8 November 2019

Countering Violent Extremism Online

Submission to the Department of Internal Affairs

Table of contents

Introduction

InternetNZ supports effective responses to online extremism

1. InternetNZ stands for an Internet that is open, secure, and for all.

2. For most people in the world, the Internet enables positive and useful ways to connect and share. But that same connectivity can be abused by people seeking to promote and incite acts of extremist violence. The Christchurch terrorist attacks which targeted mosques and killed 51 people, were also livestreamed online. Immediate and later online sharing of this footage, and of material written by the perpetrator, increased harms from these attacks.

3. InternetNZ has worked to address violent extremism online, participating in and supporting efforts by the Government which led to the Christchurch Call to eliminate terrorist and violent extremist content online. We stand with our community, and with New Zealanders, in condemning violent extremism and those who promote it online.

We support the goal of enabling a more effective response

4. We welcome work to effectively address the abuse of the Internet by violent extremists, including through updates and reforms of New Zealand law.

5. This is a complex area for policy, as the Prime Minister highlighted in her October 2019 keynote speech at NetHui. It is true that online services using the Internet were abused by a terrorist attacker, but they were also used by people to check in on family, friends and whole communities. As she said, effective policy responses will require working with civil society and the Internet community, to implement the Christchurch Call in a manner that is consistent with a free, open and secure Internet, and with international human rights law.

6. Our comments in this submission relate to the specific proposals put forward by the Department of Internal Affairs (DIA) in targeted consultations, as detailed in written materials shared through those targeted consultations.

7. These materials set out an overall goal of enabling a more effective response to violent extremist content online, through a series of policy processes.

We think proposed law changes require a phased approach

8. The immediate focus of consultation is to advance changes to the legal framework for objectionable content under the Films, Videos and Publications Classification Act 1993 (FVPCA).

9. Our view, which we heard echoed across the consultation meetings, is that some of the proposed policy measures need more detail and time to assess. While some of the proposals can be advanced to legislation as put forward and without delay, others need work to get ready for meaningful consultation.

10. We support a staged approach to the immediate proposals, advancing some now and some through processes allowing more analysis and time.

11. We recommend that the specific proposals put forward are addressed in three stages, as below:

    Stage one: Advance the proposal for **interim Censor decisions immediately**, perhaps through a statutes amendment bill amending the FVPCA.

    Stage two: Advance the proposal for **take-down notices in 2020**, with more detail and enough time for consultation on complications.

    Stage three: Consider **other proposals as part of the broader review** of media law, and in parallel with broader work to address the issues underlying violent extremist behaviour online and offline.

12. We would welcome the chance to discuss our perspective on the issues, and on the best ways to advance the goal of effectively responding to violent extremism online.

Kim Connolly-Stone

Policy Director

InternetNZ

Effective policy requires robust processes

We welcome collaboration and clearly framed policy problems

13. As you are well aware, the starting point for good policy is knowing what problem we want to solve. The goal of this policy work, set out in the engagement materials, is:

    > To address the harms from online violent extremist content caused by (a) viewing, (b) re-victimisation, (c) radicalisation, and (d) learning how to commit terrorism.
    >
    > By enabling Government authorities and online content hosts to act with certainty and speed to remove online violent extremist content.

14. In the service of that goal, the engagement materials identify specific policy problems, framed as gaps in the current law identified after the Christchurch terror attacks.

15. The immediate focus is to advance law changes to address gaps in the FVPCA framework, ahead of a full review of the FVPCA, and a broader review on content and media regulation.

16. We understand the immediate proposals have been chosen to allow quick movement, and to deliver clarity so that the parties responding to violent extremist content online, including in emergencies, are clear on what will happen and are well-coordinated in their approach.

17. We welcome the commitment to progress all legislative changes through a collaborative, evolving partnership with online content hosts and civil society, though tight timeframes for the current process are making this difficult.

But many of the proposed changes need more policy work

18. While we appreciated the chance to join consultation workshops, these offered very limited scope to effectively discuss and address consultation questions. Providing an informed perspective was difficult, with only brief, high-level descriptions of problems and proposals in the consultation materials. We have heard this concern about the effectiveness of consultation echoed from other stakeholders.

