

December 2021

SOCIAL MEDIA PLATFORMS AND THE DANGERS OF OUTLINKING

Dr. Bethan Johnson and Dr. Bàrbara Molas

*The Centre for Analysis of the Radical Right
&
Academic Consulting Services Ltd.*



Academic
Consulting
Services Ltd.

Table of Contents

Introduction	2
Report Hypotheses	4
Methodology and Case Studies	5
Facebook	9
Recommendations for Facebook	21
Twitter	22
Recommendations for Twitter	25
Reddit	26
Recommendations for Reddit	29
YouTube	30
Recommendations for YouTube	35
Reassessment of Stated Hypotheses	36
Conclusions and Recommendations	37
General Recommendations	39
Appendix 1: On the Lookout for Outlinking	40

Introduction

Social media companies have a vested interest in positively connecting people around the world. However, for years, leading platforms have seen their good intentions hijacked by unwanted terrorists exploiting their services. Relying on a mixture of human moderators and artificial intelligence (AI), Facebook and other platforms have made notable progress in identifying, blocking, and otherwise moderating extremist and terrorist content that violates community guidelines. Yet, in finding new ways to exploit hate, extremists have been learning as well, restlessly seeking to discover loopholes to advance hateful ideologies. For the radical right in particular, a key method for circumventing these policies and content-blocking technology is known as ‘outlinking,’ or hyperlinking for extremist purposes. Outlinking makes hate just one click away.

Outlinking means that, instead of directly posting words and images to communicate extremist content (which can be more easily detected), users may share links to webpages of ‘weak’ sites in order to spread radical propaganda. These weaker sites either lack the resources to identify and bar extreme content, or have more lax policies relating to content moderation. As a consequence, extremist actors post links on larger platforms to disseminate material that would otherwise be removed. COVID-19 has only deepened this already noticeable wound. Generally speaking, outlinking relies upon a reader’s willingness to click through to more obscure and/or extreme sites online, with users potentially falling down proverbial ‘click rabbit holes’ of undiluted extremist content. Worldwide lockdown restrictions have provided billions of people, especially youth, greater means and opportunity to do this, while COVID-related feelings of isolation and disconnection may provide an increased motive ([Molas, 2021](#)). With more time spent online, both for work and for the maintenance of social interaction and community building in times of crisis, the opportunities for contact with extremist actors has grown ([Bar-On, Molas, 2020](#)). As a consequence, the possibilities for hate speech and offline violence have dramatically increased. From terrorist attacks in Germany and India to anti-mask protests across the U.S. and Europe, radical-right violence continues to escalate in the midst of circulating conspiracy theories and disinformation online ([Leidig, 2021](#)). Thus, studying the use of outlinking among the radical right can help improve our understanding of gateways to extremist, and even terrorist, networks.

For all the positive interactions and networked communities Facebook has nurtured, hate-based groups have likewise gained traction on the platform. Yet despite outlinking’s increased use and serious ramifications, research on this tactic remains

extremely limited. Although the first scholarly recognition of this technique dates back to [2018](#), both academic and public-sector understandings have remained woefully underdeveloped. Critically, existing studies of outlinking have not focussed upon employment by the extreme right; as usual, the earliest discussions on the practice being were linked to religiously inspired terrorist groups such as ISIS (Lakhani, 2018), before being expanded to include comparative works between the two forms of terrorism (Heller, 2019; Macdonald *et al.*, 2020). More recently, observers have now trained their attention exclusively on the radical right, although most of the work is as yet unpublished ([RESOLVE](#) Network, 2021).

Given the paucity of empirical materials, research to date is still quite theory driven. In response, this report presents actionable original research on radical-right outlinking using illustrative examples and a comparative approach, contrasting the use of the technique on multiple platforms. Our work also provides realistic recommendations on how to prevent the use of external links for campaigns that violate community standards (if not criminal laws), while the research behind the report can also serve as the basis for a toolkit addressing extremist outlinking on Facebook.

