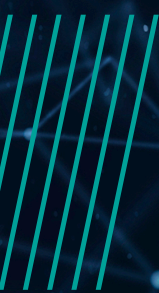




Unveiling the Complex Structure of Influence Flows in the anti-Jewish and Anti-Muslim Digital Ecosystems

Tackling Hate Lab
March 2026



Andrea Giovannetti,
Matteo Vergani



This collection of network visualisations documents the structure and evolution of anti-Jewish and anti-Muslim online hate communities in Australia, spanning the period from October 2023 to April 2025. The network is an influence network: a directed edge from node A to node B indicates that user B reacted to a post by user A with hateful content.

The underlying data, drawn from X/Twitter and Tumblr, has been filtered to retain only interactions classified as anti-Jewish or anti-Muslim by bespoke hate-speech classifiers developed in collaboration with Jewish and Muslim community members. Altogether, the network comprises 53,496 anonymised user accounts connected by 356,539 hateful influence links.

The data were retrieved from X and Tumblr using keywords capturing discussions about Israel, Gaza, Middle East politics, and Palestinians. While these discussions existed before 7 October, their volume increased sharply after that date, and the dataset includes precisely geolocated content extracted with associated metadata.

Across this research brief, we examine the relationships between anti-Jewish and anti-Muslim networks and explore possible escalation dynamics, including cumulative and reactive radicalisation processes between them. To avoid singling out any specific community, the visualisations presented in this report use colour coding only. The two networks are represented using pink and purple, without indicating which colour corresponds to which community. This choice was made for ethical reasons, as the purpose of the analysis is to understand interaction and escalation dynamics between hate networks rather than attribute attention or blame to a particular group.

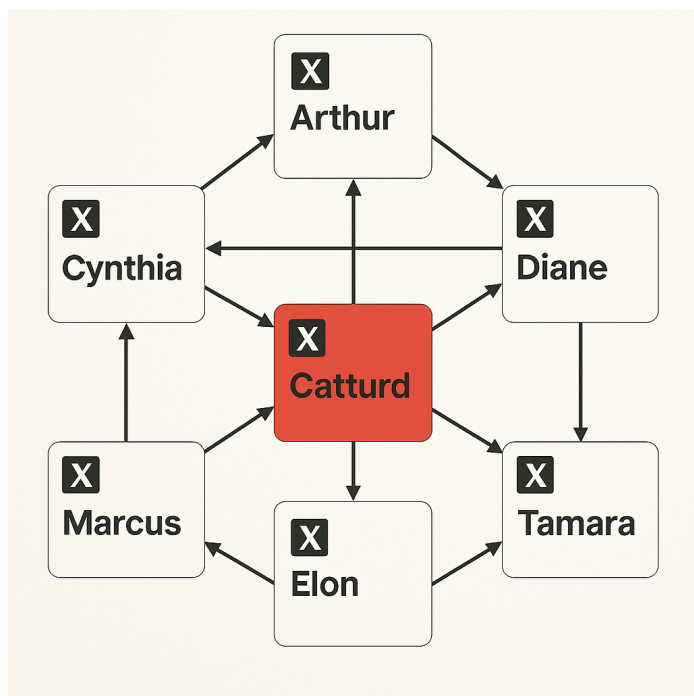
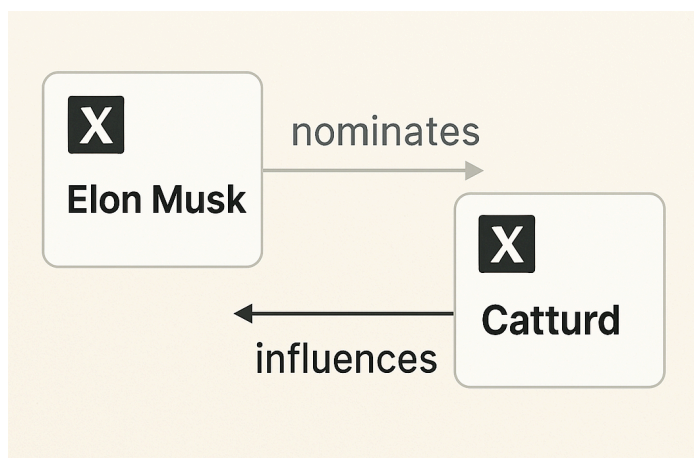


Figure 1. This figure shows an example of how an influence network can be built out of individual online interactions. In the example, one user, “Catturd”, is nominated by the account “Elon Musk”, therefore, a link is generated pointing from “Catturd” to “Elon Mask”, to indicate the influence of the former on the latter.

