



Tuesday, 28 February 2023

To:
Senate Economics References Committee
economics.sen@aph.gov.au

Dear Committee Chair,

We thank you for the invitation to make a submission to the inquiry into the Influence of international digital platforms.

Yours faithfully,

AUSTRALIAN MUSLIM ADVOCACY NETWORK

Web: www.aman.net.au



AUSTRALIAN MUSLIM ADVOCACY NETWORK

*Submission to the Senate Economics References Committee
28 February 2023*

Contents

Submission Overview	3
Terms of reference:	3
About AMAN	4
Australian impact	4
Current regulatory environment.....	7
Regulating international digital platforms	8
The effect of dehumanisation.....	11
How algorithms contribute to the spread of dehumanising information operations.....	15
The formula of dehumanising information operations	17
Access to data for research	18
Strategic use of information.....	18
How Australian policy could contend with international and domestic dehumanising information operations.....	21
Recommendations.....	21
A note about user empowerment.....	23



Submission Overview

1. AMAN has extensive experience in monitoring and regulating international digital platforms in their facilitation of hate and racism.
2. It is necessary that the Committee acknowledges the need to protect the Australian public sphere from dehumanizing speech and discourse which is spread and induced by international digital platforms. This rhetoric makes online spaces unsafe for Australians and also eventuates to discrimination, hate crimes and terrorism such as the Christchurch massacre.
3. Australian policies must be updated to protect Australians through a proactive approach to online safety that is resilient to changes in conspiracy theories, ideologies that target different groups.
4. This can be achieved through
 - a. AMAN's recommendations, such as proposed civil penalties for serial and/ or systematic dehumanising speech or discourse, reforms to the Anti-discrimination laws and the Online Safety act, which will be illustrated in this submission.
 - b. Lifting the corporate veil between the Australian and foreign arms of international digital platforms.
5. International digital platforms represent to parliament, regulators and affected communities that their Australian-based arms and employees are part of the international digital platform and capable of exercising influence and control. A different story is put forward when legal action is brought. International digital platforms use corporate structures to avoid legal accountability, including under discrimination laws.

Terms of reference:

6. On 26 September 2022, the Senate referred an inquiry into the Influence of international digital platforms to the Senate Economics References Committee for inquiry and report by the last sitting day of 2023.
 - a. the market shares of such international digital platforms across the provision of hardware and software services;



AUSTRALIAN MUSLIM ADVOCACY NETWORK

- b. vertical integration, or linking of multiple services, products and/or hardware, within such international digital platforms and resultant outcomes on users' ability to exercise choice;
- c. whether algorithms used by such international digital platforms lack transparency, manipulate users and user responses, and contribute to greater concentrations of market power and how regulating this behaviour could lead to better outcomes in the public interest;
- d. the collection and processing of children's data, particularly for the purposes of profiling, behavioural advertising, or other uses;
- e. the adequacy and effectiveness of recent attempts, in Australia and internationally, to regulate the activities of such international digital platforms;
- f. broader impacts of concentration of market power on consumers, competition and macro-economic performance, and potential solutions; and
- g. any other related matters.

About AMAN

7. The Australian Muslim Advocacy Network (AMAN) is a policy development body working to secure Australian Muslims' physical and psychological safety. Its goal is to prevent the harm caused by public expressions of hatred toward Muslims, including through violence and atrocities like the Christchurch massacre.
8. AMAN is active within the Christchurch Call Advisory Network and the Global Internet Forum to Counter Terrorism's Legal Frameworks Working Group.
9. AMAN has supported research into the techniques of information operations within the 'counter jihad' movement, an extreme right movement that propagates variations of the Great Replacement Theory and Eurabian conspiracy theory.
10. AMAN's early focus was to practically problem-solve how such information operations can be prevented from using large beacon platforms to legitimise their narratives and act as a gateway for more extreme and violent programs of action.
11. AMAN has brought a legal complaint against Facebook/Meta under the Racial Discrimination Act 1975 (Cth) and a complaint against Twitter under the Queensland Anti-Discrimination Act 1991 (Qld). The substance of both complaints pertains to international information operations that seek to dehumanise Muslims to Australian audiences.

Australian impact

Web: www.aman.net.au



AUSTRALIAN MUSLIM ADVOCACY NETWORK

12. In the 2019 federal election, approximately 12 fringe parties were running with a discriminatory anti-Muslim policy – this is the most significant number of groups we have recorded. We remain very concerned about the exportation of RWE rhetoric from the UK, Europe and the USA to Australia through coordinated exercises on international digital platforms like Facebook and its potentially devastating impacts on Australia’s democracy, social cohesion and national security.
13. In the 2022 federal election, there was a substantial contraction in fringe parties running explicit anti-Islam policies. This reflected an overall re-orientation of far-right groups toward electoral misinformation (originally, US Politics based), medical misinformation (Covid vaccines and the narrative that Covid was overstated/a hoax), and climate misinformation (portraying climate science as part of a reset global conspiracy).
14. Due to a downturn in media coverage of ISIS and political/official announcements about the threat of religiously motivated terrorism, and a surge in other mainstream rallying points, anti-Muslim disinformation operations have continued in the main pages that we observe on Facebook, but in more general far-right groups, have become less prominent.
15. A recent Australian study examined the role of mainstream and fringe media in the mobilisation of radical political movements:

The most frequently shared fringe media sources include both Australian sites (e.g. The Unshackled) and blogs as well as international (often US focused) sites, popular within far-right milieu around the world, such as Breitbart, Voice of Europe, or Jihad Watch (Lee 2015; Lima et al 2018). The most often shared URL domain in the Facebook dataset is the far-right and often anti-Muslim fringe news blog The Politics Online, which was shared almost exclusively by one prolifically posting Facebook account. This extensive sharing of far-right and often anti-Islam content from The Politics Online appears to be part of a coordinated covert scheme on Facebook that used existing far-right Facebook accounts (including one in our sample) to push high volumes of posts linked to ‘websites masquerading as news sites with generic titles like “The Politics Online”’ (Knaus et al. 2019), according to a Guardian investigation in 2019.¹

16. Facebook and Twitter continue to refuse action on these operations that dehumanise Muslims and engender violent responses, which is why AMAN is bringing legal action.
17. We include our observations about continuing information operations and the

¹ Mario Peucker, Thomas J. Fisher, Jacob Davey, *Mainstream media use in far-right online ecosystems* (Research report, 24 August 2022), <https://www.crisconsortium.org/violent-extremism/mainstream-media-radicalisation>



AUSTRALIAN MUSLIM ADVOCACY NETWORK

pages that continue to strategically benefit from them later in this submission.

18. AMAN continues to appeal to the Australian Government to remove the legal and official category of 'religiously motivated' terrorism to avoid bolstering ISIS and racist nationalist recruiters and information operations online. It is noted that Canada, New Zealand and Australia followed the United Kingdom in distinguishing between religiously and ideologically motivated terrorism in law. It is not a coincidence that these countries have actively and officially conflated Islam and terrorism and are affected by strong 'counter jihad' movements.
19. Australia's counter-jihad network was the predominant force in the organisation and evolution of white supremacist/Neo-Nazi cells in Australia.
20. The [Islamic Council of Victoria/Umar Butler report on Anti-Muslim Tweets](#) found that nearly 86% of geolocated anti-Muslim posts originated in just three places: India, the US and the UK.

We found that over 85% of all anti-Muslim posts originated in just three places: India, the US and the UK. With regard to India, we link the rampancy of Islamophobia to the fact that the BJP is currently in power. Indeed, from their refusal to condemn Ranjeet Bahadur Srivastava's assertion that 'the party will bring machines from China to shave 10–12 thousand Muslims and later force them to adopt Hindu religion [sic]'² to their passing of a law that discriminatorily deprives Muslim refugees of a pathway to citizenship,³ there are an endless number of examples of how the BJP has actively normalised hatred towards Muslims such that 55.12% of Islamophobic tweets now originate in India.

In the case of the US, although Islamophobia has long been a problem, we find that it was dramatically exacerbated by the racist, conspiratorial and inflammatory rhetoric employed by Donald Trump. We note, for instance, that Trump's response to the assassination of Qasem Soleimani triggered a massive spike in hate on 3 January 2020, with some 9,302 anti-Muslim tweets being made that day. We also observe that Trump ranked as the third most frequently mentioned user in Islamophobic posts, a great many of which were focused on defending his ban on Muslim immigration, in addition to

² Human Rights Watch, 'Shoot the traitors: discrimination against Muslims under India's new citizenship policy', *Human Rights Watch*, (online, 14 July 2022) <<https://www.hrw.org/report/2020/04/09/shoot-traitors/discrimination-against-muslims-under-indias-new-citizenship-policy>>.

³ Inamdar, N, 'Five instances that bear out the BJP's anti-Muslim stance on the campaign trail', *Quartz India* (online, 23 April 2019) < <https://qz.com/india/1602082/bjps-pragya-singh-amit-shah-guilty-of-anti-muslim-bigotry>>.



AUSTRALIAN MUSLIM ADVOCACY NETWORK

forwarding his theory that the Democrats were collaborating with ‘the Islamists’ to take over the West.⁴

As for the UK, we attribute the prevalence of anti-Muslim tweets to a multitude of factors, including the global reach of Trump’s hatred, the country’s longstanding issues with anti-migrant sentiment and, most importantly, the casual racism of former PM Boris Johnson, who once said of women who wear the niqab, ‘It is absolutely ridiculous that people should choose to go around looking like letter boxes.’⁵

21. The Australian Research Council has funded research into Transnational Hindutva Nationalism and its impact on diaspora audiences.⁶

Current regulatory environment

22. The Courts found the *Privacy Act 1988* (Cth) to apply to international digital platforms by virtue of section 5A, which extends the Act’s application to overseas entities in specified circumstances.
23. The *Online Safety Act 2021* (Cth) and *Broadcasting Services Act 1992* (Cth) do not address dehumanizing disinformation operations platformed and profited from by international digital platforms. As such, regulators like e-Safety and ACMA are not positioned to take action.
24. The Australian Code of Practice on Disinformation and Misinformation applies to international digital platforms. However, it has no effective enforcement mechanism and is self-regulatory.
25. AMAN, on behalf of the Australian Muslim community, has brought complaints under discrimination law.
 - a. In Queensland, AMAN has a complaint against Twitter Australia Holdings Pty Ltd and Twitter Inc under the *Anti-Discrimination Act 1991* (Qld). This matter is on the verge of being referred to QCAT.

