13 March 2025

Kia ora koutou,

Submission on Hara ngākau kino | Hate crime: Consultation paper

My name is Dr Sanjana Hattotuwa. As Research Director at The Disinformation Project[1] (TDP) from June 2021 to October 2024, and currently as an independent researcher focussing on online harms, Aotearoa New Zealand's media landscapes, and information integrity, I have exhaustively studied public content, and commentary targeting New Zealand's liberal democracy, information, and electoral integrity, and social cohesion in addition to multifaceted threats related to public, human, and national security.

My PhD examined the nexus between the instigation of hate, violence, and harms online – especially on social media – and offline, kinetic consequences, including targeted violence against specific communities. My doctoral research was also the first in New Zealand to study the global, and domestic discourse on Twitter after, and in response to the Christchurch massacre in March 2019. Before my PhD, I was one of the first researchers in South Asia to connect, and produce evidence around the instigation of hate on Facebook (against Muslims), and violence directed against them offline in Sri Lanka leading up to the worst race riots in the country – coterminous with how Facebook was weaponised against the Rohingya community in Myanmar. From peer-reviewed academic output to media articles, reports, podcasts, videos, and other papers, I have studied the nexus between offline violence, and online hate for well over a decade.

My submission argues for a comprehensive approach that recognises the relationship between online hate speech and offline hate crimes. I submit that the Law Commission's separation of these issues creates a problematic gap in New Zealand's legal response to hate-motivated violence.

Key points considering the Law Commission's questions in consultation document[2]:

1. My submission focuses heavily on the limitations of New Zealand's current legal framework for addressing hate crimes (**Question 5**[3]), arguing that there exists a significant disconnect between the Law Commission's narrow focus on punishing hate-motivated offences and the reality of how online hate and incitement lead to

[1] https://thedisinfoproject.org/
[2] https://www.lawcom.govt.nz/our-work/hate-crime/tab/consultation-paper
[3] Do you think there are problems with how Aotearoa New Zealand's current hate crime law is working? If so, what are those problems?

offline violence. I critiques the current model as inadequate because it fails to address the ecosystem of online radicalisation that often precedes hate crimes.
2. My submission strongly advocates for specific hate crime offences (**Questions 7[4] and 8[5]**), recommending the adoption of standalone hate crime offences rather than relying solely on the sentence aggravation model. I argue that specific offences would strengthen the response when prevention fails and violence does occur, and would send a clear signal of zero tolerance. This directly engages with Question 8's inquiry about whether a different legal model should be introduced.
3. Regarding **Question 1[6]** about hate crime in New Zealand and its impacts, my submission provides some evidence of contemporary hate crimes, citing recent research on an evangelical preacher's incitement of violence against gay, and queer communities that resulted in actual violence at Te Atatū library in Auckland. I emphasise how online hate ecosystems contribute to an increasingly volatile, uncertain, complex, and ambiguous national, and human security context.
4. I also consider which characteristics should be protected (**Question 3[7]**) by highlighting the current gaps in protection, especially regarding religion, gender, sexual orientation and disability, and emphasising the intersectional nature of hate crimes that target individuals based on multiple overlapping characteristics.
5. Finally, I engage with **Question 4[8]** about key reform considerations by suggesting that the Law Commission's current approach is too limited, proposing a more holistic framework that addresses online hate proliferation and acknowledging the complexities of balancing freedom of expression with protecting vulnerable groups.

