WORKSHOP ON TERRORISM, HATE SPEECH AND DISINFORMATION

26 February 2026, Canberra

SUMMARY REPORT

This is a summary of a hybrid workshop convened on 26 February at the City Campus of the University of New South Wales Canberra. The discussion was held under the Chatham House rule. Nothing in the report is intended to imply necessary agreement of the group with what was said. Editorial inquiries can be directed to Professor Greg Austin greg.austin@socialcyber.co.

ABOUT THE WORKSHOP

The Social Cyber Institute, in conjunction with the Crawford School of Public Policy of the Australian National University, convened a one-day Research Workshop on 26 February 2026 to analyse the linkages in Australia between counter-terrorism policy, controls on hate speech, the regulation of disinformation, and covert influence operations by foreign and domestic actors.

The Workshop particularly focused on online activities, but it was also strongly interested in other aspects of the challenge including the increase in hate crimes outside of cyberspace and resort to physical violence.

The Workshop brought together established social science investigators to present concrete research findings and to discuss them constructively and cogently. The Workshop was run as a hybrid event, both online and in-person in Canberra. The Workshop Host was Professor Emeritus Glenn Withers AO from ANU and UNSW Canberra, a former President of the Academy of the Social Sciences in Australia and Director of the Social Cyber Institute. It was co-hosted by Professor Greg Austin, also of the Social Cyber Institute and an Adjunct Professor at the University of Technology Sydney. The University of New South Wales Canberra, per Professor Deborah Blackman, kindly provided the facilities for the Workshop.

The Workshop comprised four sessions as follows:

1. Dilemmas of countering violent extremism and its impacts on Australians here and overseas.
2. Addressing disinformation operations and hate crimes in Australia against religious and migrant communities.
3. Addressing hatred and violence against other social groups: women, LGBTI, the disabled, and First Nations people.
4. Policy proposals for Australian actors (communities, business, civil society organisations and governments) to combat hate crimes and incitement to violence.

This report of the Workshop summarises the discussion, but it does not provide specific recommendations and conclusions. These are the responsibility of individual participants. Discussion operated under the Chatham House rule, where contributions to the debate are not individually identified. There were 45 enrolled participants (18 in person and 27 on-line) for the Workshop.

This report of discussion has been kindly compiled by Professor Greg Austin, with copy design by Zara Yap, but particular interpretation of discussions are Professor Austin's responsibility.

Professor Glenn Withers AO

Workshop Host

20 April 2026.

EXECUTIVE SUMMARY OF THE WORKSHOP DISCUSSION

Australian governments cannot meet today's complex threat environments for hate speech, disinformation and terrorism by assigning leadership of this effort only to security agencies or by over-reliance on legislation. Australia also needs new, durable mechanisms more firmly rooted in civil society to cultivate low tolerance for social violence in all its forms—terrorism, hate speech, misogyny, racism, antisemitism, Islamophobia, homophobia, and other forms of group hatred or subversion. These mechanisms must be designed to strengthen everyday democratic resilience rather than simply expand the punitive reach of the state after the next violent event on the political stage.

These are some of the key ideas canvassed by speakers in a workshop in February 2026 that brought together scholars, practitioners and policy thinkers to examine terrorism, hate speech and disinformation in Australia. The workshop was convened in response to the Bondi attack, new hate-speech legislation, and the establishment of the Royal Commission on Antisemitism and Social Cohesion.

One point of departure was that cultures of violence are increasingly shaped in online environments. Another was that the most effective responses to such violence need to combine more up-to-date evidence-based policy with community-led initiatives able to engage youth and vulnerable groups where they actually live, learn and socialise. Civil society organisations, local leaders and affected communities therefore need resourcing, structured roles and institutional voice in designing and evaluating responses, rather than being treated as passive implementers of centrally defined and poorly funded programs.

The discussion at the Workshop highlighted some limitations in current frameworks built around broad criminalisation, proscribed organisations and post-incident legislative cycles. Empirical evidence presented from decades of Australian cases shows that most of the individuals holding extremist beliefs actually never mobilise to violence and that additional factors such as social isolation and grievance are more predictive for mobilisation to violence than ideological label alone. One associated view was that overly broad offences and criminalised symbol bans risk diverting scarce operational capacity toward people who are hateful but non-violent, while generating chilling effects on artistic, critical and community expression—as illustrated by the Canberra café case involving satirical art in February 2026. On the other hand, a plea was made that banning Nazi symbols completely should be justified because of the gravity of the Holocaust.

Participants discussed whether Australia's counter-terrorism apparatus, heavily shaped by the early-2000s anti-jihadist paradigm, may now be misaligned with the current and fragmented landscape of mixed ideologies, online-native ecosystems and state-linked actors. At the same time, fiscal and political pressures may be pushing systems back toward narrow counter-terrorism functions at the expense of underlying prevention and social cohesion work. It was suggested that this narrow dynamic, risks doing "what terrorism wants": deepening polarisation, eroding civil liberties and weakening trust in institutions, while failing to address the drivers of extremist mobilisation.