⁴ MH Khan, F Qazalbash, HM Adnan, LN Yaqin and RA Khuhro, ‘Trump and Muslims: a critical discourse analysis of Islamophobic rhetoric in Donald Trump’s selected tweets’, *SAGE Open*, 2021, 11(1):1–16, p 5, doi:10.1177/21582440211004172.

⁵ Johnson, B, ‘Denmark has got it wrong. Yes, the burka is oppressive and ridiculous - but that’s still no reason to ban it’, *The Telegraph* (online, 5 August 2018) <<https://www.telegraph.co.uk/news/2018/08/05/denmark-has-got-wrong-yes-burka-oppressive-ridiculous-still/>>.

⁶ <https://www.deakin.edu.au/study/fees-and-scholarships/scholarships/find-a-scholarship/hdr-scholarship-transnational-hindutva-populism>



AUSTRALIAN MUSLIM ADVOCACY NETWORK

- b. AMAN also has a complaint against Meta Platforms Inc and Facebook Australia Pty Ltd under the *Racial Discrimination Act 1975* (Cth). This matter is still before the Australian Human Rights Commission.
- 26. It is noted that both companies have sought to argue that:
 - a. The Australian arms of their companies (Twitter Australia and Facebook Australia) do not have control over the platforms and, therefore, cannot be held liable for the harm these platforms cause.
 - b. The respective legislation does not apply to their overseas companies, which control the platform.
- 27. We note that these international digital platforms do not advertise this position and convey to the public, regulators, parliament and affected communities, that the Australian arms of their companies are in a position of power and control. Concerns are raised with the Australian arms and this point about a lack of control is only raised when legal action is brought.
- 28. We note it is in the interests of justice to lift the corporate veil between their Australian and foreign-based entities, to allow a Court to apply the domestic discrimination laws to social media companies.

Regulating international digital platforms

- 29. Mansted (2019) writes that activating Australian citizens to international influence and coercion is critical.⁷
- 30. This means that laws must clearly define harm to the public to help educate the public about what is unacceptable and why.
- 31. There is abundant evidence of the harm caused by leaving international digital platforms such as social media companies to regulate harmful information operations by themselves. Currently, those companies are externalising the costs to the Australian community, with harmful consequences for community safety and trust in Government and democracy.
- 32. Currently, efforts to address international influence are modelled on machinery used to counter terrorism, with an emphasis on information sharing and interoperability between Defence, Intelligence and Security and Law Enforcement.
- 33. The prevention-focused policy answers to this market failure will not emerge from security and law enforcement.

⁷ Katherine Mansted (2019), <https://nsc.crawford.anu.edu.au/publication/15715/activating-people-power-counter-foreign-interference-and-coercion>



AUSTRALIAN MUSLIM ADVOCACY NETWORK

34. The Australian Government needs to create a mechanism for joined-up social policymaking between agencies and regulators that protects the public information environment. The Australian Government has significant experience in social regulation in comparable spheres. For example, the license conditions imposed on telecommunication companies are designed to protect people with life-threatening or mental illnesses.
35. There are four main intersecting public policy goals for international digital platform regulation⁸
 - a. To prevent harm of online ecosystems and information operations to humans, democracy and democratic values, like equality before the law
 - b. To promote freedom of speech
 - c. To promote user autonomy
 - d. To promote innovation in the technology sector
36. While systems that promote safety by design are critical, we cannot escape the need for definitional clarity on harms. For example, seeking transparency on algorithms or hate speech data won't help if our framework is ambiguous on how we define harm.

UK Online Safety Bill

37. The UK online safety bill developments in December 2022 underscore the pitfalls of not providing definitional clarity, with previous efforts to address online hate and misinformation erased from the bill's latest version.
38. The UK Bill imposes various statutory duties of care on social media companies.⁹ Section 14 outlines duties to protect news publisher content, and section 15 outlines duties to protect journalistic content. The Bill defines “journalistic content” in relation to a user-to-user service, as
 - (i) news publisher content in relation to that service, or
 - (ii) regulated user-generated content in relation to that service;
 - (b) the content is generated for the purposes of journalism; and
 - (c) the content is UK-linked.
39. Currently, the UK Online Safety Bill may make it harder for targeted communities to take action on dehumanising information operations: There is no clear carveout for harmful information operations from the definition of journalistic content.

⁸ Amy Louise-Watkins (2021) *Regulating terrorist content on tech platforms: A proposed framework based on social regulation* (PhD Thesis, Swansea University)

⁹ See section 6 of the Bill.



AUSTRALIAN MUSLIM ADVOCACY NETWORK

Additionally, comment boards are not subject to the same rules as other user-to-user content.

