Te Tari Taiwhenua
Internal Affairs

New Zealand Online Crisis Response Process

Te Kāwanatanga o Aotearoa
New Zealand Government

Contents

Introduction

The lessons learned through the online response to the content associated with the Christchurch attack and others, highlighted the need for a domestic response process. Government, non-government (NGO's) and Online Service Providers (OSP's) must work together in a coordinated way to mitigate harm to New Zealanders of online content associated with terrorist and violent extremist acts. The nature of the internet is such that a response must be swift to disrupt the spread and potential for harm in such circumstances.

This document outlines the agreed voluntary Online Crisis Response Process for organisations in New Zealand to adopt in the event of a future online crisis. Furthermore, it details the roles and responsibilities of the agencies involved, the assessment process used to activate an online crisis response and the steps involved in the initiation and deactivation of an online crisis response in New Zealand. This response process is a living document and will be tested and refined on an ongoing basis.

Note the Online Crisis Response Process will not replace any existing legal requirements, mandate or standard operating procedures in place for any organisation involved.

Purpose of the Online Crisis Response Process (OCRP)

The Online Crisis Response Process aims to facilitate the rapid assessment and coordinated response to an online crisis, within the scope outlined below, and the sharing of intelligence and information in a secure and timely manner between all government, NGOs and OSPs involved. This process connects to the processes developed through the Christchurch Call Shared Crisis Response Protocol, European Union (EU) Crisis Protocol, and the Global Internet Forum for Countering Terrorism (GIFCT) Content Incident Protocol.

The Online Crisis Response Process details the coordinated response to significantly harmful content that requires an immediate response by multiple parties to mitigate the harm associated with it. This includes content that is or is likely to be objectionable and/or content that should not be visible and viewed by vulnerable members of society due to the level of harm it can cause. For the purposes of this process vulnerable members of society are defined as those who may experience increased harm, discrimination or intolerance as a result of their race, ethnicity, age, religion, disability, gender and sexual identity both in the physical and online world.

This process has been developed in response to the content associated with the 2019 Christchurch terror attacks. It is envisaged that the primary use of this process will be in relation to terrorist and violent extremist content (TVEC) online. It is envisaged that this process could also be used in a scenario where significantly harmful content such as video depicting the sexual exploitation of a child is circulated broadly and unprompted on open social media platforms for instance.

Identifying an Online Crisis

An online crisis for the purpose of this process is described as:

'A piece of significantly harmful (highly likely to be 'objectionable') online content that has broad spread both geographically and/or across multiple platforms, and is likely to create significant harm for New Zealanders who are exposed to it'

Any agency participating in the Online Crisis Response Process can raise concern about a potential online crisis and contact the nominated Crisis Manager to determine whether or not this Online Crisis Process should be activated.

- **Government** – Government agencies may identify a potential online crisis through their business as usual activity in the online space, through existing relationships or complaints from members of the public.
- **Online Service Providers** - OSP's for the purposes of this document refer to the technology industry, internet service providers, and social networking services. OSP's may notice a spike in activity associated with harmful content for instance and let the Crisis Manager know.
- **NGOs/ Other's -** Other organisations may be made aware of an online crisis from their local community, customer complaints, platform users etc.

The notification of a potential online crisis should be made to the Director of Digital Safety (or their delegate) at DIA as they will act as the Crisis Manager in the first instance. When notifying the Crisis Manager about a potential online crisis, the relevant organisation should relay (wherever possible):

- the nature of the content
- potential/perceived threat(s)
- any information they have available to assist in assessment

Determining if the Online Crisis Response should be activated

Assessing the Harm of the Online Crisis

Once online content has been flagged to the Crisis Manager, they must assess the situation and determine the how harmful the content is. This process does not replace business as usual delivery and assessment of online harm, this is reserved only for consideration of content that may reach the 'crisis or increased monitoring threshold'.

To make this assessment the following key questions should be considered in conjunction with the judgement of the Crisis Manager. These questions are intended to be a guide to inform the assessment of the need for Crisis process activation:

Key Questions	
Harm to the Public	Does the content promote, incite or glorify violence? Crisis Manager to determine the extent to which the content promotes or glorifies terrorist/ violent extremist ideologies
	Does the content depict matters of crime, cruelty or violence?
	Does the content depict a real-life incident?
	Could vulnerable members of society who are exposed to the content experience physical or psychological harm?
	Is the content appearing or likely to appear on individuals' newsfeeds and homepages? If yes, this suggests the content is being actively promoted by different platforms or by individuals, including through algorithmic settings. In this case, such indicators suggest higher / increased virality (geographically and by platform), reproducibility & resilience of such material to takedown efforts
	Do individuals or groups appear the be seeking to deliberately subvert detection and takedown/ removal systems?
	If terrorist or violent extremist content, is it perpetrator/ accomplice produced and/or live streamed?
	Does the content target or involve specific/ minority groups?