Against this backdrop, the workshop discussion therefore canvassed whether the model for controlling hate speech would be more sustainable if it was centred on partnerships and networks within civil society and multi-layered interventions based on social license, rather than on criminal sanctions delivered after the fact. There was ongoing support for continuation of targeted and well-defined criminal and civil law punishments for incitement and serious vilification, but this should, the view asserted, still need to be accompanied by improved guidance and training for frontline police and regulators; and robust platform governance to tackle explicit incitement and harassment without suppressing legitimate debate and communication. It also requires significantly more investment in community-based leadership to promote pluralism, digital literacy, and coalitions. The purpose would be to respond to antisemitism, Islamophobia, anti-immigrant racism, misogyny, gender-based and sexuality-based phobias, ableism and attacks on First Nations people as interconnected challenges rather than as competing or sectionalised grievances.

The workshop discussions included reference to the value of an explicitly multi-disciplinary, evidence-based approach in which criminologists, psychologists, sociologists, legal scholars, technologists and community practitioners co-produce policy with those most affected by violence and hate. Research from several sources suggested that women, LGTBQI, immigrants, disabled and first nations people suffer severe insecurity in Australian society on a substantial scale, but in different ways that may make their suffering less prominent in political framings than antisemitism or Islamophobia, yet no less challenging. There was a view that Australia needs to address both generalised settings for combatting violence and hate speech as well as actions aimed at protecting specific groups.

To set lower tolerance levels for hate speech and violence in general, governments and society might see a new well-funded, collaborative civil society infrastructure as a core policy need, rather than leaving most things to police and intelligence. A new collaborative policy embracing both federal and state/territory governments in collaboration with business and civil society and directed at countering grievance could be funded as a social good. It would reconcile the concepts of social cohesion and peaceful pluralism.

The proposition is that we can’t have a safe and secure social order without this higher mobilisation of whole of society engagement against hate speech and violence. Those who remain as trusted and respected leaders have a key role that is needed to cater to our better angels. Indeed, in this way, trust in our social and political organization can to an extent recover from some substantial loss over time.

TABLE OF CONTENTS

1 Context of the Workshop

From September 11, 2001 to the attack on Bondi on 14 December 2025, Australia's terrorism threat analysis evolved from discerning a small number of large, centrally organised jihadist plots to a higher frequency of smaller, more diverse and often lone-actor attacks. Plots from religiously motivated violent extremists remain prominent in the data, but there has been a marked growth in extreme-right and ideologically "mixed or unclear" cases, often fueled by conspiracy theories, anti-government sentiment and grievance politics. The March 2019 Christchurch mosque shootings by Australian Brenton Tarrant should have been a warning since he is an Australian citizen who flew to New Zealand to prepare and perpetrate the attacks. The unchecked rise of neo-Nazi groups in Australia after 2019 was one example of the changing threat, exacerbating already high anti-semitic, Islamophobic, homophobic and anti-immigrant activism. Online spaces played a central role, with terrorist and extremist actors exploiting social media, gaming and encrypted platforms to groom isolated youth and build decentralised communities that have been hard to track.

Australian entanglement with overseas wars, extremist violence and hate speech trends has aggravated or shaped domestic extremism. The country's dependence on foreign-owned internet-based platforms and the flood of disinformation by great power actors and extremists have significantly reduced the room for effective monitoring and response by Australian governments and communities. Practitioners monitoring online data report that organised networks deliberately "seed" and amplify hate across countless micro-spaces, moving people from casual prejudice to more extreme views and, in some cases, to active engagement.

Australia did build a substantial counter-terrorism (CT) apparatus after 9/11, largely geared to jihadist threats and proscribed organisations, and developed a layered countering violent extremism (CVE) architecture from 2005 onward—primary, secondary and tertiary programs, units in several states, and more intensive case-management of at-risk individuals. Yet resource constraints, shifting priorities (towards great-power competition, climate and other threats) and political fatigue have left funding for CT and CVE under pressure. Governments consistently under-invested in prevention relative to post-incident spending, and civil-society actors working on hate and disengagement reported chronic funding insecurity and marginalisation.

The 2025 Bondi attack, described as Australia's deadliest terrorist attack in terms of fatalities, triggers what several speakers characterised as one of the most toxic and polarised episodes of public discourse in recent memory. Instead of a measured search for lessons, many leading political figures engaged in mutual denunciation, imputing bad faith to opponents and amplifying grievance—precisely the second-order effects that terrorism seeks. Public debate quickly centred on "who dropped the ball", premised on expectations, perhaps unrealistic, that security agencies, courts, parole boards and social services should have foreseen and prevented the attack.

