

**CCAN Report on Anti-Dehumanization Policy**  
**Christchurch Call Summit, 2022 | 12 September 2022**

**A. BACKGROUND**

**1. Introduction**

In March 2022, the Christchurch Call Advisory Network distributed a request for information to all the members of the Christchurch Call, including both tech companies and governments.

The responses have been compiled to formulate advice to the Call Community, drawing on CCAN's diverse perspectives and expertise.

**2. The Christchurch Call**

The Christchurch Call is a global multi-stakeholder initiative to eliminate terrorist and violent extremist content online. Under the Call, supporter governments and companies have taken on specific voluntary commitments, underpinned by a shared commitment to human rights and a free, open and secure internet. Call supporters recognise the important role of civil society in helping them work on commitments and issues under the Call.

**3. Purpose of the information request**

The Call's civil society and academic advisory network, CCAN, sought information on Call supporter governments' and companies' laws, rules, policies and processes concerning dehumanizing speech and discourse.

The purpose of this exercise was: to enable the CCAN to map current approaches, understand the coverage of current laws, policies and processes, and identify where there may be gaps and weaknesses.

This work is intended to help deliver commitments under the Call, specifically:

- For supporter governments to counter the drivers of terrorism and violent extremism by strengthening the resilience and inclusiveness of our societies.
- For supporter companies to review the operation of processes that may drive users towards terrorist and violent extremist content.
- For supporter governments and companies, to work with civil society to promote community-led efforts to counter violent extremism in all its forms.

**4. Dehumanization and the Call**

Dehumanization is a common behavior of terrorists and violent extremists and a critical feature of their propaganda.

As explained in their paper published by the Global Internet Forum to Counter Terrorism (GIFCT) Risius et al. posit that

1. Dehumanization is a form of psychological violence against target groups often not recognised because frameworks overlook the experience of victim communities (p.72).

2. Serial or systematic dehumanization of an out-group is an attribute of violent extremism (p.43): 'The normative context of dehumanization establishes social preconditions within which violence by extremist instigators is likely to be perceived as justified' (p. 55).
3. Serial or systematic dehumanization of an outgroup is a 'definitory factor' that distinguishes non-violent hate groups from violent extremist groups across cognitive, behavioural and group dynamic domains (Risius et al, 2021, p.51).

The manifestos of the Christchurch terrorist and one of his main sources of inspiration, the Oslo terrorist, portray Muslims as a subhuman, animalistic existential threat through language choices and discourse (Kaldor, 2021, p.17; Lentini, 2019, p.43). The dehumanisation of groups based on protected attributes also features in ISIS communications (Mehran et al. 2022).

### ***Responding to the drivers of violent extremism***

Responding to dehumanization is critical to Call commitments and eliminating TVEC online: (1) By reducing the prominence of dehumanizing language and discourse, we remove one of the most powerful and effective tools used by terrorist and violent extremist movements to build new audiences and fast track radicalisation. We also 'strengthen the resilience and inclusiveness' of our societies. Maintaining moral and normative barriers to violence typically defeated by dehumanization is critical to maintaining resilience to violent ideology. It also maintains the inclusion of all communities, offering them protection from violence and a sense of security. As the Oslo terrorist attack and others demonstrated, targeted communities are not the only ones endangered by successful dehumanization that convinces an actor of an existential threat.

### ***How dehumanization relates to hate speech***

Dehumanization is a distinct concept from hate speech and Terrorist and Violent Extremist content (TVEC), although it often features in both. Dehumanizing language or speech (e.g., referring to a race of people as a disease) is a type of hate speech, broadly defined, and can create a heightened environment for violence. Dehumanizing discourse may be better conceived as behaviour or conduct carried out over time. Dehumanizing discourse doesn't necessarily rely on dehumanizing language and can pass dehumanizing conceptions about a group through disinformation or conspiracy theory (for example, great replacement theory).

### ***Greater capacity for definition***

Call supporters know the difficulties in defining and identifying TVEC online, whether for legislation and law enforcement or for developing and enforcing Terms of Service. Definitions of dehumanizing discourse could aid efforts to counter or disrupt the dissemination of ideology and disinformation that radicalize people over time. The 'conduct' contours of dehumanizing speech and discourse are more definable than concepts of disinformation, hate speech or extremism on their own.

