

Office of the Prime Minister



Prime Minister

Minister for National Security and Intelligence

Minister for Child Poverty Reduction

Minister Responsible for Ministerial Services

Associate Minister for Arts, Culture and Heritage

Scott
fyi-request-20735-d62cca62@requests.fyi.org.nz

28 OCT 2022

Tēnā koe Scott

Ref: NSI OIA 2022-016

Official Information Act request for copies of listed NSI briefings

Thank you for your Official Information Act 1982 (the Act) request, received on 3 October 2022. You requested:

"I would like to request copies of the following National Security and Intelligence portfolio-related briefings received by this office:

ONE: [1718NSP/033] "Foreign Interference - Brady Report and Canberra Visit" [December 2017]

TWO: [1819NSPD/064] "Our parting thoughts and wishes for the National Security System" [December 2018]

THREE: [2021NSP/014] "QAnon - Designation and Dis-information" [October 2020]

FOUR: "Russia-Ukraine Situation and National Security System Preparedness" [1 February 2022]

FIVE: "Potential Domestic Implications of Russia-Ukraine Conflict and Proposed Response" [8 February 2022]

SIX: "2022-02-25 ODESC note to PM: Russia/Ukraine ODESC #1" [25 February 2022]

SEVEN: "2022-02-28 ODESC note to PM: Russia/Ukraine ODESC #2" [28 February 2022]

EIGHT: "Policy levers for Addressing Mis/Disinformation" [9 March 2022]

NINE: "Update on Algorithmic Workstream" [8 April 2022]

TEN: "Briefing on the Declaration on the Future of the Internet" [20 April 2022]

ELEVEN: "Progress for a framework on Emerging Technology in New Zealand" [2 May 2022]

TWELVE: "Work Programme on End-to-End Encryption" [11 May 2022]

THIRTEEN: "Foreign interference targeting New Zealand communities: letter to Minister Radhakrishnan" [12 May 2022]

FOURTEEN: "Briefing to PM on Buffalo Shooter incident" [25 May 2022]"

Please find responses to the following parts of your request (bolded) in turn below:

ONE: [1718NSP/033] “Foreign Interference - Brady Report and Canberra Visit” [December 2017]

With regard to your request for the briefing “Foreign Interference – Brady Report and Canberra Visit”, this is withheld in full under the following section of the Act:

- section 6(a), as disclosure would be likely to prejudice the security, defence or international relations of New Zealand.

TWO: [1819NSPD/064] “Our parting thoughts and wishes for the National Security System” [December 2018]

With regard to your request for the briefing “Our parting thoughts and wishes for the National Security System”, this is withheld in full under the following sections of the Act:

- section 6(a), as disclosure would be likely to prejudice the security, defence or international relations of New Zealand; and
- section 9(2)(g)(i), to maintain the effective conduct of public affairs through the free and frank expression of opinion.

THREE: [2021NSP/014] “QAnon - Designation and Dis-information” [October 2020]

With regard to your request for the briefing “QAnon – Designation and Dis-information”, this is withheld in full under the following sections of the Act:

- section 9(2)(c), to protect the health or safety of members of the public; and
- section 9(2)(g)(i), to maintain the effective conduct of public affairs through the free and frank expression of opinion.

SIX: “2022-02-25 ODESC note to PM: Russia/Ukraine ODESC #1” [25 February 2022]

With regard to your request for the briefing “2022-02-25 ODESC note to PM: Russia/Ukraine ODESC #1”, please find a copy enclosed. Some information has been withheld under the following section of the Act:

- section 6(a), as disclosure would be likely to prejudice the security, defence or international relations of New Zealand.

SEVEN: “2022-02-28 ODESC note to PM: Russia/Ukraine ODESC #2” [28 February 2022]

With regard to your request for the briefing “2022-02-28 ODESC note to PM: Russia/Ukraine ODESC #2”, please also find a copy enclosed. Some information has again been withheld under the following section of the Act:

- section 6(a), as disclosure would be likely to prejudice the security, defence or international relations of New Zealand.

EIGHT: “Policy levers for Addressing Mis/Disinformation” [9 March 2022]

Please find below in Appendix A extracts from the briefing “Policy Levers for Addressing Mis/Disinformation”. The remaining information in this document is withheld in full under the following sections of the Act:

- section 6(a), as disclosure would be likely to prejudice the security, defence or international relations of New Zealand;

- section 9(2)(ba)(i), to protect information which is subject to an obligation of confidence and its release would be likely to prejudice the future supply of similar information;
- section 9(2)(f)(iv), to maintain the confidentiality of advice tendered by or to Ministers and officials; and
- section 9(2)(g)(i), to maintain the effective conduct of public affairs through the free and frank expression of opinion.

NINE: “Update on Algorithmic Workstream” [8 April 2022]

With regard to the briefing titled “*Update on Algorithmic Workstream*”, please find a copy enclosed. This is withheld in part under the following sections of the Act:

- section 6(a), as disclosure would be likely to prejudice the security, defence or international relations of New Zealand;
- section 6(b)(i), to protect the entrusting of information to the Government of New Zealand on a basis of confidence by the Government of any other country or any agency of such a Government;
- section 6(b)(ii), to protect the entrusting of information to the Government of New Zealand on a basis of confidence by any international organisation;
- section 6(d), to maintain the safety of any person;
- section 9(2)(a), to protect the privacy of individuals;
- section 9(2)(b)(ii), to protect the commercial position of the person who supplied the information, or who is the subject of the information;
- section 9(2)(ba)(i), to protect the supply of similar information in the future;
- section 9(2)(g)(i), to maintain the effective conduct of public affairs through the free and frank expression of opinion;
- section 9(2)(g)(ii), to prevent improper pressure or harassment; and
- section 9(2)(j), to enable negotiations to be carried on without prejudice or disadvantage.

ELEVEN: “Progress for a framework on Emerging Technology in New Zealand” [2 May 2022]

With regard to the briefing titled “*Progress for a framework on emerging technology in New Zealand*”, this is withheld in full under the following section of the Act:

- section 9(2)(f)(iv), to maintain the confidentiality of advice tendered by or to Ministers and officials.

TWELVE: “Work Programme on End-to-End Encryption” [11 May 2022]

We have decided to release the relevant parts of the document titled “*Work Programme on End-to-End Encryption*”. Some information has been withheld under the following sections of the Act:

- section 6(a), as disclosure would be likely to prejudice the security, defence or international relations of New Zealand;
- section 6(c), to protect the maintenance of the law, including the prevention, investigation, and detection of offences, and the right to a fair trial;
- section 9(2)(a), to protect the privacy of individuals;
- section 9(2)(ba)(i), to protect the supply of similar information in the future; and
- section 9(2)(f)(iv), to maintain the confidentiality of advice tendered by or to Ministers and officials.

THIRTEEN: “Foreign interference targeting New Zealand communities: letter to Minister Radhakrishnan” [12 May 2022]

With regard to the briefing “*Foreign Interference Targeting New Zealand Communities: Letter to Minister Radhakrishnan*”, this is withheld in full under the following sections of the Act:

- section 6(a), as disclosure would be likely to prejudice the security, defence or international relations of New Zealand; and
- section 9(2)(f)(iv), to maintain the confidentiality of advice tendered by or to Ministers and officials.

FOURTEEN: “Briefing to PM on Buffalo Shooter incident” [25 May 2022]

We understand the title as listed in your request is a working title, provided due to an administrative error. The correct title is “*Buffalo Terrorist Attack and the Christchurch Call*”. Please find a copy of this briefing enclosed. Some information has been withheld under the following sections of the Act:

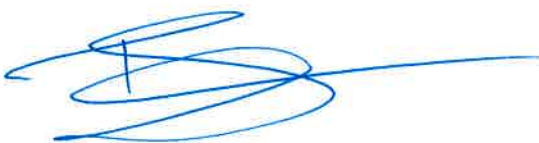
- section 6(a), as disclosure would be likely to prejudice the security, defence or international relations of New Zealand;
- section 6(c), to protect the maintenance of the law, including the prevention, investigation, and detection of offences, and the right to a fair trial;
- section 9(2)(a), to protect the privacy of individuals;
- section 9(2)(ba)(i), to protect the supply of similar information in the future;
- section 9(2)(c), to protect the health or safety of members of the public;
- section 9(2)(g)(i), to maintain the effective conduct of public affairs through the free and frank expression of opinion;
- section 9(2)(g)(ii), to prevent improper pressure or harassment; and
- section 9(2)(j), to enable negotiations to be carried on without prejudice or disadvantage.

In making our decision, we have considered the public interest considerations in section 9(1) of the Act.