Politically, the attack became the framing device for policy moves that were already in gestation. Some participants argue that the government's antisemitism and hate-speech package, a large legislative bill, and the announcement of a Royal Commission were unlikely to have been conceived and drafted purely in the short window after Bondi, yet they were presented as rapid, decisive responses. This rush to be seen to be doing something can be seen as a repeat of earlier cycles in which emotionally charged environments produce reactive, symbolically driven law-making with uncertain security benefits and significant civil-liberties costs.

2 Purpose of the Workshop

The workshop brought together Australian scholars, practitioners and policy thinkers to examine terrorism, hate speech and disinformation in Australia in the light of the Bondi attack, subsequent legislation and the establishment of the Royal Commission on Antisemitism and Social Cohesion. The meeting was conducted under the Chatham House rule. It was hosted by the Social Cyber Institute with venue support from UNSW Canberra. Participation in the workshop was broadly representative in terms of specialisations of the researchers relative to the issues, but even with 40 registered participants we were unable to attract comprehensive coverage on some issues given the short preparation time.

The workshop had several broadly equal but quite distinct aims:

- to discuss generalised approaches to combating cultures of violence and hate speech in Australia without an overly narrow categorisation of targets;
- to undertake a collective review and public airing of specialist research on particular target groups;
- to focus on ways that online environments are reshaping sentiments of perpetrators of hate and advocates for violence and intensifying the negative impact on the victims; and
- to analyse the linkages in Australia between counter-terrorism policy, controls on hate speech, the regulation of disinformation, and covert influence operations by foreign and domestic actors.

The format involved four sessions: an opening session on the changing threat environment and cultures of violence; two sessions on specific victim/target groups (including antisemitism, Islamophobia, anti-violent sentiment, hate speech targeting women, LGBTQI people and First Nations communities); and a final wrap-up session on policy directions and next steps. This report is organised around the material that emerged in presentations and discussion rather than simply around the prior stated aims of the workshop.

3 Unifying social goals

The transcript points toward a shared foundation in the scholarly community built around protecting human dignity and individual security in a pluralistic democracy alongside continuing support for urgent responses to variegated threats to specific groups.

3.1 Human Rights, Dignity and Limits on Free Speech

Several speakers noted up front that freedom of expression is important but not absolute. Approaches to hate speech need to be grounded in international human rights principles that allow restricting expression to protect national security, public order and the rights of others.

Several contributors criticised “social cohesion” as an inadequate or inappropriate organising principle. Some saw it as owing more to authoritarianism than we might want. An alternative case was made for “peaceful pluralism,” “social solidarity” and inclusive democracy as the better normative horizon.

The transcript emphasises bridging social capital—relationships across communities—over narrow “bonding” within like-minded groups, as the sociological basis for resisting hate and violence. Online hate is seen as organised, networked activity that creates “sticky spots” and permissive environments in which violence becomes thinkable. Effective responses must therefore be collective, cross-community and focused on changing relational dynamics, not just individual attitudes. There is a call to see hate speech against Jews, Muslims, migrants, women, LGBTQI people, First Nations and disabled people as interconnected forms of “violent and hateful extremism,” so that communities can be seen as ideally working together on the overarching pathologies of hate rather than competing for victim status exclusively for specific groups.

3.2 Protecting individuals from psychological harm

Several speakers argued for more attention to the impact of hate speech and online abuse—including misogynistic and racist campaigns, doxing, deepfakes and dehumanising slurs—as psychologically devastating for targets and their families, regardless of whether perpetrators “intended” harm. There were arguments for better reporting mechanisms, trauma-sensitive handling of victims’ stories, and stronger enforcement of existing laws (e.g. carriage-service offences) to signal that individuals are not expected to absorb unlimited abuse as the price of participation in society.

Trust in institutions was treated as essential: without confidence in police, regulators and courts, individuals affected by hate speech will not report, and law-based protections cannot achieve their intended function. Leadership for this was seen as essential.

3.3 Integrating regulation of hate-speech and violence prevention

Conceptually, the transcript exposes a continuum of policy response from addressing hateful but “non-violent” extremism through to violent extremism, with hate-speech environments understood as part of the enabling ecology of terrorist and other attacks. Yet some participants warn against collapsing all hate speech into “terrorism”: the unifying foundation they sketch is a human-rights and democracy frame in which (a) hate speech is constrained because of the demonstrable harms and its role in normalising violence, and (b) security tools are kept proportionate and evidence-based.

This points toward to a preference for multi-layered strategies: calibrated laws on incitement and vilification, community-based resilience work, youth-centred programs, and platform governance—anchored in a commitment to equal dignity across all targeted groups and to reducing both physical violence and the everyday psychological violence of uncontrolled hate.

