CHRISTCHURCH CALL

TO ELIMINATE TERRORIST & VIOLENT EXTREMIST CONTENT ONLINE

Christchurch Call 2022 Community Update

Table of contents

Introduction

Last year the Call Community came together and developed plans to help organise our work and realise the 24 commitments in the Christchurch Call. The work plans span a multi-year period and are intended to guide individual and collective actions, and work through partner organisations such as the Global Internet Forum to Counter Terrorism. This update is a 'mid-way check-in' rather than a final accounting of our achievements and progress.

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.

- The Community has underscored the value it sees in the Call as a collaborative undertaking, and in the multistakeholder inputs that help individual supporters achieve their commitments in a holistic, informed way, consistent with the strong commitment to human rights and a free, open, secure internet as a force for good in the world.

- There was strong demand for more regular opportunities to come together as a Community, to discuss priority or topical issues and find collaborative solutions.

- There was also recognition that work relevant to delivery of Call Commitments and elements in the four workplans occurs in many fora, with overlapping groups of Community members and others participating.

- The Call should continue to play a leadership role in taking forward key priorities under the commitments, including drawing, where feasible, on existing work. It will be important to avoid duplication, while also making progress where efforts elsewhere have become stuck or there are gaps. The Call also draws the different strands together and ensures consistency with human rights and a free, open, and secure internet.

- Broadly, Community members would like to see more action to address both the social drivers and the online interactions that contribute to terrorist and violent extremist content online.

- Across the Community, there is growing awareness and understanding of the links between terrorist and violent extremist content online, gender-based hate that targets women and LGBTQIA+ people, as well as risks to children and young people.

- There is also increasing recognition that terrorists and violent extremists use disinformation to extend their reach and build support for hateful ideologies, and that disinformation can also act as a gateway to these ideologies.

- The Community does not wish to expand the Call's remit. The focus on specific commitments and the goal of eliminating terrorist and violent extremist content online has enabled us to make such good progress. The Call should create connections and look for ways to draw from

and contribute (both in substance and approach) to work on adjacent and overlapping issues in other fora.

This report will help to set a common baseline of progress for the 2022 Leaders' Summit, and to offer advice to Leaders about the important questions they might wish to address when thinking about how to shape our work over the next year.

Reflecting on Progress: Our Work Plans

Community Work Plan

This year there was a significant focus on building a sustainable and effective Christchurch Call Community, holding more frequent dialogue on topics of interest, sharing best practice, and building knowledge and understanding about the most effective and inclusive ways of realising our shared commitments.

As a Community we committed to:

- Provide resources to the Christchurch Call Advisory Network (CCAN) for administrative support and coordination so they can more effectively engage with the work of the Call;
- Support CCAN to develop its 'advisory function' including by growing CCAN's diversity of subject expertise, a broader set of geographical perspectives, and emphasising the voices of affected communities;
- Enhance information sharing to allow the whole Community to engage on the work being done by individual Call Supporters and collectively to implement the Call;
- Increase industry support for the Christchurch Call, while ensuring that the underpinning commitment to human rights and a free, open, and secure internet is maintained;
- Developing an inclusive multistakeholder process for onboarding more supporters and partner organisations;
- Share expertise and examples of best practice and develop our collective knowledge to help deliver better outcomes in our work;
- Alongside this, develop a process to help encourage supporters to fulfil their commitments and collectively address concerns within the community about actions inconsistent with the Call commitments, and increase the overall accountability of Call Supporters; and
- Strengthen ties with the Global Internet Forum to Counter Terrorism (GIFCT), as a partner in delivering on the Call Commitments and ensure strong connections between our work and that of the GIFCT.

Progress Highlights: Community

Among the key achievements this year, we have:

- Established a Christchurch Call Advisory Network (CCAN) secretariat function to help support the advisory network's efforts, reduce the transaction costs of engaging with supporters, and help CCAN to apply itself effectively across the diverse areas in which we work. New Zealand and France provided financial support to assist with this.
- Agreed a new multistakeholder process for onboarding supporters and partner organisations to the Christchurch Call and worked through several candidate applications with full engagement from the Community.
- Brought four new industry supporters into the Call: Zoom, Roblox, Clubhouse, and Mega.

- Brought in a new partner organisation, the Global Community Engagement and Resilience Fund (GCERF).
- Hosted Call Community policy dialogues on topical issues, including effective content regulation, trust and safety practice, and algorithmic transparency.
- Continued a CCAN research project on Call Supporter approaches to dehumanising speech and discourse, with outputs to be shared with the Call Community soon.
- Begun CCAN pilot evaluation research on Call Commitment Progress with interim findings to be shared for the upcoming Summit.
- Established a Community of Practice to develop accessible descriptions of the technical underpinnings of a free, open, and secure internet and what this means for policymakers implementing the Call.