With regard to the remaining parts of your request (4, 5, and 10), we have decided under section 15A of the Act to extend the time limit for deciding on these parts by an additional 20 working days. The extension is required because further consultations are needed to make a decision on these parts of your request. Despite this, a further response will be sent to you as soon as possible.

You have the right to ask the Ombudsman to investigate and review our decision under section 28(3) of the Act.

Ngā mihi nui,



Raj Nahna
Chief of Staff

Appendix A – Extracts from Briefing: Policy Levers for Addressing Mis/Disinformation, dated 9 March 2022

Executive Summary

1. Mis/disinformation is not a new problem. It has been around in various forms throughout history. But the arrival of mass social media has enabled mis/disinformation to be scaled at a global level for the first time, delivering an asymmetric tool for manipulation of large groups of people across borders and geographical distances, with limited traceability and at very low cost.
2. New Zealand, like many of our likeminded partners, has been impacted by networks that disseminate harmful mis/disinformation. This has affected for example the public health response to COVID-19. It has also manifested in more sinister ways, leading to radicalisation of at-risk individuals, threats and intimidation of people in our communities, and as a vehicle to undermine trust in democracy, media, and the institutions of government.
3. *Withheld in full*
4. There are no easy solutions. Our approach will need to be comprehensive and long-term to build resilience within society to manage future mis/disinformation topics. New Zealand is not alone in recognising the nature of the problem and grappling with appropriate ways to address it.

Paragraphs 5-10 withheld in full

Mis/disinformation leads to a range of harms...

11. Mis/disinformation has emerged as a complex security issue around the world. The development of online tools and social media environments that enable mass participation has enabled the spread and promulgation of mis/disinformation, with implications that governments and democratic societies have been relatively slow to recognise.
12. There are potentially harmful effects from the spread of mis/disinformation. At the most acute end of the spectrum these harms include threats to public safety, incitement of criminal or violent extremist activity, breakdown of social cohesion, and efforts to undermine democratic institutions. Mis/disinformation can play a role in recruitment and radicalisation strategies of violent extremist groups. It can also be a tool of corporate or State influence or driven by ideologically motivated groups. Unwitting accomplices are often involved in the dissemination of misinformation online. Distinguishing innocent participants from the nodes of deliberate and coordinated inauthentic behaviour can be difficult.
13. The impacts in New Zealand have become significant, exacerbated by the effects of the COVID-19 pandemic. While a full analysis is yet to be carried out, mis/disinformation appears to have played a significant role in the February 2022 occupation of Parliament grounds and associated violence and intimidation towards lawmakers, the media, academics, authorities, officials, and the general public. Researchers, civil society groups, tangata whenua, and the internet community have for some time expressed concern about the impacts.
14. *Withheld in full*

15. Money also plays an important role, because of the potential to monetise mis/disinformation content as a revenue source (for extremist groups and grifters alike) and as a means of financing the production, dissemination, and promotion of mis/disinformation. The economics of mis/disinformation are largely obscure, although it appears those who profit from and finance mis/disinformation frequently operate at a safe distance from those most impacted by it.
16. Mis/disinformation networks can also be driven by a range of other factors, including charismatic online personalities, development of a sense of community or shared knowledge, relatively complex psychological phenomena (including e.g. “rabbit-holing” – the descent into belief in bizarre, improbable or conspiratorial theories; or “brigading” – where groups engage in coordinated attacks on others), and genuinely held ideological beliefs (some moderate, others less so).
17. Addressing these problems is particularly challenging, whether for governments or communities. This problem set, like the online networks that drive it, is new enough that tried and tested solutions are yet to emerge. Mis/disinformation typically fits within definitions of protected or political speech. Most mis/disinformation is legal – making it difficult to address using traditional law enforcement and intelligence tools before the harms become apparent. Regulatory tools have proved particularly difficult to develop. It can be challenging to trace the financial flows, or to keep up with the shifting patterns of mis/disinformation and their impact.
18. *Withheld in full*

In the lead up to the ‘convoy’ occupation there were clear signs of mis/disinformation networks looking to incite harm and disorder, hidden in plain sight among what seemed like more typical forms of political protest. Whilst that ambiguity made it more difficult for agencies to focus in on the threat, many community groups and academics were actively calling them out. It’s likely that security agencies will continue to struggle as such situations arise in the future.

Remainder of briefing (paras 19-43) withheld in full.



ODESC

Officials' Committee for Domestic
and External Security Coordination

To	Prime Minister
From	Tony Lynch, Chair of ODESC
Prepared by	Marika Hughes, Director, National Security Systems Directorate
CC	Relevant Ministers
Date	25 February 2022
Subject	Ukraine/Russia ODESC #1

Key points

- ODESC met for the first time in response to the Russian invasion of Ukraine that occurred on 24 February 2022.
- There are a number of workstreams already underway across government agencies. They are well-coordinated in their activities across a range of risks and issues, including diplomatic activity, economic and trade impacts, cyber threats, domestic security implications and partner expectations.
- s6(a) [REDACTED]

Background

1. The first National Security System ODESC meeting was held today (25 February 2022) to ensure agencies have a consistent understanding of the Russian invasion of Ukraine, including about New Zealand's response activities; and to consider any risks and implications arising. Chief Executives and senior delegates from 15 agencies attended.¹

Situation update

2. s6(a) [REDACTED]
3. s6(a) [REDACTED]

¹ Department of Internal Affairs, Department of the Prime Minister and Cabinet (Chair), Government Communications Security Bureau, Ministry for Primary Industries, Ministry of Business, Innovation and Employment, Ministry of Foreign Affairs and Trade, Ministry of Defence, Ministry of Transport, National Emergency Management Agency, New Zealand Defence Force, New Zealand Security Intelligence Service, New Zealand Police, Reserve Bank of New Zealand, Te Kawa Mataaho Public Service Commission and The Treasury.

s6(a)

New Zealand's response activities

5. The Chief Executive, Ministry of Foreign Affairs and Trade (CE MFAT), provided a comprehensive overview of the response activities already undertaken or underway, including but not limited to:
- The statements made by yourself and the Minister of Foreign Affairs, condemning the Russian invasion and outlining the suite of response options under consideration including travel bans, export controls and curtailment of diplomatic relationships.
 - The engagement that has already occurred with New Zealand exporters, and the plan to further engage with them (with the Ministry of Primary Industries and other relevant agencies) early next week.
 - The reinforcement of MFAT consular staff in Warsaw to support any additional consular requirements. So far, they have received two consular calls, and have deployed a small team to one of the nine reception points Poland has established at the border for the outflow of Ukrainian nationals.
 - The consideration being given to the appropriate type of support for any humanitarian assistance required, in conjunction with key overseas partners. s6(a)
6. MFAT is operating its Emergency Coordination Centre to support this work, producing twice-daily sitreps for relevant agencies to enable the situational awareness necessary for developing coordinated policy advice. The Department of the Prime Minister and Cabinet (DPMC) is also providing strategic coordination in the policy space.

Key risks and implications

7. ODESC canvassed several areas of risk and lines of effort:

s6(a)

- [Redacted]

Partner expectations

9. CE MFAT noted that New Zealand is well-aligned with partners and that there are good information flows in place, a sentiment echoed by other ODESC members throughout the meeting in relation to their engagement with overseas partners.

10. s6(a)

Energy/economic/trade/investment/domestic security impacts

11. Several ODESC members provided insights into expectations related to broader consequences for New Zealand:

- For energy and fuel, the consequences will manifest in the supplies and price of oil, which will have flow-on effects for inflation. The United States had already indicated that it would release stockpiles, which would mitigate the impacts somewhat, but prices are already rising.
- Treasury anticipates that there will be reasonably significant indirect impacts, with dampened global growth, the impact of rising oil and other commodity prices etc, as well as impacts on supply chains. There is limited data available on Russian investment in New Zealand - approximately \$40m according to Statistics NZ, but they are aware that this does not capture everything, so will continue to investigate the quantum as well as possible implications.
- Reserve Bank of New Zealand (RBNZ) noted that markets continue to function – they are volatile but there is sufficient liquidity, and RBNZ has tools available to support if help is needed. They are keeping a close eye on what will happen if sanctions are enforced that removed Russia from the SWIFT banking system.
- Police are monitoring open sources for any negative sentiment from the approximately 15,000 Russians and 1800 Ukrainians currently in New Zealand but have seen nothing of concern at this point.