### ***Maintains clear connection with human rights***

Countering dehumanization is justified in human rights law where it is linked to protected characteristics like race, religion (including lack of religious belief), gender, gender identity, sexuality. It supports the Rome Statute (prevention of genocide through the prohibition of incitement to commit genocide Art 25(3)(e)) and the International Covenant on Civil and Political Rights (non-discrimination Art 26, advocacy of hatred Art 20(2)). As a concept, it is also firmly focused on people and communities. A country, government, military, corporation or terrorist entity cannot be dehumanized, reducing the risk of this concept being weaponized against legitimate freedom of expression.

## 5. Definitions for dehumanizing language and discourse

According to Maynard and Benesch (2016), dehumanization is a technique where '[t]argets ... are described in a variety of ways that deny or diminish their humanity, reducing the moral significance of their future deaths or the duties owed to them by potential perpetrators' (80). The authors explain the historical role of dehumanizing language as follows:

Violence against victims has been variously justified by describing them as either biologically subhuman ("cockroaches," "microbes," "parasites," "yellow ants"), mechanically inhuman ("logs," "packages," "enemy morale"), or supernaturally alien ("devils," "Satan," "demons") (80).

Maynard and Benesch refer to the historical role of dehumanizing discourse and conceptions:

Dehumanising discourses and conceptions have been identified in almost all major mass atrocities, prominently including those of Nazi Germany, Stalinist Russia, Rwanda, Yugoslavia, Cambodia, Indonesia, and the Japanese occupation of China. Often, outgroup members (or victims-to-be) are even compared with toxins, microbes, or cancer, suggesting that they are polluting, despoiling, or debilitating the entire in-group—leading to particularly prominent recurring demands to 'purify' groups or societies from the supposedly toxifying elements. Multiple mechanisms explain this effect. [citations omitted] (80).

A dehumanized subject as animalistic is portrayed as 'coarse, uncultured, lacking in self-control, and unintelligent' and 'immoral or amoral' (Haslam, 2006, p. 258). A group of people portrayed as 'mechanistic' is 'lacking in emotionality, warmth, cognitive openness, individual agency, and, because [human nature] is essentialized, depth.'

Further examples of mechanistic dehumanization include attributing collective guilt for a member's crimes or constructing an entire community as an existential threat (Maynard and Benesch 2016, p.81; Bruneau et al 2018), as it essentialises the community and denies their capacity for individual agency or human warmth (Haslam, p.258).

## B. RESPONSES

Twitter was the only company to provide a response.

The following countries responded:

- France
- NZ
- Slovakia
- Ireland
- Finland
- The UK
- Japan

The Council of the European Union also responded.

A CCAN Member based in Australia was also able to contribute analysis of Australian laws to sections of this Report.

## C. ANALYSIS

### 1. Replies from Call Supporter Governments and Companies

*(a) Are there any existing laws, rules or policies prohibiting dehumanising speech or language, whether as part of or distinct from related policies on discrimination and hate speech? If yes, please specify.*

Except for Twitter, there were no existing laws, rules or policies distinctly on dehumanising speech or language.

However, there were laws, rules or policies that conceivably could cover dehumanising speech or language.

*(b) Does your government or company currently refer to or define **dehumanizing discourse**?*

There were no existing laws, rules or policies distinctly on dehumanising discourse.

However, there were laws, rules or policies that conceivably could cover dehumanising discourse.

*(c) Does your government or company have, or plan to have, laws or rules or policies or processes that aim to respond to actors which **dehumanize an outgroup (identified based on a protected attribute) to an ingroup audience over time** through the curation of information and stories? If so, please specify the policies. Please also specify any human rights considerations in your policies.*

There were no identified examples of laws, rules, policies or processes that aim to respond to this particular issue.

## 2. Comparison of Approaches

(a) Of the nine countries included in this comparison, four (4) had civil laws that could penalise dehumanising language or discourse provided it met specific thresholds set in their legislation.

	France	New Zealand	Slovakia	Council of European Union	Finland	Ireland	UK	Japan	Australia
<b>Civil Penalty Y/N</b>	NO	YES If within the scope of incitement laws and the Harmful Digital Communications Act 2015  Racial Disharmony - <i>Human Rights Act 1993</i> s 61  <i>Harmful Digital Communications Act 2015</i>	NO	YES If in the scope of <i>hate speech</i> online  Demonetisation of signatories-2022 <i>Code of Practice on Disinformation</i>	NO	NO	NO	YES If in the scope of infringement on rights of other  Claim for damages in torts  <i>Article 709 of the Civil Code</i>	YES If in the scope of anti-discrimination and anti-vilification laws, which protect some attributes

(b) Of those jurisdictions with civil penalties, the burden remained mainly with the community to bring forward complaints. New Zealand had some departmental capacity to act without community complaints under the Harmful Digital Communications Act 2015, and NZ also resources Netsafe to seek court orders against ‘harmful digital content.’