Feedback on the Community work stream

Respondents gave generally positive feedback on efforts to fulfil 'Community' objectives. Respondents gave the highest progress rating to our multi-stakeholder onboarding process for new supporters and partner organisations, with one respondent describing it as "impressive and robust."

Efforts on information sharing and best practice were also rated highly, with most contributors responding that they had made 'some' or 'good' progress in these areas. Other areas where contributors identified that good progress had been made included: broadened engagement with stakeholders; the involvement of civil society in the work of the Call; improved coordination; and improved communication with the Governments of New Zealand and France.

The Community spread itself across these quite evenly in terms of its allocation of effort to each of the priorities- perhaps reflecting in part CCAN's role coordinating civil society input.

There was a significant focus among respondents on improving ties between Community members and the Global Internet Forum to Counter Terrorism (GIFCT). Perceptions of progress in this area were mixed, with some respondents suggesting that the relationship between GIFCT and the Call Community needs to be more clearly defined. This was also raised in the Community meetings, where the need to continue driving for alignment and delivery in key areas of Call work was emphasised. There was a desire to ensure the voices of the Call's global multistakeholder community were being shared and considered as part of the GIFCT's work.

Respondents thought that there was room for ongoing improvement in facilitating interaction across the Community between civil society, governments, and online service providers. There is strong demand for more opportunities to interact, for more information-sharing to build mutual understanding and trust, and enable the Community to discuss topical issues and collaborate on solutions. One suggestion was a quarterly Community forum. A shared digital platform would also enable ongoing discussion and collaboration.

In their long form answers, respondents pointed to the value we had seen from working together collaboratively on specific projects; from focusing on the substance and the important issues; and from working across sectors. Those were the areas where respondents felt we had been the strongest and should continue to work in that way, wherever possible. "*Let's make the Call the place where we all go for important policy discussions*." Many respondents spoke of the devotion the Community has shown, the significant effort invested by many people, and the commitment to the principles behind the Call. One respondent noted that "*the Call is now an industry standard*."

Respondents considered more work was needed to bring a broader range of online service providers, including smaller platforms, and governments to the table, encouraging them to share and collaborate, and to benefit from the expertise across the Community. Respondents encouraged more firms and

governments to invest in sharing how they are working on their commitments, on responding to information requests and outreach, and to reach out to the wider Call Community in areas of interest to them.

Respondents also encouraged the Secretariat to promote more inclusive engagement, with one supporter saying they would be keen to devote more time to the work but wanted to understand where best to do so. Some respondents also noted that while the relationship between civil society and other stakeholders had improved, more work was needed to get governments and online service providers to engage with civil society and the Community in a genuine way.

Respondents identified risks and opportunities for the community work stream. Key risks identified included that efforts are diluted across a range of different tech initiatives and that CCAN's capacity to provide advice will be limited by resource constraints. Other risks included the impact of AI and the evolution of Web3.0, the rise of far-right extremism, and the proliferation of terrorism and violent extremist content (TVEC) on platforms currently beyond the reach of the Christchurch Call and the GIFCT.

Key opportunities included increasing diversity among the Community, including through improved representation of LGBTQIA+ and indigenous communities, as well as the development of tools for improved coordination among the community. Educating the Community on new technology was considered by some respondents to be a significant opportunity for the Call, particularly to ensure policy makers were equipped with the requisite understanding to make informed decisions on new technologies and the various challenges and opportunities they bring.

Algorithms and Positive Interventions Work Plan

The Community placed particular emphasis this year on examining how online user journeys might contribute to radicalisation and, in multistakeholder conversations, to identify research questions, information needs, and safe methods for exploring those questions.

As a Community we committed to:

- Devote effort to understanding online 'user journeys' and how these may contribute to radicalisation;
- Design a multistakeholder process to assess what methods can safely be used and what information is needed (while preserving trade secrets and avoiding unnecessary disclosure) to allow stakeholders to better understand the outcomes of recommendation algorithms;
- Understanding the evolving tactics of terrorists and violent extremists and how they may exploit algorithms;
- Development of comparative frameworks for the effective assessment of positive interventions and counter-narratives;
- Work in a multistakeholder context to develop a new generation of community-led online interventions; and
- Hold inclusive discussions aimed at better understanding the operation of algorithms designed to identify and remove terrorist and violent extremist content and embed complaints and appeals mechanisms as well as transparency for users.