Overview of recommendations

1. I recommend expanding the Law Commission's definition of hate crime to include online hate proliferation. While the Commission's focus (in the consultation document) on criminal offences motivated by hatred is important, we must address content that creates hostile environments, even when it doesn't constitute a criminal offence itself.
2. I strongly advocate for the adoption of specific hate crime offences. When violence occurs, I believe New Zealand's legal, policing, and response architectures needs standalone charges with higher penalties available—for instance, when someone assaults a devotee of a particular faith due to hate against that faith, and its symbols or practice. This approach not only ensures punishment reflects social harm but also sends a clear message of zero tolerance.
3. I believe the report's focus on sentence aggravation is insufficient for addressing the complex inter-relationships between online hate and offline violence. Platforms (like Facebook, YouTube, Twitter/X, TikTok etc) must implement robust moderation

[4] f specific hate crime offences are adopted what offences should they cover? Why?
[5] Should a different legal model, such as specific hate crime offences or the Scottish hybrid model, be introduced in Aotearoa New Zealand? Why or why not?
[6] Is there anything you would like to tell us about what hate crime is occurring in Aotearoa New Zealand and its impacts?
[7] What characteristics should be protected by hate crime laws? Why?
[8] What do you think about the key reform considerations we have identified for this review?

policies and develop algorithms to detect content inciting hate, to create greater friction around offline spillover effects.

4. I emphasise the need for clear guidance on balancing freedom of expression with protecting vulnerable groups. Any restrictions should be narrowly tailored to address specific harms and subject to independent oversight.
5. I propose exploring alternative dispute resolution mechanisms for addressing hate crimes.
6. I highlight the importance of recognising the intersectional nature of online hate leading to hate crimes. Effective interventions must address overlapping forms of discrimination based on race, gender, sexual orientation, and disability.
7. I insist that all approaches to deal with hate must be grounded in Te Tiriti o Waitangi.
8. Finally, I suggest that New Zealand can learn from international best practices, potentially establishing multidisciplinary threat assessment teams to evaluate cases of online radicalisation and formulate appropriate responses.

The legal architecture around online, and offline hate

New Zealand's legal toolkit for hate speech and incitement consists of: (1) targeted incitement laws in the Human Rights Act 1993 (HRA) that focus on racial disharmony (with limited scope and use), (2) digital harm laws in the Harmful Digital Communications Act 2015 (HDCA) that can capture some hateful abuse online (especially against individuals, covering more categories), (3) content regulation via the Classification Act to eliminate the most extreme hateful content (like violent extremist propaganda), and (4) general criminal law provisions for threats or incitement to commit crimes. All of these operate under the shadow of the New Zealand Bill of Rights Act 1990 (NZBORA), which means freedom of expression is always a counterweight consideration.

In response to longstanding gaps (like the exclusion of religion as a protected category) and the Royal Commission's recommendations[9], the government in 2021–2022 considered reforms. Six proposals were floated, including adding more protected groups (such as religion, gender, sexual orientation, disability) to the incitement provisions, raising the penalties (making it a Crimes Act offense with up to 3 years' imprisonment), and clarifying the legal tests[10]. After public consultation drew over 19,000 submissions (with significant opposition from *soi-disant* 'free-speech' networks in what I've studied as a disturbing example of weaponising public consultations[11]), the government opted for a pared-down change: proceeding only with adding "religious

[9] https://christchurchattack.royalcommission.nz/the-report/part-9-social-cohesion-and-embracing-diversity/hate-crime-and-hate-speech/

[10] https://www.regulation.govt.nz/assets/RIS-Documents/ria-justice-ihahra-nov22_1.pdf#:~:text=under%20the%20prohibited%20grounds%20of,incitement%20provision%20to%20match%20the

[11] The weaponisation of public consultations in New Zealand: Threats to Law Commission's consultation paper on hate crimes, https://sanjanah.wordpress.com/2025/02/06/the-weaponisation-of-public-consultations-in-new-zealand-threats-to-law-commissions-consultation-paper-on-hate-crimes/

belief" to the protected grounds[12], and tasking the Law Commission to review the broader hate speech framework. A bill to include religious communities under Sections 61 and 131 was introduced in late 2022, aiming to fill the most urgent gap after Christchurch. However, further changes to definitions or penalties have been put on hold pending the Law Commission's first-principles review. As of this writing, the HRA has not yet been amended to include religion[13] (the reform was proposed but subsequently paused amid domestic political developments in early 2023).