Some participants repeatedly distinguish between policing “bad ideas” and policing harmful behaviour, arguing that the state should focus on hate speech and incidents as conduct that damages dignity, safety and equality rather than trying to eradicate beliefs or ideas.

4 Overview of evidence and research

There is an almost 40-year Australian database on radicalisation and extremism, which supports robust findings, even across changing circumstances. For example, most radicalised individuals never mobilise to violence and risk factors for mobilisation to violence (e.g. social isolation) differ from those for holding extremist beliefs:

- Radicalisation is best understood as a spectrum of beliefs and behaviours, which may or may not lead to violence.
- Common features among radicalised individuals include a grievance, a sense of injustice, engagement with extremist material, and sometimes affiliation with groups—but these are not the same factors that readily predict mobilisation to violence.
- Factors like social isolation, while not highly prevalent in the general population (estimated around 20%), appear disproportionately among individuals who have mobilised to violent acts in the Australian dataset, suggesting different risk factors at the mobilisation stage.
- Multiple systematic reviews of CVE show “modest but overall positive” effects in interventions for CVE (for example, case-management, counselling, mentoring).
- Fewer people move from radical beliefs into violent offending, and/or there is a reduced likelihood that already-involved individuals remain engaged in or return to violent extremism.

This substantive record of evidence unfortunately contrasts with almost non-existent

evaluation evidence for some areas still, such as measures like banning slogans, symbols or possession of materials as counter-terrorism tools.

4.1 Known gaps and limitations

Despite some progress then, researchers still “can’t predict” individual behaviour with accuracy. Risk factors interact in complex, adaptive ways, and the threat picture (mixed ideologies, grievance-fueled conspiracies) is actually changing faster than validated tools can keep up.

Existing CVE frameworks and risk-assessment tools were not designed for the current hybrid threat environment, so their applicability to new constellations of grievance, anti-government sentiment and online-native movements is currently uncertain.

Several speakers observe that evidence for the effectiveness of criminalisation of phrases, symbols or possession of banned material in reducing terrorism is thin to non-existent, especially compared with the more developed CVE evidence base.

4.2 Data on hate speech and psychological harm

Some participants cited strong qualitative and quantitative evidence that hate speech and online abuse (including misogynistic and racist campaigns) cause serious psychological harm, but note that legal thresholds focused on “incitement to violence” often fail to capture this. For some groups—First Nations people, LGBTQI communities, disabled people—there are solid prevalence data on victimisation and harassment, yet these do not translate cleanly into hate-speech policy because “hate” itself is undefined in key statutes, and categories vary across jurisdictions.

Practitioners monitoring online spaces report rising volumes of hate and normalisation of extremist rhetoric, but a large and growing share of accounts are anonymous or locked, making attribution and rigorous causal inference (e.g. from normalisation to violence) difficult.

4. 3 Structural and sociological research gaps

Some speakers emphasised that extremism and hate flourish in broader crises (economic inequality, climate stress, displacement, post-Voice racism) but say research and data linking these structural drivers to online/offline hate and violence in Australia also remain under-funded and hence under-developed. Priority in grants schemes to the more seemingly apolitical STEM areas and topics exposes us socially to not adequately addressing such cultural dysfunction.
There is particular concern therefore that governments under-invest in primary prevention

and in civil-society-based research, leaving big questions about grievance, polarisation and community-level impacts of hate speech under-explored compared even with case-based security intelligence, let alone wider researched areas.

Overall, the transcript portrays a field where individual-level radicalisation and CVE are somewhat well studied, while the wider sociological ecology of hate speech, its psychological impacts, and the effects of new speech laws are only patchily understood and urgently need more rigorous, multidisciplinary research.

The workshop deliberately foregrounded hate speech targeting women, LGBTQI people, disabled people and First Nations communities, arguing that these are central to understanding contemporary cultures of violence, and are not merely "identity-based" add-ons.

Some participants pointed out that:

- Gendered and sexualised abuse, especially online, is pervasive and normalised, often dismissed as "just trolling" despite its cumulative impact on safety and participation in public life.
- Disabled people face dehumanising tropes and scapegoating, including in conspiratorial narratives about "burdens on the state" or "unproductive citizens."
- First Nations communities continue to encounter racist and colonial narratives that portray them as inherently criminal, deficient or undeserving, which intersect with contemporary debates on law and order and resource allocation.

These forms of hate were further described as part of the background noise of public discourse, shaping who feels safe to speak, organise and lead.

Some speakers noted that many individuals experience multiple forms of hate simultaneously. For example:

- A queer First Nations person may face racism, homophobia and transphobia, as well as colonial stereotypes about "traditional values" and authenticity.
- A Muslim woman may experience Islamophobia and misogyny, with specific tropes about clothing, sexuality and loyalty.