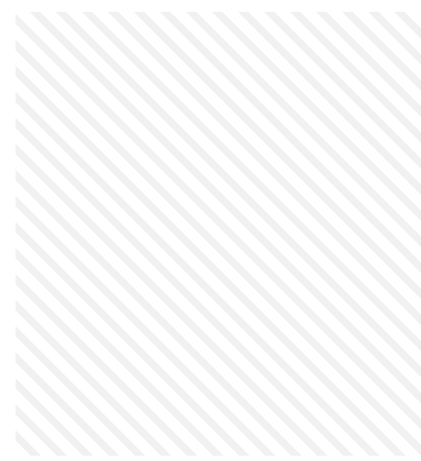


Figure 1: The Full Hate Influence Network – A Structural Overview

Figure 1. This figure presents the complete influence network of online hate targeting Jewish and Muslim communities in Australia between October 2023 and April 2025. The network comprises 53,496 nodes (anonymised social-media accounts) and 356,539 directed links constructed from 1,458,661 X interactions and 57,431 Tumblr posts.



How to read this figure

Link colour, either pink or purple, addresses the type of conveyed hate (anti-Jewish vs anti-Muslim, anonymised). Node colour follows a toxicity gradient: black nodes are accounts that provoke hateful reactions from others but do not themselves post hateful content; purple-tinted nodes have a low proportion (0–10%) of hateful posts relative to their total activity. Node size is scaled linearly by out-degree, a measure of influence — how many other users react to that account's content.

Key findings

The two major platform ecosystems, bridged by a set of connecting users, are visible. Accounts from the X/Twitter ecosphere concentrate on the bottom of the figure and connect to the Tumblr ecosystem (upper region) via a limited number of cross-platform bridges. Within the X cluster, the anti-Jewish and anti-Muslim interaction flows are spatially separated, revealing two distinct but neighbouring communities of hate. Notably, the centre of the one cloud contains a large mass of black (non-hateful) nodes. These are mainstream accounts (such as the Australian prime minister account or the account of the minister for foreign affairs) whose posts have been used as springboards for hateful commentary. This pattern underscores how hate networks parasitize broader public discourse, yet asymmetrically across topics.

Figure 2: The Hate Network Before the 7 October 2023 Attacks

Figure 2. This figure shows the influence network constructed exclusively from hateful interactions recorded before the Hamas-led attack on Israel on 7 October 2023. The same visual encoding as Figure 1 applies: node colour reflects the share of toxic content in a user's output (purple = low, progressing toward yellow and red for higher fractions), and node size is proportional to influence (out-degree).