Digital Services Act (EU)

40. The Digital Services Act (DSA) only applies to “intermediary services” provided to recipients of the service that have their place of establishment or residence in the Union.
41. The DSA does not impose obligations on intermediaries to monitor the information they transmit or store or to actively seek out illegal activity.
42. Intermediaries must, however, act against specific items of illegal content without delay if they receive a notice from the relevant national judicial or administrative authorities to do so.
43. Independent audit: Risk assessments and risk mitigation measures are both subject to annual independent audits at the platforms’ own expense. Very large platforms will be required to give auditors access to all relevant data and premises needed to perform audits properly, including data related to algorithmic systems. Where audit reports are not positive, platforms must take account of operational recommendations put forward by auditors.
44. Special transparency and algorithmic accountability measures for very large platforms: Very large platforms must:
 - a. Set out in their terms and conditions the terms used in their recommender systems and any options to influence these parameters to facilitate algorithmic transparency. At least one option must be provided that is not based on profiling.
 - b. Compile and publish information about the content, nature, origin and dissemination of advertisements until at least one year after the advertisement in question was displayed.
 - c. On request, provide the Digital Services Coordinator or the Commission access to data necessary to assess compliance with the DSA.
 - d. Appoint one or more appropriately qualified compliance officers to monitor compliance with the DSA and cooperate with the Digital Services Coordinator and the Commission.
 - e. Publish transparency reports once every six months rather than once yearly, as required for smaller platforms.
45. Section III of Chapter IV of the DSA sets out specific investigation, enforcement and monitoring powers in respect of very large online platforms, including:



AUSTRALIAN MUSLIM ADVOCACY NETWORK

- a. Enhanced supervision of platforms found to have infringed any of the provisions of section 4 of Chapter III.
- b. The Commission can intervene and open proceedings against a platform that has contravened the DSA if the Digital Services Coordinator has failed to take action.
- c. The ability to request further information from platforms and associated persons relating to infringement or suspected infringement of the DSA.
- d. Powers to conduct on-site inspections.
- e. Powers to monitor the implementation and compliance with the DSA by very large platforms, including the power to require access to, and explanations relating to, its databases and algorithms.
- f. The power to adopt non-compliance decisions and to issue fines.

The Australian context

46. Certain types of misinformation and disinformation are seriously harmful to adults and children.
47. While misinformation, disinformation, and information operations may be challenging to define, the harms that arise from them are not. Credible mechanisms can define the harm caused by misinformation and disinformation in different contexts (medical, electoral, dehumanizing discourse, science/climate science) by referring to compelling and authoritative sources. Research suggests that public awareness of the harm caused by misinformation is deficient.
48. In defining these harms, the Australian Government must open itself to civil society engagement, ensuring that this civil society is representative of Australia's cultural diversity and groups frequently targeted online.
49. AMAN's work has focused on defining the harm arising from dehumanising information operations.

The effect of dehumanisation

50. Referring to the Australian terrorist who carried out the Christchurch attack, Lentini (2019, 43) explains that,

Tarrant's solution to the crisis – indeed one on which he felt compelled to enact – was to annihilate his enemies (read Muslim migrants). This included targeting non-combatants. In one point in his 'manifesto', he indicates that they constitute a much greater threat to the future of Western societies than terrorists and combatants. Thus, he argues that it is also necessary to kill children to ensure that the enemy line



AUSTRALIAN MUSLIM ADVOCACY NETWORK

will not continue...Tarrant indicated that, when trying to remove a nest of snakes, the young ones had to be eradicated. Regrettably, children were among those whom he allegedly shot and killed.¹⁰

51. The Australian terrorist who carried out this attack was heavily influenced by international digital media from the United States and elsewhere.¹¹
52. A similar narrative inspired Anders Breivik, the Oslo terrorist who murdered 77 people in 2011. Breivik cited the author of JihadWatch, one of the information operations cited in Australian research.¹² The historical links between these two attacks, in terms of their relationship to 'counter jihad' dehumanising information operations considerable.¹³ With respect to dehumanisation, Kaldor (2021) notes,

Breivik also refers to Muslims as “wild animals,” who he argues are freely bringing about European “genocide” because “traitors... allowed these animals to enter our lands, and continue to facilitate them.” In keeping with the naturalistic theme, Tarrant’s text is also rife with mixed metaphors describing how individuals such as himself can no longer escape Western civilisation’s contamination: “there is no sheltered meadow... there is not a single place left where the tendrils of replacement migration have not touched.” Comparing immigrants to a “vipers [sic] nest”, he implores followers to “burn the nest and kill the vipers, no matter their age.” Crusius similarly bewails how those without the means to “repel the millions of invaders” “have no choice but to sit by and watch their countries burn.” The repetition of animalistic metaphors is no accident: the perpetrators intentionally dehumanise immigrants by depicting them as beastly, thereby making their complaint about Western society’s perceived decline more justifiable to their readers.¹⁴

¹⁰ Lentini, Peter. 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.

¹¹ <https://www.cnn.com/2020/12/08/youtube-radicalized-christchurch-shooter-new-zealand-report-finds.html>

¹² Abdalla, Mohamad, Mustafa Ally, Rita Jabri-Markwell. 2021. “Dehumanisation of Out-Groups on Facebook and Twitter: Towards an Assessment Framework for Online Hate Actors and Organisations.” *SN Social Sciences* (1) 9; Peucker et al (2022), op cit.

¹³ Rita Jabri Markwell, “The online dehumanisation of Muslims made the Christchurch massacre possible” *ABC Religion and Ethics*, 31 August 2020, <https://www.abc.net.au/religion/the-online-dehumanisation-of-muslims/12614148>

¹⁴ Sophie Kaldor, ‘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>.