This intersectionality means that:

- Policy and practice should not treat categories in isolation; risk assessment tools, support services and legal frameworks must account for overlapping vulnerabilities.
- Aggregate statistics that silo incidents by a single characteristic (e.g. religion only) can understate the real burden of harm and, importantly, can misdirect interventions as a consequence.

The discussion also noted that certain extremist ecosystems explicitly target women and LGBTQI people as enemies. This includes:

- Male-supremacist and "incel" communities that glorify violence against women and frame gender equality as a conspiracy.
- Ideological movements that portray LGBTQI identities as corrupting or "degenerate," sometimes linked to broader narratives about civilisational decline or globalist plots.

These narratives can serve as gateways into more explicitly political extremism, or they can function as stand-alone incubators of violence and harassment. Ignoring them in CT and CVE frameworks, some participants argued, leaves significant threats unaddressed and fails to protect those most frequently and severely targeted.

A recurring advocacy point was that women, LGBTQI people, disabled people and First Nations communities must be co-producers of policy, not merely consultees or data points, if programs and policies to assist are to be well-designed and effective.

This implies:

- Their systematic inclusion in advisory bodies, oversight mechanisms and programme design teams concerned with hate speech, extremism and CT.
- Recognition of their existing knowledge, networks and practices for managing risk and harm, which can inform more grounded and legitimate interventions.

Participants argued that failure to include these communities meaningfully in design and oversight will perpetuate blind spots, reinforce structural inequalities, and ultimately weaken Australia's overall resilience to extremist violence and hate.

The convenors and several speakers converged on a set of principles for more effective and legitimate policy, even if detailed recommendations were left for subsequent work. There was agreement among some speakers that:

- Legislation and policy in this domain must be evidence-based, drawing on the best available research on radicalisation, mobilisation and the impacts of different interventions, rather than on conventional assumptions or worst-case intuitions.
- Policy design must be multi-disciplinary, involving criminologists, psychologists, legal scholars, technologists, community stakeholders and frontline practitioners; otherwise, important risks and unintended consequences will be missed.
- Without such engagement, the system will continue to "miss" individuals who do pose a genuine threat while over-policing those whose beliefs are harmful but non-violent.

There needs to be more informed consistency over time, so that practitioners, communities and courts can understand and anticipate how laws will operate.

5 The changing threat environment and limits of current frameworks

5.1 Radicalisation to extremism or to violence

Several speakers remarked on the need to distinguish between radicalisation and extremist belief on the one hand and, on the other, a decision to resort to terrorism or violence. For this reason, it was argued that the core policy objective of government should not be the eradication of extremist belief systems per se but rather the prevention of mass casualty events.

One presenter stressed that radicalisation processes are multidimensional, adaptive and highly heterogeneous. Cases do not “present” in the same way and this makes simple risk categorisations unreliable. This complexity demands responses that are well-thought-out, evidence-informed and carefully tailored, rather than being driven by reactive political cycles.

5.2 Critique of criminalisation strategies, reactive law-making and over-reach

A central advocacy point from this speaker was skepticism toward broad criminalisation of phrases, symbols or terminology as a counter-terrorism or counter-extremism strategy. The argument was that:

Criminalising expression that is “hateful, awful, completely not appropriate” but not clearly linked to a genuine threat, risks diverting scarce police, intelligence and community resources onto individuals who may not pose a serious risk to public safety International experience, especially from the UK and US, shows that such laws can be used to charge or convict individuals for expression without robust evidence that they were likely to mobilise to violence, and such action carries uncertain security benefits.

The presenter argued for consistency and adaptability in Australia’s CVE frameworks, anchored in the empirical evidence on who actually mobilises to violence and under what conditions. In their view, the current Australian CVE initiatives lack validated risk assessment tools for the emerging mixed-ideology, grievance-fuelled, conspiratorial threat landscape, making reactive criminalisation particularly risky.

A key thread in the discussion was how terrorism and major incidents drive knee-jerk regulation, including of ongoing online hate. One speaker framed this in terms of terrorism’s strategic aim being to generate second- and third-order political effects – outrage, polarisation, and a rush to “do something” – more than to kill people per se. After Bondi, the speaker said, Australian politics delivered exactly what terrorists want: “the most un-edifying Australian public discourse I’ve ever seen”, with all sides in that debate screaming at each other and a rash of pre-drafted measures being bundled through Parliament under the guise of urgent response.

- Much of what was branded as a “Bondi response” – including wide-ranging legislation – was on the books or in drafting already; the incident served a political catalyst to push complex packages through without considered scrutiny.
- The new federal “hate” provisions are, in one view, “not really counterterrorism legislation at all” but efforts to manage social-cohesion concerns; tying them rhetorically to Bondi muddies their purpose and invites over-promising about what law can do.
- Poorly drafted, rapidly passed laws with vague concepts and poor communication provide ammunition to those who argue the state is overreaching and untrustworthy, thus undermining the very social trust they purport to protect.