Note: the above questions should be re-considered as more information comes to light during the online crisis process, it will not often be possible to have answers to all of these questions at the outset to determine if the crisis process is activated.

Other Considerations

There are a number of other factors that may help inform the Crisis Managers decision when assessing the harm of the online content to determine whether a crisis response should be invoked:

Question	Consideration
Is the content likely to be deemed objectionable?	Content that "*describes*, *depicts, expresses or otherwise deals with matters such as sex, horror, crime, cruelty or violence in such a manner that the availability of the publication is likely to be injurious to the public good"* [1]can be deemed objectionable by the Office of Film and Literature Classification (OFLC). Once content has been classified as objectionable it is prohibited to possess or distribute it.[2] When assessing whether content should be classified as objectionable, the Chief Censor will consider it in line with the definitions in the Films, Videos and Publications Classification Act 1993.
Is the content likely to be appropriate for adults but not for children?	Some content may not be deemed objectionable but may contain high impact cruelty or violence and may not be suitable for all members of the public to view i.e. bystander footage or journalistic coverage of an incident. In these instances, the Chief Censor may classify the content as R18.
If related to a real-life event, does the crisis appear to take place in New Zealand?	If a physical incident/ crisis takes place domestically and the Crisis Manager deems there to be considerable impact to New Zealanders, DIA will drive the Online Crisis Response Process and reach out to the relevant domestic and international bodies to get support in gathering intelligence and reducing the spread of the content online. The Crisis Manager will update this range of partners with relevant information from the ground level. The Online Crisis Response Process will feed into the wider DIA Internal Response Process for a Domestic Crisis[3]. DIA will work closely with NZ Police to share intelligence and evidence to support their investigation.
If related to a real-life event, and the incident/ crisis took place abroad, does it appear to include, indicate a threat to, or involve New Zealanders?	If the incident/ crisis took place overseas, the Domestic Online Crisis Response Process may be activated if the Crisis Manager deems there to be considerable impact to New Zealanders. New Zealand will adopt a supporting role and will take direction from the Christchurch Call Shared Crisis Response Protocol, EU Response Process to Online Crisis, and the GIFCT Content Incident Protocol should one or more of these protocols be activated. The Crisis Response Team will work alongside the impacted country(s) and platform(s) to gather and share intelligence on the spread of the content including by sharing hashes and URLs with international agencies to contain the spread of the content. We will also work closely with the Ministry for Foreign Affairs and Trade in instances where New Zealand citizens are among the victims.
Did the perpetrator(s) announce their intentions online before the attack?	Did the perpetrator(s) upload any content that appears to have announced their intention to carry out an attack before the incident/ crisis? Did the perpetrator(s) promote the attack(s) online i.e. providing links to the where it would be streamed or uploading a manifesto beforehand?

1 Section 3(1) of the Films, Videos and Publications Act 1993

2 Objectionable content is prohibited under section 123(1) of the Films, Videos and Publications Act 1993

3 Depending on the crisis that occurs it may be necessary to liaise with the relevant Minister(s) as well as the Prime Minister

Question	Consideration
Does the content contain identifiable information about victims or any other person?	Exposure to content that contains identifiable information or images of the victims may reignite a survivor's trauma and/or result in significant psychological effects on a range of vulnerable individuals exposed to it.

Note: The Online Crisis Response will endeavour to preserve content that is a genuine historical archive or that condemns/raises awareness of an event, in accordance with international human rights law, including freedom of expression.

Risk Consequence Guide

Once the manager has assessed the content in terms of the potential harm, they should consider the severity of the harm that the public may be exposed to. The below table provides non-exclusive guidance on how to consider consequences of content in assessing the situation – not all points have to be met at a given level to constitute, for instance, a severe consequence grading e.g. Targeting or potential targeting of a specific or minority group is not an absolute/ deciding requirement for an incident/ potential crisis to be graded/classified as severe. The Crisis Manager will take a decision based on information to hand:

Scale	Harm	Virality	Associated Risk
Severe	Content promotes, incites or glorifies violence Content that glorifies, promotes or incites terrorism/ acts of terrorism/ violent extremism Exposure may cause a high degree of physical or psychological harm The content targets a specific or minority group	Content appears on homepages or newsfeeds Content has a significant volume of shares, likes, comments and views Content may appear on multiple platforms across multiple countries	Content may be classified as objectionable Perpetrator(s) may have announced their intention to carry out the attack(s) or promoted the attack(s) online beforehand Content may include identifiable information or images of the victims