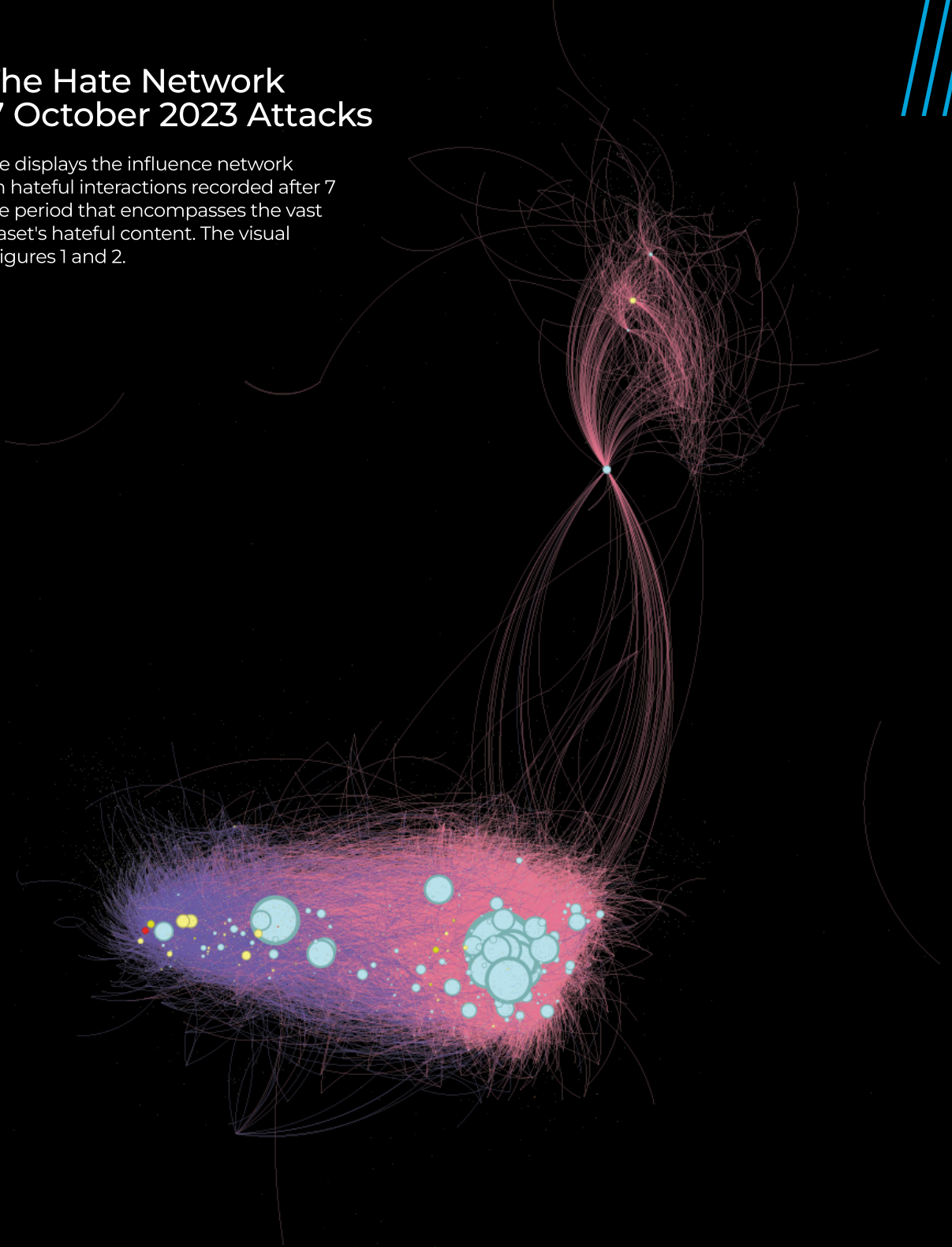
Key findings

Before the October 7 attack, the hate ecosystem was comparatively sparse. A few highly influential accounts (visible as the large light-purple nodes at centre-bottom) monopolised the influence flow, particularly within the pink component. These are high-profile or high-engagement accounts whose occasional content triggers cascades of hateful responses, despite not being prolific hate producers themselves. The purple component (left, with purple links) was also identifiable as a separate cluster, confirming that the structural divide between the two hate streams pre-dates

the geopolitical escalation. However, it remained peripheral and weakly articulated. High-toxicity accounts (yellow, orange, red) existed but had minimal structural influence, occupying the network's margins. A Tumblr cluster (upper right, scattered green/yellow dots) was present but loosely connected, linked to the main body by only a handful of long-range ties. Overall, the network exhibits a hub-dominated topology: a small number of prolific accounts shape the direction of the entire hateful conversation.

Figure 3: The Hate Network After the 7 October 2023 Attacks

Figure 3. This figure displays the influence network reconstructed from hateful interactions recorded after 7 October 2023 — the period that encompasses the vast majority of the dataset's hateful content. The visual encoding follows Figures 1 and 2.