AUSTRALIAN MUSLIM ADVOCACY NETWORK

53. 'Dangerous speech', a category expounded in detail by Maynard and Benesch (2016), is speech that constructs an 'outgroup' as an existential threat to the 'in-group,' whether this threat is real or otherwise (81).¹⁵ Dehumanisation and another technique referred to as 'threat construction' are two techniques used in dangerous speech. They are often inextricably linked: 'where dehumanization makes atrocities seem acceptable, threat construction takes the crucial next step of making them seem necessary' (82).

The unregulated online sphere

54. Researchers from Macquarie and Victoria Universities published the first study mapping the online activity of right-wing extremists (RWE) in New South Wales, Australia.¹⁶ The study identified the dehumanisation of out-groups to in-group audiences as a core component of their online socialisation.¹⁷
55. Significantly, their research found that dehumanisation existed on 'low-risk' platforms like Facebook and Twitter 'without violating platform moderation policies.'
56. While preventing dehumanization of outgroups is a widely accepted goal in the field of countering violent extremism, current social media algorithms focus on detecting individual samples through explicit language.¹⁸ Their approach completely misses purposed information operations that work to dehumanize outgroups over time with more borderline and implicit content.
57. Internal mechanisms for escalating possible breaches needs strengthening, forcing them to rely upon external actors to monitor and raise occurrences for investigation.
58. Recommender algorithms promote other pages and groups run by the same admins, which presents an issue when those accounts may be fake or examples of international influence.
59. We can point out signs of international coordinated behaviour and content-sharing by those involved in 'organised hate', but are reliant on the interest and goodwill of platforms to investigate.
60. For example, Facebook will dedicate resources to uncovering inauthentic behaviour in a region, but this depends on their will and interest to investigate a particular region.

¹⁵ Maynard, Jonathan Leader and Susan Benesch. 2016. "Dangerous Speech and Dangerous Ideology: An Integrated Model for Monitoring and Prevention." *Genocide Studies and Prevention: An International Journal* 9(3): 70.

¹⁶ Department of Security Studies and Criminology. (2020, October 9). Mapping Networks and Narratives of Online Right-Wing Extremists in New South Wales (Version 1.0.1). Sydney: Macquarie University.

¹⁷ Maynard and Benesch, op cit.

¹⁸ Abdalla et al (2021), op cit.



AUSTRALIAN MUSLIM ADVOCACY NETWORK

61. The interaction of the inauthentic policy with their misinformation and hate speech policies needs consideration. Domestic and international actors may be working in concert to spread malicious or false news to incite hatred. Still, sharing links to external sites means they can circumvent the hate speech policy, and Facebook only independently tests misinformation in a fraction of cases.

62. Facebook relies on the official terror listings of countries and its own list of white supremacist organisations (which it developed following Christchurch after realising that government lists were inadequate). Facebook uses that list to identify 'terror' or 'organised hate' groups by their names, emblems, logos, branding and slogans.

63. International digital platforms often lack explicable policies on how they assess 'hate' or 'dangerous' organisations.

64. An investigation by Guardian news revealed an overseas commercial enterprise that was

using its 21-page network to churn out more than 1,000 coordinated faked news posts per week to more than 1 million followers, funnelling audiences to a cluster of 10 ad-heavy websites and milking the traffic for profit.

The posts stoke deep hatred of Islam across the western world and influence politics in Australia, Canada, the UK and the US by amplifying far-right parties such as Australia's One Nation and vilifying Muslim politicians such as the London mayor, Sadiq Khan, and the US congresswoman Ilhan Omar.

65. In early 2020, AMAN and Birchgrove Lawyers investigated Facebook's efficacy in enforcing its hate policy standards. This was on the back of unpublished research by Victoria University in 2018, which studied over 41,000 posts in far-right Facebook groups, identifying radicalising discourse. The most prevalent mobilisation frames included five narratives about Muslims and Islam. We wanted to test whether those groups were still active 2 years later, even after Christchurch, and whether they will still propagating the same radicalising messages. The Victoria University study published the year before the Christchurch attack, found that the narratives contained in Tarrant's manifesto were largely prevalent in far-right Facebook groups.¹⁹

¹⁹ Dr Mario Peucker, Dr Debra Smith, & Dr Muhammad Iqbal, 'Mapping Networks and Narratives of Far-Right Movements in Victoria' (Project Report, Institute for Sustainable Industries and Liveable Cities, Victoria University, November 2018).



AUSTRALIAN MUSLIM ADVOCACY NETWORK

66. Our investigation of the groups we could find (some had slightly different names) revealed they were still very active. At the time Facebook welcomed our investigation, even though it was reported in national media, and acted on those pages. Facebook has since indicated that they find ‘escalation’ by civil society to be very useful in identifying harmful pages and groups; however, we have learned they are unlikely to take action to geoblock content unless there is a legal finding, placing the burden on affected communities to bring legal proceedings under discrimination laws.

How algorithms contribute to the spread of dehumanising information operations

67. There are three main categories of algorithms²⁰

Algorithms for content recommendations

Platforms may use algorithms to recommend content in ways that personalize recommendations for individual users based on their past behaviors (as well as inferred characteristics) and optimize expected value to the company by maximizing individual users’ expected engagement with recommended content.