19. Key considerations require further policy work. For example, a meaningful assessment of the balance given to human rights would require government to prepare and share a human rights assessment for public consideration.

20. While we welcome the opportunity to view an exposure draft of the Bill before the end of this year, the exposure draft process will not address the need for more policy work before the Government makes a decision to proceed with any or all of the proposed changes. This work could result in different options or approaches being chosen.

21. Designing effective regulation that works with the Internet requires consideration of the complex mix of actors, behaviours, technologies, and business incentives that can help to address harms from violent extremism online. Some proposals have potentially broad implications, which can only be identified and worked through in detailed conversations which include people who can speak to these issues from a range of different perspectives.

22. We support a robust regulatory process which can consider and address this complexity. This should start with a regulatory impact assessment (including an Internet eco-system lens) to identify and address potential unintended consequences of these proposals. We think this type of process could create robust and sustainable rules, which remain fit for purpose as new challenges emerge over time.

Good outcomes require broader consultation on the proposals

23. We think many of the proposals require broader consultation to ensure that resulting law changes will be effective and perceived as legitimate.

24. Across the targeted consultation workshops, we heard concerns that there was not enough detail or time to adequately assess the implications of proposals.

25. We also see a risk that targeted consultation did not reach people whose perspectives are vital to adequately assessing the proposals as they would operate in practice. Key perspectives that may have been missed include organisations like TradeMe and Neighbourly which operate online platforms based in New Zealand, researchers addressing the uses and abuses of online services, and independent legal experts who will interpret and apply the law.

Recommendation and next steps

We recommend a three-stage approach to proposed changes

26. A staged process has been set out for broader work addressing violent extremism online, through quick fixes to the FVPCA, followed by a broader review. We recommend extending this staged approach to the changes proposed in this consultation process.

27. **We recommend a three stage process, to allow immediate movement on Censor interim decisions, provide time to properly work through some of the more complicated proposals ahead of a broader review, and use the planned review to consider proposals with bigger implications for New Zealand.**

28. Our proposed stages under this recommendation would be:

Stage one: Immediate work on Censor interim decisions

We think the proposal for interim decisions by the Censor has only limited potential for complications, and can advance before Christmas. This could be through a statutes amendment bill if a standalone bill does not make sense for one policy change.

Stage two: Address more complicated proposals by the end of 2020

Other proposed changes raise broader issues that require wide consultation, detailed analysis, and a more holistic view to develop effective policy options and enable ministers to make informed decisions.

We think policy processes in 2020 could allow time to provide the required analysis, and to offer a reasonable consultation period on potential complications of some proposals. Though each also raises broader issues, we think this category could include:

- A potential power to make take-down orders;
- Potential changes, if needed, to avoid technical conflicts between the FVPCA and the Harmful Digital Communications Act 2015.

Stage three: Review of broader issues in media and content regulation

We think many of the proposed changes raise broader issues best addressed through the planned review of laws on media and content. We imagine this running in parallel with other policy work designed to address the underlying issues associated with violent extremism online and offline.

The definition of “publication”, enforcement of criminal penalties, the role of safe harbours, and a potential filter on Internet traffic at the ISP level all involve complicated issues with broader implications for our society.

Assessing the proposed policy measures

29. Below, we assess the six policy proposals, in terms of how well they solve identified policy problems and contribute to the overall goal of:

a) Addressing the identified harms from online violent extremist content;

b) By enabling Government authorities and online content hosts to act with certainty and speed to remove online violent extremist content.

30. We base our comments on the engagement materials, as well as on conversations with attendees from each of the four workshops we attended.

1. Enabling interim decisions by the Chief Censor

31. Currently, urgent decisions by the Chief Censor can face delay, as the current law requires written reasons for a decision within 5 days.

32. The proposal is to allow interim Censors' decisions, and to extend the time for written reasons to 20 working days after such a decision.

33. In our view, this addresses a clear, specific legal requirement, and can be implemented in a way which maintains due process, including providing reasons in a reasonable timeframe, and an appeals process. It would be helpful to see more detail on what circumstances would trigger interim decisions as part of the exposure draft process.

34. We support this proposal proceeding as part of the current process, with implementation details shared through an early exposure draft bill.

