

FROM INJURY TO FALSE FLAG

How Conspiracies Emerged on Telegram After
the Bondi Attack

MATTEO VERGANI, HUU PHUC HONG, KEWEN LIAO

1 March 2026

TACKLING HATE LAB

Real time mapping of narrative escalation

This brief presents findings from a structured analysis of public Telegram posts collected during the first 72 hours following the Bondi attack. The dataset consists exclusively of publicly accessible English-language posts mentioning Arsen Ostrovsky and related keywords. While this represents a snapshot rather than the totality of Telegram activity, it provides a high-resolution account of how narratives evolved in real time.

Using first-mention timestamp analysis, hourly stage frequency tracking, word-frequency modelling, and engagement ranking (views and forwards), we reconstructed the sequence through which reporting about an injury escalated into structural conspiracy claims.

Account names have been anonymised to protect individuals' privacy and ensure ethical compliance.



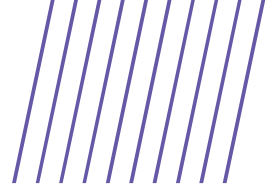
The baseline injury frame

The first posts reporting that Arsen Ostrovsky was injured included the image he took on his phone shortly after being injured (see Figure 1). This baseline “injury frame” marked the starting point of discussion. The most viewed post in the dataset containing Ostrovsky's photo,

published at 21:59 on 14 December, was a straightforward factual report of injury, receiving 148,196 views and 543 forwards. This is significant: the highest-reach content was not conspiratorial. Narrative mutation began within minutes. Identity framing (references to “October 7, 2023”) appeared 2.6 minutes after the first injury mention. The tone of some early posts including Ostrovsky's picture (Figure 1) was insinuating.

Figure 1. Image taken by Ostrovsky on his mobile phone shortly after the injury.





From image scrutiny to direct accusation

Structural conspiracy language (“false flag”) appeared 14.7 hours after baseline reporting. At this stage, AI-generated images began circulating that depicted Ostrovsky receiving make-up on what appeared to be a staged set, reframing the injury as fabricated (see Figure 3). The first recorded appearance of the AI-generated image shown in Figure 3 occurs in our dataset on 15 December 2025 at 15:46:36 AEDT.

It was posted by an Australian far-right Telegram channel (*anonymised*) and received 7,884 views, 355 forwards, and 92 reactions. The accompanying text read:

“Is this for real or fake?”

This photo of the Zionist Arsen Ostrovsky is circulating online getting the make up ready before the shooting.”



Figure 3. AI-generated image depicting Ostrovsky being prepared with make-up in a staged setting.

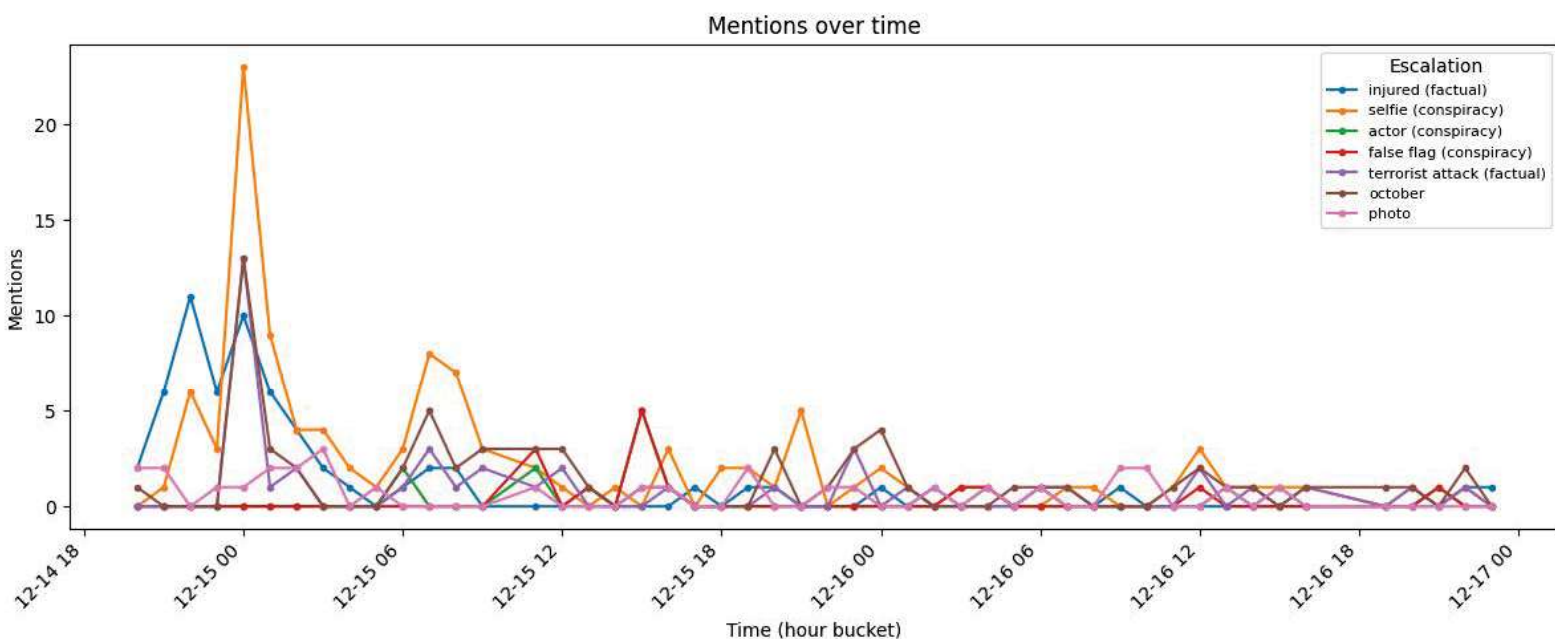
Figure 3 is AI generated. ABC NEWS Verify confirmed this using Google’s Synth ID detector, which identified the company’s invisible watermark. Visual inconsistencies also indicate fabrication, including unreadable text on the shirt, distorted tripod legs, a disappearing light stand leg, merged vehicles, and incorrectly rendered anatomical details.

