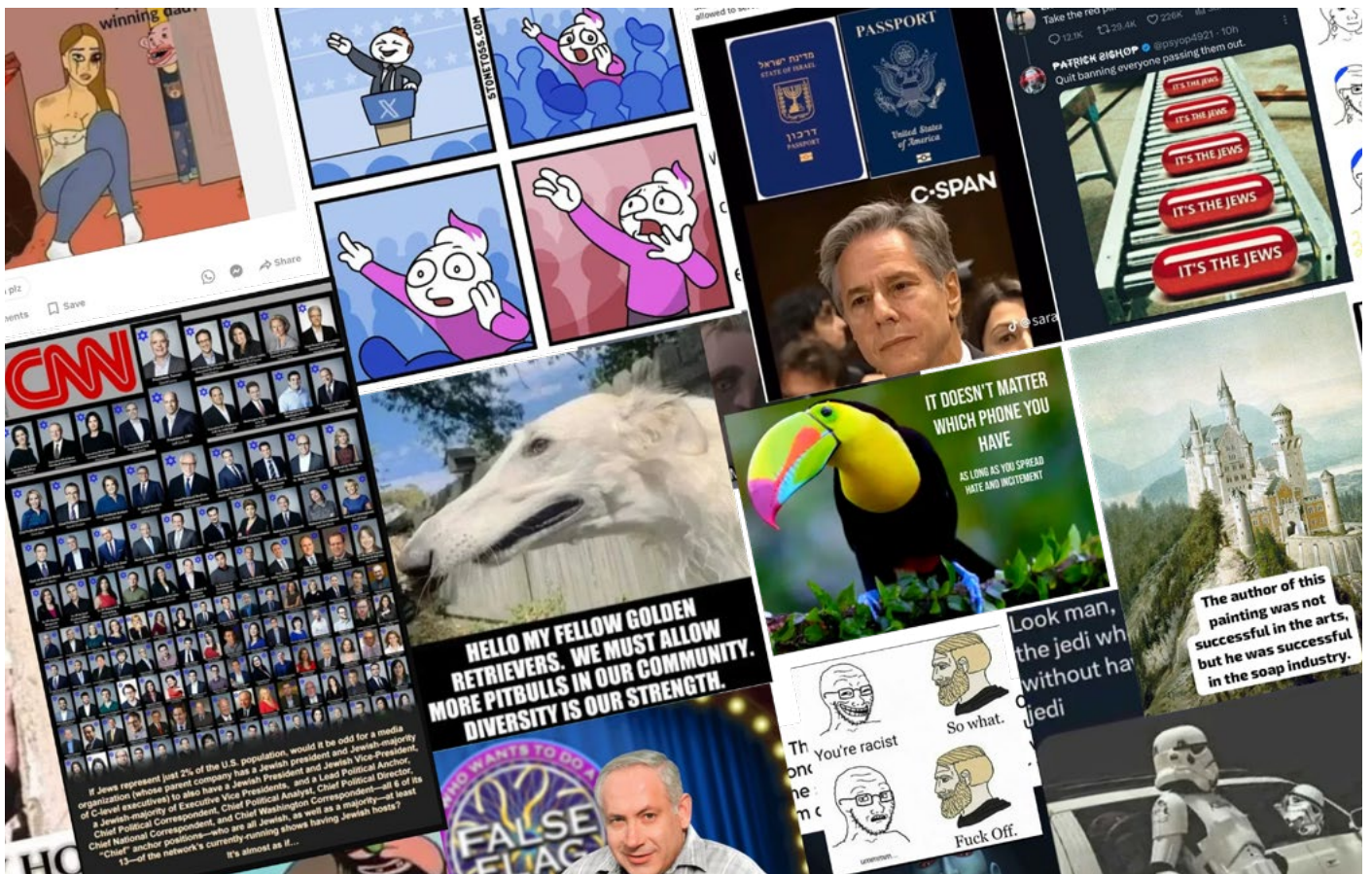


FIND YOUR MEME

HOW ANTISEMITIC MEMES GO MAINSTREAM



Introduction

The proliferation of antisemitic content on small, high-harm online services poses a significant risk to users of user-to-user safety. This includes risks of radicalisation into extremist and violent ideologies, and with serious implications for online threats, abuse and harassment. These risks are exacerbated when users are from a group with protected characteristics, which include age, race, sex and sexual orientation.

In relation to antisemitism, content on these small services tends to be more extreme than the anti-Jewish racism on large, mainstream platforms. As a result, it helps radicalise people into extreme narratives, the results of which have included violence against Jews.¹ The proliferation of antisemitism online also contributes to the rising levels of racism that divide communities. It eases the spread and amplification of conspiracy theories that undermine trust in democratic institutions and erode liberal values of tolerance and inclusion, across Europe. It also helps normalise antisemitism in both online and offline discourse.

These small platforms, including, for example, BitChute, Gab, and 4chan, often operate with minimal moderation and are also sometimes encrypted, providing safe havens for extremist content that includes antisemitic tropes, incitement to violence, and radicalising material.² Despite the harm they cause, many of these platforms manage to escape robust regulation in Britain and the EU.

This is particularly worrying, considering the major increase in antisemitism in Europe. In Britain, the Community Security Trust (CST) recorded 3,528 antisemitic incidents for 2024. This is roughly double the number of incidents recorded in 2022, and slightly less than the number recorded in 2023—when there was a sharp rise following the 7 October Hamas attack on Southern Israel.³ In the EU, some organisations across Europe reported an increase of more than 400% in antisemitic incidents following 7 October 2023.⁴ A 2024 survey found that 96% of respondents from 13 EU countries have encountered antisemitism in their daily life.⁵ Hate crimes tend to be severely under-reported, so these numbers—although high—still represent only a portion of the real occurrence of antisemitic hate crimes.

