to track which companies and governments have implemented new policies in accordance with their Call commitments.

Another industry shift since the creation of the Call has been more transparency in the way platforms defines TVEC in their terms of service and disclose moderation of those rules through transparency reporting.[524] In 2019, a few online platforms only vaguely defined TVEC in their terms of services – and many did not even do that.[525] After the Christchurch shooting, not only did companies more clearly define TVEC, they also started to report on their TVEC content moderation practices in their transparency reports. In 2022, the 15 largest online platforms that released transparency reports included TVEC information, up from only five companies in 2019.[526] It is hard to argue that the Call is solely responsible for this industry effort, as calls for increased transparency around content moderation practices have recently come from every corner of government and civil society. In fact, organisations

[520] Douek, *supra* note 21 at 75 ("Content moderation plans so far have largely been of this nature—often the announcement of a plan has been the end of a platform's external engagement with an issue, rather than the beginning. For example, the public has been left almost entirely in the dark about the effectiveness of platforms' exceptional COVID-19 misinformation rules released to great fanfare. Two years after the adoption of the "Christchurch Call to Eliminate Terrorist and Violent Extremist Content Online," there has been little public accounting of how companies have implemented their voluntary pledges. Therefore, any regulatory scheme must include an obligation for platforms to provide an annual public review of the implementation of their plans to create some measure of accountability for platforms' progress towards their goals."); *see also* Christchurch Call Advisory Network, *CCAN Report on Anti-Dehumanization Policy, supra* note 474.

[521] *Id.* at 71("Requiring platforms to publish and explain plans for how they will enforce their own rules may sound like a feeble form of accountability. But it's hard to overstate both how ineffective platforms are at enforcing their rules, and how little is known about what systems they have in place to do so. Despite being a purely procedural (not outcome-based) form of accountability, there are four main benefits of requiring platforms to have publicly available plans for rule-enforcement and that distinguish this form of systems-based transparency from the transparency theatre of aggregated information about individual cases.").

[522] Mass Violence, Extremism and Digital Responsibility: Hearing before the Senate Comm. on Commerce, Science, and Transp., 116th Cong.(8 Sep. 2019), https://www.commerce.senate.gov/2019/9/mass-violence-extremism-and-digital-responsibility (Nick Pickles of Twitter told the Committee, "We've grown that partnership, so we share URLs. So, if we see a link to a piece of content like a manifesto, we're able to share that across industry. And furthermore, I think an area that after Christchurch we recognized we need to improve, we now have real time communications in a crisis, so industry can talk to each other in real time operationally to say even, you know, not content related, but situational awareness.").

[523] *Id.* at 3 (Monika Bickert of Meta testified to the Committee, "For example, in response to the tragic events in Christchurch, we made changes to Facebook Live to restrict users if they have violated certain rules—including our Dangerous Organizations and Individuals policy. We now apply a "one-strike" policy to Live: anyone who violates our most serious policies will be restricted from using Live for set periods of time—for example, 30 days—starting on their first offense.").

[524] Ardern, *supra* note 10.

[525] OECD publishing, *Current Approaches to Terrorist and Violent Extremist Content Among the Global Top 50 Online Content-Sharing Services*, OECD DIGITAL ECONOMY PAPERS No. 296, 11(Aug. 2020), https://www.oecd-ilibrary.org/docserver/68058b95-en.pdf ("The practice of reporting information on how companies moderate and remove content based on their own ToS and policies generally, and based on their anti-terrorism and anti-violence policies in particular, is hardly widespread. Of the 23 Services profiled in this Report that issue any transparency reports at all,18 only five (Facebook, YouTube, Instagram, Twitter and Automattic) issue reports specifically about TVEC.").

[526] *Id*. at 18.

like the GIFCT, OECD and TAT have created programmes to make transparency reporting easier and standardised.[527] However, these commitments remain a priority for the Call, because the quality of transparency reporting still needs improvement. Current transparency reporting efforts are only marginally helpful, as they provide a lot of data without revealing much information at all.[528] Additionally, critics argue that aggregate content moderation enforcement numbers do not always give the full picture of trends, because the raw numbers of removals could be affected by factors that do not always reveal underlying content moderation practices.[529] Even taking into account improvements in reporting and increased attention from multiple MSIs, more work is needed for transparency reporting to meaningfully contribute to our understanding of the root causes of TVEC online.[530]

2(b). Industry-Wide Solutions

Two areas where the companies committed to industry-wide solutions involved mitigating the dissemination of TVEC online and working together on a crisis response protocol.[531] When GIFCT was restructured 2019, and began running as a distinct entity in 2020, these were two of its top priorities. To mitigate the dissemination of TVEC online, the companies further invested in the GIFCT to manage and develop the hash-sharing database. Additionally, the GIFCT started to work more closely with smaller platforms through TAT, which hosts a mentoring program to help develop capability across the sector. Like the more extensive hash-sharing database, TAT runs the Terrorist Content Analytics Platform (TCAP) which automates the detection and removal of verified terrorist content on tech platforms.[532] The TCAP primarily focuses on small tech platforms, many of which may not have the capacity to moderate TVEC or lack access to automated processes.[533] Because they do not require a financial commitment to join, TAT enables many smaller companies to learn more about terrorist misuse of internet platforms and ways to mitigate this risk on their services.

Other industry-wide commitments within the Call relate to the development of a crisis response protocol. This is an area where there are now multiple different protocols and methodologies (often based on who 'owns' or 'manages' the protocol, and for what purpose) and work is needed to ensure coordination and compatibility between them. The Call and the GIFCT – operating in different ways – each contribute to crisis response protocols to stop the rapid dissemination of TVEC and quickly remove footage from many platforms.[534] The GIFCT has developed its Content Incident Protocol, which

---

[527] Tech Against Terrorism, *Transparency Reporting Guidelines*, TECH AGAINST TERRORISM (2023), https://transparency.techagainstterrorism.org/.
[528] Douek, *supra* note 21 at 48 ("Platforms can drown observers in data while revealing little."); *citing* Sun-ha Hong, *Why Transparency Won't Save Us*, CIGI (18 Feb. 2021), https://www.cigionline.org/articles/why-transparency-wont-save-us; Mike Ananny & Kate Crawford, *Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability*, 20 NEW MEDIA & SOC. 973, 979 (2018); *see also* Nicolas P. Suzor et al., *What Do We Mean When We Talk About Transparency? Toward Meaningful Transparency in Commercial Content Moderation*, 13 INT'L J. COMM. 1526, 1528–29 (2019).
[529] Douek, *supra* note 21 at 48 ("But aggregate enforcement numbers, without more, do not explain relevant denominators or the cause of various trends. For example, when a platform reports an increase in takedowns, it might be intuitive to assume this is because that platform is doing a better job of finding violating content and removing it. But there could be many other reasons: there could be more content overall on the platform; there could be an increase in that kind of content; the platform might have lowered its confidence threshold for removing violating content; the platform might have broadened its definition of violating content; and so on.").
[530] Daphne Keller, *Who Do You Sue? State and Platform Hybrid Power Over Online Speech*, HOOVER AEGIS SERIES PAPER NO. 1902 13 (2019).; *see also* Douek, *supra* note 21 at 47.
[531] Smith and Browne, *supra* note 193 at 154.
[532] Tech Against Terrorism, *Terrorist Content Analytics Platform*, TECH AGAINST TERRORISM (2023), https://www.terrorismanalytics.org/.
[533] *Id.*
[534] Ardern, *supra* note 10.

contributes to an Incident Response Framework.[535] At the beginning of 2023, the GIFCT's crisis response systems and incident management channels had been activated 306 times to monitor and assess incidents in 44 countries. The Content Incident Protocol, which deals with crises that meet strict criteria, had been activated four times, including incidents in Halle, Germany, Glendale, Arizona, Buffalo, New York, and Memphis, Tennessee.[536] Many Call supporters have their own national or regional protocols such as the Europol Protocol. The Call has its own Christchurch Call Crisis Response Protocol, which draws on developments in the wider crisis response landscape. However, since 2022, the Call has been working to map out overlapping systems as many stakeholders have different needs when handling TVEC online.[537] For example, how should crisis response protocols respond to bystander footage, or if different protocols are needed based on regional needs.[538]

C. Future of the Call and Generative Artificial Intelligence

Alongside policymakers worldwide, the Call is turning its attention to the potential benefits and challenges posed by the development of GenAI. GenAI has recently become mainstream as millions of people around the world experiment with products like ChatGPT and Google's Bard. While the technological developments of GenAI are relatively new, the Call's focus on AI is not. One of the Call's initial commitments was to "review the operation of algorithms and other processes that may drive users towards and/or amplify TVEC."[539] This includes designing a multistakeholder process for examining the use of algorithms and automation to remove TVEC.[540] Additionally, in 2019, the tech companies' nine-point plan to implement the Call included work to "accelerate machine learning and AI." From the beginning, the Call anticipated the emerging challenges and opportunities of AI and carved out space to discuss new technologies and TVEC online. In 2022, the Call accelerated this work by launching the CCAIO, which enables accredited researchers to examine algorithmic processes and their impact on radicalization. In recent months, Call Leaders have discussed their desire to further the work the Call has already started on understanding the impact of algorithmic systems on radicalisation and consideration of emergent technologies, including GenAI.[541]

1. What is GenAI?

To understand GenAI, it is helpful to understand that an algorithm is a set of instructions given to a computer or online system that dictates how to transform a set of data into a useful informational output.[542] AI is a process that layers many algorithms and applies software code to teach computers how to understand, synthesise, and generate knowledge in ways similar to the ways in which people do it.[543]

---

[535] Global Internet Forum to Counter Terrorism, *Content Incident Protocol*, GLOBAL INTERNET FORUM TO COUNTER TERRORISM (2023), https://gifct.org/content-incident-protocol/.
[536] *Id.*
[537] *See* Christchurch Call to Action, *Christchurch Call 2022 Community Update, supra* note 461.
[538] *Id.*
[539] The Christchurch Call to Action: *Full English Text, supra* note 11 at 2 ("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.").
[540] *Id.*
[541] Christchurch Call to Action, *Four years of the Christchurch Call*, *supra,* note 494.
[542] Jory Denny, *What is an Algorithm? How Computers Know What to Do with Data*, THE CONVERSATION (17 Oct. 2020), https://theconversation.com/what-is-an-algorithm-how-computers-kno w-what-to-do-with-data-146665.
[543] Marc Andreessen, *Why AI Will Save the World*, ANDREESSEN.HOROWITZ (6 June 2023), https://a16z.com/2023/06/06/ai-will-save-the-world/ ("a short description of what AI *is*: The application of