The text accompanying Figure 3 was clearly conspiratorial. For example, one wrote:

Nobody has been able to debunk this image that appears to show Arsen Ostrovsky being made up to look like he was shot at in Bondi Beach, Australia.

Hourly frequency analysis confirms this progression (Figure 4). “Injured” spikes first. Identity and image-related terms rise shortly thereafter and dominate discussion volume. Accusatory and “false flag” language appear later and remain lower in overall frequency.

Figure 4. Mentions of tracked terms by hour. Factual and identity terms peak first; actor and false flag appear later.



The data in Table 1 suggest a clear distinction between origin and amplification. The most viewed post (148,196 views; 543 forwards) was a factual injury report. In contrast, a UK conspiracy channel known for spreading COVID-19 conspiracies (*anonymised*), appears three times among the top five combining factual or identity elements with mockery or insinuation (e.g. focus on the “selfie” or claims of being “made up”). The only post advancing a full false-flag thesis ranked third by views (38,634; 686 forwards) and was published approximately 18 hours after the initial injury report, indicating that structural conspiracy framing was neither immediate nor dominant in reach. Forward counts were higher for identity and mockery posts (2,083 and 1,161) than for the factual top post (543), suggesting that reframing content was more actively shared. Overall, the narrative evolved in layers, with different actors driving visibility and escalation at different stages.

Table 1. Top five Telegram posts ranked by views and forwards within 72 hours of the Bondi attack.

Rank	Channel	Date & Time	Views	Forwards	Content Summary
1	beholdisraelchannel	14 Dec 2025, 21:59	148,196	543	Factual report: “Arsen Ostrovsky was injured in an attack.” No identity or conspiracy framing. Most viewed post in the corpus.
2	lauraabolichannel	15 Dec 2025, 00:26	50,709	2,083	Frames Ostrovsky as “the selfie man” and October 7 survivor (“What are the chances?!”). Uses sarcasm and identity contextualisation; no explicit false-flag claim.
3	oliverjanich	15 Dec 2025, 15:35	38,634	686	Explicit false-flag framing (references to Candace Owens, crisis actor, Sandy Hook anniversary). Full conspiracy thesis; posted approximately 18 hours after the first injury report.
4	lauraabolichannel	14 Dec 2025, 23:44	36,776	904	Repeats injury fact, then questions: “Who takes a selfie in such a moment?!” Shifts focus to the image, casting doubt without using “actor” or “false flag”.
5	lauraabolichannel	16 Dec 2025, 11:02	31,983	1,161	Claims the image “appears to show Arsen Ostrovsky being made up”. Implies staging; states no one has “been able to debunk” it.

Conclusion

The evidence demonstrates a layered progression:

1. Injury reporting
2. Identity contextualisation
3. Image scrutiny
4. Insinuation
5. Accusation
6. Structural conspiracy framing

No single account originated the entire narrative arc. The conspiracy frame did not initiate the discourse; it emerged as the endpoint of a sequential mutation process.

For policymakers, the implication is clear: intervention strategies should not focus solely on explicit “false flag” language. Earlier stages, particularly identity framing and image-based doubt, which emerge within minutes and drive sharing behaviour, represent critical points for timely response.