In this report, we examine the antisemitic content that originates from these small services, and how it migrates to larger platforms, where it spreads at a greater rate and has a wider, even worldwide, reach. This report will begin with an overview of antisemitism on small services and the synergy with larger services, to explain the risks. We will then look at services to demonstrate the origins of antisemitic content on these platforms. The report ends with recommendations for policy and regulation, to tackle the harm caused by small services, urging decision-makers and regulators to apply stronger enforcement and risk-based platform categorisation to protect Jewish communities and our democracies.

Antisemitism on Small User-To-User Services

Small, high-harm platforms have become breeding grounds for a variety of extremist ideologies, including terrorism-promoting content⁶ and antisemitism, owing in part to their lax moderation. Lack of moderation is sometimes the product of financial considerations of cost to services, and other times it is a deliberate, ideologically motivated decision. As a result, these spaces appeal to communities or people who hold radical and extreme ideologies, including support for violent extremism and terrorism.⁷

These platforms amplify conspiracy theories, including Jewish control of global affairs and finance, Covid-related antisemitic conspiracies that blame Jews for the pandemic, and many war-related conspiracies that include blood-libels.⁸ A study by the Institute for Strategic Dialogue (ISD) found a significant rise in online antisemitism on both small and large platforms after Hamas's 7 October attack.⁹ There is a fluid and constant movement of extremist information from small platforms to large ones, where this content spreads and is deployed to draw users to the smaller platforms, where they may be radicalised further (this is sometimes referred to as 'breadcrumbing').¹⁰

Research by the Antisemitism Policy Trust (APT) probed antisemitism across several smaller platforms, including TruthSocial, Roblox, Rumble, Odysee and Minds.¹¹ The Trust found that antisemitism is no longer isolated to small pockets of far-right or Islamist extremist forums. Rather, antisemitism is now endemic to the entire social web, and has 'gone viral'. Social media platforms are being woven together by increasingly sophisticated and influential alt-media accounts to form what the Trust dubbed 'antisemitic superhighways' between small-scale services and large ones. These highways span multiple platforms and are used to socialise and normalise hate, conspiracies and antisemitic theories, creating a radicalisation effect. By analysing antisemitic activity and hate speech across several platforms at once, the Trust was able to see previously invisible highways emerge, with small services playing a key role. It is not only NGOs that are finding this overlap. As the UK regulator Ofcom rightly pointed out: 'We have found that a lot of terrorism content is first posted on smaller U2U [user-to-user] services and then linked to from larger, higher-reach services.'¹²

The migration of antisemitic content from small to large platforms exponentially increases its reach. This includes content that is "legal but harmful"¹³, like antisemitic conspiracy theories¹⁴, that moves from fringe platforms to mainstream ones including X, YouTube, and Instagram, exposing millions to radicalising narratives and 'funneling' them back to the small, extreme platforms.¹⁵ A 2024 APT study highlighted "thread hijacking," where extremist accounts post antisemitic comments under high-engagement celebrity posts on mainstream platforms to gain visibility and use this to direct users to fringe content.¹⁶ This cross-platform spread amplifies antisemitic memes, which often disguise hateful tropes as humour, making them harder to detect and remove.

A report by the Community Security Trust (CST) also found that far-right extremists use mainstream platforms to obtain wide audience reach, to spread their ideology in a way that circumvents those platforms' moderation, and then 'funnel' young people onto more extreme smaller platforms that have been accused of turning a blind eye to illegal terrorist and violent content, with little or no moderation, for example Telegram.¹⁷ The report found that far-right extremists have migrated from the large, mainstream social media platforms towards smaller platforms, including Gab, Bitchute and Telegram, where antisemitic content flourishes.¹⁸ This includes illegal material that calls for violence and terrorism against Jews.¹⁹

The radicalisation risk spans ideologies. Far-right groups spread "Great Replacement" conspiracies blaming Jews for immigration, while Islamist extremists, like those on Telegram's "Terrorgram," praise violence against Jews, as seen post-October 7, 2023. Far-left antisemitism can manifest as anti-Zionist rhetoric, holding Jews collectively responsible for Israeli Government action or policy, fuelling harassment, threats, and racist abuse. The APT's 2021 Instagram study with CST and the Woolf Institute found that such content can radicalise users by reinforcing existing biases, potentially increasing the propensity for violence.²⁰

The connection to offline harms is stark, and small platforms that host illegal content, such as extremist and hateful content that has been linked to radicalisation, recruitment into terror organisations and facilitation of

terror attacks.^{21,22} The International Centre for Counter-Terrorism (ICCT) identified a widespread abuse of digital technologies by violent extremists, and documented their shift to fringe platforms.²³ A briefing by the Antisemitism Policy Trust, published in August 2020, provided examples of the connection between online and offline harms, citing examples of attacks against Jewish targets (for example, the Pittsburgh Synagogue attack) and against Muslim targets (for example, the Finsbury Park and Christchurch mosque attacks).²⁴ In all of these, attackers participated in extreme online forums where they were either radicalised to the point of attacking Jews and Muslims, or inspired others to commit acts of hateful violence. The terrorist who killed eleven congregants and injured six others in a synagogue in Pittsburgh in 2018, promoted his hateful, antisemitic agenda on the social media platform Gab – where he also posted minutes before attacking the synagogue. Based on testimonies after the attack, he had consumed large volumes of racist and other material online, which had incited him to violence. Another

Regulation of small user-to-user services in the EU and Britain

Under the terms of the UK Online Safety Act, small user-to-user services are subject to core safety duties, but face a reduced regulatory burden compared to the largest platforms. All services must have systems and processes in place to deal with the sort of content in issue swiftly once on notice of it, including terrorist propaganda and racist hate, including antisemitism, once they are made aware of it. Category 1 services are subject to extra duties that operate to protect adults from some non-criminal content; small services are exempt from these obligations. All service providers are obliged to satisfy the children's safety duties if they are likely to be accessed by children – irrespective of size. Additionally, all services would have to take measures against priority content harmful to children which includes content which is abusive on the basis of race or religion amongst other grounds. However, Ofcom in its codes has imposed more far-reaching measures on services it has determined to be large. This means that in effect, niche or fringe platforms may escape more intensive scrutiny unless illegal content is reported to them, even if they are known to host extremist communities. The act's risk-based

briefing by the Antisemitism Policy Trust, on anti-Jewish misogyny, referenced a study by the American NGO Media Matters, which found a staggering 180% increase in posts containing both antisemitism and misogyny on the far-right anonymous message board 4chan between 2015 and 2017.²⁵