In recent months, several companies have released AI products that can generate new content through learning patterns from pre-existing data, including text, images, and video.[544] These GenAI products are built from large language models that are trained on an enormous amount of text to recognise patterns in language.[545] While predictive language models have been around since the 1980s, in 2017 Google researchers created a new architecture called transformers, that allowed language models to train on massive data-sets.[546] These 2017 transformer-based language models created a much richer representation of language, but were limited by the lack of computing power available to researchers.[547] As a result, initial models were expensive to build, because they required so much data to function properly.[548] However, once the data is compiled and trained, generating text or other outputs becomes relatively cheap to do and can be fine-tuned for specific tasks.[549] Given the ease of their use, it is hard to accurately forecast how the new technologies will impact content moderation processes, but a few key trends are emerging.[550]

GenAI could both positively and negatively impact the prevalence of TVEC online and its moderation in several ways. First, online platforms already heavily rely on AI models for their content moderation operations, including for the detection of spam, bots, child sexual abuse material, hate speech, TVEC, and other violating content.[551] As companies better integrate GenAI technologies into their content moderation processes, they should get better at finding and removing violating content as well as increase the accuracy of content moderation systems, because AI will be able to more closely replicate human judgment.[552] On the other hand, the widespread availability of GenAI tools will significantly reduce the costs and time it takes for bad actors to develop content.[553] Therefore, even as detection capabilities improve, the bad actors producing harmful content are likely to use GenAI to create content that can evade platform detection tools.[554] Additionally, the widespread availability of GenAI will significantly reduce the costs and time it takes for bad actors to run extensive influence operations

mathematics and software code to teach computers how to understand, synthesize, and generate knowledge in ways similar to how people do it. AI is a computer program like any other – it runs, takes input, processes, and generates output. AI's output is useful across a wide range of fields, ranging from coding to medicine to law to the creative arts. It is owned by people and controlled by people, like any other technology.").

[544] Kristen E. Busch, *Generative Artificial Intelligence and Data Privacy: A Primer*, CONGRESSIONAL RESEARCH SERVICE (23 May 2023), https://crsreports.congress.gov/product/pdf/R/R47569.

[545] *Id.*

[546] Gabriel Nicholas & Aliya Bhatia, Lost in Translation, Large Language Models in Non-English Content Analysis, Center for Democracy and Technology (May 2023), https://cdt.org/wp-content/uploads/2023/05/non-en-content-analysis-primer-051223-1203.pdf.

[547] *Id.* at 13 ("But in 2017, Google researchers released a paper on a new architecture called transformers, which allowed language models to train on lots of data at the same time, in parallel rather than in sequence. These transformer-based language models could ingest so much data simultaneously that they could learn associations between entire sequences of words, not just individual words.").

[548] Busch, *supra* note 544 at 3 ("for example, OpenAI's ChatGPT was built on a large language model that was trained on over 45 terabytes of text data scraped from the internet.").

[549] Nicholas & Bhatia, *supra* note 546.

[550] Tom Cunningham, *The Influence of AI on Content Moderation and Communication*, GITHUB (7 July 2023), https://tecunningham.github.io/posts/2023-06-06-effect-of-ai-on-communication.html.

[551] *Id.* ("AI classifiers are rapidly approaching human-level accuracy for these properties and this means that platforms (and governments) will be able to near-perfectly filter out content that violates their rules, even when content-producers have access to the same technology.").

[552] *Id.*; *see also* Alex Rosenblatt, Swanpeel Mehta, Laila Wahedi, Talha Baig, and Sandeep Abraham, *Unleashing the Potential of Generative AI in Integrity, Trust & Safety Work: Opportunities, Challenges, and Solutions*, THE INTEGRITY INSTITUTE (8 June 2023), https://integrityinstitute.org/blog/unleashing-the-potential-of-generative-ai-in-integrity-trust-amp-safety-work-opportunities-challenges-and-solutions.

[553] Josh A. Goldstein, Girish Sastry, Micah Musser, Renee DiResta, Matthew Gentzel, & Katerina Sedova, *Generative Language Models and Automated Influence Operations: Emerging Threats and Potential Mitigations*, GEORGETOWN CENTER FOR SECURITY AND EMERGING TECHNOLOGY, OPENAI, & STANFORD INTERNET OBSERVATORY (10 Jan. 2023), https://arxiv.org/abs/2301.04246.

[554] *Id.*

online.[555] As a result, it will be much easier to manipulate and synthesise media, which will make it harder for people to discriminate between real and fake media.[556] Therefore, GenAI is likely to improve the tools available for both the detection and creation of harmful content.

Another area in which GenAI could have both positive and negative impacts is in training content moderation systems to better understand local languages and contexts. The lack of non-English datasets remains one of the biggest challenges for content moderation systems because, without sophisticated classifiers, automated tools struggle to understand local contexts. While larger companies may hire teams of specialists with language expertise, smaller companies cannot hire moderators fluent in multiple languages.[557] To overcome this challenge, GenAI could assist in the creation of synthetic datasets to help train content moderation classifiers in non-English languages.[558] These generated datasets could fill in linguistic gaps and improve classifiers, which would increase the quality of content moderation and slow the proliferation of harmful content online.[559] However, these generated datasets need to be carefully trained and overseen by humans. If not, GenAI could have a negative impact because the large language models can have built-in biases which could undermine many human rights protections.[560] Therefore, it is necessary to build guardrails around this technology and establish norms. Multistakeholder forums offer promise for doing this; done well, they should enable the integration of the diverse perspectives needed to make this a safer process.

2. What is the Impact of GenAI on TVEC?

According to Brian Fishman, a terrorism expert previously employed by Meta, tech companies have been using automation and AI for years to moderate TVEC in both simple and complex ways.[561] Simple automation is used in technologies like GIFCT's hash-sharing database, which matches static information to identify TVEC online.[562] Complex automation, powered by AI, is used to build sophisticated text classifiers that can assess new material and determine the likelihood of it being TVEC.[563] Complex AI processes will become far more sophisticated, and GenAI could help create

[555] *Id.* at 8 ("Language models could drive down the cost of running influence operations, placing them within reach of new actors and actor types. Likewise, propagandists-for-hire that automate production of text may gain new competitive advantages.)".
[556] *Id.* at 8 ("Recent AI models can generate synthetic text that is highly scalable, and often highly persuasive. Influence operations with language models will become easier to scale, and more expensive tactics (e.g., generating personalized content) may become cheaper. Moreover, language models could enable new tactics to emerge—like real-time content generation in one-on-one chatbots.").
[557] Thorley & Saltman, *supra* note 149 at 7 ("Terrorist content is shared in a wide array of languages, and while larger tech platforms have the capacity to employ specialist teams with subject matter and language expertise, most companies have comparatively small moderation teams to review content and very few linguists with the appropriate mix of global dialects.").
[558] Nicholas & Bhatia, *supra* note 546, at 37 ("At once, companies are increasingly deploying multilingual language models to bridge the gap between the functionality in English and other languages across a myriad of tasks, such as harmful content detection, sentiment analysis, and content scanning. However, as we show in this paper, these multilingual systems are relatively new and perform inconsistently across languages.").
[559] *Id.*
[560] *Id.* at 6 ("Large language models' general use in content analysis raises further concerns. Computational linguists argue that large language models are limited in their capacity to analyze forms of expression not included in their training data, meaning they may struggle to perform in new contexts. They may also reproduce any biases present in their training data. Often, this text is scraped from the internet, meaning that large language models may encode and reinforce dominant views expressed online.").
[561] Fishman, *supra* note 24.
[562] *Id*. ("Simple automation matches static information to identify problematic content or patterns. This includes keyword searches, hash-matching, and various rule-based detection schemes. Sometimes these systems are extremely effective, especially when combined with intelligence collection and sharing.").
[563] *Id.* ("Complex automation, however, requires building sophisticated classifiers that not only match known bad content but also can assess novel material and determine the likelihood that it violates some predetermined rule. Using such tools to achieve policy ends is an art in itself—and in that way, social media companies are canaries

variants of known pieces of violating content and block their upload.[564] For example, one reason the Christchurch shooter's video was so hard to remove is because sympathetic extremists regularly altered versions of it, often only slightly, to evade hash-based detection systems. Using GenAI, a computer could create variants and hash them for automated detection systems. However, to protect speech and human rights, these variants should be checked before they are automatically added to hash-sharing databases. Overall, GenAI is likely to improve detection of TVEC, increase the speed and effectiveness of human processes, and provide more transparency to users.[565]

GenAI can also compound the problem of moderating TVEC online by making it easier for bad actors to create content that is more appealing to vulnerable groups which could lead to an increase in radicalisation. GenAI will drive down the costs of running TVEC influence operations by automating the production of propaganda used to radicalise extremists.[566] Additionally, GenAI could help make TVEC more compelling and persuasive by generating individualistic messages which include specific linguistic and cultural context.[567] Furthermore, GenAI could decrease the cost of recruitment by deploying GenAI chat bots that target vulnerable persons through one-on-one conversations in online environments.[568] Finally, GenAI could help influence operations to avoid detection by hash-sharing databases as they would no longer need to use copy-pasted messaging.[569] These significant risks will require technologists to work with civil society, governments, and companies to deploy safeguards and establish norms to prevent further radicalisation campaigns online.

3. Options for the Call to Address the Impact of GenAI on TVEC

A curated MSI brings together governments, companies, and civil society to address problems and propose solutions when new technologies are likely to have a profound impact on society. GenAI creates new 'tools and weapons' in the effort to combat TVEC online and the Call is strategically positioned to support solutions for problems GenAI may create as it relates to the proliferation of TVEC online.[570] The Call could tackle these challenges by expanding its ongoing efforts or by slightly restructuring its curated MSI approach. Indeed, as part of their 2022 Leaders' Summit, the Call recognised the importance of addressing new technology issues as they relate to the Call's 25 commitments, and that the Call model might assist with this work.[571] To fulfill this goal, the Call created a "New Tech" workstream, which brings together the Call's multistakeholder community to support the adoption of new technologies while promoting safety and securing against TVEC.[572]

---

in the coalmine for lawmakers and bureaucrats around the world who will increasingly need to both set policy constraining the use of AI and establish guidelines for implementing policy via AI.").

