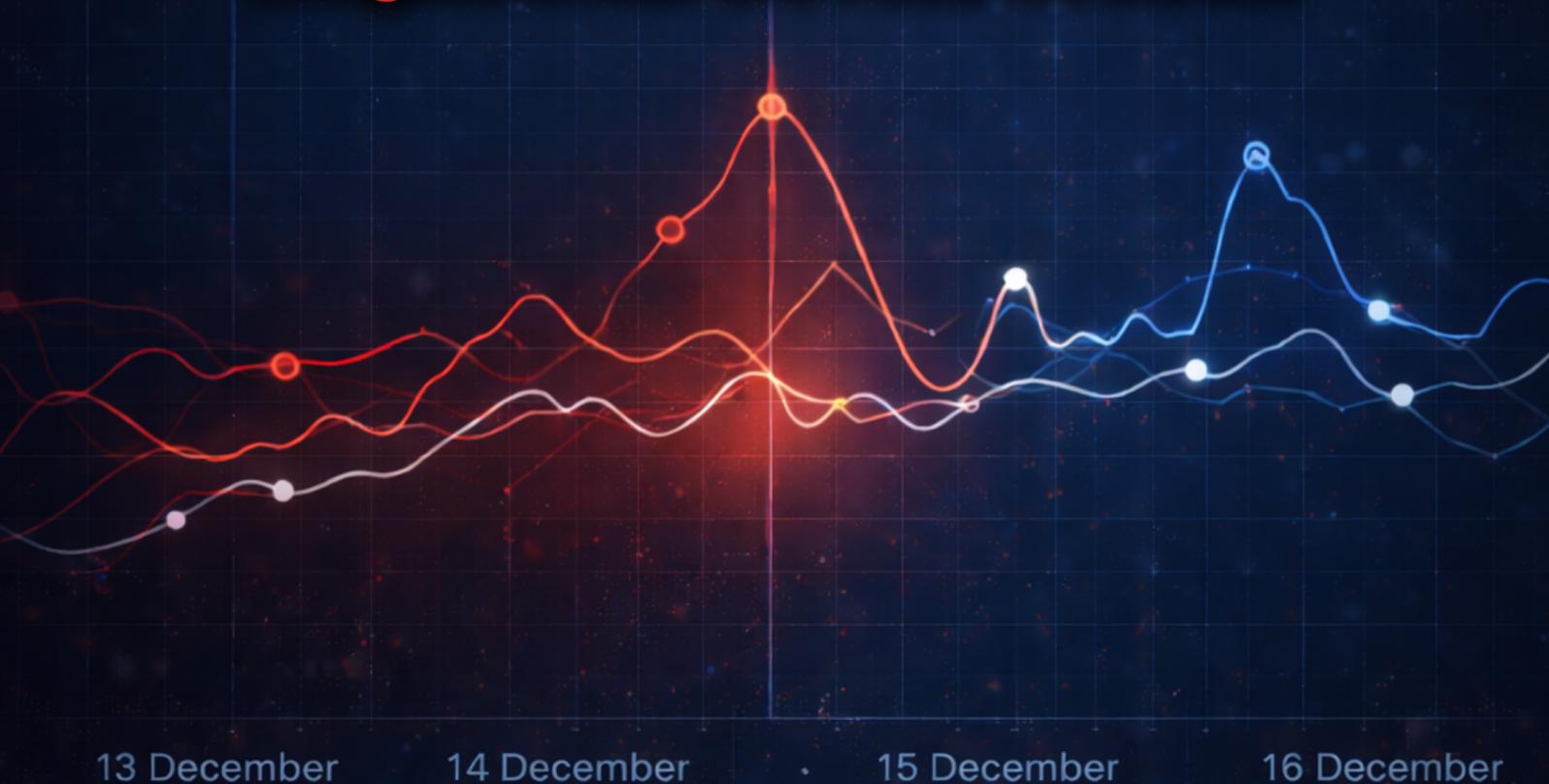


# Online Hate in the Wake of the Bondi Terror Attack:

Insights from the First 48 Hours



*Matteo Vergani, Stephanie Zi Xin Ng, Andrea Giovannetti, Haily Tran, Huu Phuc Hong, Joel Moss, Khondaker Tasrif Noor, Kewen Liao, Dan Goodhardt*

18 December 2025

Tackling Hate Lab

# Key Takeaways

- ▶ 1 in 10 posts on major platforms (X, Reddit, Instagram) in the aftermath of the attack contained toxic or hostile language, with some platforms like Reddit showing toxicity rates as high as 13%.
- ▶ Anti-Muslim hate dominated online discourse, particularly on X and Reddit. On X, anti-Muslim posts made up 18.4% of all Bondi-related content by 16 December. This is nearly 1 in 5 posts. Reddit saw a spike to 3,091 anti-Muslim posts on 15 December, up from 530 the day before.
- ▶ Antisemitic content also surged, particularly “new antisemitism” (for example, blaming all Jews for the actions of Israel). On X, new antisemitism posts rose more than tenfold, from 329 on 14 Dec to 3,855 on 15 Dec. On Reddit, similar trends saw new antisemitism jump from 279 to 1,973 over the same period.
- ▶ Conspiracy theories, especially false flag claims, played a key role in fuelling antisemitic and anti-Muslim hate, sustaining harmful narratives across platforms in the days following the attack.
- ▶ Platform response was inadequate: despite the prevalence of hate and misinformation, most harmful content remained online and was not removed by platforms.



# Introduction

At least 16 people – including a ten-year-old child – are dead after two men opened fire on a crowd of people celebrating the Jewish holiday of Hanukkah on Sunday 14 December in a public park at Sydney’s Bondi Beach. Many more are injured.

This report provides a topline overview of online reactions to the Bondi terrorist attack across four social media platforms: X (Twitter), Reddit,

Telegram, and Instagram. It focuses on mapping three key categories of harmful content: anti-Jewish hate, anti-Muslim hate, and hateful conspiracy theories circulating in the immediate aftermath of the attack.



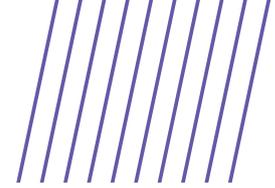
# Data and methods

In total, we analysed 71,441 posts from X (geolocated in Australia), 59,638 from Reddit, 8,138 from Telegram, and 5,684 Instagram posts and comments related to the Bondi terror attack. While our analysis of Reddit, X, and Instagram focused on capturing mainstream discourse, our approach to Telegram targeted 36 Australian extreme-right channels and 10 extreme-right users to monitor sentiment within this specific ideological community.

Data collection spanned 9 to 16 December, allowing us to capture both the first 48 hours following the attack and the preceding six days. For X, Instagram, and Reddit, data was collected using keyword-based queries designed to capture discussion of the Bondi attack (e.g. #bondiat-tacks). For Telegram, we collected all content posted in the selected channels during this period, enabling pre- and post-attack comparisons.