Additionally, on Telegram for example, CST found posts glorifying far-right terrorists including Thomas Mair and David Copeland, and posts calling for the killing of Jews.²⁶ The platform 4chan was also found to host threads containing explicit calls to kill Jews. Similar posts, containing violent, antisemitic comments, were found on Bitchute, including images with phrases such as, 'all Jews must die,' and images of people aiming weapons accompanied by threatening language. The information is easily accessible, and CST concluded that: 'the quantity and spread of this incitement poses an urgent and ongoing terror threat to Jewish communities.'²⁷

approach prioritises intervention based on scale and reach, which may leave smaller but high-harm platforms under-policed in practice.

Much of the antisemitic content online does not reach the threshold of illegal speech, although it has the power to radicalise, and to dehumanise and demonise Jews, and enhance antisemitism. Additionally, the obligations by small services to implement 'safety by design'²⁸ features that will help limit illegal and legal but harmful content, are scaled back in comparison with the duties imposed on large services.²⁹ Additionally, the OSA does not directly regulate the use of VPNs, which can easily be used to access terrorist and other extremist and harmful content.³⁰ For example, two small services known for hosting extremist content – BitChute and Gab – announced that they are no longer available to UK users. However, they remain easily accessible using VPNs.

The EU Digital Services Act also applies to platforms of all sizes but explicitly scales down obligations for smaller or medium-sized services. While all platforms must

respond to notices about illegal content, for example terrorist incitement or antisemitic hate speech, they are not obliged to undertake systemic risk assessments, independent audits, or apply algorithmic transparency unless they are designated as Very Large Online Platforms (VLOPs) – which do have risk assessment obligations. The DSA therefore assumes that the risk to public safety is proportional to audience size, even though some smaller services may be disproportionately harmful due to their niche appeal to extremist users. Smaller platforms must maintain a user-friendly notice-and-action system and cooperate with national authorities, but their internal moderation policies and content removal mechanisms are largely self-governed unless flagged by regulators or civil society. In which case, oversight can be escalated if risks are identified; however, this is a reactive approach that relies primarily on users or civil society groups reporting risks. This tiered approach creates a regulatory light-touch environment for smaller platforms, even when they are used to spread harmful ideologies. The DSA also has reduced requirements for ‘safety by design’ and transparency³¹ compared with large platforms.³² The DSA does regulate VPNs at a basic level under ‘conduit services’, but not comprehensively, and enforcement is likely to be slow and ineffective.³³

The EU Digital Services Act (DSA) and UK Online Safety Act (OSA) impose lighter obligations on smaller platforms compared to large ones, assuming risk correlates with size and reach. However, this tiered approach does not appropriately regulate niche platforms that host extremist content like antisemitic or terrorist material, which can radicalise users despite smaller audiences. While both laws mandate removing illegal content upon notice, smaller platforms are exempt from proactive monitoring, systemic risk assessments, or transparency reporting, allowing harmful ideologies to persist unless reported.

This creates a regulatory gap, as high-harm platforms can evade scrutiny. Tighter regulation, based on risk rather than just size, is needed to address these platforms’ disproportionate impact and ensure effective oversight. Our hope is that this report will assist policy-makers in understanding the risks posed and underline the case for tougher risk-based regulation to limit the harm caused by small services that host extremist and racist content.

Methodology

The platform 9gag was selected for memes selection. This service has the characteristics of familiar social media platforms, as users can create an account, upload content, comment, and share content. The platform offers categories and tagging of content, therefore separating it from a simple image uploading repository. We cannot provide real-time or proprietary website traffic statistics for 9gag.com, as this data is typically owned by the website itself or available only through paid analytics services like Similarweb, SEMrush, or Statista.

However, based on publicly available estimates from third-party traffic analytics platforms (as of late 2023 to early 2024), here’s a general overview of 9gag.com’s traffic:

Country	% of Traffic	Approx. Monthly Visits
United States	22%	~26.4 million
India	12%	~14.4 million
United Kingdom	6%	~7.2 million
Germany	5%	~6.0 million
Canada	4%	~4.8 million
Brazil	3.5%	~4.2 million
France	3%	~3.6 million
Australia	2.5%	~3.0 million
Philippines	2.5%	~3.0 million
Mexico	2%	~2.4 million
Rest of the World	~37.5%	~45 million

Table 1: Total monthly visits to 9gag.com: ~120 million (as of early 2024). These percentages are rounded and based on aggregated public data from Similarweb, Alexa (historical), and other traffic estimators. Actual figures may vary.

The platform offers open view access, namely, visitors who do not have an account can view most content, unless it is marked as “not safe for work” (NSFW), which requires an account to be viewed. For this project, we did not create an account; therefore, all the memes we collected are publicly available to peruse. In addition, to avoid any recommendation algorithm or any other preferential content treatment, we collected 535 memes from the “Trending” page of 9gag. Essentially, we simulated a user who would not search specifically for controversial or hateful content, but would browse the top-voted, commented on, or otherwise trending posts of the platform, and such content would reach the user just by browsing the page: supply, rather than demand.

The ability to select such a dataset from a non-personalised timeline is a main reason why we chose 9gag as a proxy for small social media (SSM).