I understand the Law Commission's hate crimes review as one that will likely influence any future legislative overhaul, ensuring that any expansion of the law is carefully calibrated and justified in light of free expression concerns.

The gap between ground realities and the Law Commission's focus on hate crimes

While the Law Commission's hate crime paper is timely, and valuable in addressing the penalisation of hate-fuelled offenses, it both creates, and leaves a significant gap regarding prevention and incitement. The paper's limited scope means it doesn't fully reckon with how online, and social media hate ecosystems, and dangerous speech[14] contribute to offline (often called 'real-world') harms – a critical issue if we want a comprehensive strategy against extremist violence in New Zealand.

I've already provided detailed submissions to Parliament on the Treaty Principles Bill[15], and to the Ministry of Justice on the Arms Act rewrite[16]. Both are public documents. Amongst other things, they clearly illustrate, based on years of grounded, mixed-methods research how the growth, and incitement of violence, hate, and harms online has in New Zealand contributed to an increasingly volatile, uncertain, complex, and ambiguous national, and human security context in which there's a very high likelihood of offline violence, but with little to no ability to determine perpetrators in advance. New Zealand's security, and intelligence services concur[17] with this assessment. The country's already witnessing the spillover effects of hate instigated on social media. In

[12] Human Rights Act enhanced to protect religious communities, https://www.beehive.govt.nz/release/human-rights-act-enhanced-protect-religious-communities#:~:text=%E2%80%9CCurrently%2C%20under%20the%20Human%20Rights,provisions%2C%20to%20cover%20religious%20belief

[13] Chris Hipkins' 'policy bonfire': Government cops criticism for refocus with more changes to come, https://www.nzherald.co.nz/nz/politics/govt-cops-criticism-for-policy-bonfire-with-more-changes-to-come/WG4RSZ36EJBSXOZ7GFC23IQWOM/#:~:text=Chris%20Hipkins%27%20%27policy%20bonfire%27%3A%20Government,it%20to%20the%20Law%20Commission

[14] What is Dangerous Speech?, https://www.dangerousspeech.org/dangerous-speech

[15] Submission to New Zealand Parliament on Principles of the Treaty of Waitangi Bill, https://sanjanah.wordpress.com/2025/01/06/submission-to-new-zealand-parliament-on-principles-of-the-treaty-of-waitangi-bill/

[16] Submission to New Zealand Ministry of Justice on the Arms Act rewrite, https://sanjanah.wordpress.com/2025/02/26/submission-to-new-zealand-ministry-of-justice-on-the-arms-act-rewrite/

[17] New Zealand's Security Threat Environment, https://www.nzsis.govt.nz/assets/NZSIS-Documents/New-Zealands-Security-Threat-Environment-2024.pdf

February, research on a highly influential Christian evangelical preacher's incitement of hate, and violence against New Zealand's gay, and queer communities illustrated the direct, and near immediate causation between what was implicitly, and explicitly called for in sermons posted on social media, and the violence enacted by followers at the Te Atatū library in Auckland against children, parents, and attendees of a Pride event[18].

Based on robust academic research[19] (which maps on to New Zealand's domestic context, and discursive landscapes on social media[20]), the lived experience in austere contexts defined by violence, and my own research for over a decade, I submit that it is impossible to talk about hate crimes in isolation, without embracing the concert of vectors, factors, and actors contributing to online hate, and harms.

The Law Commission's paper is primarily focused on how the law should punish and deter hate-motivated offenses (assaults, vandalism, etc), in response to the Royal Commission's recommendations after Christchurch. It explores whether New Zealand should move beyond the current sentencing aggravation model and create specific hate crime offences, what characteristics should be protected, how to record and prosecute such crimes, and so on. Crucially, though, it does not deal with "hate speech" or incitement provisions. In fact, the Commission explicitly carved those topics out of scope: "The review will not consider hate speech, including the offence of inciting racial disharmony in the Human Rights Act 1993," noting that hate speech was previously on their agenda but was withdrawn by the Government. Put simply, the discussion paper confines itself to crimes that are already criminal (assault, harassment, etc.), and does not examine conduct that is purely speech or online expression which incites hatred.