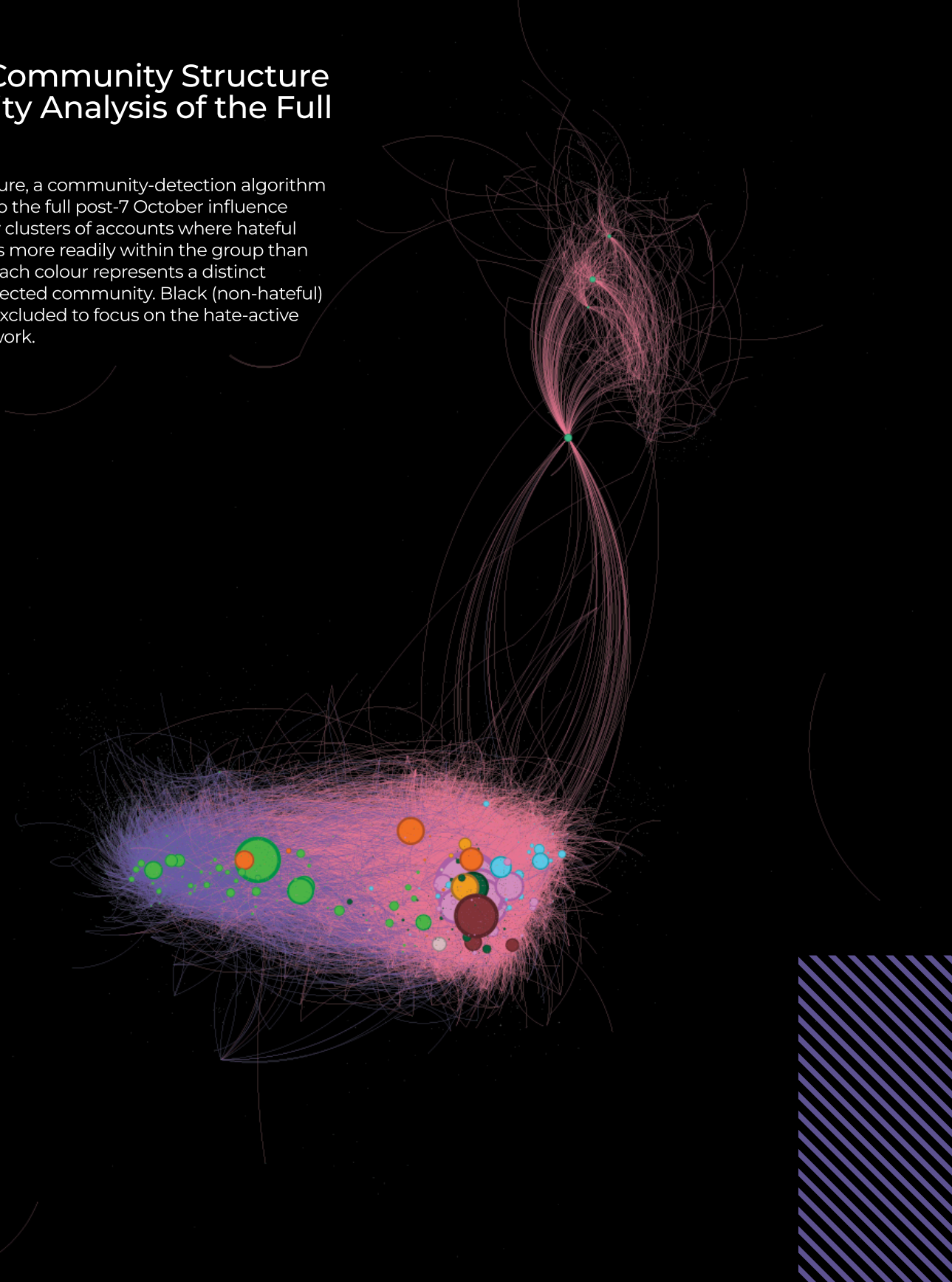
Key findings

The post-attack network is dramatically larger and denser than its pre-attack counterpart (Figure 2). The number of active nodes and links has increased substantially, and — critically — the distribution of influence has become more polycentric. The pink sub-network (pink links, right side) retained a broadly similar influence structure, anchored around a core of high-influence accounts, with several new mid-sized accounts having emerged alongside the previously dominant hubs, thus producing a more distributed influence architecture. On the other hand, the purple component (purple links, left side)

underwent a striking transformation. It became more articulated and internally heterogeneous, with new nodes of intermediate and high toxicity (darker colours: yellow, orange, and red nodes) emerging at the periphery. These are accounts that devote a substantial share of their output to purple hate. The Tumblr ecosystem (upper cluster) also expanded significantly, with visible bridge nodes pivoting between the X and Tumblr platforms. The contrast with Figure 2 illustrates how a single geopolitical event can catalyse the rapid proliferation and structural complexification of online hate.