Report Hypotheses

1. Outlinks feature patterns in their URLs that may allow for algorithmic detection. The majority of outlinking is concentrated on a few websites.
 - Anticipated recommendation if this hypothesis is supported: incorporate flagged websites into existing databases for additional vetting *vis-à-vis* hate speech, radical online networks, and extremist propaganda.
2. A large proportion of those producing outlinks are individuals with affiliations to proscribed radical-right groups.
 - Anticipated recommendation if this hypothesis is supported: monitor all legal outcomes for strategic engagement in litigation and research. Likewise, incorporate government designations of terrorist entities and hate groups.
3. Outlinking is a strategy used to help promote hate-filled narratives, but also to instil a sense of community and belonging amongst those disseminating outlinked extremist material.
 - Anticipated recommendation if this hypothesis is supported: identify who is most susceptible to extremist narratives and seek advice on counter-narratives to hate from knowledgeable parties (policy practitioners, think tanks, former radicalised individuals); promote online programmes designed to fight isolation and support inclusion and support liberal democratic values.
4. Anonymity on social media platforms contributes to outlinking.
 - Anticipated recommendation if this hypothesis is supported: generate new requirements about users providing their full, legal names (potentially as shown in national identification documents or the like) during account registration or update.

Methodology and Case Studies

Outlinking allows extremists to propagate hate-filled messages online, and yet the covert nature of this technique hinders easy close readings. Nevertheless, this report employs a research methodology that allows for a clear understanding of the phenomenon. It uses Open-Source Intelligence (OSINT) methods to draw outlinking out from the level of abstraction and minutia to show its role in promoting otherwise banned content (such as violence and terrorism). It does so through the implementation of a unique research approach, a hybrid of elucidatory case studies and cross-platform comparisons. Patterns and discontinuities in outlinking that violate corporate community standards are highlighted across four platforms: Facebook, Twitter, Reddit, and YouTube.

While there are many faces of the radical right, this report focuses upon those promoting racial, religious, and gender-based hatred, as well as those who promote violence and/or a violent worldview. Specifically, the work explores networks that primarily articulate white-supremacist, male-supremacist, neo-fascist, accelerationist, Islamophobic, and/or antisemitic views. The subjects were chosen to highlight the range of radical-right narratives circulating today, as well as formats (this study analyses group pages rather than individual accounts), activity levels, and followings on the various platforms. Moreover, efforts were made to select subjects whose actions could be tracked across multiple sites in order to examine whether radical actors approach platforms differently.

Another central element to our selection criteria is endorsement of violence and/or ideological links to government-designated terrorist organisations. In the case of the former, this report includes sites encouraging audiences to discard political passivity and embrace hate-based activism, including terrorism. Further, accounts have been qualitatively selected to *mirror* the narratives and activities of official designated terrorist entities such as Atomwaffen Division, The Proud Boys, and The Base. While government designation may expedite the process of removing content with clear links to a terrorist group ([Feldman, Johnson, 2021](#)), other accounts can sidestep these strictures by using covert rhetoric to express the same sentiments as stated by official terrorist entities ([Molas, 2021](#)). This report seeks out groups that have been overlooked by current moderation mechanisms despite promoting extremist or terrorist ideas.

Format was an additional consideration, including comparing public and private groups, or group pages versus individual accounts. In the first instance, we limited the work to public groups for two reasons: time constraints (entering private groups requires the establishment of a sock puppet able to withstand scrutiny from administrators); and second, public group usage of outlinks better illuminates the broad spectrum of radicalising content. Whereas private groups usually serve as echo-chambers for the already radicalised, public groups typically attempt to influence both radicalised and non-radicalised audiences. In the case of evaluating group and individual accounts, we elected to consider a mixture of the two (though weighted towards groups) in order to investigate whether propagandists are better at utilising outlinking in different accounts.

It also warrants noting that this report’s definition of outlinking does *not* include cross-posting from other extremist pages. Were it to have included such ideological cross-pollination, the figures would have been significantly higher than they are. Much like the issue of public vs private groups, the topic of shared content between pages warrants additional explanation. This content was seemingly used to open the door to more vitriol and outlinks on secondary sites, in effect pushing users further into extremist echo-chambers. A strong example of this would be **Reject Degeneracy, Embrace Tradition** ([@RDETAmerica](#)), which appears to collate biased, hateful, or illiberal content from other accounts before propagating them to more than 40,000 followers.