Progress Highlights: Algorithms and Positive Interventions

Work on this priority took place in multiple locations, including across several working groups of the Global Internet Forum to Counter Terrorism (GIFCT), the EU Internet Forum, the ISD Digital Policy Lab, the Global Partnership on AI, among others. The Secretariat convened a Christchurch Call coordination group to help draw together the different threads and to create a common base of knowledge and oversight of this work.

The Secretariat also organised structured conversations on issues such as uncertainty in algorithmic decision making, the function of classifiers and recommendation algorithms, and algorithmic transparency.

Respondents identified common themes:

- AI and Machine Learning systems interact dynamically with users and content, and content curation involves a complex mix of social inputs, automated processes, and social outputs.
- Community members are interested in how algorithmic systems might contribute to terrorist content being recommended to users. This is a difficult question to answer because this content is removed when it is identified and generally has a very low prevalence on mainstream platforms. Researchers must be creative in finding experimental methodologies to test this question.
- There are multiple theories around how amplification might function, and much of that work rests either on so-called ‘borderline content’, or on external harm metrics that may have less direct relationship to identified terrorist content.
- Some of the interventions deployed online involve promoting ‘high quality’ content and promoting positive pro-social user engagement rather than demoting bad content.

- There are still several policy and legal hurdles to cross for researchers wanting to conduct independent studies of social media algorithms. While the nature of these issues has been thoroughly discussed, the community has yet to come up with a unified 'fix' for this issue, and this has impeded progress.
- There is a high level of interest in conducting 'pilot studies', including those identified in the 2022 output of the GIFCT's Technical Approaches Working Group. These pilot study concepts pick up on work done at Brookings Institution, in the Global Partnership on AI, and elsewhere.
- The EU's Digital Services Act will provide a new legal framework for independent investigation of algorithmic outcomes for some online service providers. There is a high level of interest in discussing what methodologies could be applied to the questions raised by the Christchurch Call Community, and how these might be carried out ethically and in ways that protect user privacy and prevent unnecessary disclosure.
- For both positive interventions and algorithmic studies, our understanding is often restricted to the frame of a single user environment. The Community would value a broader approach where it can draw common lessons from looking across and between programmes and platforms to see what works best for people at risk.

Feedback from the Community: Algorithms and Positive Interventions

The high level of Community interest in this priority was reflected in survey responses, with more than 77% of respondents saying they had devoted at least some effort to understanding user journeys and radicalisation. Over 80% said they had devoted at least some effort to looking at how terrorist and violent extremist groups might exploit algorithmic systems. Almost half of respondents said they had devoted either the highest level of effort, or 'significant' effort to the user journeys question. Across the board, respondents assessed our progress as 'some' or between 'limited' and 'some'.

The survey responses showed Community members valued discussions hosted by the Call Secretariat and the efforts made to create a common knowledge base. They noted for instance that the Community had worked hard to 'put this issue on the agenda' globally and had done reasonably well at identifying the issues that were hindering progress.

The desire to move beyond dialogue to action was a common thread in responses, as was interest in not just how, but where, these discussions would take place. Should the Call act as a 'catalyst' or an 'amplifier' for the work, or should it try to create new dedicated vehicles for that work to take place? This feedback was also reflected through some responses on 'community' where respondents noted they had difficulty deciding where to allocate effort across the Call and other initiatives dealing with algorithms.

One respondent noted it was hard to collaborate in this space but that they were working hard to improve on these issues internally and to communicate effectively and transparently about that work. Others felt the Call had an important role in drawing together work going on in 'hundreds of different places' and integrating useful ideas to help drive progress.

Some respondents pointed to new regulation, including the EU Digital Services Act as a necessary step in achieving progress. There was also a desire to look at what was happening through the Digital Services Act and whether it had wider application. Others worried that a rush to regulate by Governments might dilute important efforts at achieving globally consistent and technologically sound approaches. An important component of making real progress on the algorithmic workplan will be to

ensure the Call facilitates the sharing of best practice so regulators and researchers around the world can benefit from the Community's collective experience in this space.

There was a strong focus in replies on the importance of incorporating diverse perspectives in the work on algorithms, and in avoiding a 'western' bias in the evaluation of algorithmic outcomes and models for understanding and intervening in radicalisation and amplification. Other respondents noted that while it was important for the Call to continue working on better understanding recommender algorithms and the role they may play in radicalisation, focusing also on prevention and positive interventions to address indicators of radicalisation remains equally critical.

Looking ahead, respondents were most worried about governments or firms acting unilaterally and without transparency to apply 'solutions', without reference to civil society. That ranged from, for example, technologically enabled online interventions, through to overly restrictive policies that have unintended consequences for some users. There is also a useful role for pilot studies within individual companies, as these are a low-cost, lower-risk way to explore issues and potential solutions.

Respondents were enthusiastic, however, about the potential for technology to empower users and stakeholders, including, for example, through third party quality assessments, interventions, and early threat detection.