Reddit was used as a proxy for large social media due to a) its popularity (billions of monthly visits), b) its position as a training dataset that further affects AI at large,³⁴ and c) its history with antisemitic content.³⁵

Data collection took place from January to August 2024. Memes were stored locally, while, through a random selection, the comments were additionally stored for 37 of the 535 memes. A thematic analysis was performed on the memes to find out: their relation to antisemitism; innuendos and/or relativism; fake news and other information warfare; cognitive warfare; if the content was religious, political, or historical in type; if it was posted on large social media platforms using Reddit as a proxy; if the meme had appeared on Reddit repeatedly (and at which subreddits); if it had been flagged or taken down by Reddit, and notes to otherwise help with the thematic analysis.

The Reddit identification was performed by an algorithm that was scripted to use results from common commercially available reverse image search tools, such as Google Image and TinEye, while the resulting links from Reddit were analysed for their content by the Qwen-Max Web Search functionality.

Additionally, some quantitative calculations were performed to assist with context. For example, the ratio of how many memes from 9gag were also found on Reddit, compared to the ones that were only posted on 9gag, was calculated to identify if smaller platforms are used to prime content for larger ones, or the number

of memes one had to view before receiving antisemitic content (“first contact” frequency) which is a contextual indication about how fast a user might become used to radicalising and hateful content.

Results

In terms of the quantitative results, only 5 of the 535 antisemitic memes that were collected from 9gag were not also found on Reddit. This translates to 0.9% of the 9gag data set being absent from or undetected in Reddit, whilst 99.1% of the memes from 9gag were reposted on Reddit. This finding showcases small social media (SSM) platforms as communication vessels with large social media, where the impact of hateful content is tested before it is further disseminated.

Concerning the number of memes, one had to browse before encountering antisemitic content; the average per browsing would be 6 memes, even though the variation across different days was quite large, from 1 to 20. However, in the comment section of these instances of antisemitic content, additional antisemitic content was encountered 100% of the time, thus facilitating radicalisation for users who engaged with this material, the variation notwithstanding.

Cognitive warfare

SSM memes showed a high frequency of cognitive and secondary antisemitism, even though directly hateful content had a significant presence too. Misinformation was less frequently posted directly, as posters relied mostly on cognitive warfare methods, for example insinuations and alleged “humour”. That is, instead of focusing on promoting or alleging a specific falsehood and trying to persuade the users directly, the intention most often has shifted to indirect methods of cultivating uncertainty and fear in users by methods such as “identifying patterns”, “just asking questions”, and “just saying”.

Hateful memes were repeatedly posted on 9gag to maximise their popularity there, but also on related larger platforms (e.g., the 9gag to Reddit pipeline). This method also intends to avoid content moderation by larger platforms. Repeated posting was threefold: a) re-posting a meme on a specific channel/subreddit if it is reported and removed, b) re-posting a meme on different channels/

subreddits if it is reported and removed, and c) re-posting a meme on various channels or subreddits even if it has never been removed, but a significant period of time has passed since the original post. Therefore, content moderation is weaponised in cases a) and b) as a way to advance content repetition, or even to create new content that promotes the removed-content owner as a righteous victim of unjust persecution. Regardless, even in case c) where content moderation is absent, repetition has the role of simple propaganda promotion.³⁶

Both methods mentioned above have a role in plausible deniability. Take, as an example, the image below (cropped in the original).



The image makes a reference to Adolf Hitler in a way that purports to be a) humorous, b) promoting an 'in-group' of people who 'know' and recognise the cropped image, and c) posted in a way that invokes plausible deniability—a viewer who is not observant may miss the reference. Additionally, the poster has not uploaded a screenshot of the exchange, but rather a photo taken by a smartphone, therefore distancing themselves from the exchange and from knowledge of the anonymous parties of the exchange. Thus, a planned promotion and relativisation of Nazism is presented as online happenstance.

Multi-locality

The topics of the memes show considerable variability, and antisemitism is not limited to one domain. The analysed memes showed a "multi-local" effect, that is, even though content about the Middle East definitely

increased, antisemitic content concerning other locales did not appear with less frequency than before (i.e., "great replacement" in the UK, "tunnel Jews" in the USA³⁷, conspiracies centred around government or banking control).

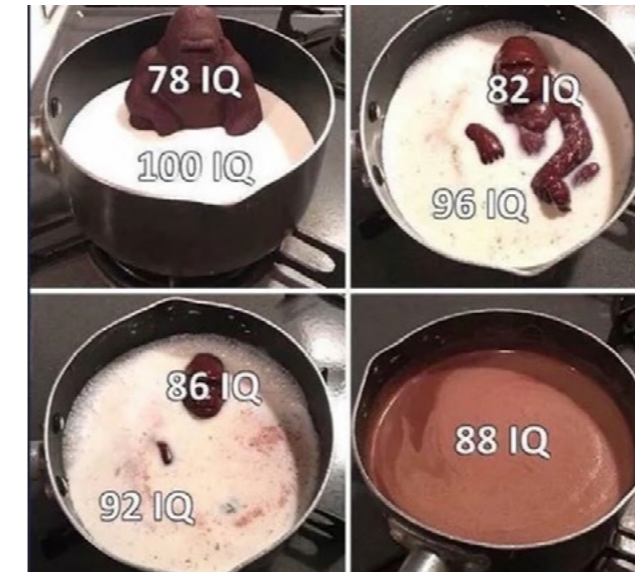
Take the image below as an example of a meme promoting the "great replacement" conspiracy. The text and imagery of the meme is cryptic, and easily understood by people who already know its content, therefore showing that memes are not only used to promote conspiracies but also to increase cohesion of an in-group by rewarding viewers who are knowledgeable about it. The conspiracy alleges that the recent increase in immigration is neither related to the economic conditions or political situation in destination countries nor to wars, droughts, or famine in source countries, but is rather a premeditated and fabricated crisis by "(((the Jews)))" in order to somehow damage 'White Europe'. As with many conspiracies, it is very vague; therefore a motive or logic behind this alleged "replacement" is not presented or is based on unscientific eugenics notions, lazily repeating the century-old Kalergi Plan conspiracy³⁸ (Race mixing would allegedly reduce IQ, therefore the general population would be easier to rule somehow).