Scale	Harm	Virality	Associated Risk
Significant	Content may promote, incite or glorify violence Content may cause some physical or psychological harm Content that glorifies, promotes or incites terrorism/ acts of terrorism/ violent extremism The content may target a specific or minority group	Content appears on homepages or newsfeeds Content has a high number of shares, likes, comments and views Content may appear on more than one platform across multiple countries	There may be propaganda material relating to the crisis appearing online Content may be classified as objectionable or R18 Content may include identifiable information or images of the victims
Moderate	Content that glorifies, promotes or incites terrorism/ acts of terrorism/ violent extremism Content may not be appropriate for children or vulnerable individuals	Content may appear on a limited number of homepages or newsfeeds Content has some shares, likes, comments and views Content may appear on multiple platforms across multiple countries	Content may be classified as R18 or objectionable
Minor	Content that glorifies, promotes or incites terrorism/ acts of terrorism/ violent extremism Minimum physical or psychological harm	Content has limited number of shares, likes, comments and views Content may appear on multiple platforms across multiple countries	Content may be classified as R18
Minimal	Content that glorifies, promotes or incites terrorism/ acts of terrorism/ violent extremism No measurable physical or psychological harm	Content has very few shares, likes, comments or views Content may appear on multiple platforms across multiple countries Content is not likely to have any increased interaction online due to only being present on a smaller platform with limited reach and use	Content may be classified as R18 or objectionable

Assessing the Virality

Once the Crisis Manager has considered the harm associated with the content, they need to make an assessment about the spread / potential spread of the content. In order to determine the risk of virality of the online content, the Crisis Manager should consider the content in light of the below questions:

Key Questions
Has the content appeared on a range of platforms/ across multiple countries?
Has the content received a significant number of shares, likes, comments and views?
Has the content been uploaded online by a perpetrator(s) or apparent accomplices as part of the overall strategy of the attack to increase virality and harm?

Crisis Assessment Matrix

The matrix below can be used as a rough guide to assist in making the decision to activate the online crisis response. This is based on risk of virality or spread of the content and the consequence/harm associated with it.

Risk of Virality						
	Yes	BAU	Increased Monitoring	Increased Monitoring	**Crisis**	**Crisis**
	Unsure	BAU	Increased Monitoring	Increased Monitoring	Increased Monitoring	**Crisis**
	No	BAU	BAU	Increased Monitoring	Increased Monitoring	Increased Monitoring
		Minimal	**Minor**	**Moderate**	**Significant**	**Severe**
		Risk Consequence Scale				

Response

Once an online crisis has been assessed, the crisis should have the appropriate response assigned to it depending on the severity of the crisis:

Risk Level	Response
Crisis	If the assessment falls between within the crisis section of the matrix then an online crisis response should be activated. **Example:** The Christchurch Terrorist Attacks. Before the video was live streamed, the perpetrator shared his intentions and extremist views online. The video features racist slurs as well as the depiction of extreme violence. The video was torrented and shared across multiple platforms in the hours after the attack. It also received significant attention in traditional media and social media with 12 million tweets about the video being posted. Due to the level of violence in the video, the perpetrator's intentions and the high virality of the video, an Online Crisis Response was activated.

Risk Level	Response
Increased Monitoring	If the assessment falls within the increased monitoring section of the matrix then normal DIA BAU operating procedures should apply along with ongoing monitoring. In this instance the Crisis Manager will send out an update to nominated contacts with any relevant information. Online Crisis Response participants that have monitoring and/ or complaints systems in place may be able to feed intelligence and trends regarding the online crisis i.e. increased volume of incoming calls to helplines, they can then share this insight with the Crisis Manager. This information will be used to remain up to date on the status of the online crisis and determine whether an online crisis response needs to be implemented. In this instance, the Crisis Manager will send out regular communication with an update on status and an overview of the current landscape. **Example:** The Halle Attack. Although the video was deemed objectionable by The Classification Office, the content did not receive significant views, shares or likes across social media platforms. In this instance New Zealand chose not to activate an online crisis response however they continued to monitor the situation online for a number of days.
BAU	If the assessment falls within the BAU section of the matrix then normal BAU operational procedures and relevant agency / industry and government actions apply. This will apply for most content on the internet that does not meet the criteria for a crisis.

Types of Response to an Online Crisis

When an online crisis is categorised as a crisis, there are a number of tools that may be used to minimise the harm of the online crisis. This section describes some of the tools that may be used and is not intended to cover all possible options – during an Online Crisis Response other suggestions / options will be considered. The intent is to use the full range of levers available to minimise harm. The responses could include, but are not limited to:

Response	Example of when the response may be used
Adding sites to Family Filters*	Content that is considered R18 and harmful but not objectionable may be added to OSP's family filters. If individuals do not have or choose to opt out of the filter, they will be able to view the content. This response may be appropriate for viral content that should not be viewed by vulnerable New Zealanders