When platforms use algorithms to maximize engagement, they cannot fully prevent harmful third-party contents from being recommended to users if those users have consumed similar contents in the past.

Algorithms for content moderation and safety

Platforms use algorithms to prevent and reduce harms by semi-automating the process of flagging, removing, and re-ranking third-party contents likely to violate platform policies or laws. When this process is performed at scale, the algorithms cannot perform perfectly and are continuously optimized to balance between precision and accuracy.

²⁰ Integrity Institute, *Summary: Integrity Institute’s Amicus Brief on Gonzalez v. Google*, 9 December 2022, <https://integrityinstitute.org/amicus-brief-summary-sheet>



AUSTRALIAN MUSLIM ADVOCACY NETWORK

If a platform prioritizes accuracy over precision in using algorithms for content moderation, its process would have a high false positive rate. Most large platforms therefore choose to prioritize precision over accuracy, which allows most users to post contents but can sometimes lead to extensive harm when false negatives are shared widely.

68. However, the ability to assess dehumanising information operations with accuracy and precision is more possible than identifying violations of hate speech policies at large, because there is a visible and distinct formula that such operations use in order to dehumanise an outgroup to an ingroup audience.²¹

Algorithms for advertising and commerce

Platforms use algorithms to serve targeted ads to individuals through “retargeting,” which relies on expressed and inferred information about those individuals that the platforms had already compiled.

²¹ In the second half of 2020, AMAN completed a study of five actors producing significant amounts of blog or pseudo-news content that triggered explicitly dehumanizing and violent responses by users on Facebook and Twitter. That study identified the following markers that were common to all five actors’ information operations:

1. Dehumanizing conceptions or conspiracy theories on the actor’s website (where applicable) in relation to an identified group (“the out-group”) on the basis of a protected characteristic;
2. Repeated features of the headlines and images that are curated for a specific audience, including:
 - Essentializing the target identity through implicating a wide net of identities connected to the protected group (e.g., “Niqab-clad Muslima,” “boat migrants,” “Muslim professor,” “Muslim leader,” “Iran-backed jihadist,” “Ilhan Omar,” “Muslim father”);
 - High degree of hostile verbs or actions (e.g., stabs, sets fire) attributed to those subjects;
 - Significant proportion of actor’s material acting as “factual proofs” to dehumanizing conceptions about out-group;
 - Potential use of explicitly dehumanizing descriptive language (e.g., frothing-at-the-mouth) or coded extremist movement language with dehumanizing meaning (e.g., invader, a term used in RWE propaganda to refer to Muslims as a mechanically inhuman and barbaric force). However, for the most successful actors, dehumanizing slurs were avoided to maintain legitimacy and avoid detection; and
 - Where there was no dehumanizing language, there was a presence of “baiting” through rhetorical techniques like irony to provoke in-group reactions; and
3. Evidence in the user comment threads of a pattern of hate speech against the outgroup.

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, p 60.

Web: www.aman.net.au



AUSTRALIAN MUSLIM ADVOCACY NETWORK

Algorithms that are used in techniques like “retargeting” primarily benefit companies, and this encourages companies to collect more and more data about users.

69. The above is important to note as it applies to state and non-state-sponsored information operations.

The formula of dehumanising information operations

70. Abdalla et al. (2021) studied the operation of five (5) online information operations located within the extreme right ‘counter jihad’ movement. The leading actor studied conveyed a similar demographic invasion narrative in Tarrant’s manifesto. They found that Facebook and Twitter’s automatic detection tools could not detect explicit dehumanising slurs or violent fantasy in the fantasies threads, meaning that online communities could react together to information towards a targeted group without disruption.²² In this phenomenon, we see a combination of the cognitive, behavioural and social license granted to participants in the comment threads to erase the humanity of the target group and indulge in violent fantasy.

71. It is contended by Abdalla et al. that,

the marshalling of stories to create an overwhelming sense of crisis and disgust does not always rely on explicit dehumanising descriptors, verbs, or coded language in the headlines. Where an audience had been primed over time, implied properties in text are capable of triggering entire sub-texts.

72. Further, actors often attributed subhuman actions to Muslims in the headlines to dehumanise all Muslims overtime. Platforms could not detect this technique as they were focused exclusively on dehumanising comparisons, synonyms and adjectives (e.g., disease, filth, cancer, weeds, insects).

73. Abdalla et al note,

Given its established links to atrocities and genocide, dehumanisation offers a widely accepted measure. However, operation through discourse (not language alone) must be analysed by platforms to competently assess actors and their information operations.



AUSTRALIAN MUSLIM ADVOCACY NETWORK

74. Abdalla et al's research points to several predictors that could be used to make competent and consistent assessments of hate actors running purposed information operations.