The image above invokes the "great replacement" antisemitic conspiracy, while adding an additional antisemitic conspiracy of "dual loyalty" or duplicity as an entry barrier to further understanding the meme. That is, the image depicts a long-nosed dog [drawing on stereotypes about Jewish facial characteristics] with a traditional Jewish kippah on, referring to "fellow labradors": the allegation is that this dog pretends to be a labrador while it is not, in order to manipulate them. Additionally, the meme has an intersectional racist approach, as immigrants

from African or Muslim backgrounds are presented as pitbulls, potentially dangerous and violent, especially towards children.

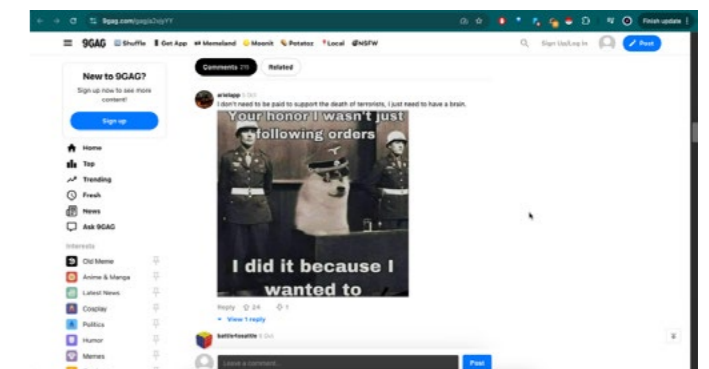
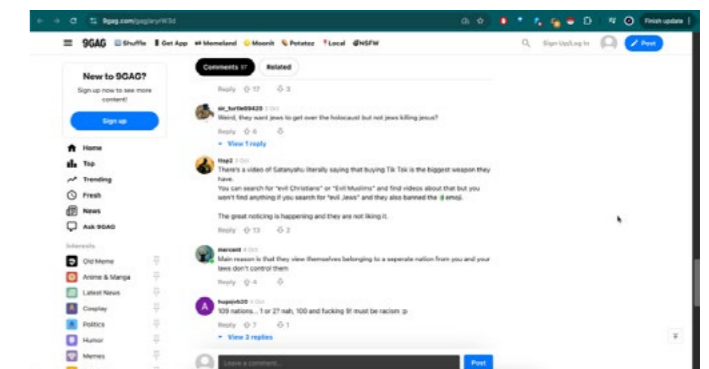
Other memes, presented below, corroborate this interpretation. However, presented in a cryptic way, these have the aforementioned cohesion effect and employ the tactic of trying to avoid content moderation by not mentioning any ethnic groups by name.

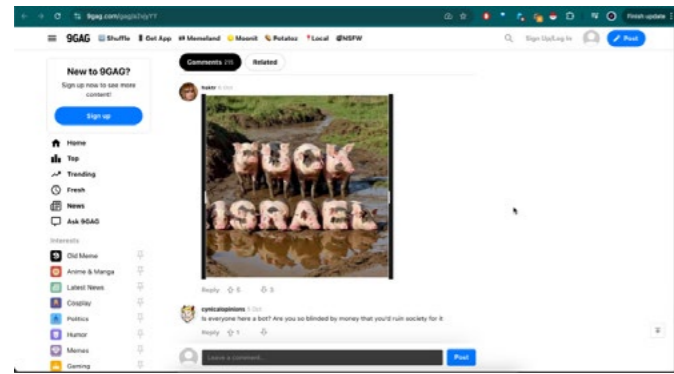


"It's always them"

The comments related to these memes consist of content attached to similar techniques, but to a lesser extent, rather, directly hateful content is more frequently encountered there. 9gag's methods are not entirely transparent, and the exact content moderation it imposes on trending topics is unknown; however, in practice in the comments section, almost exclusively, comments that are downvoted by other users get suspended, while other moderation seems to be absent. Occasionally, a warning that some content is sensitive may appear, but rarely will a comment that has not been downvoted be removed, even if hateful.

Additionally, users can find posts and comment on them by searching or clicking the tags of a post. The most common 9gag tag of antisemitic posts is ("it's always them"), while others, ("sand people"), ("small hat people"), and ("jooz") are not infrequent. Comments generally appear as text, but replies containing further memes are encouraged. Recently, AI-generated images have appeared in the comments too. Many of the most vile or hateful ideas are disseminated in the comments section due to the aforementioned lax content moderation.





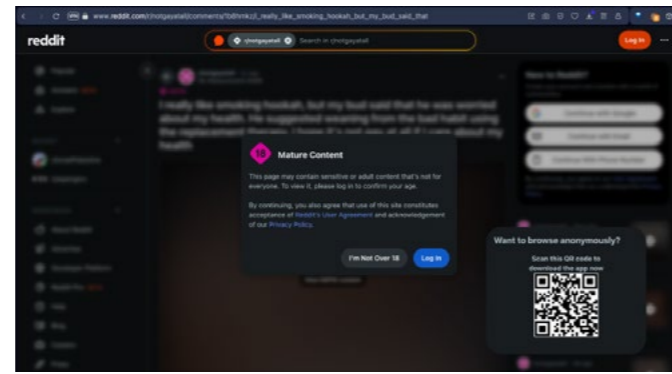
Communicating Vessels: 9gag to Reddit

Concerning re-posting 9gag memes to Reddit, 351 memes of the 535 downloaded memes produced useful search results. The rest of the web searches either received an error in the response, or the image was in a format that the search could not process correctly (corrupt file). Out of the 351 readable search results, 343 memes were found to be uploaded multiple times (97.72%), 3 were found only once on Reddit (0.85%), while 5 were not found on Reddit at all (1.43%).