Incident Response Work Plan

The Community has made solid progress against its work plan for crisis response. Since May 2021, we have reviewed and updated the Christchurch Call Crisis Response Protocol, drawing on Community feedback and developments in the wider crisis response landscape. We have supported GIFCT as it developed its Content Incident Protocol into a three-tiered Incident Response Framework, extend the hash-sharing database, strengthen 24/7 response capability, and improve its communications. We have organised and participated in international and domestic tabletop exercises, to raise awareness, test, and strengthen all protocols and ensure interoperability. We have contributed to work in the GIFCT's Crisis Response Working Group to map protocols and identify overlaps and gaps, and to develop human rights indicators for all crisis response practitioners. It was clear at the Call Community's debrief on the Buffalo, NY, attack in May that the Community has made progress but there remains room for improvement.

As a Community we committed to:

- Reviewing and updating our shared Christchurch Call Crisis Response Protocol, including robust processes on information sharing;
- Mapping the various protocols that are in place across the Call Community and identifying gaps, overlaps, and redundancies, and using table top exercises to rigorously test the overall system for incident response;
- Establishing due process and human rights protections across our crisis protocols;
- Support and engage with the GIFCT on its incident response framework; and
- Establishing an effective multistakeholder community of practice to enhance our capability to respond to incidents, and continued refinement of our protocols and systems.

Progress Highlights: Incident Response

The Christchurch Call Unit reviewed and updated the **Crisis Response Protocol (CRP)** in late 2021, building on feedback from a December 2019 workshop, Community consultations in 2020, and the planning sprint in the lead-up to the 2021 Leaders' Summit. The update was implemented in December 2021 after further consultation with the Community. The updated CRP:

- Sets out roles and responsibilities of each sector in readiness, response, and recovery phases, and acknowledges the role of CCAN in each.
- Clarifies the purpose of the CRP alongside other protocols and establishes processes to enhance executive-level information-sharing among Call supporter governments and industry during an active response.
- Provides guiding principles around due process and human rights.
- Sets out good transparency and accountability practices, including by incorporating the debriefing framework developed by GIFCT's Crisis Response Working Group in 2021/22.

There is a focus in the work plan on making sure the CRP and other crisis protocols are interoperable, with clear lines of communication and robust process and human rights protections. Community members organised and participated in various tabletop exercises develop our collective ability to work in this manner. Europol hosted an exercise involving all multiparty and domestic protocols in November 2021. GIFCT hosted two exercises in April, one on human rights and the other on communications with member companies. The European Commission, as Chair of the EU Internet Forum, and the Radicalisation Awareness Network – Policy Support (RAN-PS) held a crisis

communications training exercise in May to develop guidance for the EU Crisis Protocol. Most recently, Australia hosted a tabletop exercise for domestic agencies, other governments, and industry to increase awareness of its Online Content Incident Arrangement and improve coordination with international partners.

In 2021/22, the New Zealand Government led work in GIFCT's Crisis Response Working Group (CRWG) to **map all the protocols and identify overlaps and gaps**. The Call Community, including CCAN, was consulted. GIFCT published a summary in July 2022 that gives an overview of the purpose, aims, scope and activation criteria of the different protocols and contains recommendations for the GIFCT and multistakeholder community. These recommendations include bringing more countries and companies into crisis response, subject to appropriate criteria and safeguards, and continuing to develop the role of civil society and researchers.

Also in 2021/22, CCAN's co-Chair led work in the CRWG on the lifecycle of a terrorist incident online and **human rights indicators**. This work establishes a generic framework to help all protocol owners and participants identify the individuals and groups whose rights may be at risk at different points in the "lifecycle", to prevent and mitigate negative impacts. It is a work in progress and can be further refined and tailored for application in specific circumstances.

Through these initiatives and others, the Call Community has **supported GIFCT to strengthen incident response**. GIFCT has:

- Developed the Content Incident Protocol into a three-tiered Incident Response Framework and extended the hash-sharing database.
- Increased transparency around incident response. Information is shared with the Independent Advisory Committee (IAC), CRWG members and other protocol owners in exercises, notifications, updates and debriefs. Some information is also publicly available on the GIFCT website, in transparency reports and blogposts.
- Improved tools and processes, e.g. for situational awareness.
- Strengthened 24/7 response capability.

In doing so, GIFCT has drawn on the advice of its Working Groups and lessons learned in multistakeholder tabletop exercises and debriefs.

The work plan includes a longer-term goal of establishing a **Community of Practice** for crisis response. This Community already exists, if somewhat loosely, with multistakeholder discussions and activity taking place in the Christchurch Call, GIFCT, and EU Internet Forum. These overlapping groups are involved in ongoing efforts to update contacts, strengthen relationships, test and refine protocols, adapt to changing technology and threats, develop best practice and improve coordination.