To identify hate, we employed a range of classification approaches. We used the Toxicity measure from the Perspective API, which flags language likely to disrupt conversations, defined as “rude, disrespectful, or unreasonable comment[s] that are likely to make people leave a discussion.” To capture antisemitic and anti-Muslim hate, we applied two RoBERTa-based machine learning classifiers developed in collaboration with Jewish and Muslim communities. These models were trained on datasets annotated by community-based experts, ensuring cultural and contextual sensitivity. Following classification, we conducted a series of statistical analyses to examine patterns across platforms and over time. We also carried out a qualitative thematic analysis to contextualise the findings and develop a typology of the dominant narratives, which is presented in this report.

**NOTE:** The figures we present in this report do not represent the total volume of hateful content on each platform, but only the posts captured through our specific data collection queries.



# Toxicity

On X, 755 posts out of 9,743 (approximately 7.8%) were flagged as toxic on 14 December. This rose to 2,757 out of 33,926 (about 8.1%) on 15 December, before slightly decreasing to 2,417 out of 26,968 (around 9.0%) on 16 December. These figures indicate that the proportion of hostile or disruptive language increased slightly over the 48-hour period, rising by approximately 1.2 percentage points. This suggests that not only did toxic content surge in volume, but it also became more concentrated within the broader conversation.

On Reddit, toxic posts accounted for 1,126 of 8,540 (approximately 13%) on the day of the attack. This rose to 4,575 of 38,967 (around 12%) on the following day, and then dropped to 1,178 of 11,929 (about 10%) on 16 December. While the overall volume of posts declined significantly from nearly 39,000 to 12,000 between the second and third day, the proportion of toxic content remained rela-

tively stable with only a 3% decline between the 14th and the 16th of December.

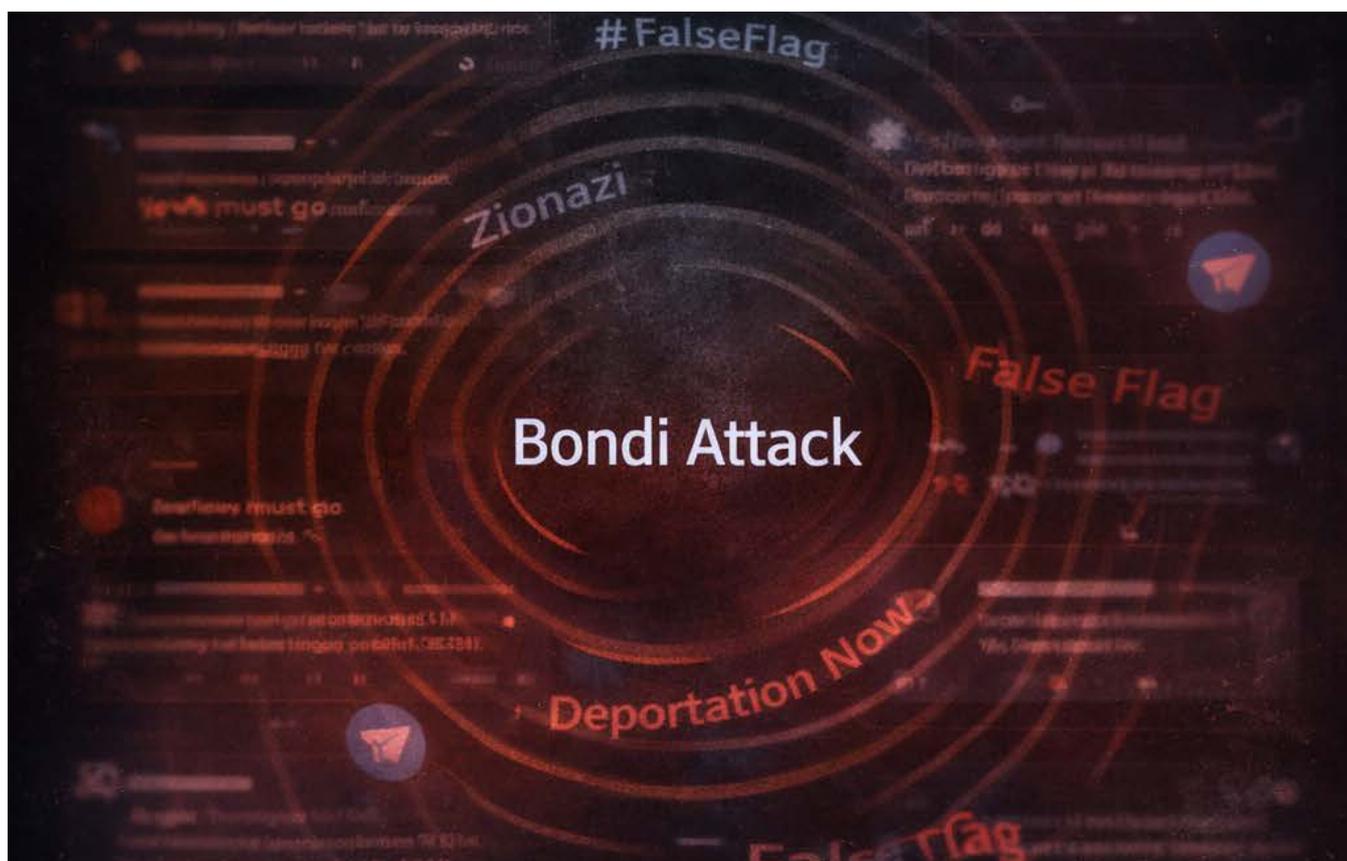
On Instagram, toxic posts made up 147 of 1,183 (approximately 12.4%) on the day of the attack, 115 of 1,503 (about 7.7%) on the following day, and 184 of 2,642 (around 7.0%) on 16 December. This suggests that although discussion volume increased over time, the proportion of toxic content declined, dropping from 12.4% on the first day to 7% on the third day. This may reflect an initial spike in emotional or hostile reactions, followed by a partial return to more neutral or moderated discourse.

Overall, the data shows that while the volume of conversations about the Bondi attacks declined between 15 and 16 December on mainstream platforms (X, Instagram, and Reddit), this suggests a drop in public attention. However, the average level of toxicity remained stable, with approximately 1 in 10 posts contain-

ing hostile or disruptive language. Differences in absolute levels across platforms partly reflect the volume of posts retrieved on each platform, rather than direct differences in user behaviour alone.