Other contributors picked up the non-crime hate incident debate from the UK. One speaker asked how far down the spectrum of online expression police and regulators should reach, given UK experience of officers “chasing Facebook posts while bicycles are stolen”. Another used the Canberra café case (satirical anti-hate art investigated under new symbol ban laws, then dropped) to illustrate the danger of front-line police being forced to interpret intention behind symbols and the chilling effect this produces on legitimate expression and art.

The overall picture, according to some, is of a politics of symbolic reassurance: governments use spectacular crimes to push through visible regulatory measures on speech and symbols, often without strong evidence of efficacy, while creating new tensions between law enforcement and communities and new legal uncertainties online.

Australia is trying to regulate online hate in a constitutional environment without a formal right to free speech, but with an implied freedom of political communication that restrains government only at the margins. Against that backdrop, we need to be alert to “bad faith” law-making. Recent protest anti -“hate” laws might have seemed politically expedient, effectively using legislation as signalling, but leaving courts and the legal system years of clean-up.

Federally, the heavily advertised new “hate offence” was in fact not included in the final January package, but this nuance has not been communicated, generating a chilling effect because people do not actually know what is and isn’t illegal.

This communication failure is itself political: the confusion around the content and scope of the laws is “eroding social trust” and, in one view, pitting police and community against each other at the very moment trust is needed most.

Discussion of the incident around the satirical art in the Canberra cafe was used to make several points during discussion:

- Frontline officers were placed in the impossible position of assessing whether a symbol’s intent (satire vs endorsement) met the threshold of hateful display, with little guidance and high reputational risk
- The main consequence was a “chilling effect” on expression, as other artists and

small businesses now have to second-guess whether their work might be interpreted as hateful by police, regardless of artistic or critical intent.
- This illustrates what happens when well-intentioned laws are drafted and passed without sufficient consideration of how they operationalise “when the rubber hits the road.”

The broader advocacy line was that such cases show the costs of reactive law-making in the immediate emotional aftermath of high-profile incidents, where cortisol and media pressure drive political behaviour.

5.3 Evolution of the CT apparatus

A subsequent contribution examined the broader counter-terrorism (CT) system and its fit for the contemporary threat environment. This speaker contrasted today’s situation with the early 2000s: terrorism is no longer the centrally coordinated Al-Qaeda of 2001, but a vastly more complex, fragmented and dynamic field—attracting less political attention and resourcing than during the peak post-9/11 years.

Examples used to illustrate the shifting landscape included:
- State-linked organisations like the Iranian IRGC, which have been designated as terrorist organisations in Australia despite being arms of a state security apparatus and this designation challenged traditional non-state definitions
- Extreme-right organisations such as “The Base,” allegedly operating out of Russia with tacit state tolerance, supporting operations elsewhere, and complicating attribution and legal response
- Online-only ecosystems such as “Terragram” and networked communities that target children for ideological purposes but have no clear organizational structure, making them hard to fit into prescribed-organisation based legal frameworks.

The speaker argued that Australia’s CT apparatus was largely built to address jihadist terrorism and remains structurally tuned to that paradigm:
- Legal architecture leans heavily on the concept of proscribed organisations, which makes support and preparatory offences relatively straightforward where an organisation is listed, but much harder when threats arise from diffuse, networked or “leaderless” milieus
- As the threat mix has diversified (extreme right, mixed ideologies, grievance-fueled conspiracies), the system’s reliance on proscribed organisations creates friction and blind spots.

This analysis led to a call for deep reconsideration—without wholesale dismantling—of whether the current CT system is optimised for the “threat environment that it is currently facing,” and, it was argued, for parallel discussions among Five Eyes partners about similar structural misalignment.

5.4 Re-prioritisation of counter-terrorism

The workshop also discussed the blurring of objectives across counter-terrorism (CT), countering violent extremism (CVE), and prevention work directed at radicalisation and harmful ideologies. One contributor questioned whether the system is clear on "what exactly it is we're countering"—terrorism, violent extremism, extremism, or radicalisation—and noted that sometimes we are attempting to both prevent and counter simultaneously, with different tools and mandates.

Given mounting resource constraints, one expectation expressed was that:

- CT, as a narrow law-enforcement and intelligence undertaking focused on
- stopping those already committed to violence, will be re-prioritised ahead of the more diffuse, upstream CVE and P/CVE work
- CVE agendas that expanded significantly in the Obama era—e.g. via the Global CVE Summit—risk being relegated to secondary status because they are harder to defend in strict cost–benefit terms when budgets tighten.

Some participants highlighted that while CT is inherently intrusive and will always impinge on civil liberties to some extent, the key question is whether such intrusions are effective and proportionate to the threat, especially when laws are passed in "knee-jerk" fashion.