[564] Cunningham, *supra*, note 550 ("The prevalence of variations of known-violating content will decrease. E.g. content that is a match against databases of illegal sexual media (PhotoDNA), IP-protected content (ContentID), or terrorist recruitment content (GIFCT). Obfuscation will become harder as AI models get better.").

[565] Rosenblatt et al., *supra* note 552.

[566] Goldstein et al., *supra* note 553 at 3 ("For malicious actors looking to spread propaganda—information designed to shape perceptions to further an actor's interest—these language models bring the promise of automating the creation of convincing and misleading text for use in influence operations, rather than having to rely on human labor.").

[567] *Id.* at 4 ("Generative models may improve messaging compared to text written by propagandists who lack linguistic or cultural knowledge of their target.").

[568] *Id.*

[569] *Id.* at 4; *noting* that propaganda will become less discoverable because ("[e]xisting campaigns are frequently discovered due to their use of copy-and-pasted text (copy-paste), but language models will allow the production of linguistically distinct messaging.").

[570] Ardern, *supra* note 10.

[571] Ardern & Macron, *supra* note 484.

[572] *Id.* at 4 ("Launch a new stream of work to understand how we can support the adoption of new technologies while promoting safety and securing against terrorist and violent extremist content.").

This workstream is considering a range of issues, including the development of immersive, augmented and virtual reality environments, the impact of the decentralised web, the use of new AI tools, and how terrorist and violent extremists use gaming platforms. A second area where the Call could expand its work to address GenAI is through its Algorithms and Positive Interventions Workstream, where the Call prioritises action to better understand the impacts that algorithms and other processes may have on TVEC.[573] Through this workstream the Call could explore ways to improve research insights into GenAI that could provide technical, political, and social assurance for governments, companies, and users.

Additionally, the Call could expand the work of the CCIAO to research how GenAI will impact the distribution of TVEC online. One way to do this would be to empower researchers to use the CCIAO to test safety features and develop guardrails for GenAI. In this way, the CCIAO could act as a tool to allow researchers to experiment with new products in a controlled setting. In the area of technology governance, this type of environment is frequently referred to as a "sandbox." In recent years, many stakeholders have deployed developmental sandboxes when experimenting with new technology, as they provide a conducive, contained space where governments, companies, civil society and other stakeholders can test technologies before launching them at scale.[574] Additionally, a sandbox would provide a controlled environment for stakeholders to work together to develop technologies in a responsible and ethical way.[575] A CCIAO development sandbox could have four key functions. First, researchers could study how users are exposed to TVEC and how a person could be redirected or otherwise disengaged from TVEC using GenAI. Second, researchers could explore the accuracy of the systems detecting and removing TVEC and concerns around bias. Third, researchers could test ways to create a healthier, safer online information environment that reduces radicalization and the risks of harms relating to TVEC. Finally, this sandbox could help foster multistakeholder solutions that support human rights and a free, open, secure internet. This project would leverage the existing work of the Call and provide a sustainable solution to addressing new and emerging technologies.

As explored above, the moderation of TVEC online is an area where there is a broad consensus among stakeholders on what should and should not be allowed online in line with human rights principles. Additionally, stakeholders are highly motivated to find solutions to the problems created by TVEC online because it can lead to offline violence. Moreover, for stakeholders considering how to moderate GenAI content online, starting with a (relatively) uncontroversial type of content like TVEC can provide a framework for other areas of content moderation. The Call could bring its multistakeholder approach to the GenAI and rapidly scale up. Additionally, as discussed above, one of the greatest threats from the development of GenAI is the potential to radicalise individuals towards terrorist and violent extremism. Therefore, the Call should consider ways to deploy its resources and scale up its impact on policy governance relating to the development of GenAI.

---

[573] Christchurch Call to Action, *Christchurch Call Initiative on Algorithmic Outcomes, supra* note 488.
[574] United Nations Department of Economic and Social Affairs, *Sandboxing and experimenting digital technologies for sustainable development*, UNITED NATIONS FUTURE OF THE WORLD POLICY BRIEF NO. 123 (Dec. 2021), https://www.un.org/development/desa/dpad/wp-content/uploads/sites/45/publication/PB_123.pdf.
[575] Wolf-Georg Ringe, *Why We Need a Regulatory Sandbox for AI*, UNIVERSITY OF OXFORD FACULTY OF LAW BLOGS (12 May 2023), https://blogs.law.ox.ac.uk/oblb/blog-post/2023/05/why-we-need-regulatory-sandbox-ai ("A regulatory sandbox promises a number of advantages. First, it promotes innovation: AI is a rapidly evolving technology, and the regulatory environment has struggled to keep up. A sandbox allows for the development of new AI technologies in a controlled environment reducing the risk of violating laws or regulations. This has proven to reduce the so-called 'time to market' for innovations, giving new businesses increased legal certainty and thereby leading to more innovation.").

D. The Future of the Call, Building a Self-Sustaining MSI and Addressing GenAI

As the Call takes stock of its work over the past four years and considers ways to expand its mandate, this report concludes by applying the framework created in Part II to the Call. The Call has already firmly established its purpose and mission, but there are steps it can take to ensure its longevity and sustainability. The Call was set up in a matter of weeks, developed processes over time, and achieved a great deal. However, the Call should consider formalizing some of these processes as it looks to expand its work on GenAI. Given the tremendous possibilities for expanding work streams and development of the CCIAO, the Call may want to accelerate its work with additional funding and staffing resources. This could be done within the existing work of the Call. But Call leaders may wish to restructure to ensure that it can receive additional funding to expand impact and scale solutions. This section explores how the Call could develop to build a self-sustaining MSI and, potentially, play a larger role in governance of GenAI issues related to TVEC online. Should any restructured MSI be considered as part of this process, it will be important to ensure objectives are properly scoped, an appropriate set of stakeholders is involved, clear terms of reference are established, and funding options are explored.

1. Step 1: Decide if an MSI is Necessary

The Call's current governing documents clearly outline why an MSI was necessary in the wake of 15 March 2019 to find solutions to eliminate TVEC online while promoting the free, open, and secure internet. Therefore, Step 1 looks the necessity of an MSI to find solutions to problems related to GenAI and TVEC online.

1(a). Define the Problem

As described above, GenAI could have a profound impact on distribution of TVEC online. It has the potential to significantly reform content moderation practices, by improving classifiers, which will more accurately remove TVEC online. However, GenAI could also super-charge the creation of TVEC and allow terrorist and violent extremists to mass produce more engaging and dangerous content. Stakeholders need to examine how to increase the benefits of GenAI while ensuring the technology does not create undue harm. The challenge will be balancing the positive aspects of GenAI that support innovation in the area while also creating safeguards to minimise the harm. Clearly, efforts to address this problem could benefit from multistakeholder input from governments, technologists, civil society, terrorism experts, and a wide range of actors.

1(b). Map the Landscape

Over the past six months, GenAI has dominated conversations related to technology policy around the world. There is consensus that development of GenAI will bring a wide range of societal challenges, from protecting copyright owners' intellectual property to combating an increase in fraud. No one is quite sure how to address these challenges because the technology's future application is unknowable, and regulation may impede technological advancement. However, that has not stopped a multitude of stakeholders, including tech companies, governments, and civil society, from proposing ideas to provide guidance on development and safety.[576]

[576] Sam Altman, Greg Brockman, & Ilya Sutskever, *Governance of Superintelligence*, OPENAI (22 May 2023), https://openai.com/blog/governance-of-superintelligence ("we need some degree of coordination among the leading development efforts to ensure that the development of superintelligence occurs in a manner that allows us to both maintain safety and help smooth integration of these systems with society. There are many ways this could be implemented; major governments around the world could set up a project that many current efforts

To better understand the current landscape, it is helpful to map current stakeholder proposals. First, there are existing MSIs that have been examining AI more broadly for years, including the OECD, which published AI principles;[577] the Global Partnership on AI;[578]and the industry-led Partnership on AI.[579] Next, several AI companies (big and small) have presented various blueprints for governing GenAI from the perspective of the industry, including OpenAI's Sam Altman,[580] Microsoft's Brad Smith,[581] and Google's Sundar Pichai.[582] Despite the many proposed options, the industry appears divided on whether it should speed up or slow down development of AI.[583] Additionally, various individuals, governments, and multilateral bodies have floated proposals in the past six months to address the challenges posed by GenAI, including:

- the UN Secretary-General, who has backed a proposal to create an international AI watchdog body like the International Atomic Energy Agency[584]

- the British Prime Minister Rishi Sunak, who has supported the idea of Britain being home to global AI safety regulation and plans to host a summit this year on how coordinated international action can tackle the risks of AI[585]

- G7 leaders, who have published a "Hiroshima process" which includes a cabinet-level discussion between G7 countries on AI governance; a report will come out later this year[586]

- the EU, which crafted legislation for regulating AI two years ago and is now working through the final text of its AI Act, which now has draft language related to GenAI.[587] The EU AI regulation includes creation of industry codes and participation in global MSIs

- the EU and US governments, which have worked together through their Trade and Technology Council on an AI Code of Conduct, to be a set of voluntary standards on GenAI until legislation is enacted[588]