This narrow scope creates a gap, and therein lies the rub.

By excluding the likes of online hate propaganda from consideration, the Law Commission's analysis doesn't directly grapple with how online radicalisation translates into offline violence. For example, the paper discusses how an offender's hateful words during the commission of a crime (say, shouting slurs while assaulting

---

[18] Library protest crossed a line, Police say, https://www.police.govt.nz/news/release/library-protest-crossed-line-police-say

[19] Müller, Karsten and Schwarz, Carlo, From Hashtag to Hate Crime: Twitter and Anti-Minority Sentiment (July 24, 2020). Available at <SSRN: https://ssrn.com/abstract=3149103>

Kunal Relia, Zhengyi Li, Stephanie H. Cook and Rumi Chunara "Race, Ethnicity and National Origin-based Discrimination in Social Media and Hate Crimes Across 100 U.S. Cities" (31 Jan 2019) New York University https://arxiv.org/pdf/1902.00119.pdf

Matthew L. Williams, Pete Burnap, Amir Javed, Han Liu and Sefa Ozalp "Hate in the Machine: Anti-Black and Anti-Muslim Social Media Posts as Predictors of Offline Racially and Religiously Aggravated Crime" (2020) The British Journal of Criminology, Volume 60, Issue 1, January 2020, Pages 93–117 <doi:10.1093/bjc/azz049>

[20] The Christchurch massacre as 'false flag': New Zealand's flailing, and failing attempts to address online harms, https://sanjanah.wordpress.com/2024/05/24/the-christchurch-massacre-as-false-flag-new-zealands-flailing-and-failing-attempts-to-address-online-harms/

someone) could prove hate motivation in court. But it does not address situations where hateful words precede or encourage the crime – as in the case of online posts or sermons inciting others to commit violence. The result is that the Commission's otherwise thorough examination of hate crime law largely omits the upstream role of digital hate campaigns and extremist communities in fomenting those crimes.

I fully appreciate this omission was intentional, and due to administrative reasons flagged in the consultation document. However, from a substantive standpoint, this separation is extremely problematic. My research, and public writing has clearly established how the incitement of hate online is often is the precursor or catalyst for (offline) hate crimes. This February alone, the inflammatory sermons and incendiary posts by a leading evangelical preacher, defined by what Prof Susan Benesch calls 'dangerous speech'[21] led directly to crimes like vandalism, assault, indecent assault, and harassment (in Auckland). By treating hate speech and hate crime in isolation, New Zealand's current approach risks missing this crucial, enduring linkage.

The Law Commission's discussion paper does acknowledge public concern that hate crime laws might impinge on free speech, and it takes care to note that thought and expression are not being criminalised *per se* – only when tied to criminal acts. But it stops short of examining whether current offences for incitement are adequate. In fact, the paper reiterates that the only specific incitement law on the books – the Human Rights Act's provision against inciting racial disharmony – is outside its definition of hate crime because it covers conduct that isn't otherwise an offence. This creates a gap in assessment: the Commission doesn't evaluate how well (or poorly!) that incitement law is working, nor how online hate targeting other groups (religion, sexual orientation, gender, etc) might be addressed, since those fall outside the outdated 1993 provision. The discussion paper focuses on addressing violence after it occurs (through charges and sentencing), but not on the incitement that helps cause such violence in the first place.

This is a significant shortcoming.

Online incitement is a driving force behind hate-motivated threats and acts in New Zealand. Yet the Law Commission's consultation document, by design, doesn't analyse whether legal tools to combat online incitement (or to intervene early against radicalisation) need reform. For instance, the paper doesn't discuss extremist online content, platform responsibilities, or the phenomenon of stochastic terrorism facilitated by social media. Nor does it touch on the use of 'dangerous speech' by local influencers to mobilise aggression (as an array of hate entrepreneurs in New Zealand increasingly do, on a daily basis).