The Reddit (re-)posters are frequently flagged by other users, their content is removed, and then re-posted. A common tactic is for users to post hateful memes with disclaimers such as “no anti semitism” [sic], or “not endorsing hate”, in order to preempt and deflect criticism.

The most common Reddit forums (called “subreddits”) where 9gag antisemitic memes were found were r/memes, r/PoliticalHumour, r/dankmemes, r/IsraelPalestine, r/Conspiracy, r/News, r/ProgrammerHumour, but occasionally also on r/UKPolitics, r/Israel, r/HistoryMemes.

While Reddit has been known to host antisemitic content^{39,40}, although its rules against hate state that it prohibits this sort of content and the platform aims to remove it. The platform has also partnered with the Antisemitism Policy Trust in order to provide training to its safety staff. Reddit offers a minimum barrier to entry, as for content that is flagged as mature it does require logging in. However, posters of antisemitic memes are unlikely to flag their content as mature.

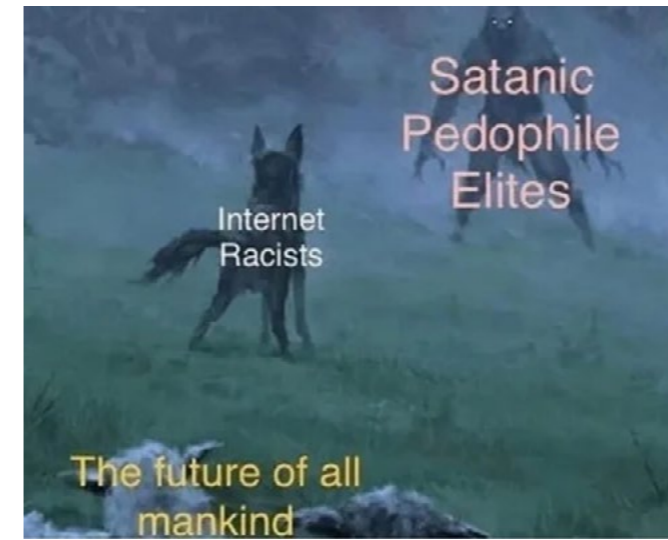


While there is no apparent hierarchy or direct command to post 9gag memes to Reddit, it is our observation that the Trending page of 9gag and the community upvoting/downvoting of comments effectively functions as a selection mechanism for the porting of hateful material from the smaller platform to the larger one.

Identity and Internet Racism

A part of the dataset, small in size but important in kind, was self-referencing memes. These made claims about the identity of the posters and their relationships to racist internet culture. It is of interest that these memes were found amongst those that were shared multiple times from 9gag to Reddit.

These memes do acknowledge the hateful and inciting nature of the material they have disseminated. In addition, they engage in a victimhood mentality, where the posters present themselves as brave content creators who face persecution for the content they post (content moderation). Finally, to promote their brave persona, they like to think of themselves as serving a greater good, which the “uneducated masses” or “normies” do not understand. This method allows them to increase in-group cohesion in the face of content moderation and community guidelines violations, attract new members, and promote themselves as pioneers.



When they put our memes up on the big screen at our trial.



Findings: How Antisemitic Memes Circulate Between Small and Large Platforms

Our analysis clearly demonstrates that small social-media platforms can and do act as incubators and accelerators of antisemitic content, with memes functioning as one possible vehicle for radicalisation. The dataset demonstrates that these services do not operate in isolation; instead, they form part of an interconnected ecosystem in which hateful content is created, refined, tested for engagement, and then exported to larger, higher-reach platforms.

Across 535 antisemitic memes collected from 9gag's publicly accessible "Trending" page, 99.1% were subsequently reposted on Reddit, indicating an almost complete migration pipeline from a small, high-harm service to a large mainstream platform. This "communicating vessels" effect means that even when antisemitic content originates in relatively niche spaces, its eventual reach is global. Trending algorithms on small platforms – combined with minimal moderation and active user coordination – serve as a selection mechanism for content that is then amplified on Reddit and similar platforms, where users seek to game the system.

This cross-platform movement is not random. Hateful memes are deliberately repeated, reframed, and recirculated across time and channels to maximise visibility and evade moderation. Their format – ironic humour, innuendo, ambiguous symbolism, and "just asking questions" – is designed to undermine detection systems while making antisemitic ideas socially acceptable. Many examples also weaponise plausible deniability, enabling users to spread harmful narratives whilst denying intent.

The study also shows how quickly users encounter harmful content: on average, every six memes in a casual browsing session. Once a user engages with these memes, especially through comment sections, which are almost entirely unmoderated, the likelihood of encountering further antisemitic content rises to 100%, creating a highly efficient radicalisation loop.

Furthermore, the analysis confirms a multi-locality effect: antisemitic narratives span multiple geographies and contexts, from "Great Replacement" conspiracies in Europe and the US, to Middle East-related tropes, to globalised cognitive-warfare techniques aimed at fostering uncertainty, fear, and mistrust.

Taken together, these findings reveal that the harm posed by small platforms is not constrained by their size. Their role as origin points, testing grounds, and amplifiers of antisemitic content means they exert influence far beyond their own user base. Yet under current UK and EU regulatory frameworks, these high-harm platforms face comparatively light duties and limited oversight. As long as regulatory thresholds remain tied primarily to scale, rather than risk and demonstrated harm, antisemitic content will continue to spread unchecked between platforms.

This evidence underscores the urgent need for a more robust, risk-based regulatory model – one capable of intervening at the source, constraining cross-platform amplification, and addressing the unique characteristics of memes, coded content, and coordinated reposting behaviour. These insights directly inform the policy framework that follows.

Suggested Regulatory Framework to Tackle Harm in Small Platforms

A. Designation & Core Safety Duties (Means)

1. Risk-based designation of Small High-Harm Services

Regulators should introduce a "High-Harm Small Service" category for platforms that disproportionately host or disseminate extremist or antisemitic content despite having relatively small user bases. Designation should be fast and evidence-driven, based on indicators which might include repeated circulation of extremist material, cross-platform funnelling into mainstream services, persistent takedown evasion, or heavy use of anonymity/VPNs to avoid enforcement. Evidence from civil-society

organisations, law-enforcement agencies, trusted flaggers, and platform data should support rapid escalation.

