TE MANA WHAKAATU
Classification Office

15 November 2022

[redacted] By email: [redacted]

Tēnā koe [redacted],

Official Information Act request

Thank you for your request under the Official Information Act 1982 (OIA) about Te Mana Whakaatu—Classification Office's Countering Violent Extremism team strategy, which we received on 31 October 2022.

You said:

> I would like to request a copy of the Countering Violent Extremism team strategy for countering digital extremism and dangerous disinformation.
>
> The strategy I am requesting is referred to in OFLC material published on the Parliament website earlier this year:
>
> https://www.parliament.nz/resource/en-NZ/53SCGA_EVI_118632_GA21083/13ad81f62ca6fc8638a750d02b996cde65d7030f
>
> "In mid-2020 the CVE drafted an integrated strategy for addressing digital extremism and dangerous disinformation." (Quote on page 4).

In responding to your request, we have weighed up the factors in the OIA, including the purposes in section 4 and the principle of availability in section 5.

Response to your request

A draft document named 'Digital Extremism and Dangerous Disinformation Integrated Strategy' is attached to this response. This document remains in draft and is not public policy.

This document was prepared by the Classification Office's Countering Violent Extremism (CVE) team in 2020. In surveying the CVE landscape of Aotearoa, the team recognised that a strategic framework would be needed to guide supportive efforts behind core initiatives such as the Christchurch Call. The team pulled together several key areas of growth and opportunity including the role of dangerous mis- and disinformation in radicalising potential extremists.

FREEPHONE: 0508 236 767 **PHONE:** +64 4 471 6770
EMAIL: info@classificationoffice.govt.nz
Level 1, 88 The Terrace, PO Box 1999, Wellington 6140, New Zealand
www.classificationoffice.govt.nz

This approach has guided the CVE team's outreach and education priorities. The document has also informed the Office's research output, including the Edge of the Infodemic report released in 2021.

The Countering Terrorism and Violent Extremism Strategy was released in June 2021. This document details how government undertakes its responsibility to keep Aotearoa safe for all New Zealanders. The strategy sets out how agencies work together to prevent terrorism and violent extremism of all kinds in New Zealand, while ensuring the systems and capabilities are in place to act early and respond when needed. The strategy includes a work programme that focuses on reducing the threat of terrorism and violent extremism.

A Strategic Framework for Preventing and Countering Violent Extremism is being developed by the Department of the Prime Minister and Cabinet. This framework is set to be released at the end of 2022.

Publication of response

This response may be published on the Classification Office's website. If it is published there, your personal information will be redacted.

Right of review

You have the right to make a complaint and seek a review by the Ombudsman of this response under section 28(3) of the OIA. Information about this process is available at ombudsman.parliament.nz or freephone 0800 802 602.

Thank you for your interest in our mahi.

Ngā mihi nui,
Te Mana Whakaatu—Classification Office

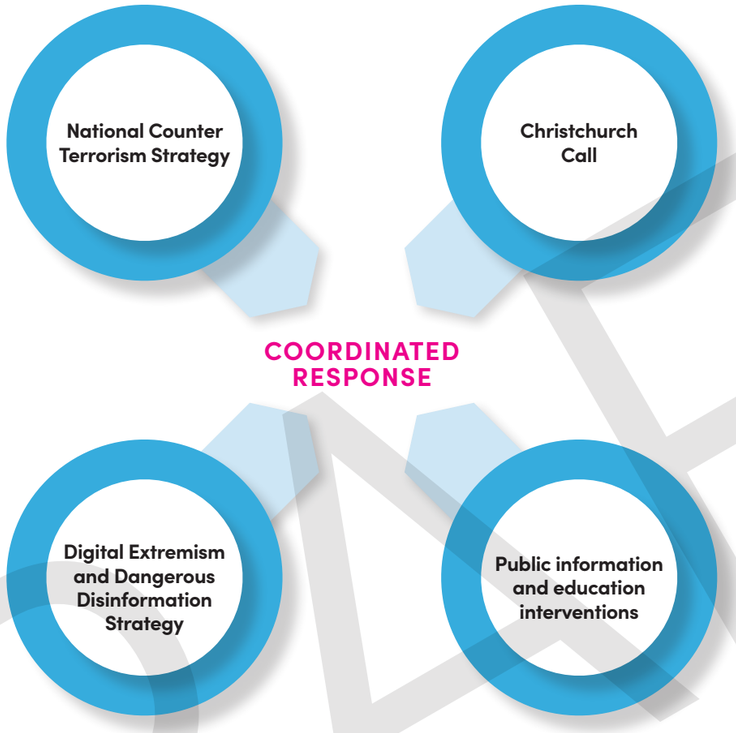
DIGITAL EXTREMISM AND DANGEROUS DISINFORMATION STRATEGY

Answering the call:

A strategy to disrupt violent extremist content and dangerous disinformation

Preventing extremist attacks is traditionally considered through a law enforcement and prevention lens as well as a national security perspective (critical for countering organised terror groups such as ISIS). Post Christchurch, the need for greater international collaboration and the importance of having governments and digital industry working together at a national/industry level to address/prevent distribution of viral, extremist material resulted in the unprecedented Christchurch Call initiative.