Figure 4: Community Structure – Modularity Analysis of the Full Network

Figure 4. In this figure, a community-detection algorithm has been applied to the full post-7 October influence network to identify clusters of accounts where hateful influence circulates more readily within the group than between groups. Each colour represents a distinct algorithmically detected community. Black (non-hateful) nodes have been excluded to focus on the hate-active portion of the network.



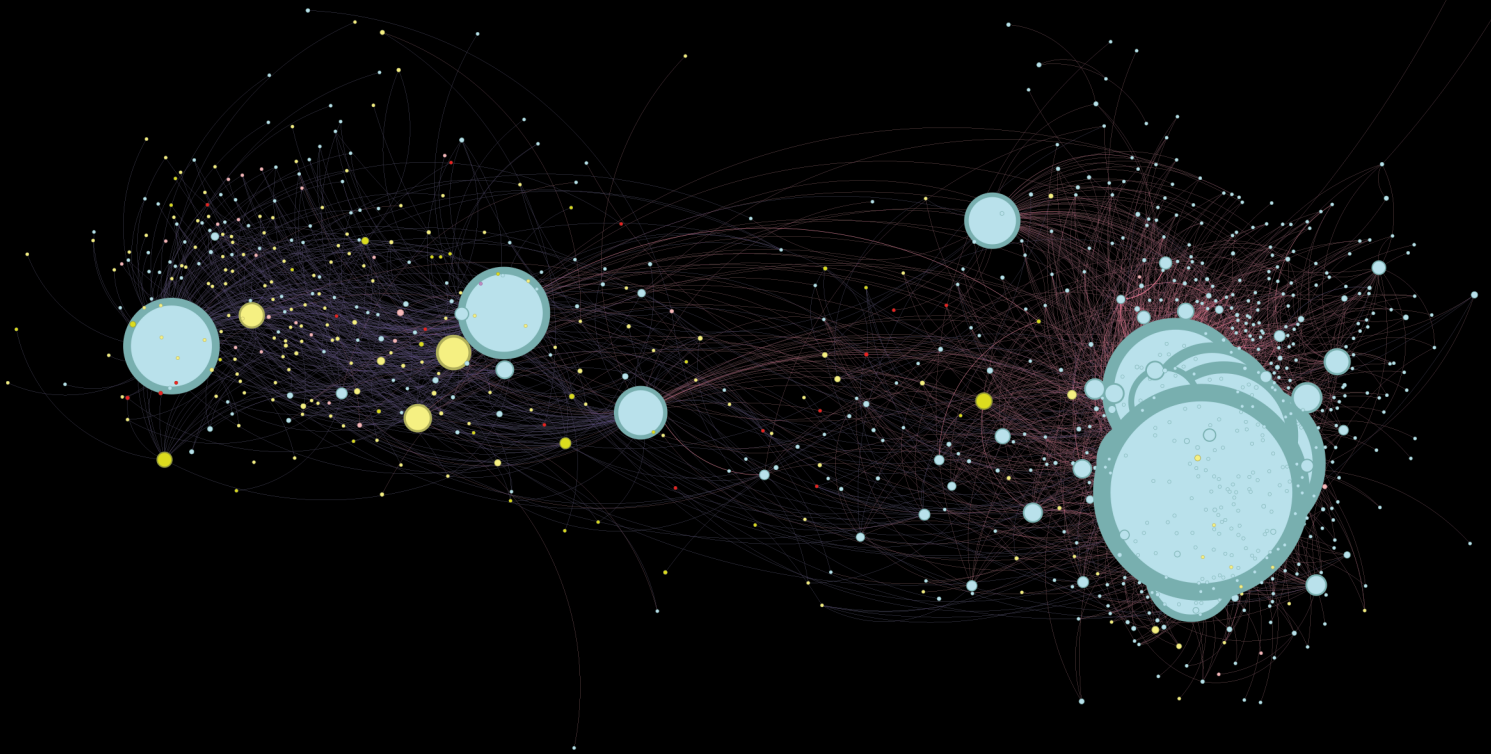
Key findings

The modularity decomposition confirms and deepens the structural divide observed in the earlier figures. The Purple component (left, predominantly green nodes) forms a single, highly cohesive community: its members principally influence and are influenced by one another, suggesting a tight echo-chamber dynamic. In contrast, the pink component (right) fractures into multiple distinct comm

unities (orange, pink, cyan, brown), reflecting a more diverse and topically heterogeneous discourse, hate in this ecosystem is not monolithic but spans a range of sub-narratives and user coalitions. One notable node bridges the two major components, engaging substantially with both purple and pink discourse. Such bridging accounts are critical for understanding cross-community influence contagion.