2. Enabling DIA to issue take-down notices to content hosts

35. We support ways to advise online services how to comply with New Zealand laws, to help online services reflect and protect the expectations of people in New Zealand.

36. The consultation materials frame the proposal as a power for Censorship Inspectors to make take-down requests. Based on the proposals it is unclear:

    a) Whether this would be limited to publications recorded on the register of classification decisions as objectionable;

    b) Whether this would apply to material which may be "objectionable" under FVPCA s 3, but which has not been the subject of a decision.

    c) Whether there is any intention to offer a way for members of the public to seek a take-down request as part of this proposal;

    d) How "online content hosts" would be defined for the purpose of this provision. This definition is essential to ensure the appropriate actors are in or out of scope, and would take further consultation to clarify.

37. We think a well-designed framework for notice and take-down could help authorities and content hosts to act with certainty and speed, and encourage responsible actions, but designing such a framework will require more work.

38. Even without a formal power, it is our understanding that censorship enforcement officers can informally advise online content hosts of potential breaches of the law, which would often be enough to trigger action under a provider's terms of service or community standards.

Experience in copyright shows that it takes time to test a notice process

39. From the proposal, it is unclear how take-down requests would fit within the framework of the FVPCA, where possession or distribution is a strict liability offence, even without a formal decision about a publication.[1]

40. The proposal is framed as a take-down notice, which may be modelled on the notice frameworks under copyright law. Copyright law provides notice-and-takedown processes within provisions on Internet service provider liability, and as a part of the infringing file sharing framework.

41. The experience of developing these processes shows that it is a difficult and slow process to craft a notice-and-takedown framework which balances relevant interests, and which people will actually use in a way that achieves the intended outcomes.

42. Even where frameworks are meant to limit liability, there is a risk of compliance costs that disproportionately impact smaller players. Detailed consultation is important to test the design and impacts of such a system.

43. We recommend that this proposal is considered through processes that provide the analysis and time needed for more detailed consultation in 2020.

[1] See FVPCA, s 3 and s 123.

3. Including livestreaming within the definition of 'publication'

44. The definition of "publication" is central to the FVPCA, so this proposal is a substantial change to the legal framework it provides, rather than a simple measure to reduce uncertainty.

45. We see particular concerns and complications arising from extending the FVPCA to cover livestreaming, namely:

    a) The FVPCA was designed for static publications. Extending it to interactive conversational media raises new free expression concerns which deserve independent analysis

    b) With live video, what counts as a publication in one context may count as a private conversation in another. This definition might inadvertently include video conferencing through apps such as Skype, Facetime and Zoom, some of which permit anyone to chat or view video by opening a link in their web browser

    c) People sharing live video might unfairly face strict liability offences for actions by others, which result in the providers of a stream sharing objectionable material they did not anticipate or intend.

46. Due to the potential impacts on free expression, and the complexity of this change to the overall framework of the FVPCA, we think this change should not proceed without consideration of broader impacts.

47. The planned review of content and media law will have the scope to consider the definition of publication, including whether it should address livestreaming, as part of broad review to deliver law that is fit for purpose in the Internet era.

48. If this proposal were to proceed despite these concerns, we would want to see narrowly-scoped provisions which limit resulting liability under the FVPCA to violent extremist content, and to situations where the person liable knows or intends that an act of distribution will reach thousands of people.

49. With time and consultation, we see potential to address livestreaming issues as part of a broader framework of carefully designed liability rules for intermediaries, including the potential for take-down notices on livestreaming.

50. We recommend that this proposal be considered as part of the broader review of media and content regulation.

4. Enabling penalties for non-compliance with New Zealand law

51. We welcome consideration of how New Zealand law can be more effectively applied to online services.

52. We agree that penalties could have a role in motivating responsible behaviour by online services used by New Zealanders. However, crafting effective penalties to apply to online services is a complex exercise, requiring

consideration of extraterritoriality, modes of legal service and enforcement, and a proper assessment of the roles and responsibilities of online intermediaries under the FVPCA.

53. As currently drafted, the FVPCA creates strict liability offences for possessing or distributing an objectionable publication. A publication is objectionable under s 3 regardless of whether the Censor's office has ruled it to be so.