Ultimately, these conditions led to the selection of 9 different extremist entities. Having met the basic criteria *vis-à-vis* ideology and promotion of violence, we selected these specific groups because they were widely cited and/or extremely active online, with particular notoriety within radical-right circles. As such, we felt that if we found these entities utilised outlinking, they could be shown to present a serious threat to society both via the specific outlinking that they utilise and for the outlinking that other sites seeking to emulate them might likely do.

With respect to evaluating outlinking among our subjects, this report draws largely from examining the top 50 posts of each selected account released within the last three years (01 September 2018 - 28 August 2021) - that is, the posts per account generating the highest amount of engagement (and presumably the largest impact). Some of the accounts examined in this report have been active for many years and/or have a propensity to release dozens of messages a day, meaning that the entire activity log for each group would have produced an avalanche of data. This would have hindered the potency of the findings, as sheer volume would have swamped important qualitative nuances. By spotlighting the top 50 posts per account per case study, this report can assess the extent to which outlinking is a *popular* technique, both in relation to the author and the audience.

While the authors have looked beyond the top 50 posts in order to understand the broader context of the individuals and groups in question, we selected the three-year timeframe due to a mixture of practical and scholarly considerations. This time frame ensures that findings provide sufficient evidence of the phenomenon before and during the COVID-19 pandemic. In the design phase, we hypothesised that the realities of national lockdowns, restricted movement, and social distancing may have impacted the behaviour of radical actors online, and thus wanted to ensure robust conclusions pertaining to both pandemic and non-pandemic conditions.

With regards to structure, this report evaluates outlinking through subsections on each of the selected platforms, allowing for comparisons between accounts within a single platform, as well as additional sections highlighting cross-platform trends. Each section also begins with an elucidation of existing community standards and corporate policies of each platform, making clear how outlinking simultaneously subverts censorship and promotes hate-based content. From there, each section features quantitative and qualitative analysis of outlinking. In addition to each section concluding with platform-specific recommendations *vis-à-vis* outlinking, this report's final section provides recommendations that may be implemented by platforms in order to better tackle the phenomenon of outlinking.

Finally, this report features an appendix in the form of a spreadsheet that chronicles the essential characteristics of the top 50 posts of each of the selected accounts. Divided into subsections based on platform, information provided includes not only whether a post includes outlinking (and, if so, where), but also the date of the post, tags, likes, comments, views, and the topic(s) discussed in the given post. Moreover, at the end of each account's analysis is a brief 'meta-analysis,' wherein additional observations are made about the frequency of outlinking and commonly outlinked sites.

Facebook

In recent years, Facebook has devised a complex ecosystem of [community standard policies](#). With regards to radical-right extremism and terrorism, the pertinent regulatory policies as articulated by Meta largely fall within the ‘Violence and Incitement’ and ‘Dangerous Individuals and Organisations’ categories. Using a three-tiered system of classifying and addressing content, these sections’ language focuses on mitigating “real-world harm” (or “offline harm”).

The ‘Violence and Incitement’ regulations [pledge](#) to take down content that provides detailed explanations of weapon-building; encourages “high-severity violence” and/or outright pledges to commit “high-severity violence.” Meanwhile, the ‘Dangerous Individuals and Organisations’ policy [opens with](#) “we do not allow organizations or individuals that proclaim a violent mission or are engaged in violence to have a presence on Facebook.” In both instances, threat credibility is an essential component in the moderation decision, while demonstrated links between on- and offline hate seemingly expedite the moderation process.

‘Blanket bans’ are imposed upon those whose real-world impacts lead to their assessment as dangerous individuals or groups. Alongside bans on users outright, content moderators regulate posts they characterise as proffering “praise, substantive support, and representation” of any such dangerous operator, and/or support for some of their foundational ideas. These traits include actions such as: sharing video/photos/audio related to terrorist attacks or otherwise or endorsing them; mass casualty events, or hate crimes; backing “ideologies that promote hate, such as nazism and white supremacy”; encouraging material support for a dangerous entity; facilitating recruitment, such as serving as a contact person for a dangerous entity; and/or expressing narratives or beliefs deemed to be “[l]egitimizing the cause of a designated entity by making claims that their hateful, violent, or criminal conduct is legally, morally, or otherwise justified or acceptable.”