On Telegram, we observed different dynamics. Pre-attack toxicity was already relatively high, with an average of 66 toxic posts out of 727 per day (approximately 9.1%). Post-attack, the absolute number of toxic posts rose to 113 out of 4,150 daily posts, but the proportion of toxic content dropped to around

2.7%. This suggests that the post-attack increase in volume was largely driven by the reposting of media articles and related content, rather than a rise in user activity and toxic language. However, from a qualitative perspective, the toxicity observed in these extreme-right Telegram channels was often more violent, explicit, and ideologically charged than what appeared on mainstream platforms.



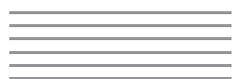
# Antisemitism

In our analysis, we distinguished between two main types of antisemitism: old and new. Old antisemitism targets Jews as Jews, drawing on long-standing myths and stereotypes that portray Jewish people as inherently dangerous, corrupt, or conspiratorial. Examples include explicit calls to “kill all Jews” or claims that “Jews are paying to destroy Australia.” It also includes Holocaust denial or distortion, such as suggesting the Holocaust was fabricated or exaggerated.

New antisemitism, by contrast, shifts the focus from individual Jews to the state of Israel, often blaming Jewish people collectively for the actions of the Israeli government. This includes slurs such as calling Australian Jews “baby killers” or “Zionazi f—wits” during discussions about Israeli government actions, where Jewish individuals are targeted simply for being Jewish, regardless of their personal views or involvement.

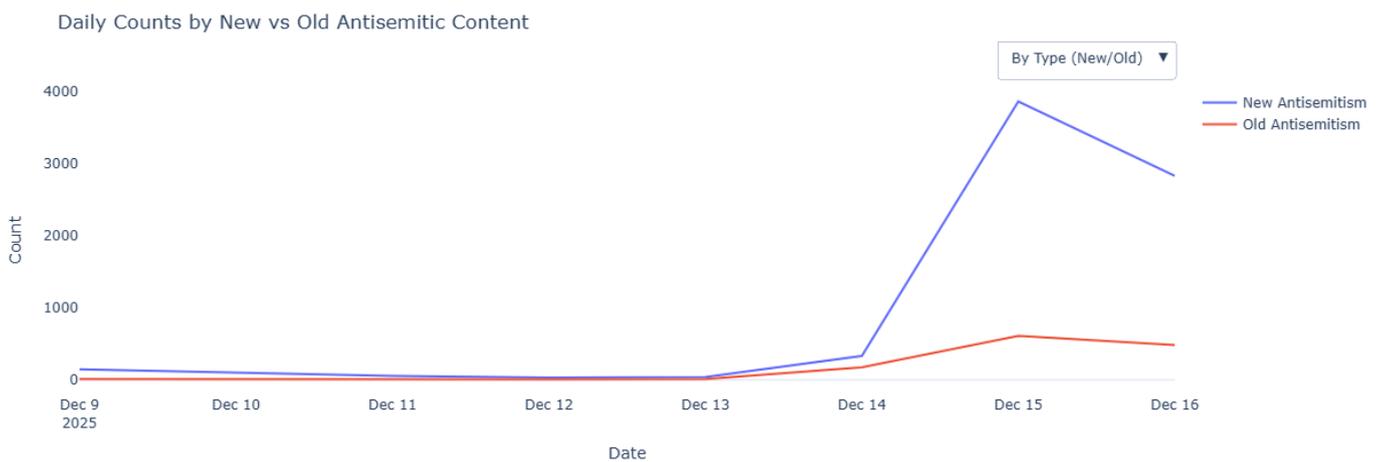
While this category remains contested, with some viewing it as legitimate political critique of Israel and others recognising it as a continuation of longstanding antisemitic tropes, it comprised the majority of antisemitic content in our dataset. Importantly, in the case of the Bondi attack, where Jewish civilians were targeted during a religious celebration with no connection to Israeli state actions, much of the antisemitic commentary portrayed the violence as retaliation for Israel. This reflects a broader pattern in which Jewish identity is conflated with the Israeli state, allowing antisemitic narratives to frame Jews globally as collectively responsible for the actions of a government they may not support or be connected to.

In the X data (Figure 1), old antisemitism increased from 172 posts on 14 December to 606 on 15 December, before decreasing to 483 on 16 December. New antisemitism showed



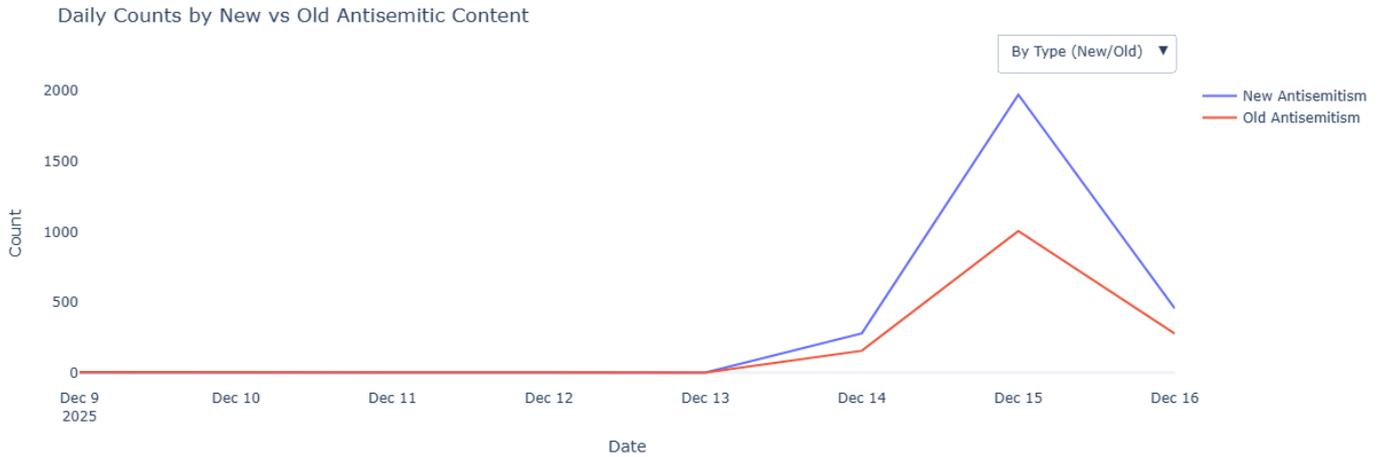
a sharper rise, jumping from 329 posts on 14 December to 3,855 on 15 December, then falling to 2,825 on 16 December. This pattern indicates a sharp surge in antisemitic content, particularly new antisemitism, in the immediate aftermath of the Bondi attack, followed by a partial decline. It suggests that antisemitic narratives, especially those linking Jewish people to the actions of the Israeli state, gained rapid traction online in the days following the attack.