54. In our view, time is needed to consider how this framework, designed for publishers or importers of static publications, can or should apply to online intermediaries which facilitate dynamic interactions between people. For example, there may be a need to recognise a difference in responsibilities between static content hosts, and online services that more actively curate content or which enable new modes of interaction which create new risks.

55. With the video of the Christchurch attacks, the problem was not the motivation or willingness of major services to remove the video, but the technical difficulty of doing so in the face of thousands of people attempting to upload and share copies of it. Even months later, we have heard that very well-resourced, highly-motivated services have struggled to totally remove every copy of the video or manifesto, though in practice no-one may be accessing these copies.

56. It is far from clear that legal penalties in terms of the FVPCA would have made any meaningful difference in March 2019, or whether they would help in future. More consideration is needed to make that assessment.

57. We recommend that penalties for online intermediaries be considered as part of the broader review of media and content law.

5. Clarifying that HDCA safe harbours do not apply to the FVPCA

58. We think the issue raised with section 24 of the Harmful Digital Communications Act 2015 (HDCA) needs more consideration.

59. To allow that consideration, we recommend:

    a) Pausing this proposal and testing the legal analysis through detailed consultation;

    b) Incorporating a more holistic consideration of liability limits into the anticipated review of laws governing social media.

60. This would provide a chance to test ideas with stakeholders in the HDCA, including the Ministry of Justice, NetSafe, lawyers operating within the framework, online service providers, and community members.

61. The wording of s 24 resulted from late amendments during a third reading debate.[2] Providing analysis and time for consultation is the best way for this process to avoid reforms which may have equally unintended effects.

62. Internationally, the approach to intermediary liability and safe harbours is currently shifting. As a pragmatic matter, we think it makes sense to consider this broader context in framing New Zealand liability limits.

Does HDCA s 24 have the stated effect?

63. We agree that section 24 of the HDCA offers an opt-in procedure, through which an online content host can be protected from direct liability for specific content posted by a user. However, this procedure can only be triggered by the online content host receiving a valid notice of complaint under s 24(3). It is unclear whether this requirement is likely to be met in real-world complaints about violent extremist content.

64. We are not sure that the problem identified is a real one. As we understand it, even if the s 24 procedure is triggered, there is a requirement that an online content host remove content as soon as practicable. Under s 24, a content host must:

    a) Issue a notice to the posting user as soon as practicable, but within 48 hours at the latest;

    b) If the content host cannot reach the posting user, or if posting user does not respond, the content host must remove the content as soon as practicable, but within 48 hours of sending the notice.

65. Section 24 has no effect on the availability of injunctions, which a Court can issue to require action by a content host (HDCA, sections 24(7), 25(5)).

Liability limits for online services raise broader issues

66. We see well-crafted liability limits as one way that the law can enable and encourage responsible behaviour by online services. We would welcome consideration of liability limits as part of a broad, coherent review of laws affecting online services used by people in New Zealand.

67. Currently, New Zealand law establishes a range of different liability limits in different statutes. In general, these limits allow service providers to avoid direct liability for user-controlled content or connections, in exchange for voluntarily complying with procedural requirements set out in the law. These are found in different statutes, for example:

    a) Copyright Act 1994: sections 92A-E address liability where an online service is used by third-parties to store or transmit infringing copies, without the knowledge or intent of the service provider.

[2] Supplementary Order Paper 91, http://legislation.govt.nz/sop/government/2015/0091/latest/whole.html#whole

b) Harmful Digital Communications Act 2015, s 24 as discussed above.

c) Films, Videos, Publications, and Classifications Act 1993: under section 122, merely providing network services does not count as "distribution" for liability purposes.

68. A goal of this process is to enable speed and clarity in decisions by online services. We support that goal, but think it requires a broader view of the frameworks governing the behaviour of those online services.

6. Consider creating a web filter at the ISP level

69. We set out policy and technical considerations on the use of Internet filtering in our paper "To block or not to block" (attached and available online).[3]

70. Those considerations tell us that filtering at the ISP level is not a viable policy option. It is at best a blunt tool, which would impair Internet use for people in New Zealand, while failing to help with the problem of violent extremist content online.