Those classified in lower tiers of extremity (promoting violence against non-civilians or advancing certain conspiracy theories) are also subject to moderation, particularly given repeat violations. Additionally, new “warning screens” can initially shield content from the user (such as graphic content), behind a statement identifying potential misinformation, while providing reporting mechanisms for fellow users. These features

can be seen most recently with regards to COVID-19 misinformation. Finally, Facebook's [Oversight Board](#) assist in the appropriate adjudication of content-moderation decisions.

The results of these policies are clear. In raw data, Facebook reported a rising number of responses to terrorist content: **9.7 million pieces between July and September 2020** as compared to **5.2 million the year before**. Moderation of organised hate, likewise, peaked between January and March 2020 at just under 5 million posts before levelling around 4 million posts in the following two quarters. In 2020, Facebook [increased the amount of content removal](#) on hate speech by 400%, with **97.1% of designated organised hate content removed automatically without it being reported by users**. Further, in relation to terrorism, Facebook's internal figures show that as of Q2 2021, terrorism accounts for less than 0.07% of all instances of alleged policy violations, while **moderators have been responsible for flagging at least 99.6%** of such content since Q2 2020.

In spite of these figures, Facebook has received criticism for its handling of radical-right extremism. For instance, in April 2021, Australian civil rights group Muslim Advocates sued Facebook and its top executives over anti-Muslim hate speech. They argued that Facebook has misled the public about how safe the platform is. [In the words of Muslim Advocates](#): "Facebook dupes lawmakers, civil rights groups and the public at large when it makes broad claims that it removes content that spews hate or incites violence and yet does not." The complaint, filed in US federal court, argues that this "failure has amplified the volume of anti-Muslim hate bombarding Facebook users." Their main contention was that "Facebook only takes action when the community does the heavy lifting in documenting the violations and is prepared to escalate through media." In Germany, by contrast, Facebook's commitment to swiftly taking down radical-right content has been such that, in July 2021, [the Federal Court ruled that Facebook acted illegally](#) in taking down racist posts and blocking the account of their author because the social network failed to inform the user.

Perhaps the most high-profile critique of these policies occurred in relation to Facebook's approach to the ban of former President Donald Trump. After initially removing Trump on 07 January 2021 in response to his comments on the Capitol Insurrection the previous day, Facebook's Oversight Board reduced the ban to [a period of no less than two years](#) (with the option to extend following expert evaluations presented to the Board). The initial ban *and* its subsequent reconsideration have drawn ire from both ends of the political spectrum, with supporters of the former President stating that any ban is unjustified and a restriction on Trump's speech rights, whereas [critics](#) state that no user should be above Facebook's rules (an accusation derived from

Facebook's statement that speech by political figures is inherently newsworthy). Going forward, the potential lifting of the ban before the 2024 election leaves Trump able to use the platform to rally supporters for another White House run.

More broadly, one of the most common sources of criticism has been Facebook's definition of hate speech itself, which undergirds much of its overall [moderation policies](#). Community standards ban attacks on *people* based on protected characteristics: religion, race, gender, or sexuality. This, however, may not get hate off Facebook. For example, a user would be free to attack the idea of homosexuality as long as members of the LGBTQI+ community as a group in itself was not being targeted. This is because Facebook defines hate speech as "a direct attack against people—rather than concepts or institutions—on the basis of what we call protected characteristics: race, ethnicity, national origin, disability, religious affiliation, caste, sexual orientation, sex, gender identity and serious disease." This definition leaves ample room for propagandists seeking loopholes. Take Islamophobia as another instance. Theoretically, Facebook moderation policies bar radical-right groups promoting Islamophobia in the form of violence against Muslims as a function of their faith. However, policy gaps permit the publishing of many common refrains used by right-wing extremists to justify those Islamophobic ideas that eventually lead to violence. Apparently not constituting a "direct attack[s] against people," users could, and do, claim that Islam is anathema to Western civilisation; that Islam is an inherently and fundamentally violent religion; and/or that Islam is waging a subtle jihad, with Muslim migrants and refugees coming to Europe to take over and establish a caliphate (i.e. [The Great Replacement Theory](#)). While stopping short of calling for violence, these eliminationist ideas are foundational for radical-right violence. The Great Replacement Theory was a driving force behind the 2011 attacks in Norway, the 2018 Tree of Life synagogue attack, and the 2019 El Paso shootings, leaving dozens dead or injured. It was even so central to the 2019 Christchurch attacker that he named the manifesto used to justify the murder of 51 Muslims at Friday prayers, *The Great Replacement*.