**Figure 1. Daily volume of old and new antisemitic content on X in Australian-based conversations about the Bondi terror attack.**



In our Reddit dataset (Figure 2), new antisemitism followed a sharp escalation–decline pattern, rising from 279 posts on 14 December (noting that the attack occurred in the late afternoon) to 1,973 on the 15th, before falling to 456 on the 16th. Similarly, old antisemitism increased from 155 posts on the 14th to 1,005 on the 15th, then declined to 277 on the 16th. This indicates a clear downward trend in both old and new antisemitic narratives within Reddit discussions in the days immediately following the Bondi attack.

**Figure 2. Daily volume of old and new antisemitic content on Reddit in conversations about the Bondi terror attack.**



In our Instagram data (Figure 3), new antisemitism shows a steady increase across the three days, from 70 to 84 to 121. This pattern suggests progressive narrative building. As the event was discussed and reframed, new antisemitism interpretations appear to have gained traction. Old antisemitism remains more stable, moving from 96 to 67 to 79. This points to the reactivation of existing tropes, with fluctuations likely driven by visibility rather than coordinated amplification.

**Figure 3. Daily volume of old and new antisemitic content on Instagram in conversations about the Bondi terror attack.**

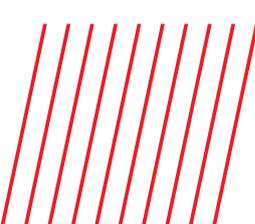


In the 36 extreme-right Telegram channels we analysed (Figure 4), antisemitic content was already present in notable volumes prior to the attack, particularly old antisemitism. Old antisemitism increased from 71 on 9 December to 160 on 16 December, showing a steady rise throughout the week. New antisemitism also grew consistently, reaching a peak of 98 on 15 December before falling slightly to 51 on 16 December.

**Figure 4. Daily volume of old and new antisemitic content in 36 Telegram channels associated with the Australian extreme right.**



Importantly, the hashtags most commonly associated with both old and new forms of antisemitism across the four platforms referenced conspiracy theories. Examples include #satanicworship, #nwo\_puppet, #mason-icsymbols\_signs, #freemason, and #falseflag.



# Anti-Muslim hate

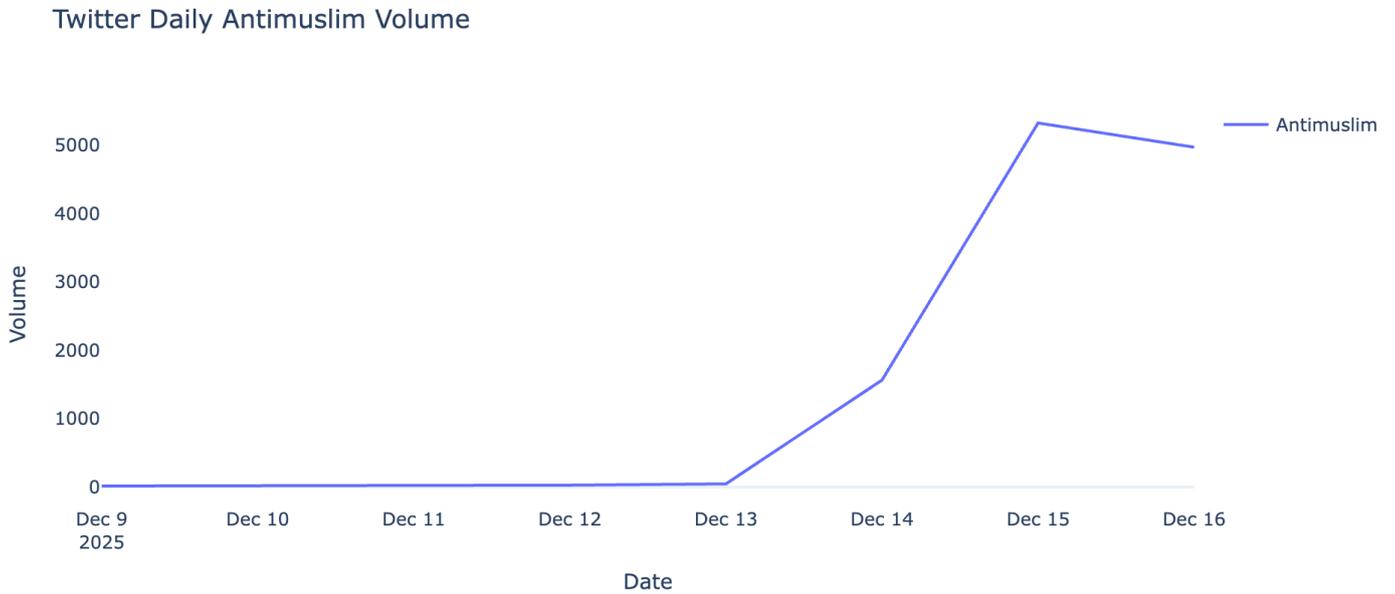
In the aftermath of terrorist attacks carried out by individuals linked to groups such as Al-Qaeda or ISIS, sharp increases in anti-Muslim sentiment and behaviour are well-documented. This pattern has been observed following events like the 9/11 attacks in the United States, the 7/7 bombings in London, and numerous others. In this case, the attackers reportedly displayed the ISIS flag, and ongoing investigations are exploring their connections to extremist networks in Australia and abroad. Despite the complex nature of these affiliations, the association between the attackers and Islam has been widely emphasised in media coverage. As a result, we observed a rise in hateful content targeting the broader Muslim community, including posts blaming all

Muslims for the attack, as well as calls for violence and mass deportations.

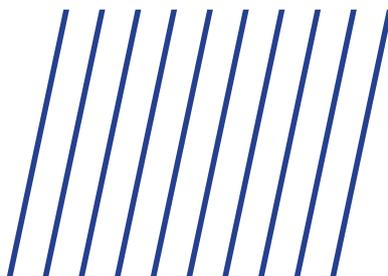
In our X data (Figure 5), anti-Muslim posts rose sharply from 1,564 on 14 December to 5,325 on 15 December, and slightly declined to 4,971 on 16 December. This represents the highest volume of hate content in absolute terms across all categories. The increase is particularly striking when viewed proportionally: anti-Muslim posts made up 16.1% of all content on 14 December (1,564 of 9,743 posts), 15.7% on 15 December (5,325 of 33,926 posts), and 18.4% on 16 December (4,971 of 26,968 posts). These figures highlight the scale and persistence of anti-Muslim sentiment in the days following the Bondi attack on X.



Figure 5. Daily volume of anti-Muslim content on X in Australian-based conversations about the Bondi terror attack.



In our Reddit dataset (Figure 6), anti-Muslim hate rose sharply from 530 posts on 14 December to 3,091 on 15 December, before dropping to 583 on 16 December. This reflects a major surge in activity the day after the attack, likely driven by highly active threads, cross-posting, and collective framing within specific subreddits. The subsequent decline on 16 December suggests a tapering of both engagement and hate-related content as the initial wave of reaction subsided.