• s6(a)

Next Steps

12. s6(a)

13. We will keep you apprised of any further significant risks or implications if they arise.



ODESC

Officials' Committee for Domestic
and External Security Coordination

To	Prime Minister
From	Tony Lynch, Chair of ODESC
Prepared by	Marika Hughes, Director, National Security Systems Directorate
CC	Relevant Ministers
Date	28 February 2022
Subject	Ukraine/Russia ODESC #2

Key points

- ODESC met for the second time in response to the Russian invasion of Ukraine that occurred on 24 February 2022.
- Agencies understand the need to move at pace in this situation, to ensure that they are taking all measures available to remain aligned with partners, and to think innovatively about what options may be available within current settings to enhance New Zealand's response.
- s6(a)

Background

1. ODESC met today (28 February 2022) at the request of the Ministry of Foreign Affairs and Trade (MFAT) as lead agency to ensure agencies have a consistent understanding of the lines of effort underway; to encourage thinking about what other response options may be available, and to consider any other issues arising. Chief Executives and senior delegates from 16 agencies attended.¹

Situation update

2. s6(a)
3. s6(a)

¹ Crown Law, Department of Internal Affairs, Department of the Prime Minister and Cabinet (Chair), Government Communications Security Bureau, Ministry for Primary Industries, Ministry of Business, Innovation and Employment, Ministry of Foreign Affairs and Trade, Ministry of Defence, Ministry of Transport, National Emergency Management Agency, New Zealand Defence Force, New Zealand Security Intelligence Service, New Zealand Police, Reserve Bank of New Zealand, Te Kawa Mataaho Public Service Commission and The Treasury.

s6(a)

[Redacted]

[Redacted]

New Zealand's response activities

5. The MFAT Chief Executive (CE) highlighted the need for agencies to be moving at speed, and thinking about what more can be done in response within current legislative frameworks and policy settings. The Chair noted that while the issue of autonomous sanctions is for political consideration, agencies should be considering all possible tools available within the following lines of effort already underway:

- Travel bans and other immigration-related settings;
- Industry risks and other economic exposure;
- Foreign investment in New Zealand, both now and into the future;
- Trade restriction options;
- Humanitarian support; and
- The impact of some Russian banks being removed from the SWIFT banking system.

6. s6(a)

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

Released under the Official Information Act 1982

Key issues

7. ODESC canvassed a range of other issues, related to both possible response options and to impacts of the conflict for New Zealand. These included:
 - The need to consider New Zealand's position and the implications related to anyone located in New Zealand wanting to travel to Ukraine to fight (despite the "Do Not Travel" advice in place).
 - s6(a) [REDACTED]
8. ODESC also agreed that government should consider how it publicly communicates about New Zealand's response activities, ensuring there is a consistent narrative about what actions New Zealand is taking. The narrative needs to highlight even the smallest actions, which cumulatively demonstrate New Zealand's commitment to providing a response aligned with its key partners. Work will be done to consider what a possible outward-facing mechanism might look like, to keep the public and partners informed.

Next Steps

9. s6(a) [REDACTED]
10. We will keep you apprised of any further significant risks or implications if they arise.

Briefing

UPDATE ON ALGORITHMIC WORKSTREAM

To Rt Hon Jacinda Ardern, Prime Minister			
Date	8/04/2022	Priority	Routine
Deadline	6/05/2022	Briefing Number	DPMC-2021/22-1916

Purpose

To update you on the Christchurch Call algorithmic work, highlight some of the problems the Community is working on, describe a sample of approaches from researchers, tech companies, and policymakers, ^{s 9(2)(f)}

Recommendations

- Note** ^{s 9(2)(f)} the Call Unit has made some progress identifying blockages, finding ways forward, and concentrating effort across the Community; and
- Consider** ^{s 9(2)(f)} to help set the direction of Call work on algorithms leading into the next Summit in September. **YES / NO**

^{s 9(2)(g)(ii)}
Paul Ash Christchurch Call Coordinator
8 / 4 / 2022

Rt Hon Jacinda Ardern Prime Minister
...../...../.....

Contact for telephone discussion if required:

Name	Position	Telephone	1st contact
Paul Ash	Christchurch Call Coordinator	s 9(2)(a)	✓
s 9(2)(a)	Chief Advisor	s 9(2)(a)	

Minister's office comments:

- Noted
- Seen
- Approved
- Needs change
- Withdrawn
- Not seen by Minister
- Overtaken by events
- Referred to

Released under the Official Information Act 1982

UPDATE ON THE CHRISTCHURCH CALL'S ALGORITHMIC WORKSTREAM

Summary

The Call's algorithmic commitments are the most complex and sensitive for the tech sector. In addition to the technical complexity of this work, tech leaders are juggling proprietary interests, competitive dynamics, and regulatory uncertainty. ^{s 9(2)(f)}

[Redacted]

^{s 9(2)(g)(i)} [Redacted]

^{s 9(2)(f)} [Redacted]

We have suggested some options that can be developed further, subject to your feedback:

^{s 9(2)(f)} [Redacted]

Purpose

To update you on the Christchurch Call algorithmic work stream, highlight some of the problems the community is working on, describe a sample of approaches from researchers, tech companies, and policymakers, ^{s 9(2)(f)}

[Redacted]

Report

1. As you know, the Call contains several commitments related to the important role algorithmic processes play in the management of terrorist and violent extremist content (TVEC) online:

Online Service Providers (OSPs)	OSPs and Governments together
6. Take measures to prevent the upload and dissemination of TVEC, including through cooperative technology development	11 and 14. Facilitate the development and deployment of interventions and the redirection of users
11. Review the operation of algorithms that may drive users towards TVEC, and implement changes, and mechanisms for reporting	15. Accelerate technical development of tools for detection and removal. 18. Support smaller platforms through sharing of technical solutions.

2. There has been progress against these commitments, including the continuous improvements made to the GIFCT hash database, the broadening of hash sharing members, and the deepening of the taxonomy. There have also been continuous improvements made to platforms' algorithmic moderation systems. As a result, ^{s 9(2)(f)}

[Redacted]

3. This has meant a shift in focus, from the issue of algorithms helping to drive users towards TVEC, to the broader question of how the user environment facilitates radicalisation and amplification of non-violative and legal content. As part of the mandates developed for the May 2021 Leaders' Summit, the Community prioritised work to assess data and information needs of researchers to understand user journeys, and the impact of platform features on radicalisation and amplification.
4. Behind all this sit two core policy questions:
 - a) Do content recommendation systems based on artificial intelligence and machine learning (AI/ML) have the potential to 'amplify' or drive users towards TVEC, or towards radicalising ecosystems or networks of users that connect with TVEC off-platform?
 - b) Could recommender systems increase the exposure of 'at risk' individuals to content and user networks which supply the ideological justification and means/strategies for achieving violent ends?
5. Meta, the European Commission, and the Institute for Strategic Dialogue (ISD) work with New Zealand as 'co-leads' of a regular Christchurch Call algorithmic working group meeting with around 70 people from across the Call Community, plus a few outside experts. This group helps oversee work taking place across the Community and advises where efforts are needed to realise our work plan objectives.

Getting answers is difficult, because of the nature of the systems being tested

6. AI/ML recommender systems are a **moving target** because they adapt over time from their interactions with users, as well as through iterative 'training' carried out by their owners. A system consists of multiple AI/ML algorithms. For example, one group of algorithms might identify particular characteristics (e.g. offensive language, graphic violence, manipulated imagery) and another might decide what to do with them based on a combination of identified characteristics (e.g. promote, demote, remove, send for human moderation). Multiple sets run in parallel serving different objectives (e.g. elevate high quality content, remove spam, delete harmful content etc.) Each of them is constantly adapting in response to new inputs including those that result from other algorithms running alongside. That means that the conclusions that can be drawn about a system at a particular point in time aren't necessarily reflective of the system as it will configure itself down the track.
7. AI/ML systems operate with **uncertainty**, which means they may be more or less certain about something they've been trained to identify and will make decisions based on what is deemed an acceptable level of uncertainty. False positives and false negatives are a feature of that system, as are potential biases and perverse consequences. There are real trade-offs to consider in managing this, as reducing uncertainty may come at the cost of bias (i.e. an algorithm can be more certain about content relating to a particular racial or language category than another). This can vary between content types. For instance, AI/ML systems are very good at classifying nudity, or finding copyright music, but much worse at assessments that involve context, e.g., bullying, satire, or irony.
8. Due to **personalisation features**, the outputs of AI/ML recommender systems can vary depending on the characteristics and behaviour of the individual user. This can include who they are friends with, their previous activity, or other characteristics the system has learned to associate, which even the programmer may not be aware of. This makes repeatable simulation difficult and creates a range of potential privacy law and data protection concerns as well as ethical questions for researchers around testing of users.

- 9. It is also difficult to test theories around amplification of TVEC on major platforms because **TVEC isn't supposed to be there**. Social media platforms participating in the Christchurch Call have measures to prevent upload and to remove TVEC. Content that sneaks through these systems can't easily be tested because it's hard to identify, and once identified is immediately removed.