Once designated, these services must meet specific transparency requirements, including reporting volumes of extremist/antisemitic content removed, data on notices from trusted flaggers, and details on moderation capacity (e.g., human reviewers per 100,000 users).

2. Mandatory systemic risk assessments

Designated small, high-harm services must conduct and publish regular systemic risk assessments. These should map extremist content flows; identify pathways into violent extremism; analyse amplification dynamics; and evaluate how material from the platform spreads into mainstream networks.

3. Baseline proactive safety duties

Recognising that full obligations applied to major platforms may be disproportionate, small, high-harm services should still meet minimum proactive requirements. These include basic automated detection and human review for illegal content; safety-by-design measures, including friction that limits virality; default-private settings for newly created channels; and removal of algorithmic recommendation boosts for known extremist communities.

4. Shared illegal-content hash database

A mandatory, cross-industry hash database – similar in concept to StopNCII – should be established for content verified as illegal (e.g., extremist propaganda, antisemitic incitement). Both major platforms and designated high-harm small services must participate. Content stored in the database would be automatically removed on upload across all participating services.

B. Accountability, Enforcement & Incentives (Ends)

5. Public watchlist and positive incentives

Regulators should maintain a public watchlist identifying platforms repeatedly involved in distributing antisemitic or extremist material. Alongside penalties, the system should also offer incentives, for example, reduced regulatory reporting burdens or tax benefits, for demonstrable improvements in safety standards.

6. Sanctions for repeated reposting after takedown

Repeated re-uploading or reposting of content previously removed as illegal should be defined as an aggravating factor in enforcement. Platforms

must demonstrate the preventative systems they have put in place; failure to do so would trigger escalating sanctions.

C. Evidence & Research Access (Enabler)

7. Privacy-minimised data access for researchers

Designated high-harm small services should provide time-limited, safeguarded researcher access under NDAs and strict ethical oversight. This enables civil-society groups and academics to map extremist networks, track cross-platform pathways, and generate evidence for monitoring, designation, and enforcement.

D. Technical Toolbox for Platforms & Regulators

8. Coded-meme and dogwhistle identification

Investment is needed in classifier research specifically targeting coded antisemitic content, including humour-framed material, dogwhistles, and linguistic or visual innuendo that currently evade detection.

9. URL and referrer-chain tracking

Tools should be developed to detect when content originating from small, high-harm services is being reposted en masse onto mainstream platforms, allowing earlier and coordinated interventions.

10. Virality dampening

When a post or account shows patterns of coordinated cross-platform amplification, temporary throttles on reach or reposting should be deployed to disrupt rapid spread.

E. International Cooperation

11. EU-UK platforms and civil society working group

A permanent cross-border working group involving regulators, law enforcement, civil society and platforms should coordinate responses to spikes in extremist content, particularly when content spreads simultaneously across jurisdictions.

12. Modernised mutual legal assistance

Mutual legal assistance processes must be updated to enable faster evidence preservation and data sharing while protecting privacy and due process. This is essential for the timely investigation of extremist networks operating across borders.

F. Proactive & Educational Measures (Prevention)

13. Safety-by-design for users and creators

As part of a broader safety-by-design regime, small, High-harm services should implement brief and clear onboarding that explains platform rules, the illegality of extremist content, and the consequences of violating them. This can include age-appropriate examples of banned content to reduce accidental rule-breaking and increase user awareness of risks.

14. Mandatory community standards literacy for moderators

Designated platforms should train their moderators in recognising extremist narratives, coded antisemitic and other racist symbols, and cross-platform mobilisation tactics. Training should be updated regularly and audited to ensure moderators can identify emerging tropes and dogwhistles.

15. Educational counter-messaging partnerships

Platforms should be encouraged to maintain partnerships with civil-society organisations that specialise in antisemitism, online extremism, or digital literacy.

16. Youth-targeted digital resilience programmes

Many high-harm small services tend to attract adolescents or young adults, as such, regulators should require platforms to integrate age-appropriate digital literacy prompts, including recognising extremist recruitment content and patterns, evaluating the validity of sources, and resisting pressure to share harmful memes or join extremist groups. However, the majority of media literacy efforts should be carried out on a national level in educational institutions.

Conclusion

Antisemitism is no longer confined to fringe online spaces. As this report demonstrates, small but high-harm services play an outsized role in creating, refining, and exporting hateful content into the mainstream digital sphere, where it reaches millions and contributes to real-world harm. Memes, with their blend of humour, ambiguity, and virality, have become one of the most effective vectors for spreading antisemitic and other hateful narratives, shaping perceptions, and radicalising vulnerable users.

The current regulatory regimes in the UK and EU do not adequately account for this reality. By focusing primarily on platform size rather than demonstrated risk, they leave significant blind spots in the online safety system gaps that extremists exploit. Meanwhile, the ease with which content flows from small to large platforms ensures that harm cannot be contained at its point of origin.

A more agile, risk-based approach is essential. Regulation must reflect how antisemitism is produced and replicated online: through coded memes, coordinated reposting, cognitive warfare, and cross-platform pipelines. The policy framework outlined in this report offers a path to more effective oversight – one that strengthens enforcement, improves transparency, enhances technical capability, and embeds preventive and educational measures across the ecosystem.

Protecting Jewish communities and safeguarding democratic values more broadly requires confronting the harms that begin on the smallest platforms but reverberate across the entire digital landscape.