More remains to be done, including to bring more countries, companies, civil society organisations and community groups into the Community of Practice. As one of the next steps, and as envisaged in the work plan, the Call Secretariat will develop an onboarding document for new supporters so they are aware of the CRP and able to participate in a way that protects human rights and a free, open and secure internet.

Feedback from the Community: Incident Response

Community members have had varying levels of involvement in crisis response work, with some allocating considerable effort to this area and others little or none.

The majority sees that some progress has been made on all areas of the work plan, with the strongest perceptions of progress – and positive comments on - updating the CRP; mapping the different protocols, analysing gaps, and improving overall coordination through tabletop exercises; and

supporting the GIFCT. Respondents commented on the increased opportunities for multistakeholder input, and the progress made on notifications and debriefing on incidents.

Respondents also acknowledged the work done in CRWG on human rights indicators and the GIFCT's dedicated tabletop exercise. There was slightly weaker satisfaction with overall progress in this area.

In general, those few who perceived only limited progress across the workplan also tended to have allocated limited or no effort to these areas of work, which may suggest they had limited visibility of the results – something for the Call Secretariat to work on.

Overall, the Community sees we are doing well to deliver our objectives for crisis response. There is still a lack of clarity around when the CRP can and should be activated, and interest in further developing some dimensions of it (e.g. human rights protections, data preservation).

Survey respondents would like to see work continue to maintain and update the various protocols, to clarify roles and to improve coordination, particularly through regular tabletop exercises. They would also like to see more work in relation to mainstream media.

There is also a need for the Call Secretariat to reflect on how to keep Community members not involved in e.g. GIFCT's CRWG across developments and make sure they can contribute to specific pieces of work and participate in tabletop exercises.

The key opportunities and risks that Community members identified correspond well to the findings and recommendations of the mapping and gap analysis, and Buffalo debrief. In particular, Community members point to the need to integrate a diverse range of (cooperative) online service providers into crisis response, and to find ways to counter adversarial tactics involving modification of content and use of unmoderated and 'alt-tech' services. There is also strong support for developing the role of civil society in crisis response, particularly through the CRP. Finally, the Community should also reflect on success indicators for crisis response and find ways of measuring and articulating this to the media and public.

Transparency Work Plan

The Community has made progress in achieving greater transparency, particular among online service providers. There is ongoing work across many fora, linked to the Call and generally, to develop and disseminate best practice on transparency and reporting. Countries are progressing legislative frameworks and reporting practices to meet government transparency commitments. Respondents from all sectors called for greater transparency from companies and governments alike, with greater focus on government transparency.

In addition, amongst other things, the Community calls for greater alignment between transparency work in different fora, taking opportunities to work on transparency with smaller companies, and risks to ensuring transparency as the internet continues to evolve.

As a Community we committed to:

- Ensuring meaningful diversity and inclusion in our work on transparency, and removing barriers to participation;
- Raise awareness of the Christchurch Call and broaden understanding of our work;
- Gather resources and best practice guides to help Supporters improve their transparency and reporting practices;
- Better identify the information needs of different audiences, address trust gaps, and improve the accessibility of information for stakeholders;
- Identify and mentor (with the help of tech against terrorism and the GIFCT) a diverse range of companies to improve their transparency practices; and
- Identify critical gaps in our knowledge and understanding.

Progress Highlights: Transparency

The transparency workplan covers two dimensions of transparency. The first is transparency around the Call itself: what the Community is doing; how supporters are delivering their commitments; and impacts. The second focuses on commitments to transparency regarding terms of service and content moderation. (Algorithmic transparency and crisis response are dealt with under those two workplans.)

As co-founders and Secretariat of the Call New Zealand and France, working with the wider Community, have lifted efforts to raise awareness of the Call and broaden understanding of our work. Some of this work is Community-focused. For example, the Secretariat has regular monthly meetings with CCAN and sends monthly updates to the wider Community. Outreach and wider awareness and understanding of the Call has improved, in particular through panels, speeches and presentations, including at RightsCon, IGF, and the Tech4Democracy Summit. The Secretariat has also created a factsheet about the Call, helped GIFCT develop a factsheet for companies and, on 1 September 2022, launched a new website, with more detail on the Call and work in specific areas.

In the lead-up to the 2022 Summit, the Community survey and CCAN's pilot evaluation have helped the Community to assess how well supporters are delivering Call Commitments and to understand the impact of the Call. The CCAN is learning much in the process about how to do this kind of evaluation for the Call, which is important for transparency, accountability, and establishing priorities for future work. This in turn will contribute to enhanced transparency about the Call and its work.