The Call Community has developed some ideas around how to get past these difficulties

s 9(2)(f)

- 10. A number of stakeholders have been working on the basis that the indirect route to address amplification and radicalisation is through identifying and suppressing so-called 'borderline content' or 'grey-zone content.' That content is not illegal and does not violate companies' terms of service but can be identified in various ways as being close to the 'policy line'. Facebook founder Mark Zuckerberg in his 2020 testimony to Congress suggested that engagement-based systems drive users closer to the policy line, which could be a 'gateway' towards more extreme ideas and beliefs.
- 11. There have been some empirical studies on Facebook and YouTube to test this theory, with a range of conflicting and inconclusive observations. A key challenge is how to measure 'extremeness' (i.e. agreed metrics). s 9(2)(ba)(i)

[Redacted]

- 12. s 9(2)(f) s 9(2)(g)(i) Previous attempts by the global advertising industry, through their 'Global Alliance for Responsible Media' (GARM), to develop uniform harm and prevalence metrics ss 9(2)(f); 9(2)(g)(i)

[Redacted]

- 13. s 9(2)(ba)(i)

[Redacted]

- 14. s 9(2)(f)
- 15. The United Kingdom is considering approaches in its Online Safety Bill that could restrict certain categories of harmful but legal content. s 6(a)

[Redacted] s 9(2)(f)

s 9(2)(f)

16. The Integrity Institute – a non-governmental organisation composed of former and current trust and safety employees at social media firms – has made a range of suggestions about recommender systems. They suggest the best approach is to move away altogether from an engagement-based model which risks promoting highly engaging adversarial narratives, whether or not these happen to be captured by an agreed definition of 'borderline'. Firms should instead optimise their recommender systems for 'quality' with the relevant metrics defined according to their corporate values system. These would need to be simpler and more values-based than existing terms of service or community standards.

17. ss 9(2)(j); 9(2)(ba)(i)

18. ss 9(2)(j); 9(2)(ba)(i)

...audits and reporting mechanisms

19. The only way to reliably understand whether changes to AI/ML-based systems are helping reduce the risk from amplification and radicalisation is through independently verified research. As part of a package of measures in the draft EU Digital Services Act, certain AI/ML systems including recommenders operating on large social media platforms would need to be opened for 'audit' by authorities or independent researchers. If retained in the final version of the Act, and once in force (i.e. in several years), this could provide the means to identify and potentially correct harmful biases.

20. Such audits are easier prescribed than done. Separating the outputs of an AI/ML based system from the inputs of human users and human moderators who are fully integrated within these systems requires some innovative approaches to experimental design, s 9(2)(f). While we are some way away from having a universal approach to multistakeholder reporting mechanisms across the whole Christchurch Call Community, there have been promising developments. The Call's Algorithms working group has narrowed down some of the questions and information needs in this area.

21. The GIFCT is working towards identifying viable and credible experimental approaches as part of its technical approaches working group. A/B testing is a method that can measure the impact of algorithmic tweaks across large samples of users. s 9(2)(ba)(i)

22. s 9(2)(i) [Redacted]

There are two to three emerging regulatory approaches

23. There are a range of emerging regulatory approaches in this area. s 6(a) [Redacted]

24. In anticipation of the EU Digital Services Act, France has lifted provisions from the draft EU bill introducing a 'duty of care' principle and a range of transparency provisions into French law. s 6(a) [Redacted]

25. In the US, several proposals have been tabled, s 6(a) [Redacted]

26. China has passed a law on social media algorithms which sets out a range of requirements, including transparency for users about recommendations and the ability to 'opt out' of recommendation features. It requires algorithmic systems to follow a range of ethical guidelines including to "prioritise mainstream values" and "create positive energy". s 6(a) [Redacted]

27. ss 9(2)(i); 6(a) [Redacted] Others, including the UK, Canada, Ireland and Australia are at different stages of developing legislation in this area, with a range of common features reflecting the desire to manage 'harms' and harmful behaviour that forms part of the user environment but which doesn't map directly onto illegal content.

28. In New Zealand, our current regulatory system has no specific requirements on transparency, nor algorithmic risk management or oversight. ss 9(2)(i); 9(2)(i)(iv) [Redacted]

We can improve on this work through greater company engagement

29. As reported in our briefing ^{ss 6(b)(1); 9(2)(g)(i)} of 17 December 2021 and aide-memoire of 9 February 2022, ^{ss 6(b)(1); 9(2)(g)(i)} on some of the outstanding issues identified by the Christchurch Call algorithms group. For instance:

^{ss 6(b)(1); 9(2)(g)(i)}



30. Our assessment is that over time, and ^{ss 6(b)(1); 9(2)(g)(i)} in forging a baseline consensus on access to data for researchers.

^{s 9(2)(f)}



31. Algorithmic amplification and the role of online user experiences in radicalisation are both important questions in the context of work on dis-/mis-/mal-information. ^{s 9(2)(f)}

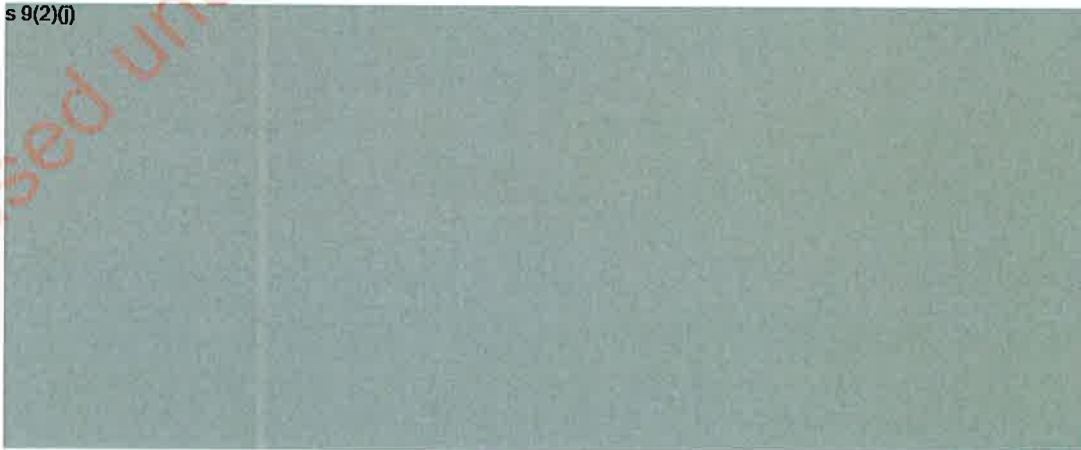


^{s 9(2)(f)}



32. ^{s 9(2)(f)} We suggest the following for your initial consideration:

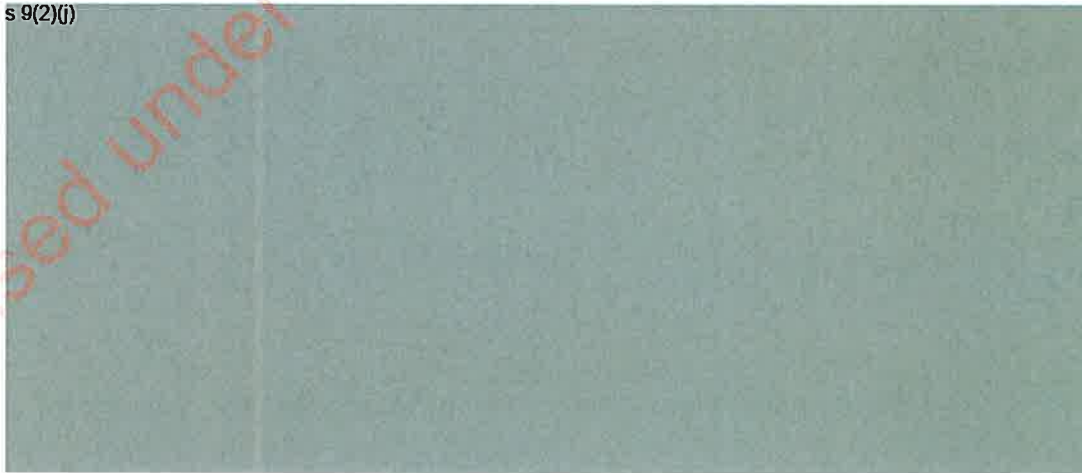
a) ^{s 9(2)(f)}

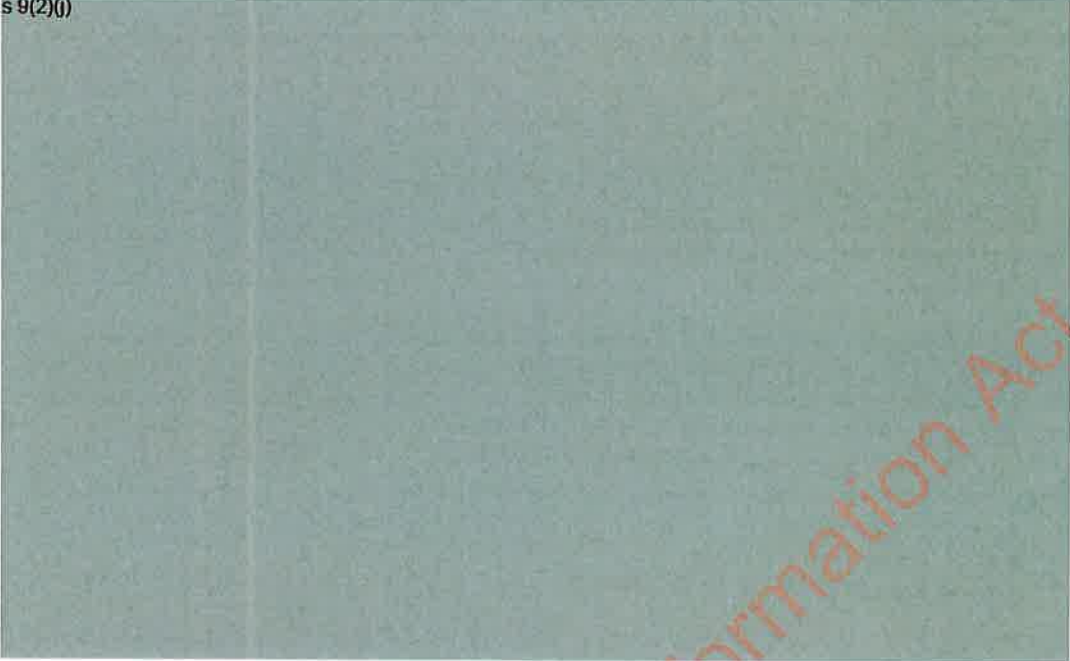


s 9(2)(f)

b)

c) s 9(2)(f)



d) ^{s 9(2)(f)} 

Next Steps

33. The Call Unit welcomes your feedback on the options outlined in paragraph 32 above and subject to your views, ^{s 9(2)(f)}

 ^{s 9(2)(f)}

The Call Unit would be available to discuss these with you prior to your visit, if helpful.

Attachments:	Classification:	Title:
Attachment A:	Unclassified	Table of Relevant Christchurch Call Commitments and Work Plan Objectives

Attachment A:

Relevant Call Commitments and Work Plan objectives

Pillar 1: Recommender algorithms and User Journeys

Christchurch Call Commitments		2021 Work Plan Objectives
11	<p>Review the operation of algorithms and other processes that may drive users towards and/or amplify terrorist and violent extremist content to better understand possible intervention points and to implement changes where this occurs.</p> <p>[...] This may include building appropriate mechanisms for reporting, designed in a multi-stakeholder process and without compromising trade secrets or the effectiveness of service providers' practices through unnecessary disclosure.</p>	<p>The Call Community will devote effort and resources to better understanding the "user journey" and the role this may play in the broader radicalisation process.</p> <p>We will design a multi-stakeholder process to establish what methods can safely be used and what information is needed - without compromising trade secrets or the effectiveness of Online Service Providers' practises through unnecessary disclosure- to allow stakeholders to better understand the outcomes of algorithmic processes and their potential to amplify terrorist and violent extremist content.</p>

Pillar 2: Detection and Removal of Terrorist & Violent Extremist Content

Christchurch Call Commitments		2021 Work Plan Objectives
6	<p>Take transparent, specific measures seeking to prevent the upload of terrorist and violent extremist content and to prevent its dissemination on social media and similar content-sharing services, including its immediate and permanent removal, without prejudice to law enforcement and user appeals requirements, in a manner consistent with human rights and fundamental freedoms. Cooperative measures to achieve these outcomes may include technology development, the expansion and use of shared databases of hashes and URLs, and effective notice and takedown procedures.</p>	<p>This year the Community will host an inclusive discussion on developing a framework to continuously review and improve the efficiency of [complaints and appeals processes] and support greater transparency and explainability in this area.</p>
15	<p>Accelerate research into and development of technical solutions to prevent the upload of and to detect and immediately remove terrorist and violent extremist content online, and share these solutions through open channels, drawing on expertise from academia, researchers, and civil society.</p>	<p>2019 9-Point Plan announced by Amazon, Google, Microsoft, Facebook, and Twitter</p> <p>"We commit to working collaboratively across industry, governments, educational institutions, and NGOs to develop a shared understanding of the contexts in which terrorist and violent extremist content is published and to improve technology to detect and remove terrorist and violent extremist content more effectively and efficiently.</p>
18	<p>Support smaller platforms as they build capacity to remove terrorist and violent extremist content, including through sharing technical solutions and relevant databases of hashes or other relevant material, such as the GIFCT shared database.</p>	<p>This will include:</p> <ul style="list-style-type: none"> • Work to create robust shared data sets to accelerate machine learning and AI and sharing insights and learnings from the data. • Development of open source or other shared tools to detect and remove terrorist or violent extremist content. • Enablement of all companies, large and small, to contribute to the collective effort and to better address detection and removal of this content on their platforms and services.

Pillar 3: Positive Interventions

Christchurch Call Commitments		2021 Work Plan Objectives
11	<p>Review the operation of algorithms and other processes [...]</p> <p>This may include using algorithms and other processes to redirect users from such content or the promotion of credible, positive alternatives or counter-narratives.</p>	<p>[We will...] Empower a new generation of community-driven online interventions</p> <p>This year the Call Community working with the GIFCT will seek to identify and empower the next generation of digital interventions against radicalisation, working to build a consistent framework for comparative evaluation...</p>
14	<p>Develop effective interventions, based on trusted information sharing about the effects of algorithmic and other processes, to redirect users from terrorist and violent extremist content.</p>	<p>Governments will work in an open multi-stakeholder context to identify information that could be shared to assist with positive interventions.</p>



Briefing

WORK PROGRAMME ON END-TO-END ENCRYPTION

To: Minister of National Security and Intelligence (Rt Hon Jacinda Ardern)
Minister Responsible for the NZSIS and GCSB (Hon Andrew Little)
Minister of Internal Affairs (Hon Jan Tinetti)
Minister of Justice (Hon Kris Faafoi)
Minister of Police (Hon Poto Williams)
Minister of Customs (Hon Meka Whaitiri)
Minister for the Digital Economy and Communications (Hon Dr David Clark)

Date	13/05/2022	Priority	Routine
Deadline	27/05/2022	Briefing Number	2122NPS/111

Purpose

This briefing outlines a Work Programme for end-to-end encryption, which includes collection of data to weigh up the impact of end-to-end encryption on law enforcement and public safety against the broader privacy, human rights, security, and economic benefits.

Recommendations

1. Note the National Cyber Policy Office is leading a cross-agency Work Programme on end-to-end encryption.

 Marika Hughes Acting Deputy Chief Executive National Security Group
11/05/2022

 Rt Hon Jacinda Ardern Minister of National Security and Intelligence
/ / 2022

Hon Andrew Little
Minister Responsible for the NZSIS
and GCSB

/ / 2022

Hon Jan Tinetti
Minister of Internal Affairs

/ / 2022

Hon Kris Faafoi
Minister of Justice

/ / 2022

Hon Poto Williams
Minister of Police

/ / 2022

Hon Meka Whaitiri
Minister of Customs

/ / 2022

Hon Dr David Clark
Minister for the Digital Economy and
Communications

/ / 2022

Contact for telephone discussion if required:

Name	Position	Telephone		1st contact
s9(2)(a)	Principal Policy Advisor, National Cyber Policy Office, Department of the Prime Minister and Cabinet	DDI s9(2)(a)	Mobile s9(2)(a)	✓
Halia Haddad	Manager, National Cyber Policy Office, Department of the Prime Minister and Cabinet	DDI s9(2)(a)	Mobile s9(2)(a)	

Minister's office comments:

- Noted
- Seen
- Approved
- Needs change
- Withdrawn
- Not seen by Minister
- Overtaken by events
- Referred to

Released under the Official Information Act 1982

WORK PROGRAMME ON END-TO-END ENCRYPTION

Purpose

1. This briefing outlines a Work Programme for end-to-end encryption, which includes collection of data to weigh up the impact of end-to-end encryption on law enforcement and public safety. s9(2)(f)(iv)
2. The objective of the Work Programme is to inform recommendations on how law enforcement can access the information it needs to ensure New Zealanders are protected from harm, while also protecting individual rights to privacy and security.
3. The work will consider the potential impact of end-to-end-encryption on law enforcement and public safety (e.g. enabling criminal activity, and impeding access to content for legal processes) and the benefits of end-to-end encryption to New Zealand businesses and individuals (e.g. increased privacy, security, and economic prosperity). This approach aligns with the Cyber Security Strategy vision that New Zealand is confident and secure in the digital world, and economic growth is enhanced.