These omissions mean that the discussion paper under-appreciates the radicalisation glide paths leading to hate crimes: from an online or social media post, meme or video, to an individual's embrace of violent extremism, to an eventual violent act or terrorism.

[21] What is Dangerous Speech?, https://www.dangerousspeech.org/dangerous-speech

In short, New Zealand's existing legal framework does not fully address the symbiosis, and synergies between online incitement, radicalisation, and offline violence. The incitement law remains too narrow; we lack robust hate speech prohibitions for online content; we have no specific mechanism to pre-emptively disarm those brewing in online hate forums; and much reliance is placed on general terrorism or criminal laws that kick in late in the process.

The proposed reforms contemplated in the discussion document either don't touch this nexus or risk widening gaps. This underscores a need for more holistic approaches, comprehensive reform, and a grounded appreciation of hate dynamics in contemporary New Zealand.

Recommendations

Meaningfully online hate proliferation effectively requires a multi-faceted approach, acknowledging the complexities outlined in the Law Commission's report and building upon its existing framework.

The following recommendations offer concrete strategies to strengthen the consultation document's presentations, respond to existing and emerging challenges, and align with broader legal, and social objectives.

1. The Law Commission's definition of hate crime, focused on underlying criminal offences motivated by hatred, needs expansion in contemporary domestic, and international online contexts. While the report acknowledges the issue of 'hate speech' as distinct from 'hate crime', the pervasive, persistent, and persuasive nature of social media, and online platforms necessitates addressing hate proliferation that doesn't necessarily constitute a criminal offence itself but contributes to a hostile environment. Strengthening the framework requires a broader definition of 'online hate proliferation' to include the systematic dissemination of content that promotes prejudice, incites violence, or normalises discrimination against protected, and vulnerable groups. This broadened definition would facilitate more targeted interventions offline, and encourage online platforms to take proactive measures against harmful content that risks spilling over to kinetic violence.

2. Consider the adoption of specific hate crime offences. Following through on creating standalone hate crime offences will strengthen the response(s) when prevention fails and violence does occur. If someone assaults a devotee of a certain faith tradition or vandalises a place of worship due to hate, a specific hate crime charge (with higher penalties) should be available, as recommended by the Royal Commission. This not only ensures the punishment reflects the crime's social harm, but also sends a clear signal of zero tolerance. It can improve reporting and recording of hate crimes, helping authorities understand the seed, spread, and scope of the problem. Legislation should make it explicit that online activities (like posting terrorist manifestos or hateful screeds) by an offender can be used as

evidence of motive. By expressly recognising the hate element, the law affirms the link between the ideology and the act, which can aid in public awareness and rehabilitation efforts for offenders.

3. The report's focus on sentence aggravation, while a valuable tool, is insufficient to address the scale and speed of online hate. As the report states, the creation of specific hate crime offences could help to deter such offending in the future, increase reporting of hate crime, increase prosecutions and facilitate the creation of tailored rehabilitation programmes for hate crime offenders." However, specific offences alone are inadequate. New Zealand can borrow from the eSafety Commissioner in Australia, who has powers to compel social media platforms to take down harmful content, subject to necessarily high thresholds. There should be similar laws which cover hate proliferation, and an independent body setup to receive complaints, investigate, make recommendations and have powers to compel social media platforms to follow recommendations as a way to stymie the growth of (offline) hate crimes. To this end, and reflection the inextricably entwined nature of online, and offline hate, the Law Commission should emphasise the need for proactive measures by online platforms, including robust content moderation policies, transparent reporting mechanisms, and the development of algorithms that detect and remove content inciting hate, and violence. Furthermore, platforms must be held accountable for failing to address online hate proliferation effectively.