---

become part of, or we could collectively agree (with the backing power of a new organization like the one suggested below) that the rate of growth in AI capability at the frontier is limited to a certain rate per year.").
[577] *See* OECD.AI, *OECD AI Principles overview*, OECD.AI POLICY OBSERVATORY (May 2019), https://oecd.ai/en/ai-principles.
[578] *See Home*, THE GLOBAL PARTNERSHIP ON ARTIFICIAL INTELLIGENCE (2023), https://gpai.ai/.
[579] *See Our Work*, PARTNERSHIP ON AI (2023), https://partnershiponai.org/work/.
[580] Sam Altman, *Written Testimony before the U.S. Senate Committee on the Judiciary Subcommittee on Privacy, Technology, & the Law*, SEN. COMM. JUDICIARY SUBCOMM. PRIVACY, TECHNOLOGY & THE LAW (16 May 2023), https://www.judiciary.senate.gov/imo/media/doc/2023-05-16%20-%20Bio%20&%20Testimony%20-%20Altman.pdf.
[581] Brad Smith, *How do we best govern AI?,* MICROSOFT ON THE ISSUES (25 May 2023), https://blogs.microsoft.com/on-the-issues/2023/05/25/how-do-we-best-govern-ai/.
[582] Jennifer Elias, *Google CEO Sundar Pichai warns society to brace for impact of A.I. acceleration, says 'it's not for a company to decide*', CNBC (17 April 2023), https://www.cnbc.com/2023/04/17/google-ceo-sundar-pichai-warns-society-to-brace-for-impact-of-ai-acceleration.html.
[583] Future of Life Institute, *Pause Giant AI Experiments: An Open Letter*, FUTURE OF LIFE INSTITUTE (22 March 2023), https://futureoflife.org/open-letter/pause-giant-ai-experiments/; *noting* as of 15 July 2023, the letter contained 33,002 signatures.
[584] Michelle Nichols, *UN chief backs idea of global AI watchdog like nuclear agency*, REUTERS (13 June 2023), https://www.reuters.com/technology/un-chief-backs-idea-global-ai-watchdog-like-nuclear-agency-2023-06-12.
[585] Natasha Lomas, *UK's AI safety summit gets thumbs up from tech giants*, TECHCRUCH (8 June 2023), https://techcrunch.com/2023/06/08/uks-ai-safety-summit-gets-thumbs-up-from-tech-giants/.
[586] Marika Katanuma*, G-7 Leaders Agree to Set Up 'Hiroshima Process' to Govern A*I, BLOOMBERG (20 May 2023), https://www.bloomberg.com/news/articles/2023-05-20/g-7-leaders-agree-to-set-up-hiroshima-process-to-govern-ai#xj4y7vzkg
[587] Natasha Lomas, *EU lawmakers eye tiered approach to regulating generative AI*, TECHCRUCH (22 April 2023), https://techcrunch.com/2023/04/21/eu-ai-act-generative-ai/.
[588] Natasha Lomas, *EU and US lawmakers move to draft AI Code of Conduct fast*, TECHCRUCH (1 June 2023), https://techcrunch.com/2023/05/31/ai-code-of-conduct-us-eu-ttc/.

- in the US, both the White House[589] and the Department of Commerce's NTIA have started consulting stakeholders on how to boost AI accountability.[590] Additionally, US Senate Leader Chuck Schumer has outlined a process to address GenAI through regulatory solutions[591]

- over 19 African governments that have started developing national strategies on AI, including by establishing commissions, setting up task forces, undertaking public sector reform, and building capacity.[592]

However, none of these efforts combine technical solutions with a multistakeholder framework that could specifically address the impacts of GenAI on TVEC online. While there are many proposals out there relating to how companies should moderate GenAI content online, addressing issues raised in a relatively discrete problem-set, such as TVEC online, could serve as a model for other types of content. Therefore, the Call community could fill a gap in the current landscape and kick off a broader conversation. While many institutions are working on ethical principles and proposals for oversight of GenAI, few are linking theories to practice in the way the Call could through expansion of its work streams and further development of the CCIAO. Additionally, in the proposals recently floated it is difficult to discern efforts to connect with a global multistakeholder community to address the problem holistically, rather than regionally or nationally, or form the perspective of one sector or another.

1(c). Question whether an MSI is the Best Approach

As established in best practices set forward by the Internet Society, an MSI is the best approach when:
- decisions impact a wide and distributed range of people and interests
- there are overlapping rights and responsibilities across sectors and borders
- different forms of expertise are needed, such as technical expertise
- the legitimacy and acceptance of decisions directly impact implementation.[593]

All four factors are clearly present when considering the issues surrounding GenAI and TVEC online. First, the spread of TVEC online is a problem which is distributed globally and affects stakeholders in every country as it presents a wide range of challenges, from national security concerns to the mental health of victims.[594] Second, as this report discussed in Part I, moderating user-generated content on online platforms creates a tangled web of overlapping rights and responsibilities. To ensure these rights and responsibilities are upheld, companies and governments need to work alongside stakeholders from a wide range of sectors and backgrounds to strike the right balance. This challenge will only accelerate with the use of GenAI. Third, to address the proliferation of TVEC online, stakeholders need to look to a variety of experts including human rights lawyers, technologists, community advocates, and government regulators. Finally, because GenAI can dramatically impact how TVEC is moderated

[589] Ashley Gold, *White House pushes forward on "responsible AI,"* AXIOS (24 May 2023), https://www.axios.com/pro/tech-policy/2023/05/23/white-house-pushes-forward-on-responsible-ai.
[590] NTIA Office of Public Affairs, *NTIA Seeks Public Input to Boost AI Accountability,* NATIONAL TELECOMMUNICATIONS AND INFORMATION ADMINISTRATION (11 April 2023), https://www.ntia.gov/press-release/2023/ntia-seeks-public-input-boost-ai-accountability.
[591] Reuters, *U.S. Senate leader schedules classified AI briefings*, REUTERS (7 June 2023), https://www.reuters.com/technology/us-senate-leader-schedules-classified-ai-briefings-2023-06-06/.
[592] Jake Okechukwu Effoduh, *Seven ways that African states are legitimizing artificial intelligence*, OPENAIR AFRICA INNOVATION RESEARCH (https://openair.africa/7-ways-that-african-states-are-legitimizing-artificial-intelligence/.
[593] Internet Society, *Internet Governance: Why the Multistakeholder Approach Works, supra* note 215 at 2.
[594] *See* Christchurch Call to Action, *Christchurch Call Second Anniversary Summit Co-Chair Statement 2021, supra* note 483, *noting* Christchurch-inspired attacks including in Poway, El Paso, Dayton, Halle, Glendale, Nakhon Ratchasima, Nice, and Vienna.

online, a multistakeholder approach will help to build legitimacy and acceptance of the solutions proposed.

Much of this approach was encapsulated by former Prime Minister Jacinda Ardern in her June 2023 op-ed:

> I see collaboration on AI as the only option. The technology is evolving too quickly for any single regulatory fix. Solutions need to be dynamic, operable across jurisdictions, and able to quickly anticipate and respond to problems. There's no time for open letters. And the government alone can't do the job; the responsibility is everyone's, including those who develop AI in the first place. Together, we stand the best chance to create guardrails, governance structures and operating principles that act as the option of least regret. We don't have to create a new model for AI governance. It already exists, and it works.[595]

2. Step 2: Establish the Objectives and Functions of the MSI

2(a). Set Objectives

The objectives of the Call are outlined in the 25 commitments that governments and companies adopted. While partner organisations and CCAN are not formal supporters of the Call commitments, they agree to support the objectives of the Call and delivery of its commitments.[596] One reason the work of the Call has made a positive impact towards eliminating TVEC online while supporting a free, open, and secure internet is because it established clear objectives through the Call commitments and supporters publicly committed to implementing them.[597] This type of accountability created common principles that guided future policy planning, through Leaders Summits and work streams, and supported a bottom-up multistakeholder approach.[598] Any expansion or restructuring of the Call's work to address GenAI should be similarly based upon clear objectives, to create accountability and facilitate broader policy development across industry. Work related to GenAI and TVEC should be based closely onthe objectives set out in the Call text that provide for work on new technologies.

2(b). Determine the Function

MSIs can perform a wide variety of functions; these may evolve over time, depending on strengths and opportunities. Three overarching approaches for the function of an MSI including standard-setting, policy-oriented initiatives, or project-oriented initiatives.[599] The function of the Call is a mix of a policy-oriented initiative and a project-oriented initiative. The Call commitments provide the set of objectives, and its function has focused on convening stakeholders during the Leaders' Summits and working-level multistakeholder workstreams between summits. The Call's main function has been to help facilitate

---

[595] Ardern, *supra* note 10.
[596] Christchurch Call to Action, *Our Community – Partners, supra* note 507 ("A partner is an organisation that is not a government or online service provider but is committed to supporting the objectives of the Call and contributing to the realisation of the Call commitments through their work.").
[597] Douek, *supra* note 21 at 73; *noting* that positive developments in content moderation practices happen when companies ("hav[e] an ongoing planning and review process would allow for this consultation to become more consistent, transparent, and prospective, with stakeholders engaging on an iterative basis at each review rather than only on an ad hoc basis."); *citing* Robin Kundis Craig & J. B. Ruhl, *Designing Administrative Law for Adaptive Management,* 67 VAND. L. REV. 1, 43 (2014) ("formulation of the plan itself is clearly an adaptive management moment that lends itself to public input").
[598] *See* Christchurch Call to Action, *Our Work: Leaders' Summits, supra* note 485.
[599] Gleckman, *supra* note 209.

large-scale collaboration that can drive bottom-up efforts by all stakeholders.[600] Many MSIs evolve over time to meet new challenges, and the Call has demonstrated its ability over the years to respond to a developing technology landscape.[601] To address GenAI issues through the expansion of current work, or restructuring the MSI, the Call should maintain a policy- and project-oriented function, with a focus on convening, which is foundational to the work of the Call. Resisting the temptation to become a standard-setting initiative will help differentiate the work of the Call from many of the regulatory initiatives that many are considering.

3. Step 3: Decide who is a stakeholder

To increase the self-sustainability of the Call and to expand its work, the Call needs to prioritise bringing together the right group of stakeholders. The Call's formal supporters now include 56 governments, representing virtually all the world's liberal democracies and 14 online service providers: Amazon, Meta, Google, YouTube, Zoom, DailyMotion, Microsoft, Qwant, JeuxVideo, Line, Twitter, Roblox, Mega, and Clubhouse. Additionally, the Call has four partners: Council of Europe, Global Community Engagement and Resilience Fund, TAT, and UNESCO. Finally, CCAN has 46 members.[602] Despite its wide range of stakeholders, the Call lacks a strong focus on GenAI experts and technologists. Diversifying and increasing the number of supporters of the Call with this background will be necessary before the Call can expand its work on GenAI and build legitimacy for its proposed solutions.

3(a). Criteria to Join

Like most curated MSIs, the Call sets different criteria to join for different stakeholder types. To become more sustainable and address GenAI issues, the Call will need to rapidly on-board new stakeholders. It should therefore update its criteria to better reflect current practices. For governments, companies, and partners the process to join is the same and consists of an application to the Call Secretariat and a formal disclosure of baseline information. [603] Next, that information is assessed to consider the suitability of the applicant to join the Call community, and advice is provided to members of the community. [604] Should the application proceed to the next stage, the potential supporter or partner meets with the Call community for a 75-minute virtual community engagement session where the applicant discusses their work and take part in a question and answer session with Call participants, in which the applicant is able to ask and answer questions.[605] This practice aligns with the objective of the Call, which stresses transparency and multistakeholder collaboration.