Response	Example of when the response may be used
DNS Blocking/ Poisoning*	Internet Service Providers can block specific websites in New Zealand for customers using DNS servers. This approach may be used in instances where it is deemed appropriate to block entire websites such as those dedicated to hosting the most egregious TVEC. In the event that any blocking activity is agreed to, DIA will manage, review and validate content to protect staff in ISP's. The data regarding sites to be blocked will be managed and updated by DIA through a managed spreadsheet
Media Campaign to the Public	Content that is spreading online, whether or not deemed objectionable, may benefit from a coordinated media campaign to inform the public about the potential harms from the specific content and advice about where to go for help if concerned and staying safe online
Content blocking for School Networks*	Network 4 Learning may block sites to prevent content being viewed by children in schools in New Zealand
Targeted Engagement with Media Outlets	The Crisis Team may engage with media outlets in New Zealand to ensure any reporting is not inflammatory or targeting a particular community. Media outlets may also be used to provide up to date messaging to the public and information about where to go for advice and support
Geo Location Blocking*	Social media platforms may use this option to restrict individual's access to certain online content based on user's geographical location for example this would reduce the risk of vulnerable New Zealander's being exposed to harmful online content emerging from another country
Interstitial's*or trigger warnings	The Crisis Team may request the use of Interstitials by social media companies where this is possible, for content that is deemed harmful but not illegal. Social media platforms may reduce individual's exposure to content by limiting the visibility of harmful content. Content containing sensitive or graphic information appears with a warning informing people about the content before they view it. This gives people the option to uncover and view the content at their discretion or not see it at all. This is intended for newsworthy content, historical/freedom of expression content that condemns/raises awareness of an incident/ potential crisis in accordance with international human rights law
Content Take Down	The Online Crisis Response team may request the removal of specific content such as videos, posts etc. from platforms. This can help reduce the spread of harmful content online and prevent individuals from being exposed to graphic or sensitive material.

Response	Example of when the response may be used
User Profiles*	Social media platforms may pause, block or delete specific users accounts from their websites for non-compliance with their terms of service/ equivalent user requirements. This could be used in instances where individuals or groups are spreading terrorist or violent extremist content online. It is important that online service providers work closely with NZ Police to ensure that if users accounts are paused or deleted that this will not impact the investigation efforts

* These actions may be adopted where appropriate and will be dependent on the technology available for each organisation involved in the response

The Online Crisis Response Process

This section outlines:

- The roles and responsibilities that will be stood up during the online crisis response process and the crisis deactivation process;
- The online crisis process; and
- The deactivation of the online crisis process.

Roles and Responsibilities

When an Online Crisis Response is initiated, there will be a number of key roles fulfilled by New Zealand government officials as the core Crisis Team:

Role	Responsibilities
Crisis Manager	This position will be held by the Director of the Digital Safety Team (or their delegate) in the Department of Internal Affairs (DIA). This individual is responsible for activating and deactivating the Online Crisis Response Process. They will oversee the management and coordination of the response including tasking, driving the virtual operation room, decision making and communication with global crisis process representatives (where needed)[4]. The Crisis Manager will be responsible for updating government stakeholders as part of the broader response[5] for both 'Crisis' and 'Increased Monitoring' incidents. The Crisis Manager will be the primary point of contact for global protocols and other crisis responses in New Zealand and will share information / updates with relevant international partners as required. Once the process is deactivated, the Crisis Manager will oversee the completion of a report detailing the actions taken by the agencies involved

4 This may include other governments, international law enforcement, OSPs, the GIFCT etc.

5 Government stakeholders refers to the Minister of Internal Affairs, other Ministers as relevant, Department of Internal Affairs Chief Executive and other executives as necessary.

Role	Responsibilities
Communications Lead	Responsible for writing updates and maintaining key messages relating to the actions being taken as part of the online crisis response process. All agencies will be responsible for organising their own comms however the Communications Lead will provide regular updates to the group which representatives can use to inform their own media releases. The Communications Lead will coordinate the release of information relating to the online crisis on both social media and traditional media for the DIA. It will be preferable for other agencies to point to this central information source as much as possible rather than recreate material.
Reporting/ Intelligence Lead	Responsible for receiving and compiling information from multiple sources to provide clear situation reports for each update call. This individual will oversee the coordination of intelligence into the online crisis. They will oversee the maintenance of a central repository where information will be shared with relevant agencies and Ministers, on responses considered and used, agencies involved, and the nature, distribution and spread of the content online. The Reporting/ Intelligence Lead will work closely with the Communications Lead/Crisis Coordinator to ensure relevant actions/next steps are captured and pushed out in a timely manner to relevant domestic and international audiences. The Reporting/ Intelligence Lead will share any data/ intel/ situational updates and reports with relevant international partners.
Technical Lead	Oversees the fulfilment and feasibility of any technical solutions proposed as part of the response process. They will also ensure that all necessary ICT infrastructure is in place to ensure the efficient operation of the response process once it is activated. This includes the establishment of a secure conference line and file sharing system
Crisis Coordinator	Responsible for the documentation of the actions and decisions taken as part of the online crisis response process. They will document all discussions and decisions made in the virtual operation room and on conference calls. They will issue invitations to conference calls, prepare and share notes after each update with all relevant parties
Logistics Support	Responsible for coordinating access to the DIA building, refreshments and supporting the Crisis Coordinator throughout the crisis response process

The Online Crisis Response Process

The steps involved in the Online Crisis Response Process are detailed below, including what is required for the thresholds below 'crisis' where increased monitoring will take place:

Start

Notification to Crisis Manager of harmful (possibly illegal) content that has 'gone viral' or has the potential to go viral/ or Crisis Manager identified content that may need a crisis response

Crisis Manager reviews the content using the crisis criterial

Is a crisis response, increased monitoring or BAU approach required?