The workshop pointed again to a need for a comprehensive and meticulous system-level review of whether Australia's CT apparatus, largely designed for early-2000s jihadist threats, is properly configured for the current, more variegated landscape. While no one advocated dismantling existing structures wholesale, there was support for:

- Testing whether proscribed-organisation laws and associated institutions can adequately address diffuse and online-native threats
- Rebalancing investments between CT and CVE in a way that reflects both resource constraints and the evidence on long-term efficacy

Embedding considerations of civil liberties and social cohesion into CT decision-making, to avoid delivering strategic wins to terrorist movements by undermining democratic norms.

6 Online challenges

The regulation of online hate can be deeply political on multiple levels: how problems are defined, whose harms count, what gets criminalised, and how governments use crises to legislate quickly without adequate community consultation and deliberation. The discussion did not focus on the relationship between hate speech in online environments and related content that might have originated in a live environment as opposed to having been created initially online.

- In 2025, the federal criminal code was amended to expand the coverage of groups covered by hate speech to insert an offence of advocating force or violence against

groups distinguished by “sex, sexual orientation, gender identity, intersex status, [and] disability”. The complete list of groups in the amended legislation therefore included nationality, national or ethnic origin, political opinion, race, and religion[1].

- By contrast, the highly politicised 2026 “Combating Antisemitism, Hate Crimes and Violence” amendments focus only on race, national or ethnic origin in their new hate-related prohibition mechanism; they explicitly do not extend to gender, sexuality or disability. So, organisations advocating hate against people on other criteria apart from race, national or ethnic origin, cannot be proscribed on that basis.

One speaker observed that this exclusion had generated resentment and criticism from women’s, LGBTIQ+ and disability advocates, who see their exposure to hate and violence as well documented but still not recognised at the same legislative level as race. This speaker offered some grim statistics:

- High prevalence and lethality of violence against women, and well-established literature linking misogyny to violent extremism and mass violence.
- Substantial, documented levels of harassment, abuse and violence against LGBTQI+ people, with many reporting that hate speech impacts the entire community, not just individuals.
- Disabled people being significantly more likely to experience violence and crime, with disabled women at particularly high risk.

Advocates in these sectors, the speaker noted, often argue that the threshold of “urging violence” in current federal law is too high to protect them from the everyday hate and harassment they face online and offline. Thus, the politics of regulation is not just about overreach, but also about under-reach for certain groups.

The Workshop also included some discussion of definitional politics.

- The new federal mechanism for prohibited “hate” organisations does not define “hate”; instead, it cross-references “hate crimes” scattered through federal and state law and then (controversially) sweeps in behaviour that was not an offence under state law at the time it occurred.
- This retrospective extension is described as “wild” and “quite shocking”, and is almost certain to end up being litigated; in the meantime, it creates legal uncertainty.

[1] It appears that no convictions have yet been publicly recorded under these amended federal hate-crime provisions. If any charges have been laid, they do not appear to have resulted yet in reported final convictions.

A discussion of definitions of antisemitic content in surveys of online hate speech showed how classification choices can be both methodologically complex and politically contested. One speaker offered a taxonomy of antisemitism that has multiple primary categories and 27 subcategories, including a “Nazism inversion” category (e.g. Netanyahu depicted with Hitler’s reflection, Israeli flag with swastika elements). There was discussion of the link between moral outrage against Israel for its actions as a state and hatred against Jews. The question was posed whether such coding risks classifying legitimate criticism of Israel as antisemitism. One answer was no, since the position taken in the methodology and the published research was their working definition explicitly distinguishes between criticism “similar to that levelled against any other country” (not antisemitic) and using Israel as a proxy for “the Jews” (antisemitic). The speaker continued that some claims that “you can’t criticise Israel” is a narrative that is weaponised by people who want “free scope” to engage in antisemitism while claiming racism is just politics.

This exchange illustrates how regulating online hate inevitably intersects with contested politics around Palestine/Israel, Islamophobia and antisemitism, and how the same ` definitional line (e.g. when does critique become hate?) is fought over in both activism and law.

6.1 Platforms, profit and global information politics

Some participants were highly skeptical of platform governance and referenced the political economy and business models for online activity as the main obstacles to a more effective regulatory environment. The challenge was summed up by the statement: “Allowing more hate is more profitable.”

- One contributor argues that the United States is now a leading global source of disinformation, a situation aggravated or facilitated by the concentration of platform ownership and related technologies in the United States. Some of these platforms have undertaken political censorship of user content.
- There is concern that Australian elites and media (e.g. Sky News After Dark) have imported US-style culture-war narratives, including anti-immigrant sentiment and Great Replacement rhetoric, and that this is then amplified online in Australia through YouTube and other platforms.