A national digital extremism and dangerous disinformation strategy is needed to support and reinforce these pillars, while providing an opportunity for national state, industry and community groups to work together to address these serious and growing threats.



Freedom and democracy first:

'Designing in' Bill of Rights protections to get the balance right.

Countering digital content harms is inherently complex and challenging, as the cyber environment is vast and ever changing (there are 500 hours of fresh video content put on YouTube per minute). Determining what is 'dangerous disinformation' as opposed to legitimate free expression is challenging.

Even seriously disturbing footage of a violent death can have important social value or can support just outcomes. It is vitally important that citizens can openly engage with and debate ideas and issues of all kinds. In order to navigate this complex space, **key principles** must be 'designed in' to any policy/major operational change:

- Freedom of expression
- Privacy
- Transparency
- Evidence-led
- Harm prevention

WHY do we need a strategy

There are indications that the rising tide of extremist, violent material and ideologies, along with potentially harmful disinformation is becoming a 'hard problem' that is beyond the capabilities of government agencies alone to fix.

Real progress demands an integrated strategy – so we can co-ordinate efforts across not just government agencies, but also incorporate the efforts of digital platforms, mainstream media organisations and community groups while supporting and engaging the public.

New Zealand

We have been uniquely impacted by the 'weaponisation' of social media and the viral propagation of extremist and abhorrent material. The livestream broadcast of the mosque attacker's acts on March 15 exploded across all major internet platforms and was viewed by many thousands of New Zealanders – both adults and children.

In the COVID-19 environment, the available evidence suggests there has been an increase in engagement with extremist content and potentially dangerous disinformation. A significant number of cell towers have been subject to attacks in NZ –that appear to be correlated with a rise in conspiracy theories linking 5G with the pandemic.

Globally

The 'livestream video' and 'Great Replacement ' documentation from the March 15 attacks have been referenced in at least five subsequent lethal, racially motivated attacks around the world.*

Internet analysts worldwide are observing that levels of extremist content created, exchanged and distributed on the internet are climbing since the onset of the pandemic.

Threats are being identified from new groups, often domestically-based and which do not fit traditional profiles and models of operation for terrorist organisations. These new challenges demand new ways of thinking, connecting and mobilising government, industry and the community.

THE ONLINE HARM PREVENTION GROUP (OHPG)

DIA | NETSAFE | TE MANA WHAKAATU Classification Office | Police | CERT | EDUCATION | INTERNET NZ | PRIVACY COMMISSIONER | NETWORK 4 LEARNING

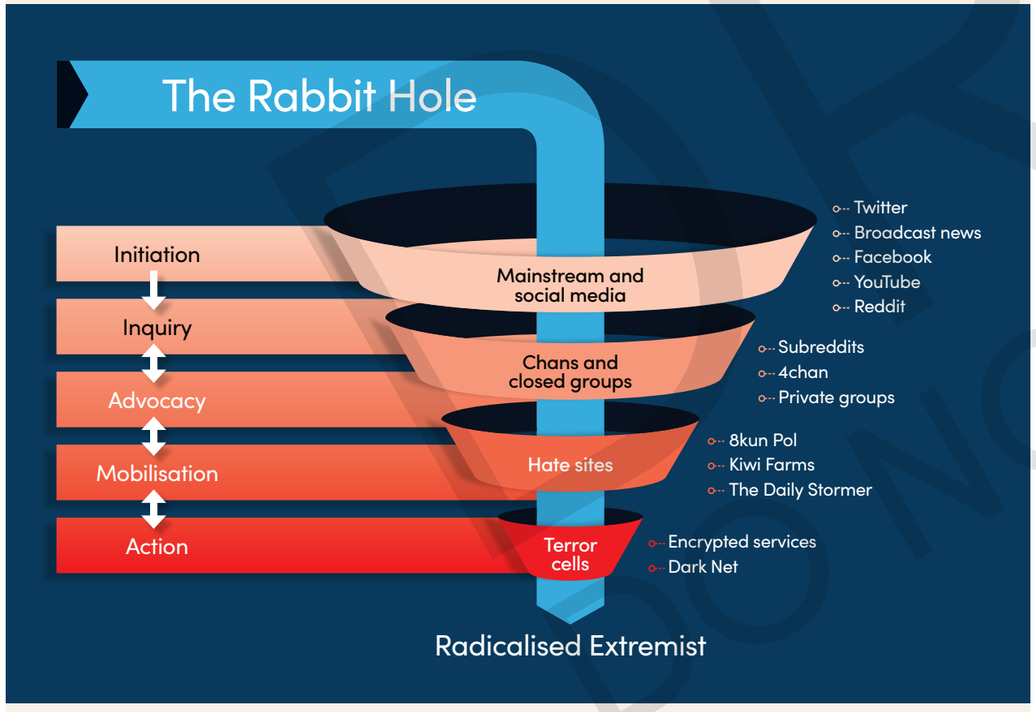
The Digital 'Rabbit Hole'

It is increasingly evident that online platforms can play a significant role in developing radicalised attitudes and distributing promotional extremist videos and literature.