No Response Required - BAU

Incident is determined to have minimal/ minor harm and low risk of virality. BAU processes continue

End.

Increased Monitoring Required

Crisis Manager triggers online crisis response process and contacts the relevant representatives

- Crisis Manager notifies all primary points of contact (via email or phone) that an online incident has occurred and that it has been categorised as requiring increased monitoring. The Crisis Manager will request that agencies with complaints/ monitoring systems in place and are in a position to share intelligence and trends about the online incident feed that to the Reporting Lead. This information will be used to remain up to date on the status of the online incident and determine whether a crisis response needs to be implemented. The Reporting/ Intelligence Lead will compile incoming intelligence and provide situational reports.
- The Crisis Manager will also address what public guidance is required and where members of the public should be directed to for support, information and complaints/queries.
- The Crisis Manager will advise DIA's General Manager of Regulatory Services and the Deputy Chief Executive of Regulatory Services of the incident and any actions being undertaken

The Crisis Manager will send out regular communication to the crisis response team based on incoming information.

The Crisis Manager will review the increased monitoring status every two hours to assess whether it is the appropriate categorisation. Once the situation is no longer deemed to be a 'crisis' the Crisis Manager will issue a final comms notifying agencies that the crisis response process has ceased. The representatives will be responsible for filtering this notification down internally through their organisations.

The Reporting/ Intelligence Lead will generate a final intelligence report to be shared with relevant authorities and key contacts in the Crisis Process. This report will give a final view of the situation and articulate actions undertaken.

End.

Crisis Response Required

Crisis Manager triggers online crisis response process and contacts the relevant representatives

OFLC will review the content and determine if it meets the standards to be classified as objectionable

DIA establish a physical war room in the DIA office. This will act as the base of operations during the response

Crisis Manager assigns following roles: Comms Lead, Reporting/ Intel Lead, Crisis Co-ordinator, Logistics Support and Technical Lead

The Crisis Manager will advise DIA's General Manager of Regulatory Services and the Deputy Chief Executive of Regulatory Services of the incident and any actions being undertaken

A secure teleconference line is established by DIA and all agencies are invited to join no later than 30 mins from crisis activation

Crisis Manager will provide
- Overview of the nature of the incident to date and its online threat
- Each representative on the call speaks to their view of the issue and the immediate action they can take
- Agencies who cannot offer any support are free to opt out of the call/ response
- Crisis Manager addresses what public guidance is required and where members of the public should be directed to for support, information and complaints.

Crisis Manager sets actions and a time for next call. All agencies do not need to dial into each call if not required.

Regular call will repeat every 2 hours 24/7 until 'crisis has stopped'. At each update call.
- Reporting/ Intelligence Lead will compile incoming intelligence and provide situational reports
- Communications Lead will provide regular updates to the group to help agencies inform public messaging
- On each update call the key representatives will provide updates on their actions, shared view of situation and any further steps required
- Crisis Manager will confirm if the crisis status is still in place
- Crisis Coordinator will circulate the key notes and actions to the wider group immediately following each call

The Crisis Manager will review the crisis status every two hours to assess whether it is the appropriate categorisation. Once the situation is no longer deemed to be a 'crisis' the Crisis Manager will issue a final comms notifying agencies that the crisis response process has ceased

Once the situation is no longer considered a crisis, the deactivation process can begin

Engagement with OSPs and global processes during crisis

The focus of the New Zealand Online Crisis Response Process is about what efforts can be taken by agencies and service providers to minimise harm domestically and protect human rights.

This Online Crisis Response Process will be activated where the situation is deemed a crisis in New Zealand. In cases where the content has a global concern, responses such as the Christchurch Call Shared Crisis Response Protocol[6] and/or the Global Internet Forum to Counter Terrorism Content Incident Protocol (CIP)[7] and/or EU Crisis Protocol[8] may be activated. Where the Christchurch Call Protocol or other protocols are activated, the Crisis Manager will act as one of the conduits receiving and sharing information, working closely with Department of the Prime Minister and Cabinet (DMPC) , Ministry of Foreign Affairs and Trade (MFAT) and NZ Police representatives. In each instance, the Crisis Manager will be responsible for being the New Zealand contact point for the broader response and feeding intel and information into the global protocols and from the global protocols into the domestic process.

There may also be instances where a global crisis response is not activated, yet the domestic online crisis response may be activated (i.e. when an assessment shows the need to mitigate harmful content that is focused on New Zealand).