The evidence and analysis so far cannot justify considering a filter

71. Internet filtering inherently interferes with free expression on the Internet. Before even considering filtering as an option, we would expect to see:

a) Robust evidence that the type of filter proposed could be effective in addressing harms from violent extremist content online; and

b) A comparison of filtering with other options, informed by a credible assessment of relevant human rights interests and impacts.

72. Without this analysis, consideration of an Internet filter should go no further.

A filter at the ISP level creates more problems than it solves

73. In the consultation materials, a filter is put forward to address two different policy problems:

a) **A technical problem** that there is no way for Government to block websites which repeatedly fail to comply with removal requests;

b) **A legitimacy problem** that ISPs voluntarily blocking face legal risk and accusations of being 'censors'.

74. We think a filter at the ISP level cannot solve these problems.

The technical problem: targeting, accuracy, reach, blocking, and cost

75. To do its job, a filter at the network level has to block connections to targeted content, while allowing other connections to proceed normally.

[3] InternetNZ, "To block or not to block: Technical and policy considerations of Internet filtering", (September 2019), <internetnz.net.nz>.

76. As highlighted by the Christchurch attacks, violent extremist content presents particular challenges, in that:

    a) This content is shared by motivated people who want to reach a target audience, whether to cause harm or serve their ideology,

    b) Its meaning and impact is highly contextual, using otherwise benign symbols as "dog whistles" with no meaning to outsiders, as was the case with the so-called manifesto of the Christchurch attacker,[4]

    c) Messages and venues used can shift quickly, making it hard for any filtering system to keep up,

    d) This content is primarily spread through major social media services and in private chat apps, which a web filter will not handle. This in itself is a major limitation on effectiveness (even if you discounted other effectiveness arguments concerning the work arounds of motivated users).

77. The model of the Digital child exploitation filter system (DCEFS) is unlikely to effectively address these challenges, as it is designed to target material that is clearly identifiable based on content without context, which troubled people are motivated access but not to share with the media or broader public, and which is reviled almost universally.

78. Part of the effectiveness of the DCEFS system is that its block page refers people to specialist therapy services that address the underlying psychological and social issues that drive a demand for targeted material. It is unclear what comparable intervention pathway is proposed for people attempting to access extremist material.

79. There is a risk that any additional filter system will motivate more people to use VPN services, reducing the effectiveness of the existing DCEFS system.

The legitimacy problem: who decides how blocking works?

80. The second policy problem relates to the legitimacy of Internet filtering, and how it is perceived by stakeholder groups across New Zealand society. Though raised in relation to ISPs, these legitimacy concerns apply to any filter.

81. In targeted consultation, a range of community groups expressed concerns about the risk of scope creep and potential overreach from a filter. Concern was expressed that a filter would harm many of the same socially vulnerable groups most directly harmed by violent extremism online.

[4] Office of Film and Literature Classification, "The Great Replacement: Reasons for the classification decision" (March 2019), <classificationoffice.govt.nz>.

82. These legitimacy concerns create their own design requirements for any filter. Assuming it could ever be justified, such a system must:

    a) Clearly operate in a way that is independent from central government, with credible, independent oversight,

    b) Address the potentially contentious, contextual nature of targeted material in a way that is transparent and accountable to a range of stakeholder communities,

    c) Be openly developed and implemented, with regard to definitions of targeted content, the filtering rules which result, the technical design, and the effects the system has in practice,

    d) Provide usable avenues for appeal and correction of mistakes,

    e) Be subject to structural protections that would prevent its misuse if the same model were to be adopted in less democratic societies.

83. Given the sometimes contentious, dynamic, and contextual nature of violent extremist content online, it is not clear that the DCEFS model which bases operation, filtering decisions, and appeals within DIA would offer the level of independence from Government needed to ensure public trust in the legitimacy of a filter system of the type proposed.

84. Ensuring effective and independent oversight grounded in the broader community would be essential, including civil society voices as well as the technical, ISP, and government voices represented on the current reference group for the DCEFS filter.

85. We recommend that consideration of a web filter not proceed. If it were to proceed, we see a need for much more policy work to inform consultation, and enable an assessment of technical and legitimacy concerns.

Conclusion

86. We welcome the chance to engage on these proposals. As above, we support the overall goals of this process, but think more consideration is needed to deliver policy which achieves those goals in a sustainable and effective way.

87. We welcome the chance to further discuss the process from here.