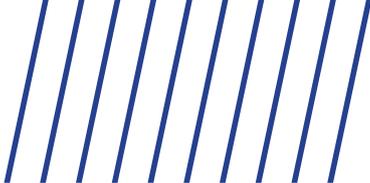
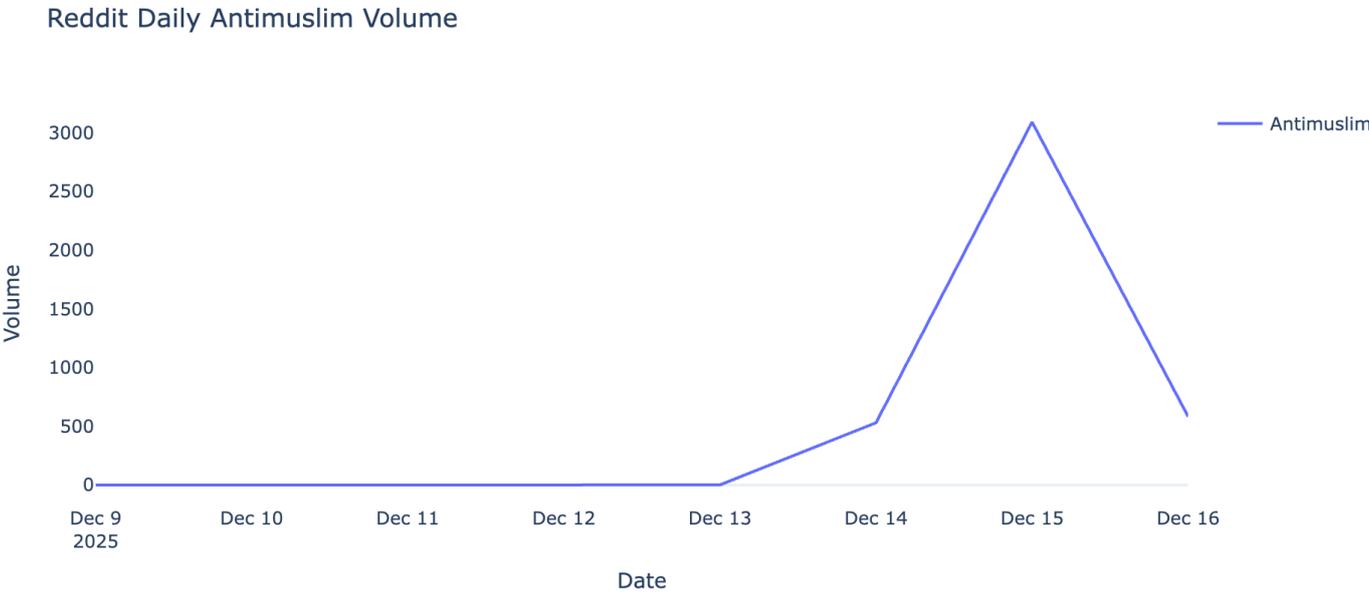
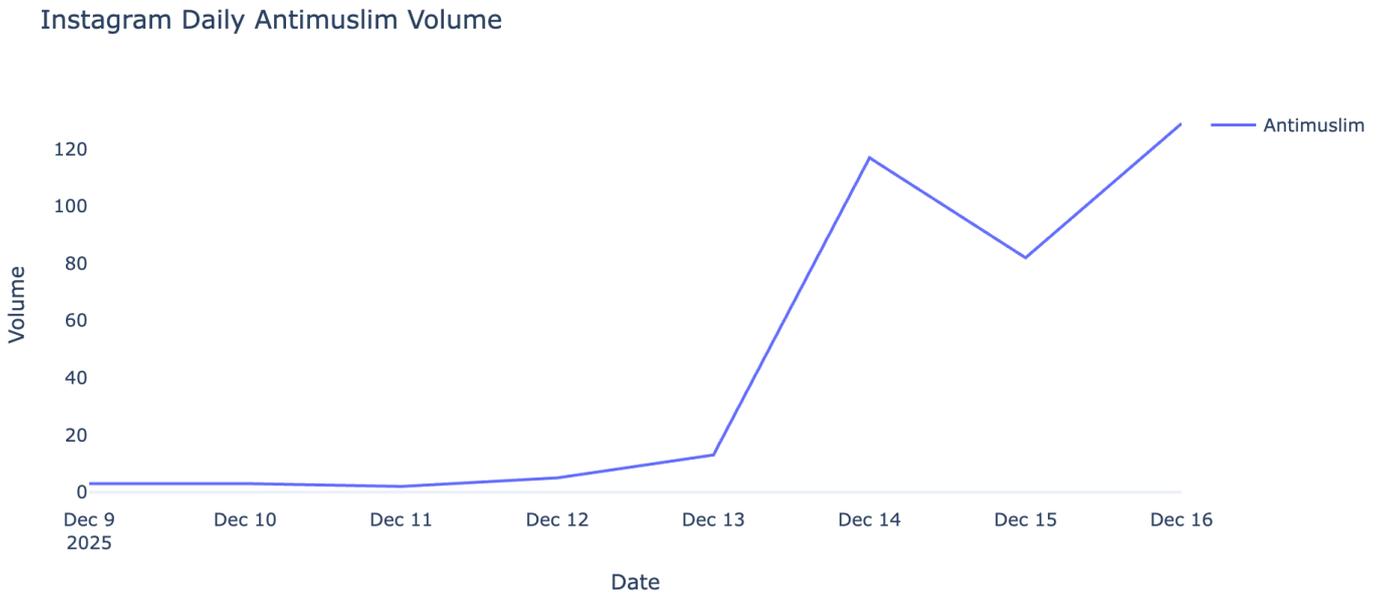


Figure 6. Daily volume of anti-Muslim content on Reddit in conversations about the Bondi terror attack.