This makes regulation of online hate inherently transnational and geopolitical: laws passed in Canberra or Sydney operate in an ecosystem dominated by US-based platforms and content flows, and foreign state actors seeking to inflame grievances and distrust.

One of the more challenging aspects of analysing the power of online platforms mentioned in one presentation was that of distinguishing normalisation vs representativeness, that is how to read the political significance of online hate.

On one hand, it was argued that there has been a normalisation of antisemitic content online at a far higher level than for other hates. On the other hand, information was offered as a counterpoint that in one study, only 0.2% of tweeters were classed as

prolific, suggesting that a tiny cohort generates a large share of the visible content. This was interpreted to mean that platform data cannot be read as a simple proxy for majority public opinion. Another speaker noted that social media captures who is loud and unrestrained, not necessarily the distribution of attitudes in the population, and that policy needs to distinguish between fringe activists and the broader public.

At the same time, even those skeptical of representativeness arguments recognise the political impact of normalisation. Two people noted that repeated elite cues (from people like Trump, Boris Johnson, Pauline Hanson and Fraser Anning) are followed by spikes in hate crimes or hate speech, suggesting that politician rhetoric and media frames help "license" online and offline hate. Leadership matters, negatively as well as positively. Another speaker tied this argument to the extreme right's deliberate strategy to shift the Overton window[21] via online culture, memes and "borderline" content, such that talking points once confined to neo-Nazi circles become mainstream commentary about immigration or "cultural Marxism".

The politics of regulation here is double-edged: ignoring hateful online currents allows them to shape discourse and risk; over-reading them as "the public" risks securitising politics around a noisy minority.

7 Civil society norms and solidarity

The transcript also contains proposals to re-orient the politics of regulating online hate. Several speakers argued for:

- Focusing on behaviour rather than ideas: policing hate incidents and organised hateful extremism (including online, via better incident data and tracking of organised networks) while avoiding outright thought-policing.
- Building robust, long-term data systems on hate incidents, including online hate, as a shared national resource rather than ad hoc, fragmented exercises.
- Investing in digital literacy and civic education from an early age so citizens can critically assess online content, rather than relying primarily on takedowns and bans.

Two speakers added that we must:

- Be transparent and proportionate in takedown and platform-regulation efforts, or we risk fueling conspiracy narratives about censorship and anti-government distrust – the very mal-information dynamics foreign actors seek to exploit.
- Reframe the normative project away from a thin "social cohesion" doctrine and toward peaceful pluralism and social solidarity, explicitly acknowledging ongoing conflict and difference while protecting those most targeted by online and offline hate.

The discussion included a call for a funded "communities collaborative" – a permanent,

[2] The Overton Window is the range of policies or ideas acceptable to the mainstream public at any given time, often termed the "window of discourse".

cross-community structure that would work both against hate within communities and ease tensions where communities target each other, including in digital spaces, as part of a broader nation-building and resilience project.

In this reframing, regulating online hate is not just a matter of tweaking offences or forcing platforms to take posts down. It becomes a sustained and public contest over what kind of political community Australia wants to be, which harms and histories it is willing to name, and how it balances free political communication with the safety and dignity of those most exposed to digital hostility. This is a challenge to be grounded in civil discourse, embraced under the vision of peaceful pluralism and not wished-away under the pretext of social cohesion.

LIST OF ENROLLED PARTICIPANTS

Greg Austin, University Technology Sydney and SCI

Australian Federal Police (5 persons)

Greg Barton, Deakin University

Emma Belton, University of Queensland

Deaprtment of Home Affairs (6 persons)

Mark Duckworth, Deakin University

Justin Ellis, University of Newcastle

Asher Flynn, Monash University

Gerard Gill, Curtin University

Andrew Goldsmith, Flinders University

Patrick Scolyer Gray, National Instiute of Strategic Resilience

Shandon Harris-Hogan, Victoria University

Michael Humphrey, University of Technology Sydney

Danielle Ireland-Piper, Australian National University

Andrew Jakubowicz, University of Technology Sydney

Noelle Martin, University of Western Australia

Lisa Materano, Social Cyber Institute

Katina Michael, University of Sydney and SCI

Andre Oboler, On-Line Hate Prevention Institute and La Trobe

University Mathieu O'Neil, University of Canberra

Mario Peucker, Victoria University

Dominique dalla Pozza, Australian National University

Imogen Richards, Deakin University

Rachel Sharples, Western Sydney University

Anjalee de Silva, University of Melbourne

Simone-Maree Spano, Australian National University

Chris Taylor, Australian Strategic Policy Institute

Pascale Taplin, Australian National University

Amanda Third, Western Sydney University

Akshit Tyagi, Australian National University

Levi West, Australian National University

Glenn Withers, Australian National University, SCI and UNSW Canberra

Andrew Zammit, Victoria University