Much of the Community's transparency work plan is focused on achieving the longer-term goals of having more companies and governments providing meaningful transparency and reporting around their policies, tools, and actions to detect and take down TVEC, and how they respect and protect

human rights in the process. This covers day-to-day operations and incident response efforts (see separate Incident Response section).

Community members have expended significant effort to improve company transparency and reporting on TVEC. GIFCT and Tech Against Terrorism have continued incentivising companies of different sizes and types to start transparency reporting and improve over time, through their respective membership criteria and Tech Against Terrorism's TrustMark programme. Tech Against Terrorism also provides practical support, through its Transparency Reporting Guidelines for Tech Companies (publicly available), Knowledge Sharing Platform (available to registered users), and active mentoring (including for companies seeking GIFCT membership). In some instances, CCAN members such as the Center for Democracy and Technology have also provided advice directly to individual tech companies on ways to improve existing transparency and reporting, generally and in relation to TVEC.

Community members have also worked to identify, describe, and promulgate elements of best practice. Important developments over the last year include the OECD's launch of its pilot Voluntary Transparency Reporting Framework (VTRF) and portal, and the update and extension of the Santa Clara Principles – the latter led by the Electronic Frontier Foundation and involving other civil society organisations in CCAN. In GIFCT's Transparency Working Group, Dr Courtney Radsch published work on transparency good practices and lessons learned from other sectors, and Dr Chris Beall of the Centre for International Governance Innovation is building on that work to elaborate guidance on the specific purposes, audiences, and information needs underpinning tech company transparency on TVEC. Under Denmark's Tech for Democracy initiative, the Global Network Initiative has spearheaded the creation of an Action Coalition for Meaningful Transparency that involves other CCAN and Community members. This Coalition aims to map existing efforts, improve coordination, identify gaps, and produce recommendations for companies and governments, and offers the prospect of drawing together key strands of this work. The updated Santa Clara Principles include new principles for governments and state actors. Tech Against Terrorism also published Transparency Reporting Guidelines for Governments focused specifically on efforts to eliminate TVEC online.

The OECD's second benchmarking report noted improvements in the number and quality of transparency reports with TVEC-specific information being issued by the global top 50 tech companies. Much of the progress on company transparency has been made on a voluntary basis. In addition, some governments have developed legislation to tackle online harms, often including transparency and reporting requirements. Over the last year, for example, the EU has passed its Digital Services Act and the UK has continued to progress work on its Online Safety Bill. Community members from the tech industry and civil society have contributed advice. As regulators move to implement these requirements, we expect to see further guidance and support for the tech sector – a good example is Australia's safety-by-design assessment tools and resources - and improvements in transparency and reporting across the board.

There has also been progress on government transparency, albeit this has been more limited. Europol's Internet Referral Unit issued its fourth annual transparency report in December 2021, providing high level information on its work to assess and refer TVEC online, and to improve implementation of the EU Crisis Protocol. Other Call supporter governments include some information about their legal mandates, policy frameworks and operations on their websites and in annual accountability documents. See, for example, the Australian e-Safety Commissioner's Annual Report. However, as noted by Tech Against Terrorism early this year, it remains the case that globally governments and law enforcement agencies publish "little to no information […] on how they identify terrorist and violent extremist content, how they work with tech companies to have such content removed, and on what legal basis identification and takedown is justified". New Zealand's Department of Internal Affairs' published its first transparency report on digital violent extremism in April 2022,

which Tech Against Terrorism assessed as meeting many of the criteria in its guidelines and setting an example for others to follow.

Feedback from the Community: Transparency

Some members of the Call Community have been very active over the last year on transparency; others have devoted little or no effort to this area.

Survey respondents perceived some progress has been made across the work plan. They indicated that the Call has put transparency "on the table", the Community is developing its understanding of meaningful transparency and a range of resources for companies, and "reporting is improving", including among small companies as an outcome of the "valuable" mentoring and membership work of Tech Against Terrorism and GIFCT.

Given that so much company-focused transparency work is happening in other fora, it is difficult to assess how much progress is attributable specifically to the Call. It is important in setting the agenda, though. Survey respondents underscored that the Call should continue to harness and contribute to other efforts and avoid duplication.

Overall, 90% of respondents consider the Community is doing fairly well or better to deliver our transparency objectives.

Looking forward, respondents indicated the Community should allocate more effort to distilling what meaningful transparency means for Call supporters (including in terms of data access); providing clear guidance for companies and governments; turning ideas, principles, and recommendations into action; benchmarking and monitoring progress; and understanding the overall benefits and costs of transparency.