However, the criteria each type of stakeholder group must meet to apply is slightly different.

- *Government stakeholders*. Governments must commit to protecting a free, open, and secure internet. This requirement has been a strength to the initiative because some of the biggest challenges surrounding TVEC come when authoritarian governments seek to label dissidents as "terrorist" or "violent extremist", in the absence of human rights-respecting criteria. One way many curated MSIs verify this commitment is by requiring membership of groups such as

[600] Ardern, *supra* note 10.
[601] Strickling & Hill, *supra* note 205 at 46.
[602] Christchurch Call to Action, *Our Community – Partners, supra* note 507.
[603] Christchurch Call to Action, *Our Community, Joining the Christchurch Call Community*, CHRISTCHURCH CALL (May 2023), https://www.christchurchcall.com/our-community/joining-the-christchurch-call-community/.
[604] *Id.*
[605] *Id.*

the Freedom Online Coalition.[606] Given the large number of governments currently supporting the Call, recruitment in this area does not appear to be necessary, and the criteria does not need updating.

- *Company stakeholders*. Companies must comply with the Call's commitments which could be a barrier to entry for many smaller companies, which may not have the resources to do extensive public reporting. However, the Call community has worked closely with smaller companies throughout the on-boarding process to help encourage an increase in company membership. As the Call begins to recruit companies with expertise in GenAI, it may want to examine how these companies provide public reporting about their products in less traditional ways. For example, a company like Anthropic uses a "Trust Portal", which would not neatly map onto company commitments, but does encapsulate the objective of transparency.[607] Call leaders should work collaboratively with interested companies to refine or interpret these terms in ways that facilitate wider participation.

- *Partner stakeholders*. The partner category includes organizations that are not companies or governments and do not meet the criteria for CCAN membership or prefer to participate in the Call in their own right. This category has potential for building new relationships with trade associations, research projects, and academic institutions.

Regarding the participation of civil society, technical experts and academia, the Call has two approaches: one formal and one informal. The formal process is through CCAN, which has its own criteria and processes for inclusion of new members, codified in 2022 terms of reference.[608] While many of the underlying criteria for joining CCAN overlap with the Call's broader objectives, CCAN also has created several eligibility requirements for interested participants. These requirements include that applicants cannot receive more than 25 per cent of their operating budget from governments and corporations and that their work is not "strongly influenced by a government or private sector company."[609] These requirements are limiting for many stakeholders including academic and research institutions that would have otherwise wanted to support the Call. As a result, the Call informally consults a broader group of civil society organisations through its workstreams.[610] If CCAN plans to keep these restrictions, the Call should consider on-boarding more stakeholders as partners to ensure wider participation. Ideally, all stakeholders would be formally affiliated with the Call, to increase the legitimacy of the MSI.

3(b). Ensuring Inclusivity

For multistakeholderism to be successful, an initiative needs to include a diverse group of stakeholders working on solving the problem defined in Step 1. A curated MSI must balance the desire to have the largest coalition possible with one that ensures key stakeholder can find consensus to drive solutions. The Call has brought together an inclusive group of government supporters who support a free, open,

[606] Freedom Online Coalition, *Get Involved – Government Members*, FREEDOM ONLINE COALITION (2023), https://freedomonlinecoalition.com/government-members/ ("Aspiring members' applications are assessed based on their domestic record when it comes to respecting human rights online, the countries voting record in international fora on Internet freedom issues, and the degree to which the country takes a proactive role on furthering Internet freedom in its foreign policy. The Coalition further looks to ensure wide geographical representation.").
[607] Anthropic, *Trust Report*, ANTHROPIC (2023), https://trust.anthropic.com/.
[608] Christchurch Call Advisory Network, *Terms of Reference, supra* note 503.
[609] *Id.*
[610] Ardern & Macron, *supra* note 484.

and secure internet. However, when it comes to being more inclusive of company, partner, and civil society stakeholders, the Call still has work to do – particularly if it is going to address GenAI issues.

Here are five ways the Call could improve outreach to ensure inclusivity and add stakeholders to contribute to the work on GenAI:

- *Streamlining process*. The process established by the Call for joining is critical to maintaining openness and building trust. However, in practice, the process for joining the Call can take months, in part due to scheduling calls and meetings. To ensure scalability, the Call should find ways to expedite the member on-boarding process by batching together applicants or focusing more on the written responses.
- *Dedicated staffing*. There is a lot of interest in supporting the work of the Call, but this has not always translated to formal commitments. To ensure potential supporters are on-boarded, the Call should designate someone within the Secretariat to lead recruitment efforts and rapidly expand the organization.
- *Social media presence*. The Call website provides relevant information on how to join and stakeholder expectations. However, many potential stakeholders want to understand the impact Call supporters have on the broader discussion around TVEC online. Creating a social media presence could help prospective participants become more familiar with the work of the Call and more comfortable with joining.
- *More linguistic diversity*. The Call's working language is English, and it also publishes its documents in French. To recruit a more diverse group of stakeholders, the Call should explore translating documents into other languages, particularly Spanish and Arabic to bring onboard stakeholders from Latin America and the Middle East. While this can be an expensive undertaking for any MSI, it will help expand global reach.
- *Funding stakeholders*. To ensure a wider diversity of participation, the Call should consider pooling resources to help support costs of civil society members.[611]

3(c). Meaningful Representation

It is essential to prioritise meaningful representation from all stakeholder categories within an MSI to build legitimacy. To ensure meaningful representations, there are several groups the Call should continue to prioritise for recruitment:

- *Global South countries*. As mentioned above regarding linguistic capabilities, the Call does not have many supporters from industry or civil society in Latin America, the Middle East or Africa. These regions are crucial to the discussion of TVEC online. In the Middle East, ISIS and other terrorist organisations continue to exploit online platforms to spread their message. In Latin America, countries have experienced a rise in terrorism and violent extremism, including an attack on the Brazilian capital building. Additionally, inclusion of African voices

[611] Strickling & Hill, *supra* note 205 at 50 ("Multi-stakeholder processes are generally quite resource-intensive, both in terms of time and money. A single initiative focusing on a specific policy issue can take months from start to completion. Many multistakeholder organizations hold multiple meetings a year, often in far-flung places across the globe. For stakeholders with limited resources, in-person attendance can be prohibitively expensive. While most venues try to provide remote participation opportunities for stakeholders who are unable to travel, there is a sense that stakeholders who participate in person can have more impact on the group decision than those who engage remotely.").

could help steer governments away from more problematic attempts to censor or manipulate online discussion.[612]

- *Smaller companies*. The Call is working closely with TAT and others to address this gap through current workstreams. Bringing on smaller companies will be necessary to properly address GenAI issues.

- *LGBTQIA+*. As with the smaller companies, the Call is working to be more inclusive in this space. Conveners should continue to prioritise this work as they look to address the drivers of radicalisation.

10. Step 4: Set up Terms of Reference

Setting up clear terms of reference should be a top priority for the Call as it considers restructuring the MSI or expanding its work to address GenAI and TVEC. The unprecedented speed with which the Call was established meant that many of structural elements were retrofitted to the Call commitments. These processes have developed into a de facto term of reference through Leaders' Summits and consensus building. Curated MSIs do not need overly complicated terms of reference; however, a formalised agreement will help the organisation scale and build legitimacy. Additionally, if the Call is to restructure its MSI to accept outside funding, establishing clear terms of reference will help the organization comply with tax laws and attract investment. This section will explore the evolution of the de facto practices and how the Call could codify them.

4(a). Leadership and Agenda Setting

Call leadership resides within the Call Secretariat and is led by the Prime Minister's Special Representative on Cyber and Digital (Paul Ash) in Aotearoa New Zealand, and the Ambassador for Digital Affairs (Henri Verdier) in France, who work with their respective teams on the Call.[613] Because there has been continuity in these roles, and Ash and Verdier have been dedicated to the work over the past four years, and there has been no need to determine who comes next. However, to ensure longevity and sustainability of the Call, the terms of reference should design a process for choosing and appointing new leaders. This is particularly important where MSIs are led by democratic government actors who could face electoral changes.

In addition to leadership, the Secretariat sets the Call's agenda. The Call's leaders "work closely with civil society representatives, government officials, partner organisations and online service providers across the Call Community, reflecting our commitment to the Call's multistakeholder approach."[614] Agendas for Leaders' Summits fall to the Call Secretariat to develop. Over the years, they have been set in different ways, but overwhelmingly they have been crafted by consensus through an informal process. Terms of reference could better articulate this process, which would provide legitimacy to community-wide decisions and ensure inclusion of all voices.

---

[612] Vincent Obia, *What can African countries do to regulate artificial intelligence?*, LONDON SCHOOL OF ECONOMICS BLOG (13 June 2023), https://blogs.lse.ac.uk/medialse/2023/06/13/what-can-african-countries-do-to-regulate-artificial-intelligence/?ref=everythinginmoderation.co.
[613] Christchurch Call to Action, *The Secretariat*, CHRISTCHURCH CALL TO ACTION (2023), https://www.christchurchcall.com/about/the-secretariat/.
[614] *Id.*

4(b). Decision-making Authority

The Call is a curated MSI because it makes decisions by consensus. This type of decision-making strengthens the credibility of potential solutions by ensuring they benefit the greatest number of stakeholders.[615] The Call's decision-making happens in the lead-up to Leaders' Summits, where the community decides which areas it will prioritise over the coming years.[616] The Call currently lacks specific terms of reference, and therefore does not identify how the community would make decisions if there was no consensus. This has not yet been a problem and the informality has helped build trust between stakeholders when working on joint statements. However, terms of reference could set out working methods for reaching consensus which may be necessary if disagreements between stakeholders arise in the future. One possibility for the Call's terms of reference would be to mirror CCAN's, which state decisions are made by "modified consensus", and that if that cannot be reached, members take a vote.[617] The Call community may not want the same structure as that of CCAN, but formalisation of process will help ensure trust between stakeholders if (and when) decisions are contentious.