Background

4. In considering the Budapest Convention Cabinet paper in October 2020, Cabinet directed officials to examine whether legislation, tools, and related settings relevant to cybercrime are fit for purpose and the challenges of enforcing the law in a digital age [Cab ref: CBC-20-MIN-0129].
5. The Budapest Convention Cabinet paper and subsequent advice to the Minister of Justice also confirmed the commitment for officials to coordinate a multi-agency process to identify cross-cutting challenges relating to public safety and law enforcement in the digital age, including the increasing use of end-to-end encryption.
6. The use of end-to-end encrypted communications has generated a difficult policy discussion since the 1990s, with regular bouts of the "crypto-wars" seeking to weigh the commercial and privacy benefits of encryption against impediments to law enforcement access to content data or plaintext.
7. The scale of the potential challenges to law enforcement agencies in New Zealand is unclear; this problem suffers from inadequate data across all crimes. Having datasets that plainly highlight the operational challenges end-to-end encryption creates for law enforcement will enable a better understanding of the impact on public safety, and better policy recommendations to decision-makers.

Encryption debate in context

8. End-to-end encryption excludes third parties from accessing the content or plaintext shared between communicating users. This encryption means that when a sender wants to communicate with a receiver, they share a unique key to decrypt the message. No one else can access that key – not even the service provider. End-to-end encryption for messaging apps and data storage that is, for practical purposes, unbreakable, is now openly available for anyone that owns a smartphone.

End-to-end encryption has significant benefits...

9. Strong encryption is necessary to ensure any communication or data transfer that requires privacy and confidentiality can occur over the public internet.¹ An open and secure internet that protects human rights, ensures fair economic competition, delivers secure digital infrastructure, promotes pluralism and freedom of expression are essential tenets of the 2022 Declaration for the Future of the Internet.²
10. The importance of encryption in protecting human rights is exemplified in repressive states, where end-to-end encrypted messaging apps allow human rights defenders to seek, receive and impart information securely and associate online.³
11. The war on Ukraine has demonstrated the value of end-to-end encrypted messaging apps to securely receive up-to-date information and advice on personal safety and public health and to convey information about the situation. In Ukraine, end-to-end encrypted messaging app Signal's downloads surged 1,075 percent between 24 February and 20 March 2022; in Russia, the app experienced 286 percent growth.⁴ This increased use of end-to-end encrypted messaging apps offers a glimpse into consumer demand for private communications when mainstream options are limited.
12. Encryption also provides economic value and utility to governments, the private sector, and end-users alike. It delivers critical privacy and cyber security controls that support the vision of a resilient and future-focused economy, as outlined in the draft Digital Strategy for Aotearoa.
13. As the tech sector expands, market trends have led to consumers and businesses preferring strong encryption for their digital products and services.⁵ Service providers across the wider technology industry now deliver a wide range of products and services with strong encryption as a core part of their offering. Enabling exceptional access to encrypted data for law enforcement purposes increases consumer concerns and is one driver of demand for encryption.

...But it also creates challenges for law enforcement and public safety

14. End-to-end encryption can cause two distinct problems when used in applications such as messaging and communications platforms. While the cause of these two problems is the same, their impact and proposed solutions differ:
 - a. End-to-end encryption enables criminal offending by shielding from investigators communications between suspects. For example, without the ability to examine content data suspected to contain child sexual abuse material (CSAM) or terrorist and violent extremist content (TVEC), investigators will be unable to determine if any images or videos are objectionable material contrary to the Films, Videos and Publications Classifications Act 1993.⁶

¹ The top five industries that use encryption are: Technology and software, transportation, healthcare, and pharmaceuticals, financial services and manufacturing. <https://www.enrtrust.com/-/media/documentation/reports/2021-global-encryption-trends-exec-summary-re.pdf>

² https://www.whitehouse.gov/wp-content/uploads/2022/04/Declaration-for-the-Future-for-the-Internet_Launch-Event-Signing-Version_FINAL.pdf

³ Freedom Online Coalition Joint Statement on Defending Civic Space Online <https://freedomonlinecoalition.com/wp-content/uploads/2021/06/FOC-Joint-Statement-on-Defending-Civic-Space-Online.pdf>

⁴ Growing Demand for Messaging Security in Ukraine and Russia <https://www.statista.com/chart/27161/pre-post-invasion-downloads-signal-telegram-ukraine-russia/>

⁵ The New Zealand tech sector starred in the local M&A market during 2021, which saw an overall increase in deal volume of 55 per cent. <https://www.reseller.co.nz/article/694713/tech-sector-delivers-quarter-kiwi-m-deals-bumper-2021/>

⁶ See section 3 of the Films, Videos and Publications Classifications Act 1993 <https://www.legislation.govt.nz/act/public/1993/0094/55.0/DLM313407.html>

- b. End-to-end encryption can make it more difficult for service providers to identify and manage harmful and illegal content (such as CSAM or TVEC) on their services, which makes it much harder to prevent, and report, criminal activity on their networks.

s6(a), s9(2)(ba)(i)



New Zealand's current regulatory framework does not fully address end-to-end encryption

19. New Zealand has a range of regulations and powers that enable, with a warrant, lawful access to data held by individuals, telecommunications providers, and other businesses. When properly obtained in this manner, this data can be used as evidence at trial.
20. The primary issue for law enforcement in using these powers is that if content data or plaintext is not provided, then the data accessed will be of little or no value. In some cases, useful access would only be possible with the consent of an individual suspect, in a situation where there may be no incentive for them to provide access to their encrypted communications and data.
21. Current legislation¹¹ on law enforcement access to data is largely based on the provision of data in a usable format. With end-to-end encryption hindering access, even for service providers, this is becoming increasingly infeasible.
22. Metadata, or the “outside the envelope” information (such as sender and receiver data, file size, time sent etc.), can assist law enforcement to identify perpetrators and victims. Over

⁷ <https://www.telegraph.co.uk/business/2021/11/20/people-shouldnt-have-choose-privacy-safety-says-facebook-safety/>

⁸ All U.S. service providers have a statutory obligation to report CSAM to the National Center for Missing and Exploited Children

⁹ <https://www.missingkids.org/gethelpnow/cybertipline/cybertiplinedata>

¹⁰ Ibid.

¹¹ The Telecommunications (Interception Capability & Security) Act 2013, the Search and Surveillance Act 2012, the Intelligence and Security Act 2017, the Telecommunications Information Privacy Code under the Privacy Act 2020, the Policing Act 2008, the Mutual Assistance in Criminal Matters Act 1992, the Customs and Excise Act 2018 and the Films, Videos and Publications Classifications Act 1993

time, however, as relevant metadata is encrypted (and this is an increasing trend) the identification of offenders and victims will also become more challenging.

Proposed Work Programme

23. s9(2)(f)(iv)

We don't currently have data on the impact of end-to-end encryption on investigations

24. s6(c)

s6(c)

The scale of the use of end-to-end encrypted apps for criminal purposes will be difficult to determine, due to lack of accessibility. This notwithstanding, data on how many investigations law enforcement agencies commenced, but have not been able to continue due to a lack of access, will be important and useful in understanding the problem. Collecting this data across all agencies will enable decision-makers to be better informed about any appropriate next steps.

25. A cross-agency working group¹³ will be established to agree how and what data is collected, and to monitor progress. The working group will analyse the data collected to inform recommendations on how law enforcement may be able to access the information it needs to protect individuals from harm, while protecting the rights of individuals to privacy and security and applying the principles of Te Tiriti o Waitangi. This will include exploring solutions for proportionate, reasonable, and necessary law enforcement access and reviewing whether agencies are taking full advantage of existing regulations and capabilities

s6(c)

26. The working group will also undertake an analysis of the benefits of end-to-end encryption, including whether any potential solutions to access information would undermine the security end-to-end encryption provides to New Zealand businesses and individuals.

27. A collaboration forum will be established to consult on the draft recommendations by the working group. To ensure a wide range of opinions are heard, the collaboration forum will include experts from the private sector, civil society, community groups and academia.

28. The working group will report to the CSSCC on the findings of the analysis, including any gaps, inconsistencies, or areas where interagency cooperation could be improved, and responses from the collaboration forum. The report will consider options to achieve the government priorities in the Cyber Security Strategy and to address the impact of end-to-end encryption on law enforcement and public safety agencies.