Access to data for research

75. The US Congress is considering a Bill²³ in relation to social media company transparency on data in 2021. Until recently, Twitter was one of the most helpful companies for providing API data to researchers. However, there have been signs free access is being withdrawn. Twitter should return to its original API sharing model, which was minimum standard of what we ought to expect from social media companies.
76. The European Digital Services Act has provisions in relation to data sharing.²⁴ Very large platforms must provide vetted researchers with access to key data for the purposes of conducting research on “detection, identification and understanding of systemic risks in the EU” and “assessment of the adequacy, efficiency and impacts of the risk mitigation measures” on request from a regulator.
77. Researchers must be affiliated with academic institutions, be independent from commercial interests, have proven expertise, and commit to data security and confidentiality requirements. Specific data access conditions will be set out in delegated legislation, which will be developed over the next 18 months. Companies may request that the Digital Services Coordinator amend a request if it does not have access to the data, or if giving access to the data will lead to vulnerabilities for the security of its service or protection of confidential information.

Strategic use of information

78. Sibley (2022) writes about the strategic use of narratives by the counter-jihad/ anti-Islam movements, namely, The For Britain Movement (TFBM), the Democratic Football Lads Alliance (DFLA), and Patriotic Europeans against Islamisation of the Occident UK (PEGIDA UK).²⁵
79. Sibley writes these movements ‘represent a new type of far-right politics, based on the anti-Islam master frame, which actively opposes fascism, thereby distancing itself from the traditional right (Berntzen, 2019; Jackson and Feldman, 2014).’ She explains

²³ Social Media Disclosure And Transparency of Advertisements Act of 2021
<https://www.congress.gov/bill/117th-congress/house-bill/3451/text?r=35&s=1>

²⁴ Vermeulen, M. (2022). Researcher Access to Platform Data: European Developments. *Journal of Online Trust and Safety*, 1(4). <https://doi.org/10.54501/jots.v1i4.84>

²⁵ Sibley, A. (2023). Behind the British New Far-Right's veil: Do individuals adopt strategic liberalism to appear more moderate or are they semi-liberal? *The British Journal of Politics and International Relations*, 0(0). <https://doi.org/10.1177/13691481221148330>



AUSTRALIAN MUSLIM ADVOCACY NETWORK

Resulting from World War II, certain ethnic-based prejudices such as biological racism were deemed unacceptable (Kallis, 2013). Far-Right parties that use an ethnic national identity narrative, therefore, are often categorised as racist and xenophobic. By focusing on the perceived values of Islam rather than Muslims, individuals avoid convictions for hate speech. By focusing on Islam and not restricting an individual's ability to practice Islam, freedom of religion laws are not violated (Howard, 2017).

80. AMAN considered the diverse ideological and social dynamics of 30 Facebook pages that repeatedly post content from identified international information operations to contextualise the use of this content.
81. Most pages display common elements of one or more of the following characteristics:
 - a. The dehumanisation of Muslims as socially, culturally and or ethnically inferior
 - b. The veneration of Western Civilisation and Christianity by association
 - c. The romanticism of perceived traditional Western values, such as freedom of speech
 - d. Sympathy to Far-Right conspiracy theories and reactionary political counternarratives
 - e. A limited pool of news or factual data to reinforce confirmation bias
 - f. Zero-Sum Game attitude to an Islamic presence in society
82. However, there were also significant differences, outlined below in categories we have observed.
83. We share this to highlight that international dehumanising information operations provide material to a **wide range of groups**, and with enough frequency, may provide a social ramp towards progressively more extreme and violent worldviews.

Civic Traditionalists

84. Based upon an imagined shared community of common values, the civic conservative illustrates a well-established and rigid sentiment toward Muslims. While not as directly hostile to Islam as other groups, the narratives surrounding the proliferation of Islam matches the fear of erosion to the civic society they believe once existed.
85. While sympathetic to many of the narratives of the more populist or ethnically focused counterparts in their circles, the Traditionalist is brought to the discussion by a looming sense of nostalgia and regret for any perceived emergence of Islam. They are concerned with protecting the state they believe they know.



Neo-Crusader ideologies

86. Neo-Crusader pages firmly focus on the ongoing conflict between Islam and Christianity as the primary combatants. The ideological drivers reflected a religiously charged rhetoric, often relying on selectively quoted Christian scripture or Christian orthodoxy for political, social and cultural guidance against Islam. A by-product of this hostility is sympathy towards Israel and uniting other Abrahamic religious groups against Islam. Any notion of religious pluralism or secularism is deliberately cast aside.

Far-Right Populists

87. A group steeped in the momentum of reactionary political thought, the Populists function much similar to other groups but with one key deviation: their easy transition from one target of Far-Right hostility to another. With a focus on generating traffic and outrage as a rallying cry for their followers, this group has the greatest ease in pushing for political outcomes.

88. Proclaiming a narrative of standing against emerging threats, this group largely fosters and maintains anger as a political tool.

89. While having a broad range of targets, the agenda's frequency and intensity can create a feedback loop of specious reasoning to draw a constant sense of purpose. Due to their ability to rapidly adapt and integrate different political narratives into their worldview, they can keep themselves as a constantly active movement with the greatest potential for action. While not often sophisticated in messaging, the phenomenally high speed of narrative integration, ideological replication and fostering of hyper-scepticism make them highly resistant to differing outlooks.

Ethno-Nationalists

90. While displaying similar fundamental values to other civic conservatives, the Ethno Nationalist takes the discussion of race and culture one step further. They base much of their ideology on ethnic/racial distinctiveness and hierarchy. Inherently 'Western' traits are believed to exist almost genetically, justifying a vehement opposition to allowing anything that might upset a perceived natural order to race relations.