4(c). Transparency

Terms of reference for the Call should address external and internal transparency. Currently, the Call's external transparency is found on its website, which provides access to a range of documents and extensive reporting on the work of the Call.[618] The public-facing documents make clear what the Call has accomplished, its working methods, and its plans. As a result, external transparency has largely been demonstrated through the many public-facing documents the Call has produced over the years. Terms of reference could codify these best practices and confirm that all information is easily accessible to stakeholders.

Additionally, the terms of reference could help establish better working methods to create more internal transparency. Internal transparency between stakeholders is necessary to build trust and cohesion – particularly when stakeholders can be inherently suspicious of each other.[619] For the Call, open communication can be challenging to maintain, as trust in technology companies is low, national security concerns are high, and human rights interests are paramount. The terms of reference could help foster internal transparency by mandating that stakeholders are given access to working documents, notification on when meetings are happening, and opportunities to contribute to the decision-making process. The Call could facilitate this work through the creation of a community platform which could serve as a portal for Call community members to securely access this information and share ideas. The Call's current website lacks the capability to host this portal, but further technological solutions should be explored and implemented.

4(d). Accountability

The legitimacy of an MSI is directly tied to its accountability of the stakeholders, and therefore the terms of reference should specify accountability mechanisms.[620] The Call has demonstrated the

[615] Internet Society, *Internet Governance: Why the Multistakeholder Approach Works, supra* note 215 at 2.
[616] Ardern & Macron, *supra* note 484.
[617] Christchurch Call Advisory Network, *Terms of Reference, supra* note 503.
[618] Christchurch Call to Action, *Our Community, Joining the Christchurch Call Community, supra* note 603.
[619] For a full discussion of how to create internal transparency to build trust between stakeholders *see* Baumann-Pauly et al., *supra* note 32 at 22; van Huijstee, *supra* note 323 at 57-58; The Stanley Center, *supra* note 324 at 6.
[620] The Stanley Center, *supra* note 324 at 3.

accountability of government and company supporters to the Call commitments over the years through stock-taking exercises and public events.[621] Some MSIs, such as GNI, do annual assessments of company members, to hold them accountable.[622] The Call does not need to undertake extensive review to achieve broader accountability, but could follow the lead of GNI and set out in its terms of reference that governments and companies are making good-faith efforts to fulfill their Call commitments over time. If the Call looks to advance work related to GenAI, it will be important to establish how participants establish accountability for commitments they make, and which commitments apply to which community members.

4(e). Funding

It is essential for any terms of reference to address how an MSI is funded and discloses its funding – both to build legitimacy and because it is legally required to do so in most jurisdictions.[623] The Call's work is currently funded by the New Zealand and French Governments; this funding pays the staff of the Secretariat and supports event costs.[624] If the work of the Call expands to cover GenAI, the Call Secretariat may want to consider new potential funding sources, including from non-government contributors. Any terms of reference should clearly define what types of funding the Call will accept and how it will disclose funding.

To accept additional funding, the Call may want to restructure the organisation as a non-profit entity rather than a government-funded initiative. One common legal instrument MSIs use worldwide to accept outside funding is to register in the United States as a non-profit organisation under s 501(c)(3) of the Internal Revenue Code. A non-profit organisation can qualify as a 501(c)(3) if it is created for educational or public benefit purposes and this status provides the non-profit a federal tax exemption to receive donations.[625] MSIs such as the Internet Society,[626] GNI,[627] and ICANN[628] are all 501(c)(3)s that receive funding from governments, charitable trusts, and for-profit companies. This tax structure is widely used, easily established, and often a requirement before parties will donate to an MSI, because it allows contributions to be tax-exempt. To establish a 501(c)(3), the Call would need to set up bylaws that would closely mirror the terms of reference suggested in this report, and complete the necessary paperwork in the United States.[629] Other jurisdictions have similar provisions including New Zealand which provides this type of legal mechanism through its Charities Act 2005.[630] If the majority of funding

---

[621] *See* Christchurch Call to Action, *Christchurch Call Community Consultation, Final Report, supra* note 437; *see also* Christchurch Call to Action, *Christchurch Call Community Insights 2022,* CHRISTCHURCH CALL TO ACTION (Sep. 2022), https://www.christchurchcall.com/assets/Documents/Christchurch-Call-2022-Community-Insights-Discussion-guide.pdf.
[622] Global Network Initiative, *Company Assessments*, GLOBAL NETWORK INITIATIVE (2022), https://globalnetworkinitiative.org/company-assessments/.
[623] The Stanley Center, *supra* note 324 at 5; *see also* van Huijstee, *supra* note 323 at 31.
[624] Christchurch Call Advisory Network, *Terms of Reference, supra* note 503.
[625] United States Internal Revenue Service, *Exemption Requirements – 501(c)(3) Organizations*, INTERNAL REVENUE SERVICE (6 Feb. 2023), https://www.irs.gov/charities-non-profits/charitable-organizations/exemption-requirements-501c3-organizations.
[626] Internet Society*, Form 990, Return of Organization Exempt from Income Tax*, INTERNAL REVENUE SERVICE (Rev. Jan. 2020), https://www.internetsociety.org/wp-content/uploads/2020/12/ISOC-2019-Form-990-Pub-Insp-Copy.pdf.
[627] Global Network Initiative, *Form 990, Return of Organization Exempt from Income Tax*, INTERNAL REVENUE SERVICE (Rev. Jan. 2020), https://globalnetworkinitiative.org/wp-content/uploads/2022/03/2020-990-GNI.pdf.
[628] International Corporation of Assigned Names and Numbers, *Form 990, Return of Organization Exempt from Income Tax*, INTERNAL REVENUE SERVICE (Rev. Jan. 2020), https://www.icann.org/en/system/files/files/icann-irs-990-fy19-12may20-en.pdf.
[629] Internal Revenue Service, *Applying for 501(c)(3) Tax-Exempt Status*, PUBLICATION 4220 (Rev. 3-2018), https://www.irs.gov/pub/irs-pdf/p4220.pdf.
[630] Charities Act 2005 (NZ), https://www.legislation.govt.nz/act/public/2005/0039/latest/DLM344368.html.

is coming from US-based donors, a 501(c)(3) structure can reduce overhead costs. However, if the Call opts to restructure into a charitable entity, it will be important to carefully assess a wide range of jurisdictional risks.

Once a legal entity is created, the Call should establish processes for evaluating any offers of funding, as sources can have a significant impact on the credibility and longevity of the work. Potential funding sources for the work of the Call include charitable donations, corporate partnerships, government grants and membership-based or event-based funding. The Call should consider the following important aspects before accepting funding from each.

- *Charitable donations*. Many MSIs are funded by charities that believe that the mission of the MSI aligns with their own charitable goals and objectives. Because TVEC online is an intractable problem that causes widespread harm, many charitable organisations may want to support the Call. The Call could look to partner with similarly aligned organisations such as the Ford Foundation, which advances justice worldwide;[631] the Knight Foundation, which promotes freedom of expression and democracy;[632] or the Hewlett Foundation, which helps shape internet policies that promote new technologies.[633] Ideally, the charitable organisation would support a multistakeholder approach and all 25 commitments of the Call.
- *Corporate partnerships.* MSIs frequently accept funding from corporations because this signals "buy-in" from key stakeholders. This could be an option for the Call if the terms of reference ensure a multistakeholder decision-making process to maintain legitimacy. The creation of a non-profit entity would increase the potential for company funding, as many countries have anti-corruption legal restrictions which prevent companies from directly funding government projects.[634] However, in 2023, tech companies have cut back their corporate funding, as the industry is facing widespread lay-offs and slashed budgets for charitable projects.[635] The calculations may be different for projects working on GenAI, because in that context the industry is directly asking for multistakeholder input and facing an onslaught of regulation. The Call's multistakeholder work on GenAI may be one area where companies are willing to support non-profit entities in 2023.
- *Government grants*. For the past four years, the Call has relied on government funding from New Zealand and France. Other governments have funded adjacent projects; the United States government has supported the CCIAO, and the Canadian government has funded the TCAP.[636] The Call should continue to look to government supporters for funding. One way to enable more governments to commit funding would be to structure the Call's work into specific

[631] Ford Foundation, *Work, Our Grants*, FORD FOUNDATION (2022), https://www.fordfoundation.org/work/our-grants/awarded-grants/ ("We are committed to our mission of advancing equality and justice around the world. Over the years, our grantmaking has evolved to meet the challenges facing our society.").
[632] The Knight Foundation, *Program Areas*, THE KNIGHT FOUNDATION (2023), https://knightfoundation.org/programs/ ("We are social investors who support a more effective democracy by funding free expression and journalism, arts and culture in community, research in areas of media and democracy.").
[633] The Hewlett Foundation, *Cyber*, WILLIAM + FLORA HEWLETT FOUNDATION (2023), https://hewlett.org/programs/cyber/#our-grantmaking ("The Hewlett Foundation makes grants to proactively define, research, and manage the burgeoning intersections between people and digital technologies. The Cyber Initiative seeks to cultivate a field that develops thoughtful, multidisciplinary solutions to complex cyber challenges and catalyzes better policy outcomes for the benefit of societies around the world.").
[634] Public Finance Act 1989 (NZ), *A Guide to the Public Finance Act*, THE TREASURY (26 Nov. 2019), https://www.treasury.govt.nz/publications/guide/guide-public-finance-act-html.
[635] Alyssa Stringer, *A comprehensive list of 2023 tech layoffs*, TECHCRUNCH (12 July 2023), https://techcrunch.com/2023/06/05/tech-industry-layoffs-2023/.
[636] Ardern & Marcon, *supra* note XX.

projects, rather than seek general support funding. One specific project could be work on GenAI, which many governments are keen to better understand.

- *Membership-based or event-based funding*. Many MSIs are funded by membership or event fees. This type of funding does not appear practical for the Call as it does not have membership fees like the GIFCT, nor does it charge stakeholders to attend events like the Leaders' Summits. One project where it might be possible for the Call to consider cost-recovery would be for researcher access to the CCIAO. This could discourage initial use of the technology, but it could be something the CCIAO team could consider as the project develops.

11. Step 5: Sustaining Forward Momentum

Like many curated MSIs, the Call was created in reaction to tragic events that inspired stakeholders to commit their resources to solving an underlying problem. After four years, the Call has proven that it can sustain stakeholder interest in addressing the challenges posed by TVEC online. In part, the success of the Call has been due to a laser-focus on the problems of TVEC online and its dedication to building a multistakeholder framework. To sustain momentum and branch into new areas, the Call must take care to maintain stakeholder trust, deliver results, stay relevant, and fight burnout.