	France	New Zealand	Slovakia	Council of European Union	Finland	Ireland	UK	Japan	Australia
<b>If there is a civil penalty, does it specify whether the regulator automatically enforces or the community</b>	N/A	Netsafe - community reporting organisation that can seek court orders against harmful digital content  Department of Internal affairs team to counter violent extremism online; which will find and remove	N/A	Demonetising purveyors of disinformation.  Community reporting and obligations on signatories of the 2022 code of practice on	N/A	N/A	N/A	N/A	The community must bring a complaint under discrimination laws and be prepared to litigate against the social media company.

must bring a complaint?		content and, on the lower scale of content, notify online service providers to investigate a breach of TOS		disinformation requiring reporting on their implementation					
-------------------------	--	--	--	--	--	--	--	--	--

(c) With criminal laws, the responsibility lies with the police and the government. Of the nine countries included in this comparison, all had criminal laws that could penalise dehumanizing language or discourse if it met the thresholds set out in their legislation. However, the thresholds vary greatly, and some laws only apply to **direct attacks on individuals** online rather than **direct or cumulative attacks on group identities**.

	France	New Zealand	Slovakia	Council of European Union	Finland	Ireland	UK	Japan	Australia
<b>Criminal Penalty Y/N</b>	<p>YES</p> <p>If in the scope of hate speech laws (term of imprisonment and or fine)</p> <p>24 June 2020 introduced laws to observe online hate speech in order to share the information to different actors.</p> <p>There is an online reporting platform for illegal content online (Pharos).</p> <p>January 2021, a national tribunal and prosecutors office was established, dedicated to reporting, investigating and enforcing online hate laws.</p> <p><i>Public incitement to national, racial or religious hatred or</i></p>	<p>YES</p> <p>If in the scope of incitement laws and harmful digital communications act 2015</p> <p>Inciting racial disharmony - <i>Human Rights Act 1993 s 131</i></p> <p><i>Summary Offences Act 1981</i></p>	<p>YES</p> <p>If in the scope of specific incentive offences, being all criminal acts inspired by hate against a social group</p> <p><i>Slovakian Penal Code 300/2005</i></p>	<p>YES</p> <p>If in the scope of Incitement to violence discrimination</p> <p><i>Council Framework Decision 2008/913/JHA on combating certain forms and expressions of racism and xenophobia</i></p>	<p>YES</p> <p>If in the scope of 'agitation against a population group' and 'defamation.'</p> <p><i>Section 9, 10, 10a Finnish Criminal Code</i></p>	<p>YES</p> <p>If in the scope of <i>Prohibition to of Incitement to Hatred Act, 1989</i></p>	<p>YES</p> <p>can be encountered in the commission of hate crimes but is not a necessary element of offences</p> <p>Stirring up offences, Incitement to hatred and violence - <i>Public Order Act (1986)</i></p>	<p>YES</p> <p>In the scope of discriminatory speech and behaviour - exclusive to race and ethnic origin</p> <p><i>Act on the Promotion of Efforts to Eliminate Unfair Discriminatory Speech and Behavior against Persons with Countries of</i></p>	<p>YES</p> <p>If dehumanising speech directed at a victim meets the harm threshold of the Online Safety Act (s7) or Criminal Code (s474.17)</p> <p>It does not cover dehumanising speech directed at the community.</p>

	<i>violence (Loi 1881, article 23 et 24.8)</i>  <i>Public, racial, national or religious defamation (Loi 1881, art 23, 29.2 et 32.2)</i>  <i>Public, racial, national or religious insult (Loi 1881, art 23 et 29.1 et 33.3)</i>							<i>Origin other than Japan Act No. 68 2016</i>  <i>Hate Speech Elimination Act</i>	
--	--	--	--	--	--	--	--	--	--

(d) Interestingly, the intent requirements were not explicit for most relevant criminal offences, but may still be a requirement.