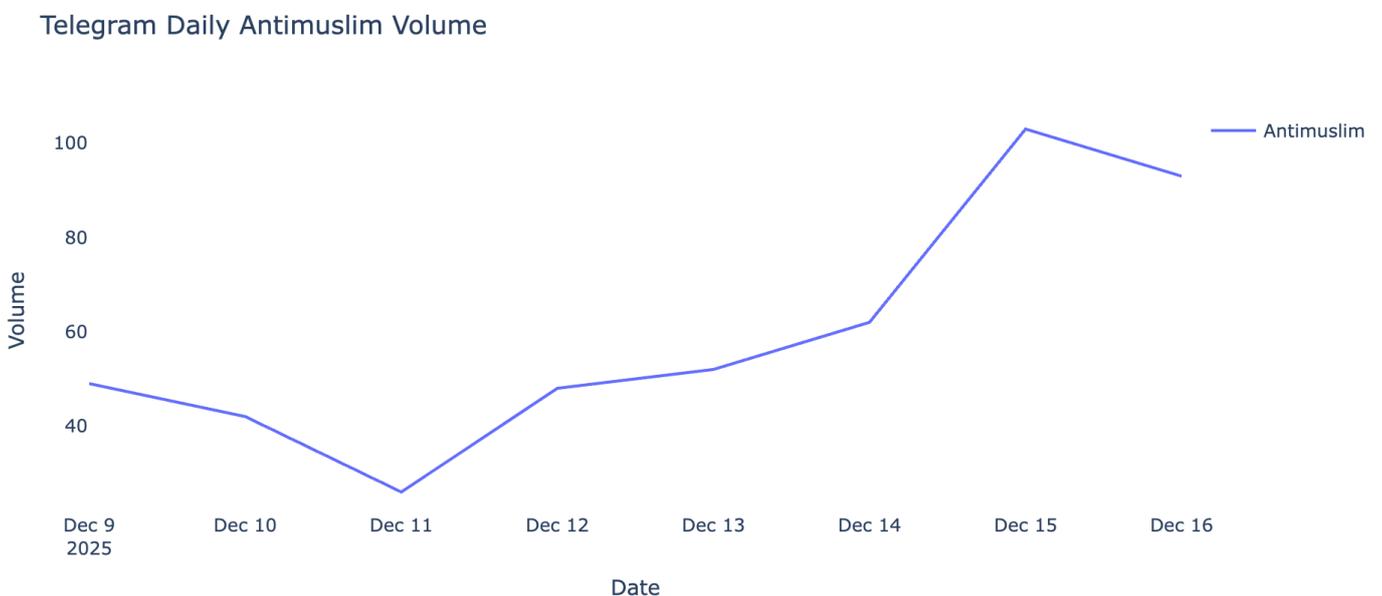
In our Instagram data (Figure 7), anti-Muslim hate peaked at 117 posts on 14 December, dipped to 82 on 15 December, and then rose again to 129 on 16 December. This pattern suggests an immediate reaction on the day of the attack, followed by a brief decline, and then a renewed surge in hateful content as discussions evolved and new content circulated.

Figure 7. Daily volume of anti-Muslim content on Instagram in conversations about the Bondi terror attack.



In the 36 extreme-right Telegram channels we analysed (Figure 8), anti-Muslim hate remained relatively stable, averaging 43 pre-attack and 43 post-attack, indicating consistent levels across the period.

Figure 8. Daily volume of anti-Muslim content in 36 Telegram channels associated with the Australian extreme right.

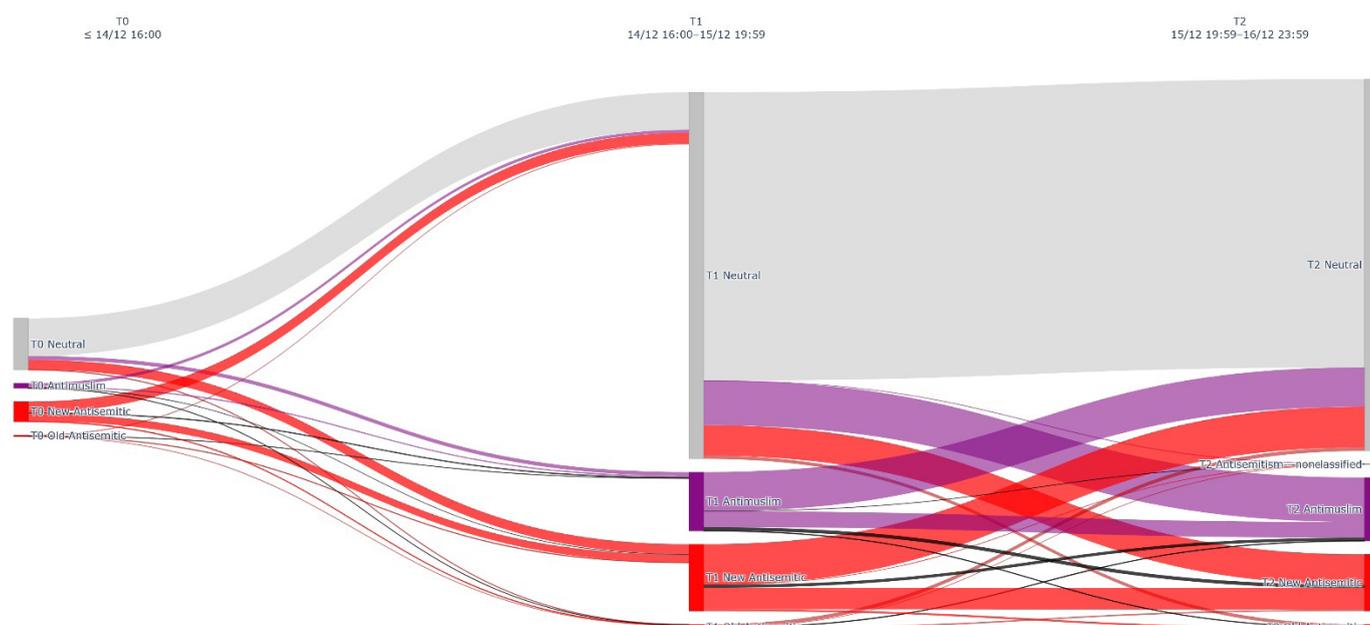




# Hate flows

Figure 9 illustrates the flows of hateful narratives across four platforms: Reddit, X, Instagram and Telegram. It tracks the emergence and evolution of anti-Muslim hate, new antisemitism, and old antisemitism from the day of the Bondi terror attack through to midnight on 16 December. While some data is available for the days preceding the attack, it originates primarily from Telegram and is of limited relevance, as this report focuses specifically on conversations that emerged in response to the Bondi attack. For details on the data collection approach, see the Data and Methods section.

**Figure 9. Flow chart illustrating user transitions into and out of anti-Muslim, new antisemitic, and old antisemitic narratives across four platforms following the Bondi terror attack.**



As expected, most anti-Muslim content following the attack came from new participants entering the conversation on the day of the Bondi at-

tack. A small portion of users in our dataset had expressed anti-Muslim sentiment prior to the attack. Notably, some individuals who had previously posted antisemitic content (both old and new) also began posting anti-Muslim content after the attack, pointing to an overlapping stream of hate. A comparable pattern was observed with old antisemitism, which also attracted users who had previously posted anti-Muslim content. Qualitative analysis suggests this was largely driven by far-right narratives portraying both Muslims and Jews as threats, and calling for mass deportations in the name of protecting “White Australians”. These narratives were already present in white supremacist circles prior to the attack and were significantly amplified in its aftermath.