29. The Work Programme will also complement the CSSCC approved horizon-scanning of existing and emerging technologies. This project, due to report to the CSSCC in mid-2023 and led by the same National Cyber Policy Office team as the end-to-end encryption Work Programme, will lift the view of agencies above "current problems with current frameworks" to think ahead to what public safety and law enforcement will look like due to the logical

¹² This includes the Government Communications Security Bureau and the New Zealand Security Intelligence Service

¹³ CERT NZ, Ministry of Business, Innovation and Employment, Department of Internal Affairs Government Communications Security Bureau, Police, Ministry of Justice, Ministry of Foreign Affairs and Trade, Customs, Netsafe, New Zealand Security Intelligence Service, Office of the Privacy Commissioner and Stats NZ (building on their experience and knowledge of data collection)

consequences of current trends, or technologies that are on the horizon but not yet ubiquitous.

30. The working group report will also support the review of the Search and Surveillance Act 2012 and inform activity under the all-of-government Transnational Organised Crime Strategy.
31. Further updates and advice will also be provided to Ministers (see further below under next steps).

Ongoing international engagement will be required...

32. New Zealand is unlikely to be able unilaterally to address any issues with encryption. For example, efforts by New Zealand to provide a legislative override of end-to-end encryption could result in a range of challenges, including the potential withdrawal of services on which businesses and consumers rely.
33. As the overarching encryption environment will be strongly influenced by international technology, and market and regulatory developments, an ongoing effort will be required to monitor and influence international activity, including engagement with our closest security partners.
34. The working group will consider these international elements when assessing and prioritising any recommendations.

s9(2)(ba)(i)

35. s6(c), s9(2)(ba)(i)

36. s6(c), s9(2)(ba)(i)

Next Steps

37. The cross-agency working group will be established in July 2022. It will ensure robust measures are implemented to support the integrity and reliability of the data collection processes.
38. Data will be collected for a minimum period of three months. The working group will monitor progress to ensure the data collection remains on track. The CSSCC will be updated to ensure governance of the collection process.
39. At the conclusion of the data collection, the working group will review and analyse the data. The findings will be used to develop recommendations on how law enforcement might be able to access the information it needs while protecting the rights of individuals to privacy and security, and the principles of Te Tiriti o Waitangi. It will also include analysis of the

economic benefits of encryption to New Zealand, and whether any proposed changes would undermine the economic prosperity or security of New Zealand in general.

40. These recommendations will also draw on any relevant information from the collaboration forum, s6(c), s9(2)(ba)(i) [redacted] and international engagement with other jurisdictions.

s9(2)(f)(iv) [redacted]
[redacted]

s9(2)(f)(iv) [redacted]
[redacted]

Consultation

43. This briefing has been consulted with New Zealand Police, the Department of Internal Affairs, the Government Communications Security Bureau, the New Zealand Security Intelligence Service, the New Zealand Customs Service, the Inland Revenue Department, Stats NZ, the Ministry of Foreign Affairs and Trade, the Ministry of Business, Innovation and Employment, the Ministry of Justice, and CERT NZ.

Annex:	Classification:	Title:
Annex A:	RESTRICTED	Work Programme on End-to-End Encryption

Released under the Official Information Act 1982

Annex A is withheld in full under the following grounds:
Section 6(c)
Section 9(2)(ba)(i)
Section 9(2)(f)(iv)



Briefing

BUFFALO TERRORIST ATTACK AND THE CHRISTCHURCH CALL

To Prime Minister Rt Hon Jacinda Ardern			
Date	25/05/2022	Priority	High
Deadline	27/05/2022	Briefing Number	DPMC-2021/22-2210

Purpose

To brief you on the Buffalo shooting, the Christchurch Call Community's response, and to set out next steps, ahead of your meetings with Tech Leaders in San Francisco and Seattle, and the Biden Administration in Washington.

Recommendations

- Note** the Call Unit is gathering information on the response to Buffalo and is working closely with US counterparts.
- Note** our initial assessment is that the structures developed by the Call Community have been helpful in blunting the online impact of Buffalo, and that more work is needed to address gaps, outlined below.
- Agree** the attached supplementary talking points as a helpful guide to discussions with US government and technology firm leaders. **YES / NO**

s9(2)(g)(ii)

Paul Ash
Special Representative for Cyber and Digital

25/05/2022

Rt Hon Jacinda Ardern
Prime Minister

...../...../.....

Contact for telephone discussion if required:

Name	Position	Telephone	1st contact
Paul Ash	Christchurch Call Coordinator	s9(2)(a)	✓
David Reid	Chief Advisor, Christchurch Call Unit	s9(2)(a)	

Minister's office comments:

- Noted
- Seen
- Approved
- Needs change
- Withdrawn
- Not seen by Minister
- Overtaken by events
- Referred to

Released under the Official Information Act 1982

BUFFALO TERRORIST ATTACK AND THE CHRISTCHURCH CALL

Purpose

To brief you on the Buffalo terrorist attack and the Christchurch Call Community's response, and to set out next steps, ahead of your meetings with technology leaders in San Francisco and Seattle, and the Biden Administration in Washington.

The Buffalo attack was carefully planned and linked to Christchurch

1. On Saturday 14 May an 18-year-old man, Payton Gendron, drove 200 miles to a supermarket in Buffalo, New York. At around 2.30pm EDT [6.30am Sunday NZT] he began shooting in the parking lot, before shooting a security guard and entering the store. He killed 10 people and injured many more, in a terrorist attack he had meticulously planned over months. Gendron targeted the busiest shopping time in a predominantly African American neighbourhood.
2. Gendron wore a helmet and bulletproof vest, covered himself in white supremacist symbols and used a weapon modified to carry an illegal amount of ammunition. He used a GoPro camera to upload a livestream of the attack on the social media streaming service Twitch, which through its parent company Amazon Group is a supporter of the Christchurch Call.
3. According to media reports, the shooter had been radicalised online over the preceding two years, drifting from sub-Reddits dedicated to hunting through to 4chan discussions dedicated to weapons and political extremism. Late in 2021, he set up a private Discord server to journal his influences, motives, and planning for the attack, accumulating more than 500 pages over the next few months.
4. Gendron distilled some of this material, including excerpts from other manifestos, racist memes and conspiracies, and screeds of tactical information into a 180-page manifesto on Google Drive. Thirty minutes before the attack, the shooter invited a small group to join a private Discord server with links to the Twitch livestream and the manifesto.
5. There are numerous links between Buffalo and the 15 March 2019 attack in Christchurch:
 - Gendron was clearly inspired by Christchurch. He says in his manifesto that the Christchurch terrorist was the person most responsible for his own radicalisation.
 - Gendron sought out the Christchurch manifesto after seeing the livestream and credited it with starting his 'real research' into the "great replacement" conspiracy theory. His Discord diary reveals moments of doubt where engagement with the manifesto content helped to spur him forward.
 - Gendron had planned his attack for 15 March to coincide with the Christchurch anniversary but cited equipment failures for preventing this going ahead. He reported being suicidal after missing his chance to attack on the anniversary.
 - Gendron also copied the tactics deployed in Christchurch, although he eschewed Facebook Live in favour of the gaming platform Twitch. This choice was evidently due

in part to his belief that Facebook safety improvements might thwart him but also because he viewed Facebook as a 'boomer platform'.

6. Gendron has left a record that may help us understand what led him to carry out his attack. The record includes many early warning signs, including an obvious fascination with firearms and shootings, active planning and preparation for an attack, and a propensity to violence. None of these seem to have been flagged or acted upon. Further research on his online history may reveal other factors that should be considered as part of the Christchurch Call's work stream on algorithms and positive interventions.

Twitch responded to user reporting and acted within 2 minutes of the first report.

7. Representatives from Amazon and their subsidiary Twitch debriefed the Call Unit following the incident. Their systems had been upgraded following the Christchurch Call s9(2)(ba)(i) [REDACTED]. The Twitch livestream started at 2:00pm local time with 22 people following the livestream. s9(2)(ba)(i) [REDACTED]

8. The Unit's understanding, from speaking to independent researchers, is that one of the 22 users recorded the livestream and uploaded it to a platform called Streamable. Others took copies of the video, and the manifesto from Google Drive, and posted them to 4chan and Kiwifarms. Supporters then used fringe services like Catbox, Kaotic and Ghostbin to make copies of the video, manifesto, and diary, sharing links on mainstream social media as well as Bitchute, Gab, Parler, Patriots.win and Telegram. That initial capture seems to be the source of subsequent manipulated copies that have been seen across the web. s6(a) [REDACTED]

The GIFCT activated its content incident protocol

9. After being alerted to the incident s9(2)(ba)(i) [REDACTED] GIFCT and its members convened to assess the situation and took the decision to activate the Content Incident Protocol (CIP) at 4:52pm local time. GIFCT notified the US Government, and then issued a public statement and notified members of the Independent Advisory Committee (including New Zealand) by 6:30pm local / 10:30am NZT. s9(2)(ba)(i) [REDACTED]

10. Activating the CIP enabled GIFCT to begin adding hashes of the perpetrator's content to the hash-sharing database (HSD),¹ to help members find it quickly on their platforms and action it consistent with their terms of service. The CIP was active for just over 24 hours and resulted in 870 visually distinct items being added to the HSD, including 740 images and 130 videos. Member companies also took independent enforcement action. s9(2)(ba)(i) [REDACTED]

¹ Outside of a CIP, content can only be added to the database if it has been produced by an individual or organisation on the UNSC Consolidated Sanction List. GIFCT is working currently to create a more comprehensive definitional framework for its work.