Survey respondents said they want to see tech companies continuing to improve transparency, with an emphasis on the largest companies providing information that shows the impacts of their policies, implementation, and decisions. There was also an interest in ensuring that online service providers engaged their technical experts as well as their policy and legal experts to help advance crucial discussions.

More focus and effort should also go into government transparency. It is less clear who is responsible for driving this work and how progress is being tracked. One respondent would like CCAN to share more information and updates on its members' work in support of the Call. There are opportunities to use the expertise in the network more effectively and to use outputs like the Santa Clara Principles more effectively to support Call work. There should be more feedback, so CCAN members understand where, how and why their advice is leading to change (or not).

Respondents identified opportunities in working with a broader range of tech companies, including fast-growing companies operating at scale (TikTok was cited as one example) and smaller companies; harnessing new technologies to enable faster and more comparable transparency outputs; and using new auditing regimes and capacities to improve trust and confidence. One respondent saw an opportunity for the wider Call community to provide advice to companies as they consider how they will comply with the new EU Digital Services Act

New technology and an enduring, effective Call Community

Respondents to the Community Survey were asked for their views on what the growth of new technologies could mean for the different work plan areas, and how we might coordinate across technology initiatives.

Respondents identified potential risks:

- Some spoke to how stakeholders might respond to technological developments e.g. through adopting unilateral or splintered approaches or responding to geopolitical drivers in a manner that undermines the Call's values-based, global, multistakeholder effort.
- Some also felt misunderstandings around the operation of algorithms and the issue of amplification could lead to poorly calibrated responses, or even opaque attempts to impose behavioural constraints or values on users. This could erode the efficacy of our work or the social license to work in that space.
- A diversifying technology sector might impact on achieving the objectives of the Call, e.g. through an increasing number of online service providers that did not share our objectives or were actively subverting efforts to respond to terrorist and violent extremist content online.
- This included concerns that the changing format of online platforms with more 'live experiences' and more decentralisation might advance more quickly than the moderation techniques and technologies, and that both might overtake the consideration of human rights and transparency.
- Respondents saw significant risks that decentralised technologies could make the online experience less transparent for users and offer potential for terrorists and violent extremists to co-opt those technologies. Adversarial behaviour was flagged multiple times as a risk.
- Some respondents were worried that well-meaning attempts at creating more transparency could lead to information overload and a potentially saturation of information that could, in turn, lead to a misplaced focus on some issues simply because there were more data points on them.

And opportunities:

- Several respondents thought the Call could act as an exemplar for other tech initiatives, including in the way it engages stakeholders, embeds human rights principles and a free, open, secure internet in its work.
- One respondent felt it was important to keep discussing what success looks like, to ensure we had a realistic perspective on what was achievable.
- One key opportunity, flagged multiple times, was the inclusion of more voices, including LGBTQIA+, other languages and ethnicities, and with young people. This could help build a focus on developing healthy communities online and mitigating the negative effects of terrorist and violent extremist content online.
- Smart use of technology could assist with more inclusive conversations, and with some of the tasks the community has set itself e.g. identifying risks, horizon scanning, and rapid

communication during a crisis; with improved data access to empower research and improve the outcomes of content curation and positive interventions, and

- E.g. use of more third-party tools (such as e.g. industry standards and independently developed quality metrics, or middleware) to help improve the user experience and reduce harm online.

On coordinating with other initiatives, respondents asked the Community to play to its strengths and its values base, to be active in looking for areas of overlap and, to look for common solutions, platforms, and systems to help build inclusive dialogue on tech.

Appendix One: Summary of quantitative responses to the Community Survey

The Community Survey required respondents to assess in a quantitative manner the progress toward and efforts devoted to each commitment under the four work plans. These responses informed the written aspects of this report. A summary of these quantitative are included here.

Response rate by sector

	Governments	Online Service Providers	Partner Organisations	Civil Society
# of replies	10	4	2	13
% of total replies	34%	14%	7%	45%

Community

The Community Survey respondents rated overall progress under this work plan a 6.8 out of 10.

Assessment of Progress by Respondents	Average out of 5 (i.e. major progress achieved)
Resources for CCAN admin support and coordination	2.8
Supported CCAN to develop its advisory function	3.3
Enhancing information sharing	3.4
Increased industry support for the Call	3.2
Multistakeholder engagement on onboarding supporters and partner organisations	3.7
Sharing expertise and best practice	3.5
Addressing concerns about the commitments and accountability of supporters	3.0
Strengthening ties with the GIFCT	3.4

Effort Devoted by Respondents	At least some effort	High or very high effort	Highest level of effort
Resources for CCAN admin support and coordination	54%	31%	8%
Supported CCAN to develop its advisory function	58%	38%	15%
Enhancing information sharing	85%	52%	22%
Increased industry support for the Call	58%	31%	8%
Multistakeholder engagement on onboarding supporters and partner organisations	67%	37%	19%
Sharing expertise and best practice	85%	50%	15%
Addressing concerns about the commitments and accountability of supporters	56%	30%	8%
Strengthening ties with the GIFCT	81%	56%	33%

Algorithms and Positive Interventions

The Community Survey respondents rated overall progress under this work plan a 5.6 out of 10.