91. Focusing on a clearly xenophobic sense of racial superiority, these ideological proponents take on an almost Darwinian mindset while exclaiming strawman

Web: www.aman.net.au



AUSTRALIAN MUSLIM ADVOCACY NETWORK

arguments to support their points. This often is used in the form of attacks on Muslims as refugees and frames the discussion in terms likening them to sub-humans. They use emotionally charged imagery and stories to promote a loaded view of ethnic integration as difficult and undesirable, as violence and genetic extinction are the only perceived outcomes.

92. Ethno-nationalism is observed in both Hindutva groups and white supremacist groups.

How Australian policy could contend with international and domestic dehumanising information operations

Recommendations

93. Ideologies and targeted groups evolve and change. A proactive and resilient architecture must be considered through existing online safety and media regulation schemes. This architecture must
 - a. adopt an 'atrocities-prevention' approach focused on maintaining collective social barriers to hatred and violence in all its forms. This will help shift the burden further upstream to digital platforms and away from communities most affected by downstream carceral, securitised and heavy-handed policing approaches.
 - b. capture content that socialises people towards violence yet limits the aperture of regulator intervention to serial or systematic vectors of **hate** that lower an audience's barriers to violence.
 - c. be resilient to changes in targeted community, ideologies, and tactics (for example, capturing harm through disinformation).
 - d. provide a framework for education that supports critical thinking by applying logic and universal principles that a person would like to see applied to them.
 - e. Provide a superior and precise definition that attracts public support and provides certainty about what behaviour crosses the line to administrators and the public.
94. It is submitted that dehumanising speech and discourse satisfy the above architecture's elements as a standard of harm.

Web: www.aman.net.au



AUSTRALIAN MUSLIM ADVOCACY NETWORK

95. The Australian Muslim Advocacy Network (AMAN) has previously proposed civil penalties for serial or systematic dehumanising speech or discourse for the Online Safety Act using a notice and action takedown model.²⁶
96. AMAN has suggested draft *definitions for dehumanising speech and discourse*.²⁷
97. Regarding possible law reform to protect the public sphere from dehumanising discourse about protected groups, the Australian Government should plan for systems that place responsibility on platforms with the resources and expertise to resolve concerns about content moderation and negative algorithmic impact.

Anti-Discrimination Laws

98. The Australian Government could consider an amendment to Anti-Discrimination laws like the provision within the Privacy Act, which clarifies their extraterritorial application to activity with a substantial link to Australia. This could be done with legislation brought into parliament as part of a discrimination law package connected to the proposed Religious Discrimination Bill in 2023.

Online Safety Act 2021 (Cth)

99. The Australian Government could amend the Online Safety Act 2021 (Cth):
 - a. Prohibit platforms from enabling social media users or webpages to
 - b. Repeatedly encourage abusive audience behaviour towards a protected group; or
 - c. Repeatedly target, insult and abuse an identifiable individual based on their intrinsic attributes.
100. Consider further civil penalties against serial actors of dehumanising speech and discourse and the platforms that enable them.
101. The Australian Government could address algorithms, particularly recommender algorithms, once the Government has discerned clear objectives for protecting the public information environment.
102. In consultation with expert civil society from affected communities, E-Safety should draft an industry standard for assessing actors that incite hatred towards groups based on protected attributes. That standard must encompass information operations that dehumanise a group based on a protected

²⁶ <http://www.aman.net.au/wp-content/uploads/2022/06/Sub03-Australian-Muslim-Advocacy-Network.pdf>

²⁷ To see AMAN's working definition for dehumanising speech and discourse: https://www.aman.net.au/?page_id=1425



AUSTRALIAN MUSLIM ADVOCACY NETWORK

characteristic. We cannot rely on industry to draft this code.²⁸ E-Safety must adopt an approach grounded in multistakeholderism, rather than individually and separately consulting each community group.

103. The Australian Government could make platforms share in costs to regulators arising from their platforms and internalise the costs (financial, legal, loss of income, psychological) to individuals and communities arising from online hatred throughout their life.
104. The Australian Government could consider introducing a duty of care on digital platforms to uphold Australian hate speech standards, which may prompt platform investment in compliance units. Currently, they do not invest in this.

Lifting the Corporate Veil

105. Further, the Australian Government must lift the corporate veil between the Australian and foreign-based arms of international digital platforms. International digital platforms must not be allowed to use corporate structures to avoid legal accountability under Australian laws.

A note about user empowerment

106. User empowerment gives users more choices in how information is presented to them.
107. The latest iteration of the UK online safety legislation (December 2022) has relied extensively on 'user empowerment' protections to address online hate and misinformation, likely because agreement couldn't be reached on other approaches.
108. We agree that user empowerment is important. However, we note the inherent limitations of user empowerment: Communities that are hyper-sceptical of hate speech controls and more likely to embrace absolutist free speech. They will not use options to remove hate speech and misinformation. This means that targeted 'outgroups' will continue to be endangered by dehumanising disinformation operations.

²⁸ Reset Australia, 'How outdated approaches to regulation harm children and young people' https://au.reset.tech/uploads/report_-co-regulation-fails-young-people-final-151222.pdf