4. The report must provide clear guidance on balancing freedom of expression with the need to protect vulnerable groups. The Law Commission acknowledges concerns about infringing on the rights to freedom of thought, expression, and association. Navigating this tension requires a careful and nuanced approach that prioritises the safety and well-being of individuals while upholding fundamental rights. The report should emphasise the importance of proportionality, necessity, and legitimacy when restricting online speech, mindful of the many ways measures to constrict the reach of hateful content, and commentary that eschew simplistic, binary approaches like deplatforming, and banning[22]. Any restrictions should be narrowly tailored to address specific harms and should be subject to independent oversight and judicial review.

5. The Law Commission should explore the potential of alternative dispute resolution (ADR) mechanisms for addressing hate crimes. While criminal prosecution may be appropriate in some cases, alternative approaches, such as mediation and restorative justice, can offer more effective and timely remedies, particularly for less serious forms of hate's instigation, and invocation. These mechanisms can provide victims – for example, those who were part of religious cults or extremist networks - with a platform to voice their concerns, seek redress, and promote reconciliation.

[22] A/74/486: Report on online hate speech, Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, delivered to UN General Assembly at its 74th session, 2019, September 2019, https://www.ohchr.org/en/documents/thematic-reports/a74486-report-online-hate-speech

6. The report must consider the intersectional nature of online hate leading to hate crimes. Online hate, and the eventual offline expressions including kinetic violence often targets individuals based on multiple and overlapping characteristics, such as race, gender, sexual orientation, and disability. Effective interventions must recognise and address these intersectional forms of discrimination, ensuring that all individuals are protected from online harm(s).

7. Any, and all approaches to deal with hate offline and/or online should be grounded in Te Tiriti o Waitangi. Māori face unique challenges regarding online hate, including the misappropriation and denigration of cultural symbols and language. Any strategy must uphold the principles of partnership, protection, and participation, ensuring that Māori voices are central to shaping solutions.

8. In implementing these recommendations, New Zealand can also learn from international best practices. The interplay between online extremism and offline violence is a global challenge, and many countries are experimenting with solutions. For example, multidisciplinary threat assessment teams in some European nations examine individuals flagged for extremist social media posts to decide on interventions (ranging from mental health support to arrests). New Zealand could establish a similar model that brings together tech experts, police, psychologists, and community leaders to evaluate serious cases of online radicalisation and formulate appropriate responses.

The data, and evidence around the rise of hate crimes against specific identities, and communities in New Zealand is already well-established[23], and also clearly highlighted in the Law Commission's consultation paper. The year-on-year rise in reported hate crimes since 2021 is significant, especially against race, and ethnicity, and increasingly around sexual orientation.

My research bears witness to the online, and social media discourses that fuel hate, violence, and harms. No legislative, legal, policing, or policy response to hate crimes focussing on downstream impacts, crimes, and offline consequences will succeed unless a more holistic approach is adopted, which also looks at the complexities around hate's generation, and spread online by specific actors, and networks who benefit from social media platform affordances. I hope the Law Commission more seriously considers this symbiotic nature.

[23] New data on violence against LGBTQ+ people makes 'grim reading' – and undermines NZ's inclusive reputation, https://theconversation.com/new-data-on-violence-against-lgbtq-people-makes-grim-reading-and-undermines-nzs-inclusive-reputation-239706

Exclusive: Racism, homophobia fuelling thousands of crimes in New Zealand each year, figures show, https://www.theguardian.com/world/2023/jun/08/exclusive-racism-homophobia-fuelling-thousands-of-crimes-in-new-zealand-each-year-figures-show

Rise in hate incidents reported to police, https://www.rnz.co.nz/news/national/509717/rise-in-hate-incidents-reported-to-police

For the safety, security, and sake of all New Zealanders, I wish the Commission well in its endeavours.

Ngā mihi,

Dr Sanjana Hattotuwa