---

6 For endorsing governments and OSPs (who support the Christchurch Call to Action to eliminate terrorist and violent extremist content online , herewith referred to as the "Christchurch Call") to respond rapidly, effectively and in a coordinated manner to the dissemination of TVEC following a terrorist event in a manner consistent with human rights protections, with a view to harm minimisation and disrupting terrorist aims. This protocol is an agreed voluntary mechanism outlining shared crisis response principles and processes for endorsing countries and OSPs to adopt in the event of terrorist and violent extremist attacks with an online component. It is for use in the context of a terrorist or violent extremist attack where a high probability exists for significant online impact given associated content depicting the event. To trigger the threshold for a crisis, the TVEC will most likely be perpetrator or accomplice produced. This is most likely to be activated by the government affected by the real-world terrorist or violent extremist event and can only a Christchurch Call supporting governments and OSPs that have formally endorsed the protocol can officially activate it. Note it does not supersede any domestic or international laws.

7 The ***GIFCT CIP*** is a process by which GIFCT member companies become aware of, quickly assess, and act on potential content circulating online resulting from a real-world terrorism or violent extremist event. No one individual or organization can activate a content incident. Rather, the protocol is based on the existence of content online relating to the real-world terrorism or violent extremism event—like Christchurch and Halle—and potential distribution of that content, including a live stream of murder or attempted murder produced by the attack's perpetrator or an accomplice. The GIFCT CIP is a standalone industry process, but was designed to be easily integrated into external crisis response procedures, including the Christchurch Call Shared Crisis Response Protocol developed in response to the commitments of the Christchurch Call to Action to eliminate terrorist and violent extremist content online and the EU's Crisis Response Protocol. See here for more: https://gifct.org/joint-tech-innovation/

8 ***The EU Crisis Protocol****: Collective response to viral spread of terrorist and violent extremist content online* may be activated only by an EU Member State or Europol where a crisis is identified within the EU. It aims to facilitate rapid assessment of the online impact of terrorist attacks, secure and timely sharing of critical information between EU Member States law enforcement (LE) and other competent authorities, Union bodies (in particular Europol), OSPs and other relevant stakeholders in accordance with relevant legislation and within the relevant mandates, and to ensure effective coordination and management of the crisis. See more here: https://ec.europa.eu/home-affairs/sites/homeaffairs/files/what-we-do/policies/european-agenda-security/20191007_agenda-security-factsheet-eu-crisis-protocol_en.pdf

Deactivating the Online Crisis Response

The categorisation of the online crisis will be regularly reviewed. Once it is no longer considered a crisis, the following process will be carried out:

Crisis manager will make the decision in consultation with the Crisis Team. This decision will be informed by the crisis criteria above as well as the below questions:
Is there anything more the Crisis Team can do at present?
Has the availability of content reduced?
Has the spread of content reduced?
Has there been a reduction in the volume of calls or complaints regarding the content?

Crisis Manager will issue final comms via email notifying all representatives involved that the crisis response process has ceased. The representatives will be responsible for filtering the notification down internally through their organisations.

All agencies will revert to BAU operating procedures with ongoing monitoring for 2 weeks following crisis deactivation.

Following a cool down period of 2 weeks, the Crisis Manager will hold a conference call with the core Crisis Team involved in the response to confirm if process can be fully deactivated.

At this point any actions taken as part of the response e.g. blocking, may stop or will move into longer-term BAU process e.g. family filtering

The Reporting / Intelligence Lead will generate final copies of the intelligence report.