	France	New Zealand	Slovakia	Council of European Union	Finland	Ireland	UK	Japan	Australia
<b>Intent Clear Y/N</b>	<b>YES</b>	<b>YES</b>  Does the material have an inciting effect assessed on multiple factors such as the mode and spread of the communication?  Human Rights Act s 131 -incite racial disharmony -'intended to excite such hostility, ill-will, contempt or ridicule.'	<b>Unclear</b>	<b>YES</b>  Dehumanising speech as an aggravating factor	<b>Unclear</b>	<b>Unclear</b>	<b>Unclear</b>	<b>Unclear</b>	<b>YES</b>

(e) Consideration of the effect of speech was a consistent emphasis in criminal laws. Dehumanisation through discourse such as disinformation or misinformation about the group may not meet the harm threshold.

Criminal laws	France	New Zealand	Slovakia	Council of European Union	Finland	Ireland	UK	Japan	Australia
<b>Effect of Material Considered Y/N</b>	<b>YES</b> Dehumanising speech falls under laws where it contributes to incitement to discrimination, hatred, or national, racial or religious violence.  Dehumanising speech is also a contributory factor to laws against defamation, insult and misinformation of a protected class.	<b>YES</b> Dehumanising speech is a factor of speech that is threatening, abusive and insulting, inciting hostility or contempt towards a class of person based on race, ethnicity or national origin.  Also encapsulated in digital communication that would denigrate an individual based on these protected classes in addition to colour, religion, gender, sexual	<b>YES</b> "Criminal acts inspired by hate against a social group, such as foundation, support and propagation or sympathy with a movement aimed at suppression basic rights and freedoms; fabrication retention and of extremist content or defamation of a nation, race, belief, incitement of racial and ethnic hate.	<b>YES</b> Dehumanising speech as a form of incitement to hatred or discrimination and is considered where public incitement to violence or hatred against persons on the basis of race, colour, descent, religion, ethnic or national origin	<b>YES</b> Makes available to the public 'an opinion or message where a certain group is threatened, defamed or insulted based on race, colour, birth, ethnic or national origin, religion and sexual orientation.'	<b>YES</b> 'Criminalises certain forms of threatening, abusive or insulting conduct that are intended or likely to stir up hatred against a group of persons on account of certain characteristics, such as race, religion and sexual orientation.'	<b>YES</b> England and Wales have incitement to hatred and violence offences where threats, abuse or insults likely or intended to stir up hatred on grounds of race, religion or sexual orientation.	<b>YES</b> 'announces in public that there will be harm to the life, body, liberty, honor or property of persons with countries of origin other than Japan and that which is extremely disparaging to persons with countries of origin other than Japan, for the purpose of facilitating or inducing a discriminatory attitude against them.'	<b>YES</b> Online Safety Act 2021 (Cth)  material was intended to have an effect of causing serious harm to a particular Australian adult;  material as being, in all the circumstances, menacing, harassing or offensive;  Criminal Code Act 1995 (Cth) – s474.17(1)  using a carriage service in a way that is menacing, harassing or offensive.

		orientation or disability.							
--	--	----------------------------	--	--	--	--	--	--	--

(f) Regulatory approaches to disinformation are still very varied and in the early stages of development and implementation.

	France	New Zealand	Slovakia	Council of European Union	Finland	Ireland	UK	Japan	Australia
<b>Does their law or policy address disinformation?</b>	<b>Disinformation is not strictly addressed.</b> However, the legal framework allows common law to address dehumanisation in disinformation where it would be considered hate speech.	<b>YES</b>	<b>YES</b> to the extent of 'fabrication that would incite hate.'	<b>YES</b> Measure 14.1: Relevant Signatories will adopt, reinforce and implement clear policies regarding impermissible manipulative behaviours and practices on their services, based on the latest evidence on the conducts and tactics, techniques and procedures (TTPs) employed by malicious actors, such as the AMITT Disinformation Tactics, Techniques and Procedures Framework.  <a href="https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation">https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation</a>	<b>YES</b> To the extent of 'defamation laws'	<b>NO</b> The new bill will If the actors are communicating or distributing communications which incite hatred, this would be prohibited under the incitement to hatred legislation.	<b>NO</b>	<b>NO</b>	<b>YES</b> Self regulatory code. Requires harm to be both imminent and serious.

**Disclaimer:** CCAN has prepared this table to the best of its knowledge on the information provided. Should any of the assumptions contained in these tables be incorrect, CCAN encourages Call Members to advise CCAN as soon as possible and it will be corrected.

## OBSERVATIONS FOR POLICYMAKERS

CCAN encourages Christchurch Call members (Governments and Online Service Providers) to

1. Work together on strategies that complement and support each other on the global internet to counter the production and dissemination of dehumanising speech and discourse about outgroups identified based on protected attributes, including race, religion (and lack of religious belief), gender, gender identity and sexuality to an ingroup audience.