Figure 5: The Hate Network in the Two Months Before the Adass Synagogue Arson (Dec 2024) – Capturing a Hate Dividend

Figure 5. This figure and Figure 6 form a critical paired comparison centred on a specific offline hate incident: the arson attack on the Adass Israel Synagogue in Melbourne in December 2024. This figure displays the influence network constructed from hateful interactions occurring in the two months immediately preceding the attack. Only accounts that were active in this window are included. The colour and size encoding follow the onion scheme: purple = low toxicity share, yellow = intermediate, orange/red = high; size = out-degree influence.



Key findings

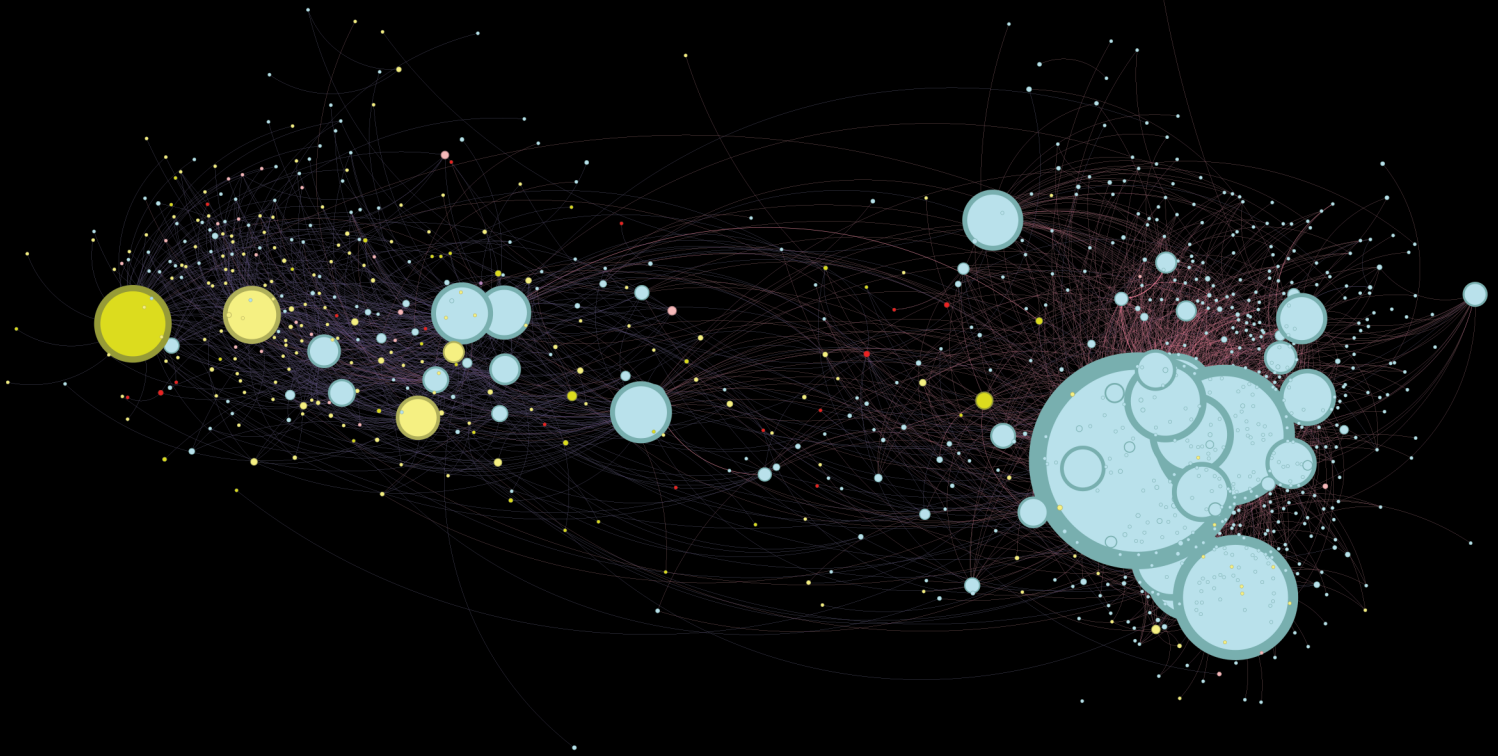
In the two months before the Adass arson, the network shows two spatially distinct communities. On the right, the pink cluster is anchored by a small number of very large, predominantly purple/low-toxicity hubs — high-influence accounts that attract hateful reactions. On the left, the purple cluster