This report along with intelligence packages will be delivered to stakeholders

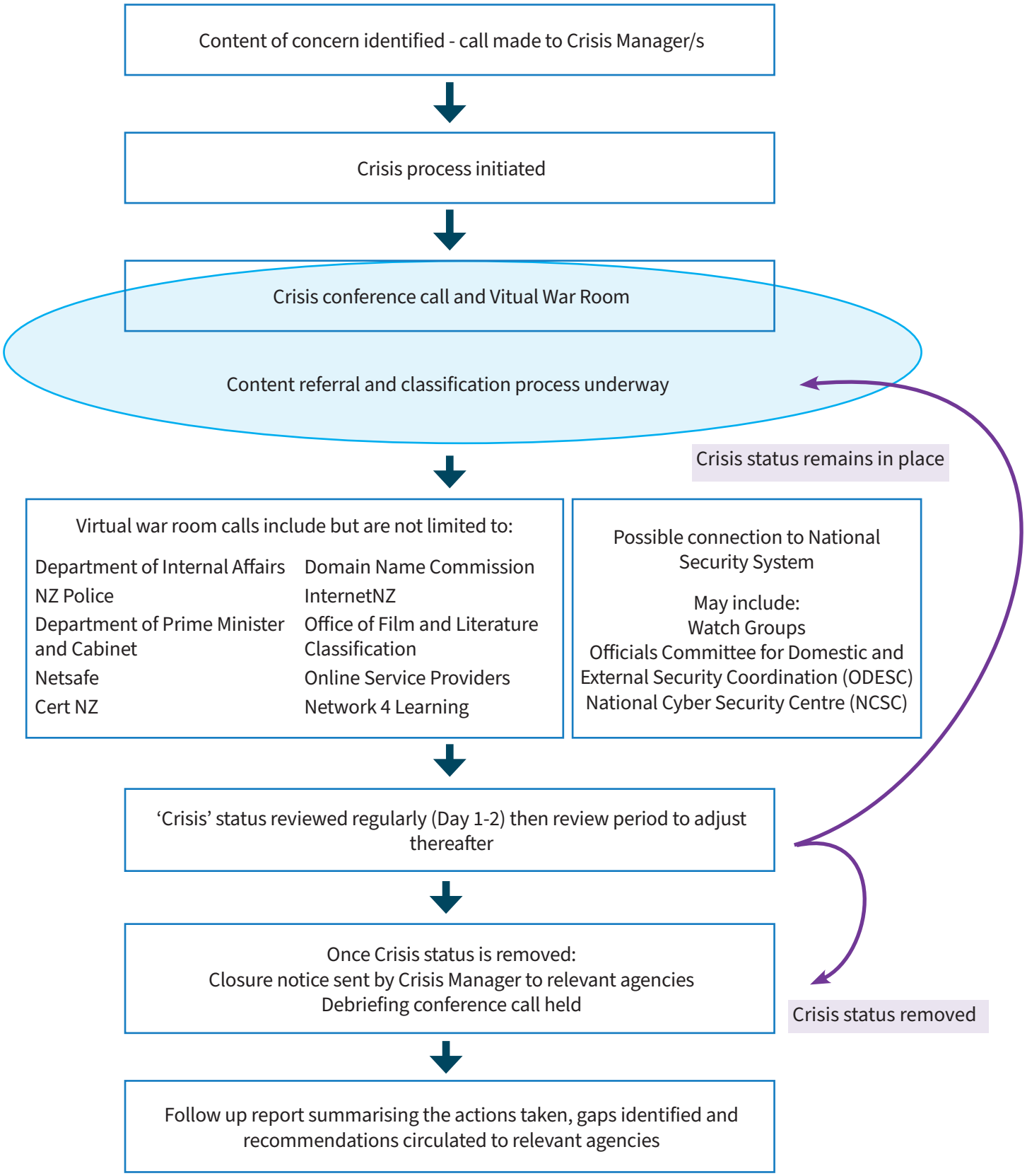
The Communications Lead will generate the final public messaging

One week after the crisis has been formally deactivated, the Crisis Manager will hold a debriefing session with the core Crisis Team to discuss the actions that were taken, any gaps or limitations identified during the repoonse, success points and recommendations going forward

DIA will generate a final report which will detail the background to the crisis, the actions taken and recommendations going forward. This report will be shared, in confidence, with all participants of the response process. A copy of the report and the final version of the intelligence spreadsheet will be saved in the shared location.

Summary of the Online Crisis Response Process

The below table illustrates the end-to-end Online Crisis Response Process detailed above:



Principles and Expectations of the Online Crisis Response Process

Principles of Operation

By signing up to the New Zealand Online Crisis Response Process, all participants agree to adhere to the following principles of operation:

- **Authenticity** - All decisions must reflect the scope of the crisis criteria and fit within the overall aims of the domestic online crisis response process
- **Transparency and Public Confidence** - The New Zealand public are entitled to request government information. Any actions taken during an online crisis response must be defendable. We will proactively release messaging to the public to ensure they remain informed of the response process throughout. This will help to keep us accountable, reassure the public and to help prevent harm
- **Speed and Agility** - Situations unfold rapidly in the digital age. During a crisis, the speed of action and our agility to respond to changes in the situation are core to how we operate
- **Intelligence** - The consistent rapid sharing of information and data on a secure platform will be necessary to limit the spread of harm online. This will include sharing URL's, hashes, situational awareness, media and metadata
- **Trust and Accountability**– The Crisis Manager is the key decision maker for the response and is required to take responsibility for any actions taken and decisions made. On the initial crisis conference call if representatives challenge whether a crisis should be called, this can be discussed, reviewed and the crisis status stepped down if required. To ensure accountability there will be appropriate measures put in place and all records will be fully maintained. If at any point during the response the Crisis Team disagree with any decision made by the Crisis Manager and the issue cannot be resolved with Crisis Manager or Crisis Team directly, then an escalation should be made to the DIA General Manager
- **Communication** - During a crisis we will communicate information and updates up to senior stakeholders, across all the agencies involved and outside to the media and public. By having a Communications Lead, DIA will ensure a coordinated and consistent flow of information. There is also onus on other agency partners and OSPs to work collaboratively in the interest of timely harm reduction and to comply with law enforcement requests/ requirements
- **Harm minimisation** - This process is intended to minimise exposure and harm resulting from a terrorist or violent extremist attack, including to society and social cohesion, as well as any victims or intended victims of the attack and to anyone exposed inadvertently to terrorist or violent extremist content
- **Respect for human rights and dignity, and fundamental freedoms** - Any response will be carried out in a rights-respecting manner, including with respect to the rights of any victims or intended victims. New Zealand government agencies must act consistently with the Bill of Rights Act 1990
- **Preserving a free, open, secure and globally connected internet** - No associated tools deployed should undermine functionality and connectivity of the global internet nor compromise its infrastructure