5(a). Maintaining Trust between Stakeholders

The Call has maintained trust between stakeholders by staying focused on its original commitments and ensuring all voices are heard. This has been particularly evident during Leaders' Summits, which seat members of civil society with heads of state and government and tech executives.[637] The Call should be applauded for its commitment to give all stakeholders a seat at the table. Additionally, between summits, stakeholders join working groups to address identified challenges which builds trust between participants with common interests. However, this trust can be fragile. CCAN, for example, stated in 2022 that engagement with civil society was often siloed and transparency around cross-Call collaboration was lacking.[638] Going forward, the Call could build additional trust among stakeholders by creating more structured transparency around decision-making and participation.[639] However, any additional process should be carefully designed to not add unnecessary friction that decreases communication. Additionally, the Call could increase internal transparency and ensure cross-Call collaboration by setting up a community platform to facilitate document-sharing and discussions between stakeholders.

5(b). Delivering and Documenting Results

Elimination of TVEC online is a complex problem that requires a whole-of-society approach through many incremental steps. The 25 Call commitments help the community break down the problem into smaller goals; over the years, the Call has delivered results on many projects. The Call has been successful because stakeholders around the world want to showcase the work they are doing to eliminate TVEC online and contribute to solving the problems identified by the Call. The Call helps amplify these bottom-up efforts by stakeholders and provides a forum where organisations can collaborate. However,

---

[637] Christchurch Call to Action, *Leaders' Summits, supra* note 485.

[638] Christchurch Call Advisory Network, *Statements, Christchurch Call Evaluation Pilot Project*, CHRISTCHURCH CALL ADVISORY NETWORK (12 Sep. 2022), https://christchurchcall.network/our-work/statements/, *noting* CCAN stated in this document that they would like more ("meaningful transparency is needed to facilitate independent evaluation of supporters' work under the Call" and "engagement with the Christchurch Call is siloed and there is little transparency around cross-Call collaboration, particularly in cases when crisis response protocols are enacted.").

[639] van Huijstee, *supra* note 323 at 31.

as this report has documented, it can be hard to know whether a company or a government has made policy changes in furtherance of their Call commitments, or for another reason.[640] Therefore, to continue to deliver results and sustain stakeholder momentum, the Call should more explicitly track and document its work. For an MSI like the Call, documenting results will strengthen the "Tinkerbell Effect", where results exist because people believe they do. Belief in results will lead to the manifestation of even more results.

To promote the results delivered, the Call could amplify its communications in three ways: providing a running list of Call accomplishments on its website, engaging more on social media, and facilitating greater cross-stakeholder communication. First, the Call should more clearly showcase results on its website. The Call currently publishes annual reports and work plans which document activities, but these reports are not always consistent or comprehensive, which makes long-term tracking difficult.[641] Instead, the Call could combine all results into one publicly available list that it updates on a regular basis. One example of a consultative MSI that documents accomplishments is Meta's Oversight Board, which works with Meta on a tracker for each of the Board's recommendations and case decisions, updated on an ongoing basis in Meta's Transparency Center.[642] Another potential tool, suggested above, is a greater social media presence, through which the Call could document updates and share announcements from stakeholders. Finally, a community platform for all stakeholders could facilitate broader sharing of accomplishments so the entire Call community could share results with each other.

5(c). Staying Relevant

The Call has stayed relevant over four years by evolving and adapting to two areas of change: the nature of online extremism and new technologies. First, unfortunately, the Call is still relevant because terrorist and violent extremists continue to be radicalised online and perpetuate violence. This threat has evolved over the past four years, as TVEC online has transformed from Islamic extremism and lone actor white supremacy to a broader threat to democracy including, but not limited to, the incitement to violent protests in Washington, DC on 6 January 2021, in Wellington on 2 March 2022, and in the "Freedom Convoy" which turned violent in Canada in 2022.[643] As these attacks happened, some in the Call community looked to the work of the Call to address rising challenges; the Call has worked with new stakeholders as they are exposed to these tragic events.[644] As the TVEC landscape has shifted, the Call has shown that its 25 commitments have broad application and current events have reinforced the need for multistakeholder solutions. As a result, addressing the changing nature of online extremism has been largely uncontroversial for the Call community, as expressed through statements at Leaders' Summits.[645] The Call has also evolved by addressing the changing technological landscape, through its work on algorithms, automation, and new technology. One way to continue maintaining relevance is to expand this work to include GenAI. As many stakeholders are looking to find solutions to amplify the benefits of GenAI and mitigate the harms, including in specific areas such as TVEC.

---

[640] *See* Douek, *supra* note 21 at 75; *see also* Christchurch Call Advisory Network, *CCAN Report on Anti-Dehumanization Policy, supra* note 474.
[641] Christchurch Call Advisory Network, *Statements, supra* note 638.
[642] Meta, *Oversight Board Recommendations*, META TRANSPARENCY CENTER (16 June 2023), https://transparency.fb.com/en-gb/oversight/oversight-board-recommendations/.
[643] Thorley & Saltman, *supra* note 149.
[644] Ardern & Marcon, *supra* note 484 at 4 ("The Christchurch Call has developed a unique model for coordinating action, bringing together affected communities, civil society, and technical experts, alongside international organisations, industry and government. By applying the distinct capabilities of each sector and forging a community with shared purpose and ambition, we are delivering results.").
[645] *Id.* at 4 ("The success of the Call is now well recognised, and some stakeholders have expressed interest in seeking to understand how the Call might work on related issues. We agreed the Call should remain focused on the scope of its Commitments building on our success and sustaining this work into the future.").

5(d). Motivating Stakeholders and Fighting Burnout

A tragedy like the events of 15 March 2019, can spur a lot of action, but the goal should be to build a long-lasting MSI that creates results over time. As the work of the Call enters its fourth year, leaders should consider what stakeholders need to stay motivated, advance the work, and avoid burnout. The needs will be different for government, company, and civil society stakeholders.

First, government stakeholders typically represent large legacy institutions which can be better resourced and staffed than other stakeholders. The work of the Call involves several different government departments, including those dealing with national security, technology policy, and foreign affairs. This is both a blessing and a curse because it means there are more people to do the work but can also lead to competing priorities within a government. The Call's strategy to be "leader-led" can help direct government stakeholders from the top. However, governmental priorities can shift quickly to address constituent concerns, or following an election cycle. The Call should try to address current or emerging trends in TVEC, to ensure government stakeholders are able to show their constituents they are responsive to their concerns. For example, a focus on GenAI could interest many government stakeholders currently examining regulatory and non-regulatory solutions to combating potential harms, while encouraging innovation in GenAI.

Second, company stakeholders in any MSI can lose motivation or burn out when a project is viewed as a discretionary cost-centre or negatively affecting the public relations function. In the case of online platforms, as discussed in Part I, companies are motivated to spend money to moderate content online for three primary reasons: to avoid legal liability, to create a hospitable environment which incentivises user engagement, and to uphold ideals of corporate responsibility.[646] In economic terms, an employee of an online platform needs to demonstrate that participation in the MSI addresses at least one of these three costs. The Call has been successful in attracting company participation because it speaks to all three of these motivations. However, there are still risks of burnout. First, the Call is one of many MSIs addressing the problem of TVEC online, leading to a drain on overall resources.[647] When a company only has one or two employees working on an issue, they often must prioritise the MSI that most impacts their bottom line. Second, as mentioned above, the tech industry has recently undergone widespread layoffs, which means there are fewer employees working on content moderation problems.[648] Third, as regulators explore new areas to legislate, companies may need to prioritise bigger legislative threats. In this case, the policy issues surrounding the use of GenAI are also where companies are prioritizing time and resources. Therefore, the Call's focus on GenAI can usefully respond to emergent threats and opportunities, support new company stakeholders, and help executives maintain a commitment to the work of the Call.

Finally, civil society organisations tend to have fewer resources, and their allocation of time and attention can be dependent on the priorities of influential donors. However, there are far more civil society organisations than there are tech companies or governments, so the pool of potential stakeholders is much larger. Regarding the Call, many civil society organisations joined because they were passionate about the mission of the organisation and were compelled to lend their time and expertise in the aftermath of 15 March 2019. However, over the past few years governments and companies have launched countless initiatives to address a wide variety of content moderation

[646] Klonick, *supra* note 36 at 1618.
[647] *See references* throughout paper to other MSIs including: GIFCT, OECD, TAT, and EU Internet Forum.
[648] *See* Stringer, *supra* note 635.

problems.[649] Unfortunately, many civil society organisations are struggling to participate in all the MSIs related to freedom of expression and content moderation. The Call must find ways to motivate civil society organisations to dedicate their time to this project over others. The Call has incentivized civil society by giving them a seat at the table as well as access to top government and company officials. Another way the Call can support civil society is by promoting thoughtful CCAN-led initiatives, such as the Report on Anti-Dehumanisation Policy.[650] Finally, Call supporters can assist with travel and other expenses for civil society, which would go a long way to keeping these stakeholders engaged.

Overall, the best way to fight burnout is to bring in new voices and topics – both can refresh the work of an MSI and make it stronger. To bring in new voices, the Call could explore setting up a recruitment committee that could prioritise on-boarding new stakeholders. This committee could create thorough on-boarding materials that can be easily distributed to potential stakeholders covering frequently asked questions and extolling the benefits of membership. Additionally, the Call could highlight the work it is doing on algorithms and GenAI, as this work has the potential to bring in stakeholders who may have never heard of the Call but could meaningfully contribute to its work.

12. Step 6: Deciding When the Work is Done

This section has outlined how the Call can sustain momentum and find new life by embracing new workstreams and new stakeholders. Therefore, it is hard to imagine the Call deciding that its work is done, given its lofty mission to "eliminate terrorist and violent extremists content online." An intractable problem like this does not lead to easy solutions, and the Call's work will likely continue well into the future. But even when there is no clear end date, an MSI should think about what comes next. The best practices outline three possibilities for deciding when to wrap up: shut down, join forces, or pivot the mission.