While we have seen terror groups such as ISIS previously use digital platforms and social media in very sophisticated ways, now new groups including extreme white nationalist organisations are evolving the same tactics. A feature of the new digital environment is the highly decentralised, digital–savvy aspects of these groups, who are highly capable in adopting/leveraging fears (such as those related to the COVID-19 pandemic) to propagate hate and violent ideology.

Modern 'digital extremists' are highly aware of the mechanics of the 'rabbit hole' staying just within the terms of use of big platforms such as YouTube and Facebook in order to reach wide audiences, some of whom progress to less moderated and more extreme platforms (like 8chan/kiwi farms) and potentially then to encrypted 'cells'.

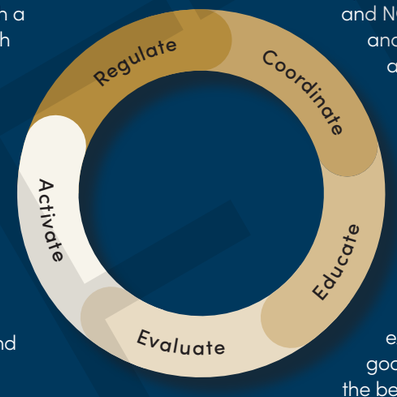
A key benefit in understanding the 'rabbit hole' means that we can start to see holistically how different techniques can disrupt the flow down the hole at different stages; good informative and accurate public information can disrupt the uptake of disinformation and how other levers can operate at different stages.



Strategy – key elements

Regulate
Current reviews of hate speech legislation and amendment of the FVPCA can be used to help reinforce and build awareness of a modern, effective strategy. Further regulation can be designed to connect with a modern, evidence-led approach to the digital word.

Coordinate
Having an integrated strategy to deal with complex online issues like digital disinformation and extremist content is vital. The OHPG represents a unique coalition of departments, agencies and NGO's with insight, responsibility and relationships to build and drive a truly comprehensive strategy, leveraging relationships with and understanding of digital industry as well as global partners.



Activate
Social media companies and digital platforms are central to the success of any moves in this area – integrating them into a planned set of staged interventions to help establish a more open, safer, transparent and less hate-filled internet.

Educate
Overseas experience indicates that education can be a key means of building resilience to disinformation/ extremism in young people. And good, trusted information can be the best antidotes for all ages.

Evaluate
Evidence-led policy is essential in this area. The understanding of what needs to be researched and evaluated is beginning to emerge. We can devise and execute research to properly inform solutions tailored for NZ.

INTERVENTIONS

There are a range of known interventions that can help counter misinformation and disrupt the flow towards extremist material.

CLEAR PUBIC INFORMATION can help counter disinformation and harmful narratives.

COMMUNITIES can rally online groups and narratives.

EDUCATORS can help build digital literacy and critical skills to inoculate against propaganda and manipulation.

DIGITAL PLATFORMS can effectively police themselves to limit the propagation of hate and deplatform violent extremists.

NGO'S AND COUNTER HATE GROUPS can identify extremists 'lurking in plain sight' and call out emerging dangerous and violent ideologies.

CROWN ENTITIES can monitor, research, report and in some cases classify material as restricted or unlawful.

ENFORCEMENT AGENCIES such as Police, DIA can help develop more sophisticated tools such as take-down notices, digital filters, nudge notifications and contact research.

LAW ENFORCEMENT/STATE SECURITY can play their traditional prevention and security roles, while being informed by and in turn informing the broader collaborative effort.

EXECUTION – FIRST STEPS

1. COORDINATE
The OHPG is well placed to coordinate and communicate an approach reflected in the digital strategy. It could work with key stakeholders, industry and community groups and could align with existing work/ strategies.

2. EDUCATE
Research is being gathered on current concerns in the education sector around disinformation – this can in turn help inform support/modules

3. EVALUATE
DIA is undertaking NZ Research (ISD) and the Classification Office is scoping a National Survey. Work with industry research and academics could be coordinated to develop a coherent evidence base.

4. ACTIVATE
Set up an inter-agency group with responsibility for developing strategy further and building networks.

5. REGULATE
One opportunity is progressing the CVE Bill. Hate speech law reforms could also be coordinated through an integrated approach. Further relevant work can be coordinated under a general media review.