Operational Expectations on Information Sharing

Classifying Content as Objectionable

In the event of an online crisis, speed is essential to minimising harm. The Chief Censor will be consulted and where possible, will make an 'interim' decision on classification. Where this is not possible this will be communicated, and guidance developed (such as using the term 'likely objectionable' as used in Christchurch response).

Alerting the Crisis Manager

When an agency is alerted to online content that may require an online crisis response, the identifying agency should contact the Crisis Manager.

ICT Information

All conference calls will be scheduled by the Crisis Coordinator. In the event of an online crisis response being initiated, they will issue an invite to all representatives which will include the teleconference details. This invite will also include details of the physical operation room for representatives who wish to attend the initial launch call, or any follow up calls in person.

Communications

A Communications Lead will be appointed for the duration of the online crisis response. This individual will work with the points of contact and comms teams from the organisations involved in the online crisis response to agree appropriate public messaging and will provide regular updates on each follow-up call. This information can be used to inform each agency's own media releases.

They will work with the points of contact and comms teams from the organisations involved in the online crisis response to ensure the public are directed to the necessary helplines and complaint portals. They will also agree on messaging to inform the public on the incident as it unfolds and how they can stay safe online.

They will develop and maintain a document outlining the public messaging each organisation is issuing in order to ensure there is a coordinated and staggered release of information on both social media and traditional media between all organisations. This will ensure that there is a consistent message being filtered out to the wider public and avoid duplication of information being provided to the public.

During a crisis it will be important to communicate information and updates up to senior stakeholders, across all the agencies involved and outside to the media and public. By having a Communications Lead we will ensure there is consistent information flow to the New Zealand Prime Minister or other relevant ministers. If an online crisis response requires us to connect with our international partners as in the case of Christchurch, the Communications Lead will work with the crisis response representatives to develop translations for the public messaging.

Expectation of Confidentiality

All agencies and OSPs who volunteer to participate in the Online Crisis Response Process will sign a Confidentiality Agreement. This will cover:

1. **Authority for decision:**

- In the event of an online crisis and if DNS blocking is utilised, the Department, in consultation with the Chief Censor, will make the decision about what content should be blocked. The Department will accept responsibility for any content that is blocked as part of the online crisis response and maintains an appeal process for these decisions to be challenged;

2. **Protection for OSPs**

- Protection for the OSP's with regards to the information they share as part of the crisis process which would not normally be shared with the Government or other industries;
- Protection for OSP's from viewing objectionable and sharing URLs and hashes for objectionable content for the purposes of the crisis process; and

3. **Confidentiality:**

- Most information shared during a online crisis response is confidential and will not be shared outside of the response, with the exception of information released as part of the online crisis response comms.

Additional Considerations

Connection to the International Response Processes

As previously outlined, the Domestic Online Crisis Response Process connects with international response protocols to ensure there is a coordinated response to an online crisis with the shared aim to reduce virality and minimise harm through timely information sharing.

The GIFCT was established by Facebook, Microsoft, Twitter and YouTube to disrupt terrorist content on their platforms and ensure ongoing knowledge-sharing and technical collaboration. It is now an independent organisation, with its mission to prevent terrorists and violent extremists from exploiting digital platforms. In the event of an online content incident, the GIFCT will appoint a representative to liaise with the platforms and the Domestic Online Crisis Response Team. This does not prevent NZ from requesting information from/liaising with individual platforms as needed and vice versa

In coordination with DPMC/MFAT, information will be shared as required to inform the Christchurch Call Shared Crisis Response, should that Protocol be activated.

The Crisis Manager will be one of the nominated contact points for New Zealand for both the GIFCT and Christchurch Call Protocols. As part of their role, the Crisis Manager will act as a conduit between the Domestic Crisis Response Process and the GIFCT and Christchurch Call Protocols, when the latter are initiated, in close coordination with DMPC and MFAT.

Law Enforcement Guidance

During the crisis response the Crisis Response Team will work closely with the NZ Police to ensure that any steps taken as part of the response do not negatively impact any investigation underway.

The Police representative(s) on the update calls should be able to inform the other stakeholders about what content (where removal is occurring) is required to be sandboxed or preserved and where this is possible. They will also inform the group about any data and intelligence that should be captured where possible. There is an important need to balance public protection and harm minimisation with evidential needs for any criminal case.

Te Tari Taiwhenua
Internal Affairs

Te Kāwanatanga o Aotearoa
New Zealand Government