New Zealand also activated early as part of the response

11. Upon seeing news of the Buffalo attack, the Christchurch Call Unit quickly convened officials across government, plugged into the US and French systems, and connected with the Global Internet Forum to Counter Terrorism (GIFCT) to support their response. The Unit worked with DPMC's Christchurch RCOI team to reach out to New Zealand communities; provided an update to the Christchurch Call community and engaged New Zealand and international media in response to extensive requests for comment. A key strategic communications priority was to encourage users who saw the livestream or manifesto not to amplify the terrorist's message and to report the content instead.
12. The acting Chief Censor 'called in' and provisionally classified both the manifesto and the video. That gave Tech Against Terrorism a basis to activate its Terrorist Content Analytics Platform, sending alerts to a wider range of firms about any content they might be hosting.
13. We subsequently learned that the Buffalo footage, on its own and, separately, edited together with Christchurch attack footage, was used to re-traumatise victims of the Christchurch attack. s6(c), s9(2)(a), s9(2)(ba)(i)

[REDACTED]

14. s6(a)

We need more information to assess the response, but early actions appear to have blunted the impact of the content from Buffalo.

15. s9(2)(g)(i)
The Call Unit has reached out to Amazon, the GIFCT, the Operating Board companies (Meta, Microsoft, Twitter, YouTube) and has scheduled discussions with OSINT researchers, the Christchurch Call Advisory Network, and the Integrity Institute to gather any insights. The information we have been able to gather so far is relatively general as reflected in this brief.

16. GIFCT is planning to hold a formal debrief on 16 June. s9(2)(ba)(i)
The Call Unit is also liaising with governments and CCAN about their information needs. We have emphasised to GIFCT that debrief processes needs to provide meaningful transparency for the community, going beyond GIFCT's actions to enable a system-wide view of the impact of the response. s9(2)(g)(i)

17. A total of three CIPs have been activated to date: Halle (October 2019); Glendale (May 2020); and Buffalo, with Buffalo being the first to take place following the establishment of an independent GIFCT secretariat. It is important to note the differences between the three cases.

- The Glendale attack was essentially ad hoc, with the perpetrator using his phone to video it and upload clips to Snapchat. There were no fatalities and, in the end, no terrorism charges.
- The Halle shooter was hindered by tactical errors and technical choices. His livestream, although 35 minutes long, was viewed live by only five people. Twitch s9(2)(ba)(i)

s9(2)(ba)(i) [redacted] s9(2)(g)(i) [redacted] six hours to take down an autogenerated archival copy of the video that more than 2,500 people viewed. GIFCT did not declare a CIP until seven hours after the attack.

- The Buffalo shooter was deliberate and thorough in planning the offline and online dimensions of his attack. It was therefore a more critical test of the crisis response system than either Halle or Glendale. Twitch and GIFCT responded rapidly. s9(2)(ba)(i) [redacted]

18. Amazon and Twitch representatives are unequivocal that the Call and associated work on crisis response in GIFCT have made a positive difference. Through the development of shared and individual processes and tools to detect, triage, and respond to incidents of this sort, the community is responding more quickly and reducing the prevalence of content on members' platforms.

19. Indeed, the reports of numbers of video views on different platforms are orders of magnitude less than after the Christchurch attack. Anecdotally, we understand that law enforcement and journalists found it difficult to track down copies of the video and manifesto on mainstream sites, a far cry from the situation after the Christchurch attack, when thousands saw this content involuntarily in their feeds. But it is hard to substantiate without good comparable data from the companies.

20. There is clearly still work to be done. One 54 second clip was reportedly viewed thousands of times on Facebook before it was taken down. Tests run by the New York Times and the Tech Transparency Project surfaced various clips with advertisements running alongside. Another video on Twitter was allegedly up for two days and attracted 261,000 views. And the platforms have clearly struggled to address out-links. One link on Facebook to Streamable was shared 46,000 times.

International media have not been complementary about the response

21. The balance of international media coverage has been critical of the response. While there are valid critiques there are also reasons to be sceptical about some of the criticism. For instance there has been extensive commentary on the length of time it took to remove the livestream from Twitch. s9(2)(ba)(i), s9(2)(c) [redacted]

22. One of the 'communications gaps' about the response has been about the adversarial nature of the content incident. The Buffalo terrorist, like the Christchurch terrorist, claims to have acted alone – a claim which is often taken at face value. However, a decentralised network of individuals supporting the terrorist in both cases ensured a resilient presence of the content online. These actions are not generally well understood or covered by media.

23. Addressing the resilience of content from Buffalo, as with Christchurch, will require conversations about measures s6(a), s9(2)(g)(i), s9(2)(f) [redacted]

Buffalo has shown that there are 'gaps' in the system and more work is needed...

24. The debriefing process will likely bring issues to light that need to be addressed. At a first look there are some obvious areas for focus, which we outline below:

s9(2)(c), s9(2)(g)(f)

[Redacted]

[Redacted]

s6(a), s9(2)(g)(i), s9(2)(i)

[Redacted]

Addressing user behaviour and drawing clearer lines regarding quotes and links

- The Buffalo and Christchurch manifestos are in some sense 'modular' and designed to be repackaged into smaller component pieces to assist with dissemination.

• s9(2)(g)(i), s9(2)(i)

[Redacted]

• s9(2)(g)(i)

[Redacted]

• s9(2)(g)(i), s9(2)(i)

[Redacted]

Helping small platforms

• s9(2)(g)(i)

[Redacted]

- s9(2)(g)(i) [REDACTED]
- The Call should play a role in assisting to ensure small platforms have access to useful tools, whether on reasonable commercial terms or through open sourcing. It would be useful to discuss with a range of companies what their challenges are and how the Call Community might help them to address TVEC more effectively.

s6(a),
s9(2)(i)

[REDACTED]

Smart approaches to dehumanising conspiracy theories and disinformation

- The dehumanising narrative that underpins Gendron's ideology, and that of the Christchurch terrorist, has a persistence beyond the realm of TVEC. s6(a) [REDACTED]
[REDACTED] It can be seen in a range of online and traditional media where it is amplified not just by social media algorithms but also by cable television.
- Counter-narratives and dialogue are among the more powerful tools that can be deployed against dehumanisation. They are also among the most difficult to orchestrate centrally. An holistic approach to disinformation will also be relevant to addressing some of the ideological drivers of violent extremism and terrorism.

Next Steps

25. In addition to offering condolences and solidarity to US counterparts New Zealand, as the co-founder of the Christchurch Call, can offer important substantive assistance in responding to what happened in Buffalo. s6(a) [REDACTED]

Among these we suggest:

- s6(a) [REDACTED]
- Working with the Call Community to identify priority areas of work for prevention and crisis response.
- Lifting our focus on algorithms and positive interventions as a means of addressing the radicalisation threat s6(a), s9(2)(i) [REDACTED]
- Lifting our focus on options to address the problem of online platforms that continue to host TVEC and allow users to port it back onto mainstream platforms.

- Engaging s6(a) on an innovation and technology fund to tackle some of these issues.
- s6(a)
- s6(a)
- Establishing a 'virtual team' to prepare for the Leaders' Summit, s6(a)
- Working s6(a) to ensure the July GIFCT Summit is substantive, meaningful, and a useful stepping stone towards the Call Community Leaders' Summit.

26. We provide supplementary talking points below which may prove useful with US government and technology sector leaders as you meet with them this week.

Attachments:			
Attachment A:	Pg10	Supplementary Talking Points for US officials	Withheld in full under sections 6(a) and 9(2)(g)(i) of the Act
Attachment B	Pg11	Supplementary Talking Points for Tech Firms	Withheld in full under sections 9(2)(g)(i) and 9(2)(j) of the Act
Attachment C:	Pg12	Comparison of GIFCT Content Incidents	Withheld in full under sections 6(a) and 9(2)(ba)(i) of the Act

Pages 10 - 14 are withheld in full under the following sections:

- Section 6(a)
- Section 9(2)(ba)(i)
- Section 9(2)(g)(i)
- Section 9(2)(j)