Endnotes

- <https://antisemitism.org.uk/wp-content/uploads/2020/08/Online-Harms-Offline-Harms-August-2020-V4.pdf>
- <https://committees.parliament.uk/writtenevidence/19538/html/>
- <https://cst.org.uk/research/cst-publications/antisemitic-incidents-2024>
- <https://fra.europa.eu/en/news/2024/jews-europe-still-face-high-levels-antisemitism>
- <https://www.bbc.co.uk/news/articles/c147w9572dvo>
- <https://techagainstterrorism.org/news/2019/04/29/analysis-isis-use-of-smaller-platforms-and-the-dweb-to-share-terrorist-content-april-2019>
- <https://www.ofcom.org.uk/siteassets/resources/documents/consultations/category-1-10-weeks/240428-first-phase-of-online-safety-regulation/responses/tech-against-terrorism?v=202111#:~:text=Tech%20Against%20Terrorism%20research%20shows%20that%20terrorist,resources%20to%20understand%20and%20tackle%20the%20threat.>
- https://www.isdglobal.org/digital_dispatches/rise-in-antisemitism-on-both-mainstream-and-fringe-social-media-platforms-following-hamas-terrorist-attack/
- https://www.isdglobal.org/digital_dispatches/rise-in-antisemitism-on-both-mainstream-and-fringe-social-media-platforms-following-hamas-terrorist-attack/
- <https://antisemitism.org.uk/wp-content/uploads/2024/12/Antisemitism-Goes-Viral.pdf>
- <https://antisemitism.org.uk/wp-content/uploads/2024/12/Antisemitism-Goes-Viral.pdf>
- https://www.ofcom.org.uk/_data/assets/pdf_file/0019/271243/volume-2-illegal-harms-consultation-1.pdf p.22
- A term used in the UK to describe content that is not illegal but carries a high risk of physical or psychological harm. It is used to refer to things like content that glorifies eating disorders, self-harm, or certain types of misinformation, especially as it impacts children. The Online Safety Act addresses this through specific duties, especially for children's safety, by requiring platforms to protect them from certain "primary priority" content that is identified as harmful, such as content related to suicide and self-harm, eating disorders, and abuse, even if it is not illegal. This can also include antisemitic content that does not cross the legal threshold into hate speech, but can radicalise into more extremist ideologies, cause psychological distress and facilitate hate crimes offline.
- The legality of specific content depends on jurisdiction. For example, Holocaust denial is not illegal in the UK in most cases, but is illegal in some European countries.
- https://www.ofcom.org.uk/siteassets/resources/documents/consultations/category-1-10-weeks/270826-consultation-protecting-people-from-illegal-content-online/responses/community-security-trust-and-antisemitism-policy-trust_redacted.pdf?v=375000
- <https://antisemitism.org.uk/wp-content/uploads/2024/12/Antisemitism-Goes-Viral.pdf>
- <https://cst.org.uk/data/file/2/4/We%20are%20Generation%20Terror.1639414296.pdf>
- <https://cst.org.uk/news/blog/2020/06/11/hate-fuel-the-hidden-online-world-fuelling-far-right-terror>
- <https://cst.org.uk/data/file/2/4/We%20are%20Generation%20Terror.1639414296.pdf>
- <https://antisemitism.org.uk/wp-content/uploads/2021/09/Instagram-Report.pdf>
- <https://antisemitism.org.uk/wp-content/uploads/2020/08/Online-Harms-Offline-Harms-August-2020-V4.pdf>
- https://home-affairs.ec.europa.eu/policies/internal-security/counter-terrorism-and-radicalisation/prevention-radicalisation/terrorist-content-online_en
- <https://icct.nl/publication/right-and-left-wing-violent-extremist-abuse-digital-technologies-global-south>
- <https://antisemitism.org.uk/wp-content/uploads/2020/08/Online-Harms-Offline-Harms-August-2020-V4.pdf>
- <https://antisemitism.org.uk/wp-content/uploads/2020/06/Web-Misogyny-2020.pdf>
- <https://cst.org.uk/news/blog/2020/06/11/hate-fuel-the-hidden-online-world-fuelling-far-right-terror>
- <https://cst.org.uk/news/blog/2020/06/11/hate-fuel-the-hidden-online-world-fuelling-far-right-terror>
- <https://www.onlinesafetyact.net/analysis/safety-by-design/>
- <https://www.legislation.gov.uk/ukpga/2023/50>
- <https://www.independent.co.uk/voices/online-safety-act-labour-safer-internet-day-b2695966.html#:~:text=Yet%2C%20this%20measure%20is%20doomed%20to%20fail,dangerous%20than%20those%20hosted%20in%20the%20UK>
- <https://www.taylorwessing.com/en/interface/2022/the-eus-digital-services-act/digital-services-act---an-overview#:~:text=Small%20companies%20and%20micro%20enterprises,of%20providers%20of%20intermediary%20services>
- <https://www.euronews.com/next/2024/01/18/dutch-watchdog-seeks-big-tech-input-for-platform-guidelines#:~:text=Any%20business%20providing%20digital%20services%20to%20Europeans,will%20face%20fewer%20obligations%20than%20the%20largest>
- <https://kennedyslaw.com/en/thought-leadership/article/2025/caught-in-the-net-understanding-the-eu-digital-services-act-s-regulation-of-user-interactivity/>
- <https://www.businessinsider.com/reddit-comments-ai-training-models-google-openai-jen-wong-huffman-2025-1>
- <https://www.adl.org/resources/press-release/adl-and-reddit-help-moderators-confronted-antisemitism>
- We have been notified by Reddit that some of the Subreddits from this report have been banned.
- <https://forward.com/opinion/575722/chabad-tunnels-antisemitism-twitter>
- https://en.wikipedia.org/wiki/Kalergi_Plan
- <https://www.spicenter.org/resources/reports/reddit-shuts-down-some-racist-anti-semitic-web-forums>
- <https://www.timesofisrael.com/after-9-months-reddit-finally-bans-group-spreading-thinly-veiled-anti-semitism>

Contact APT



www.antisemitism.org.uk



@antisempolicy



Antisemitism Policy Trust



mail@antisemitism.org.uk