The figure shows that there was a notable shift in the users expressing hate between the first 24 hours and the following 24 hours after the attack. The majority of users who posted anti-Muslim, new antisemitic, or old antisemitic content in the

first 24 hours subsequently shifted to neutral (non-hateful) language in the second 24-hour period. However, a substantial segment of users who had initially expressed neutral views in the first 24 hours went on to express anti-Muslim and antisemitic hate in the 24–48 hour window following the attack. Notably, the majority of anti-Muslim content posted in the 48 hours after the attack came from individuals who had not expressed hate in the immediate aftermath but began doing so a day later. This pattern suggests that while some users expressed a strong immediate emotional response that quickly de-escalated, others were drawn into hateful discourse more gradually, potentially influenced by the evolving online narratives and media framing in the aftermath. It highlights how hate can both spike instantly and build over time, engaging different populations at different stages of the post-crisis news cycle.

---

# Hate networks

To understand online influence, we deploy a network visualisation of 33,249 X conversations on the topic, involving 4,407 accounts. In Figure 10, each account is represented as a circle, with an arrow pointing from account A to account B if account B reacted to content posted by account A. Circle size is proportional to the amount of influence exerted on other nodes. An algorithm organises nodes geometrically by influence, positioning more influential accounts towards the centre and less influential ones at the periphery. Notably, the account @AlboMP appears at the centre of the network.

Links and nodes are coloured according to the tone of the shared content. Pink links and white nodes indicate neutral or positive tone, while dark red and pale orange indicate antisemitic and Islamophobic content, respectively. While the majority of content is flagged as neutral or supportive in tone, both hemispheres contain aversive peripheries. These are relatively symmetrical, with 1,168 accounts engaging in at least three actions (posting, replying, or sharing) involving antisemitic content, and 1,174 accounts engaging similarly with anti-Muslim content.



Figure 10. Network visualisation of 33,249 X conversations related to the Bondi terror attack involving 4,407 accounts.

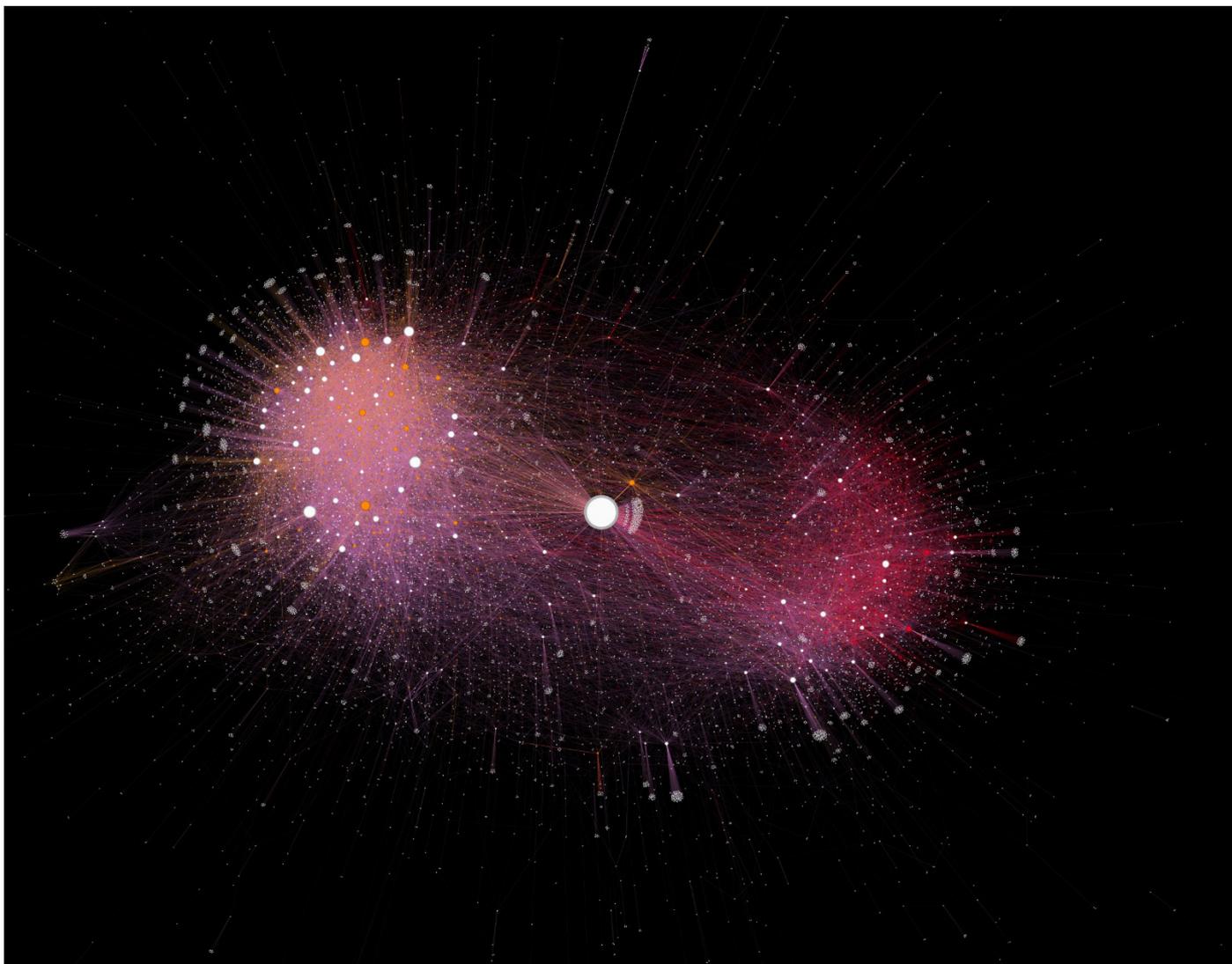
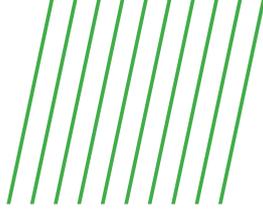


Figure 10 reveals two key patterns:

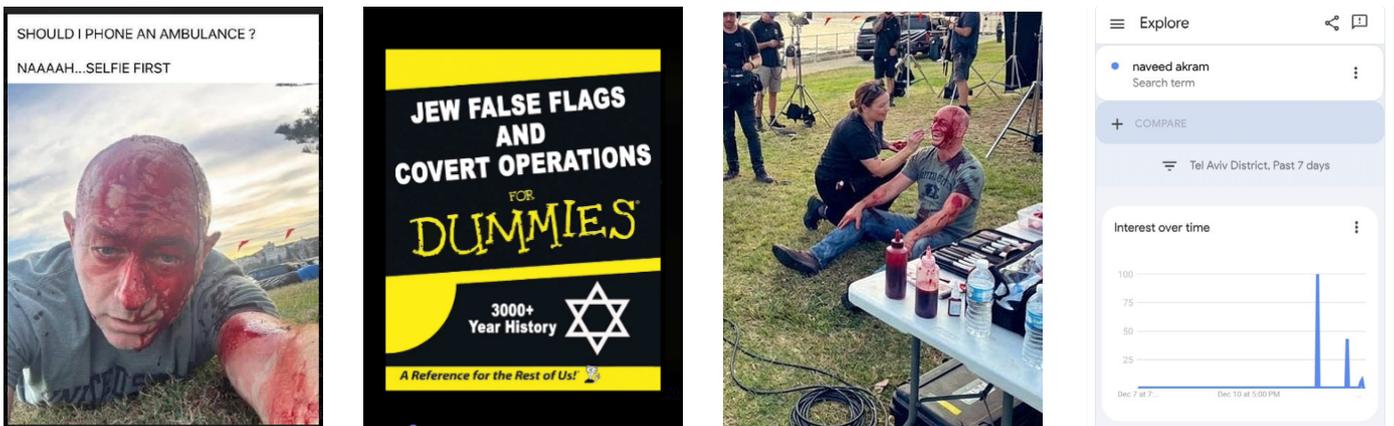
1. Conversations and activity are concentrated in two distinct hemispheres of accounts, one containing anti-Muslim and the other antisemitic content, with little or no direct interaction between them.
2. Accounts in both hemispheres react to institutional accounts, such as the Prime Minister's (@AlboMP), who serve as bridges between these otherwise separate communities.



# Hateful conspiracies

We identified two broad types of hateful narratives about the Bondi attack. The first, which we call “false flag” narrative and has a strong antisemitic element, suggests the attack was pre-planned and that someone else, such as the Australian or Israeli government, had prior knowledge. A widespread anti-Jewish conspiracy claims that the attack was self-inflicted or staged by Israel (Mossad) to regain sympathy or justify actions in Gaza. Language frequently refers to Zionism, victimhood, and global manipulation. This narrative generates high engagement and strongly polarised responses.

Figure 11. Screenshots of images featured in conspiracy posts promoting false flag narratives.



A related narrative alleges that the Bondi attacker, Naveed Akram, was searched for or tracked from Israel days before the attack. This narrative is widely circulating across the three platforms we examined: X, Instagram, and Telegram. One example includes images shared to reinforce this claim, showing Google search terms in the Tel Aviv area and suggesting that around 100 people searched for Naveed Akram. These claims have largely been debunked, often due to time zone errors

or misinterpretation of metadata. Debunking posts do circulate, but they do not always spread as widely as the original claims.

The second hateful narrative, which we call “blame attribution” narrative, focuses on attributing the blame of the attacks to a range of individual and group targets. This narrative is associated with strong anti-Muslim, anti-immigrant and anti-Jewish themes. Within this narrative, we can distinguish two main themes: one blaming individuals and falsely identifying people falsely accused of involvement in the attack. Some posts amount to doxxing or place individuals at risk of harm by falsely accusing them of involvement. The second theme is group based, and attributes blame to Muslims, Jews, immigrants or other groups. Public responses vary, ranging from harassment and threats to expressions of sympathy once misidentification is challenged. This narrative is also often associated with claims that the Australian government and police were negligent or complicit. Posts often argue that

authorities ignored warnings, particularly from Muslim communities. Some content combines institutional blame with hostility towards pro Palestine protesters or Muslim communities, while other posts include anti-Jewish narratives by questioning victims’ credibility, motives, or behaviour during the attack.

In Table 1, we provide a list of conspiratorial themes (clustered within the three main types discussed in this section) with links to examples from a range of social media platforms.



This left wing account is trying to get me murdered by falsely claiming that I was involved in the Bondi Massacre

@elonmusk can you please get this taken down I am extremely fearful for my safety. Left wingers are trying to get me killed



9:26 AM · Dec 15, 2025 · 740.6K Views

**Table 1. Conspiratorial hateful narratives circulating in the 48 hours following the Bondi attacks.**

Type of conspiracy	Theme	Examples	Social media post text (de-identified)
False flag	Retrospective sense making	Bondi Beach shooter Naveed Akram was googled in Tel Aviv on December 13	"Okay, FBI. Please explain why alleged Bondi Beach shooter Naveed Akram was googled in Tel Aviv on December 13, 1 day before the shooting. I'll wait."
	Israeli state involvement	The Bondi Beach shooting is a false flag to enforce strict law against antisemitism	"The Bondi Beach shooting is a false flag to enforce strict law against antisemitism"
	Shooters were IDF (or Western) trained	Shooter served in the IDF	"Australian Naveed Akram Served in the Israel Defence Forces IDF Gaza"
	Deep state theories	Mass shootings are Jesuit/Masonic rituals	"These staged events are designed by the Empire to divide the slaves of the world to fuel their Ordo-Ab-Chao model."
Blame attribution	Personalised blame	False accusations of being involved in the Bondi attacks	"Bondi Shooter Was Radicalized by Racist Australian Influencer Drew Pavlou"
	Group blame: anti-Muslim and anti-Jewish	True victims are 'White Australians'	"The true victims of the Bondi attack are White Australians"
	Group blame: anti-Muslim	Muslim immigration is blamed for the terrorist attack	"Australian streets were flooded with radical Muslims a few months. Today they opened opening fire on them"
	Group blame: anti-Jewish	Victims were Israeli government linked or IDF	"Arsen Ostrovsky, one of the Zionists who spent the past two years justifying Israel's genocide and supporting Israel's bombing hospitals in Gaza, was injured in today's attack at Bondi beach, Australia."

# Conclusions

In the 48 hours following the Bondi terror attack, online hate and toxic language were persistent across major platforms, with approximately 1 in 10 posts containing offensive or disrespectful content. Anti-Muslim sentiment was a dominant form of this toxicity, particularly on X and Reddit, reflecting patterns observed after previous terror attacks. Antisemitic content also featured prominently, with new antisemitism (such as blaming all Jews for the actions of the Israeli state) being more common than traditional forms. This surge in hate was fuelled by both immediate emotional reactions and coordinated far-right narratives targeting Muslims and Jews.

While some users disengaged from hateful content within 24 hours, others joined in later, suggesting that online hate can gain momentum over time. A particularly striking finding is that anti-Muslim content on X reached the highest volume across all categories, with nearly 1 in 5 posts about the Bondi attack (18.4%) on 16 December containing anti-Muslim sentiment. Conspiracies falsely accusing Jews of staging the attack and framing Muslims as inherently violent also circulated widely, especially in extreme-right Telegram channels.

Importantly, most of this harmful content remained online and was not removed by platforms. These findings highlight the urgent need for timely content moderation, targeted counter-narratives, and community engagement to prevent the spread and normalisation of hate during crises.