Assessment of Progress by Respondents	Average out of 5 (i.e. major progress achieved)
Understanding online user journeys and radicalisation	2.9
Multistakeholder processes for assessing information needs on recommendation algorithms	2.8
Understanding how terrorists and violent extremists might exploit algorithmic systems	3.1
Comparative frameworks for the effective assessment of positive interventions	2.6
Work in a multistakeholder context to develop a new generation of community-based interventions	2.5
Inclusive discussions on the algorithms designed to identify and remove terrorist and violent extremist content and embed complaints and appeals procedures for users	2.9

Effort Devoted by Respondents	At least some effort	High or very high effort	Highest level of effort
Understanding online user journeys and radicalisation	77%	44%	23%
Multistakeholder processes for assessing information needs on recommendation algorithms	81%	38%	27%
Understanding how terrorists and violent extremists might exploit algorithmic systems	85%	44%	19%
Comparative frameworks for the effective assessment of positive interventions	65%	31%	12%
Work in a multistakeholder context to develop a new generation of community-based interventions	63%	30%	4%
Inclusive discussions on the algorithms designed to identify and remove terrorist and violent extremist content and embed complaints and appeals procedures for users	62%	38%	27%

Incident Response

The Community Survey respondents rated overall progress under this work plan a 7.0 out of 10.

Assessment of Progress by Respondents	**Average out of 5 (i.e. major progress achieved)**
Reviewing and updating our shared Christchurch Call Crisis Response Protocol, including robust processes on information sharing	3.5
Mapping the various protocols that are in place across the Call Community and identifying gaps, overlaps, and redundancies, and using table top exercises to rigorously test the overall system for incident response	3.6
Establishing due process and human rights protections across our crisis protocols.	3.3
Support and engage with the GIFCT on its incident response framework	3.9
Establishing an effective multistakeholder community of practice to enhance our capability to respond to incidents, and continued refinement of our protocols and systems	3.4

Effort Devoted by Respondents	**At least some effort**	**High or very high effort**	**Highest level of effort**
Reviewing and updating our shared Christchurch Call Crisis Response Protocol, including robust processes on information sharing	60%	40%	20%
Mapping the various protocols that are in place across the Call Community and identifying gaps, overlaps, and redundancies, and using table top exercises to rigorously test the overall system for incident response	64%	36%	20%
Establishing due process and human rights protections across our crisis protocols.	64%	32%	8%
Support and engage with the GIFCT on its incident response framework	72%	48%	40%
Establishing an effective multistakeholder community of practice to enhance our capability to respond to incidents, and continued refinement of our protocols and systems	54%	46%	17%

Transparency

The Community Survey respondents rated overall progress under this work plan a 6.6 out of 10.

Assessment of Progress by Respondents	Average out of 5 (i.e. major progress achieved)
Ensuring meaningful diversity and inclusion in our work on transparency, and removing barriers to participation	3.3
Raise awareness of the Christchurch Call and broaden understanding of our work	3.2
Gather resources and best practice guides to help Supporters improve their transparency and reporting practices	3.0
Better identify the information needs of different audiences, address trust gaps, and improve the accessibility of information for stakeholders	3.0
Identify and mentor (with the help of tech against terrorism and the GIFCT) a diverse range of companies to improve their transparency practices	3.2
Identify critical gaps in our knowledge and understanding	3.1

Effort Devoted by Respondents	At least some effort	High or very high effort	Highest level of effort
Ensuring meaningful diversity and inclusion in our work on transparency, and removing barriers to participation	64%	36%	8%
Raise awareness of the Christchurch Call and broaden understanding of our work	69%	38%	12%
Gather resources and best practice guides to help Supporters improve their transparency and reporting practices	50%	31%	12%
Better identify the information needs of different audiences, address trust gaps, and improve the accessibility of information for stakeholders	60%	44%	16%
Identify and mentor (with the help of tech against terrorism and the GIFCT) a diverse range of companies to improve their transparency practices	40%	24%	12%
Identify critical gaps in our knowledge and understanding	69%	35%	12%

CHRISTCHURCH CALL

TO ELIMINATE TERRORIST & VIOLENT EXTREMIST CONTENT ONLINE

christchurchcall.com