features several yellow nodes (moderate-toxicity accounts) and is characterised by dense purple internal connections. The overall topology reflects a period of active but structurally conventional hate circulation: influence remains concentrated in traditional hubs, and the most toxic accounts have not yet gained outsized structural positions.



Figure 6: The Hate Network in the Two Months After the Adass Synagogue Arson (Dec 2024) – Capturing a Hate Dividend

Figure 6. This figure shows the influence network for the two months following the Adass Israel Synagogue arson in December 2024, forming a direct comparison with Figure 5 (the two months before).



Key findings

While the pink component (right) remained relatively stable in structure — the same large hubs continue to dominate — the purple side of the network (left) underwent a dramatic reconfiguration. Higher-toxicity nodes (visible as larger yellow and yellow-green accounts on the left) have gained substantially more influence compared to the pre-arson period. This shift reveals what we term a "hate premium": in the aftermath of a high-profile antisemitic attack, accounts that consistently produce a greater proportion of hateful content

in the purple network were rewarded with increased centrality and reach within the influence network. The finding is significant for P/CVE practitioners because it suggests that offline hate incidents do not simply amplify existing online dynamics: they structurally empower the most extreme voices in one hate community, creating a feedback loop that may escalate inter-community hostility. Monitoring these centrality shifts in the immediate aftermath of offline events could provide actionable early-warning signals for law enforcement and community safety organisations.

AUTHORS BIO

Matteo Vergani

Matteo is director of the Tackling Hate Lab and Associate Professor in Sociology at Deakin University and specialises in radicalisation and hate crime, publishing in leading international journals and securing large research grants. He collaborates with numerous institutions and government agencies in Australia and Canada. His research advances the systematic consolidation of knowledge in hate and extremism studies through large-scale systematic reviews and the development of rigorous measurement tools of online and offline hate and radicalisation. His research programme fosters multidisciplinary collaboration across social sciences, data science, econometrics and engineering, leveraging advanced technologies for analysing digital archives and social media big data.

Andrea Giovannetti

Dr Andrea Giovannetti is Co-Director of the Tackling Hate Lab, Assistant Professor of Economics at the Australian Catholic University and a member of the Violence Research Centre at the Institute of Criminology of the University of Cambridge, where he previously held a Marie Curie Postdoctoral Fellowship. His research on organised crime, contemporary extremism and social cohesion combines machine-based econometrics with advanced computational methods in network theory to support policymakers and security agencies on a large spectrum of inter-connected issues. Andrea's collaborations with public stakeholders on complex social threats include London Metropolitan Police, Merseyside Police, Home Office and Home Affairs.



TACKLING

HATE