Members may choose to treat this problem through different policy frameworks: For example, disinformation, hate speech, non-discrimination, harmful digital communications or torts.

2. Ensure the strategy:
  - a. Recognizes serial or systematic dehumanization of an out-group based on a protected characteristic as a form of psychological violence to suppress and silence those communities and make violence against those communities seem more acceptable.
  - b. Explains the risk that dehumanization by violent extremist movements create for targeted groups, society and democracies as a whole.
  - c. Acknowledges that the portrayal of groups through curated information and news as polluting, despoiling or debilitating society, or as an existential threat to society is a powerful form of dehumanization.
  - d. Defines dehumanizing speech/language and discourse.
  - e. Explains the member's approach to countering this problem.
  - f. Encompasses dehumanization of groups identified on basis of protected characteristics, not only attacks on individuals.
  - g. Contemplates how policy decisions by members, speech by political actors and traditional media are used in online ecosystems seeking to dehumanise a particular group identified on the basis of a protected attribute.
  - h. Shares responsibility across all producers of information at scale for disrupting drivers of dehumanising discourse in line with human rights principles.
  - i. Considers the levers for enforcement, in particular, that
    - i. The conduct and intent elements are precisely defined in any policy or law.
    - ii. The burden of enforcement for public harm does not rely on community complaints and private litigation. The burden of enforcement is shared by an independent public regulator (in the case of governments) and the platform (in the case of tech companies).

- iii. Penalties and enforcement processes are fit for purpose, proportionate to harm and prioritise non-carceral approaches given the risk of carceral systems to vulnerable adults and children.
  - iv. Community groups that work effectively in this space can access resources.
- j. Establishes a civil society review mechanism to ensure laws and policies, and their mechanisms for enforcement, are reviewed regularly for their fitness for purpose and human rights compliance.

## REFERENCES

- Bandura A (1999) Moral disengagement in the perpetration of inhumanities. *Personal Soc Psychol Rev* 3(3):193–209.
- Bruneau E, Kteily N, Falk E. Interventions Highlighting Hypocrisy Reduce Collective Blame of Muslims for Individual Acts of Violence and Assuage Anti-Muslim Hostility. *Personality and Social Psychology Bulletin*. 2018;44(3):430-448. doi:10.1177/0146167217744197
- Goffman E (2009) *Stigma: notes on the management of spoiled identity*. Simon and Schuster, New York.
- Haslam N (2006) Dehumanization: an integrative review. *Personal Soc Psychol Rev* 10:257.
- Kaldor, S., 'Far-Right Violent Extremism as a Failure of Status: A New Approach to Extremist Manifestos through the Lens of Ressentiment' (Research Paper, International Centre for Counter-Terrorism – The Hague, May 2021) 17 <https://icct.nl/app/uploads/2021/05/Far-Right-Violent-Extremism-as-a-Failure-of-Status.pdf>.
- Kteily, Nour & Hodson, Gordon & Bruneau, Emile. (2015). They See Us As Less Than Human: Meta-Dehumanization Predicts Intergroup Conflict Via Reciprocal Dehumanization. *Journal of Personality and Social Psychology*. 110. 10.1037/pspa0000044.
- Lentini, P. (2019) "The Australian Far-Right: An International Comparison of Fringe and Conventional Politics" in Mario Peucker and Debra Smith, eds. *The Far-Right in Contemporary Australia*. Singapore, 43.
- Maynard JL, Benesch S (2016) Dangerous speech and dangerous ideology: an integrated model for monitoring and prevention. *Genocide Stud Prev* 9(3):70.
- Mehran, W., Byrne, M., Gibbs-Pearce, E., Macfarlane, A., Minihane, J. & Ranger, A. (2022): Humour in jihadi rhetoric: comparative analysis of ISIS, AlQaeda, TTP, and the Taliban, *Behavioral Sciences of Terrorism and Political Aggression*, DOI: 10.1080/19434472.2022.2075028.
- Risius M, Blasiak K, Wibisino S, Jabri-Markwell R, Louis W (2021) Dynamic Matrix of Extremisms and Terrorism (DMET): a continuum approach towards identifying different degrees of extremisms. Report to the Global Internet Forum to Counter Terrorism. <https://gifct.org/wp-content/uploads/2021/07/GIFCT-TaxonomyReport-2021.pdf>