6(a). Shut Down

An MSI typically shuts down when most stakeholders can no longer justify the financial cost or ideologically support the work of the initiative. For example, the NETMundial Initiative shut down because stakeholders did not trust the organisers to carry forward the multistakeholder work of the NETMundial conference.[651] For the Call, this type of shut down would come if its government or company supporters decided they could no longer support the Call commitments. While the Call has support from 56 governments, its leadership and mandate remain with the New Zealand and French Governments – both of which face challenges in 2023. There will be an election in New Zealand in October 2023, which will bring leadership changes to the country no matter the outcome. While there remains strong political support for the victims of 15 March 2019, it is possible a new government may not wish to continue the work of the Call. In France, President Macron has faced severe political backlash in 2023, forcing a shift in his government's priorities to domestic issues and the war in

[649] *See examples* of MSIs exploring issues related to content moderation that are not focused on TVEC include and not already mentioned in this paper: Digital Trust and Safety Partnership, https://dtspartnership.org/; Family Online Safety Institute, https://www.fosi.org/; Transatlantic High Level Working Group on Content Moderation Online and Freedom of Expression, https://www.ivir.nl/twg/; Atlantic Council's Scaling Trust on the Web, https://www.atlanticcouncil.org/wp-content/uploads/2023/06/scaling-trust-on-the-web_executive-report.pdf; Tech Coalition (TC)'s strategic framework for combating child sexual exploitation and abuse online, https://www.technologycoalition.org/knowledge-hub/this-is-test-knowledge-3.

[650] Christchurch Call Advisory Network, *Our Work - Reports*, CHRISTCHURCH CALL ADVISORY NETWORK (2023), https://christchurchcall.network/our-work/reports/.

[651] *See also* Brown & Esterhuysen, *supra* note 234; Internet Society, *Internet Society Statement on the NETmundial Initiative, Press Release, supra* note 241; Strickling, *supra* note 242; Malcolm, *supra* note 357.

Ukraine. Losing the support of these two governments could force the Call to shut down. Additionally, the Call needs to consider its online service provider supporters, many of whom are cutting costs and may no longer have financial reasons to support working on the Call. One way to hedge against losing the support of these groups is to diversify the funding sources of the Call. Ultimately, to avoid shutting down, the Call should communicate clearly with all stakeholders on changing circumstances and ensure a high-level of trust.

6(b). Join Forces

As discussed throughout this report, there are many organisations working on ways to eliminate TVEC online around the world. One option for the Call to consider in its next steps is whether it should join forces with some of these organisations to reduce redundancies in the space. Several options include integrating the Call's multistakeholder work more meaningfully into organisations currently addressing the issue of TVEC online, including the GIFCT or TAT. The GIFCT currently has some limited multistakeholder elements through its Independent Advisor Committee and TAT works informally with many of the same stakeholders as the Call. Additionally, the Call could continue its work alongside a broader MSI addressing content moderation issues, such as GNI or I&J. Another option would be for the Call to partner with a multilateral initiative and encourage them to take on a multistakeholder format. However, any UN process will likely be untenable because the inclusion of authoritarian regimes into relevant conversations will make it difficult to uphold the principles of a free, open, and secure internet. Finally, the Call could seek to partner with an academic or research institution looking to expand its work. All options could help streamline funding sources and reduce duplication of efforts. Ultimately, stakeholders must decide whether another organisation would be a good steward of the mission of the Call and uphold the principles set out in the beginning.

6(c). Pivot

When addressing complex societal problems without clear answers, MSIs must constantly evolve and pivot towards new solutions. Countering the drivers of TVEC online while upholding a rights-respecting framework is one of those complex societal problems. The Call may want to pivot its mission as the MSI matures. One area that this report has discussed at length is a pivot towards GenAI and how that technology might impact the moderation of TVEC online. For all the reasons discussed in Steps 1 through 5 of this section, pivoting the work of the Call to focus on how GenAI will impact TVEC could be important to the future of the Call.

CONCLUSION

As the Call considers its future, this report has several suggestions to help the organisation build a self-sustaining MSI. These suggestions are based on an exploration of how single-sided and multistakeholder models have impacted the governance of user-generated content online over the years. Governmental regulatory frameworks have inherent problems balancing human rights and adapting to technical challenges and companies are struggling to draw lines around acceptable and unacceptable speech. Therefore, this report argues that MSIs are the best, most sustainable, model to protecting the freedom of expression and reducing harmful content online. As demonstrated by the success of multistakeholderism in the internet governance space, the best solutions to content moderation challenges come about when MSIs bring together a diverse coalition of stakeholders and craft consensus-based policies. By examining the history of MSIs, this report proposes a set of best practices to guide the next steps for the Call to build a self-sustaining MSI.

The Call was set up as an MSI in the wake of the tragic events of 15 March 2019 and has made significant progress towards eliminating terrorist and violent extremist content while protecting a free, open, and secure internet over the past four years. The Call has accomplished this through a mulitstakeholder approach that brings together governments, tech companies and civil society. To sustain the momentum of the Call and advance its core mission, first, the Call should explore restructuring the MSI to ensure it has a strong foundation to scale and grow the organisation. Second, the Call should further expand its work to address the challenges and opportunities posed by the development of GenAI and its impact on TVEC online. After working with the Call team for several months, this report outlines how the Call could achieve both goals.

First, to ensure the Call can grow into a self-sustaining MSI, this report goes step-by-step through the best practices set out in Part II and applies them to the work of the Call. Under this framework, in Steps 1 and 2, the Call clearly defined the problem it is trying to solve and set objectives for stakeholders in its founding documents. However, under Step 3, the Call has not yet achieved its full potential to on-board a diverse group of stakeholders. Bringing the right stakeholders into the room is critical to build legitimacy for the work of the Call. Therefore, this report suggests that the Call consider ways to streamline its on-boarding process, appoint a dedicated staff member to work on recruitment, invest in linguistic capacity in non-English languages, and explore options for funding the work of civil society participants. Next, in line with the suggestions in Step 4, the Call should adopt a formal terms of reference document to guide its work. The Call was set up in just a few weeks and its practices have adapted over time, however, without clear terms of reference, it will be challenging to attract outside funding which is required to scale growth. The Call does not need a complicated terms of reference, but should consider formalizing processes around Call leadership, consensus-based decision-making, internal and external transparency requirements, and funding opportunities. To sustain momentum, as outlined in Step 5, the Call needs to build trust between stakeholders, track and amplify its results, stay relevant, and fight stakeholder burn-out. This report suggests the Call should consider developing an internal communications platform to encourage engagement with stakeholders, update its website to better track the results of the MSI, build a presence on social media, and focus on recruitment efforts. Finally, in Step 6, the Call is not in a place where it needs to shut down or join forces with another MSI, but it could explore pivoting its work to focus more on GenAI.

Second, this report examines why the Call should further expand its work on GenAI. As detailed above, the Call has addressed the impact of AI on content moderation from the beginning and therefore, is in a prime position to become a leading MSI developing best practices. The Call should explore ways to

expand the work of the CCIAO to foster a multistakeholder effort to better understand how GenAI will impact the prevalence of TVEC online. While TVEC is only one type of content that will be impacted by GenAI, it is a good place for an MSI to start because stakeholders generally agree on foundational definitions and the harms of the proliferation of TVEC are so great. Indeed, the risk of offline harms caused by the prevalence of TVEC online has shifted in recent years from as violent extremists have attacked democratic institutions in places like Washington, DC on 6 January 2021, in Wellington on 2 March 2022, and in the "Freedom Convoy" which turned violent in Canada in 2022. Understanding how GenAI will impact TVEC online and finding multistakeholder solutions to address the problems could be foundational to all other GenAI challenges going forward. Therefore, the Call should explore how it can expand its work in this area.

This report suggests that the Call should go through each of the steps in the proposed framework to expand its work to address the challenges and opportunities of GenAI. First, the Call should clearly define the problem and map the landscape as there are many initiatives cropping up in this space. However, as stated above, the Call is the only effort combining technical solutions with a multistakeholder framework that could specifically address the impact of GenAI on TVEC online. To accomplish Steps 1 and 2, the Call must work with its current stakeholders to build consensus around scope expansion otherwise it risks losing support for its current efforts. For Step 3, the Call will want to recruit additional stakeholders, particularly those with technical expertise and GenAI companies, to increase its legitimacy on proposed solutions. Regarding Step 4, the Call should adopt terms of reference to attract new funding partners to expand the work of the MSI and establish foundational principles regarding leadership and decision-making. For Step 5, sustaining forward momentum, the Call's expansion into GenAI will significantly help the MSI stay relevant, fight stakeholder burnout, and attract a wide array of new stakeholders to join. Therefore, expanding the Call's work on GenAI is critical to maintaining legitimacy and credibility. Finally, in Step 6, a pivot towards GenAI is necessary for the Call.

As Jacinda Ardern outlined in her op-ed in June 2023, "I see collaboration on AI as the only option… Together, we stand the best chance to create guardrails, governance structures and operating principles that act as the option of least regret. We don't have to create a new model for AI governance. It already exists, and it works."[652] The Call has the foundations and by implementing these recommendations it can better ensure its future as a self-sustaining MSI and make a useful contribution to a key aspect of digital and internet governance.

[652] Ardern, *supra* note 10.

APPENDIX: FREQUENTLY USED ACRONYMS

Artificial Intelligence	AI
Christchurch Call Advisory Network	CCAN
Christchurch Call Initiative on Algorithmic Outcomes	CCIAO
Christchurch Call to Action	The Call
Digital Services Act	DSA
European Union	EU
Generative Artificial Intelligence	GenAI
Global Network Initiative	GNI
International Covenant on Civil and Political Rights	ICCPR
International Telecommunications Union	ITU
Internet and Jurisdiction	I&J
Internet Assigned Numbers Authority	IANA
Internet Corporation for Assigned Names and Numbering	ICANN
Internet Engineering Task Force	IETF
Internet Governance Forum	IGF
Islamic State of Iraq and Syria	ISIS
Multistakeholder Initiative	MSI
National Telecommunications and Information Administration	NTIA
Organization for Economic Co-operation and Development	OECD
Regulation on Preventing the Dissemination of Terrorist Content Online	TCO
Section 230 of the Communications Decency Act	Section 230
Tech Against Terrorism	TAT
Terrorist and Violent Extremist Content	TVEC
The Global Internet Forum to Counter Terrorism	GIFCT
United Nations	UN
United Nations Educational, Scientific and Cultural Organization	UNESCO
United States	US
World Summit on the Information Society	WSIS