Making matters more complex, Facebook has been accused of lacking transparency regarding its policies. Both the community standards and means of enforcing them have been identified by several organisations—including the [Brennan Center for Justice](#), [The New Yorker](#), [The Brookings Institute](#), [The Georgetown Law Journal](#), and [New York University's Stern Institute for Business and Human Rights](#)—as being either unclear or misguided. Further investigations have contended that current moderation standards fail to sufficiently approach the issue, with headlines in [Forbes](#) and [BuzzFeed News](#) among those claiming that Facebook's policies lead to daily regulatory mistakes or are subject to the "whims" of executives and a limited number of other high-ranking

employees. [Misinformation](#) is a big part of this problem, according to a group of academics from NYU Ad Observatory who researched ad transparency and the spread of misinformation on Facebook. The group aimed to uncover who paid for political ads and how they were being targeted, but Facebook banned the personal accounts of these researchers for violating the platform's terms of service by "scraping user data without permission". Laura Edelson, an NYU researcher involved in the project, claimed that Facebook "wants to end independent scrutiny of its platform". The company, she argued, lacks transparency and its terms of service are far from clear. In sum, there is a [concern](#) about a lack of "active and informed consent for mass surveillance of Internet users" via Facebook. Based on such criticism, Facebook promised to make it easier to track political ads on its network, and it recently [released a new content transparency report](#) to advance this cause. Nevertheless, experts say that Facebook's [efforts have failed](#) to change their minds.

Meanwhile, advocates are concerned about how Facebook utilises and shares users' information, as well as about company policies on the regulation of information flow across its platform. For example, human-rights groups in Vietnam, Sri Lanka, Myanmar, and the US [have called on Facebook](#) to work more directly to combat online hate speech while ensuring control over users' data. There has been a perception of [inadequate stakeholder engagement on Facebook's part](#). Generally speaking, there are concerns that there is not enough information available to make informed choices when using the platform; that is, without exposing users to "[undisclosed risks](#)."

That Facebook is now acting faster and detecting more content, and yet is still not effectively communicating these realities to the public, suggests not only that safety policies may need improvement, but transparency as well.

Within this landscape, the use of outlinking by extremist accounts on Facebook may be described as highly variable, in the case studies ranging from almost **three quarters of top posts to** being **entirely absent**. The most prolific user of outlinks on Facebook studied herein is **Arktos (@Arktos)**. The account is eponymous, as the name of the official Facebook page (managed, according to Facebook transparency information, by individuals in Poland, Hungary, and Andorra) is also that of the company to which it is linked. Arktos is arguably the foremost publisher of radical-right extremist texts globally. Since opening its doors a dozen years ago, it has published almost 200 texts in several languages. Its [website](#) asserts that it "does not seek to propagate any specific ideology, system of beliefs or viewpoint", but that it focuses on translating and publishing English editions by the European "New Right." In their own words, they "want to provide a voice for individuals and viewpoints who are often overlooked by the mainstream...We view

ourselves as offering fodder for the minds of those who envision a world that is different from the one we inhabit today, but are uncertain of exactly what shape it will take.”

In fact, it praises well-known social Darwinists, Islamophobes, fascists, and neo-fascist sympathisers. Titles it publishes include: [*Breached Enclosure: Why the West Is Being Defeated by Islam but Might Still Come Out Okay*](#); [*Return of the Swastika: Hate and Hysteria versus Hindu Sanity*](#); [*Ethnic Apocalypse: The Coming European Civil War*](#); [*Prelude to War: Chronicle of the Coming Cataclysm*](#); [*A Handbook for Right-Wing Youth*](#); and [*Why We Fight: Manifesto of the European Resistance*](#). As these titles may indicate, Arktos publishes works that set out pseudo-intellectual arguments in favour of ethno-separatism, if not white supremacy and Islamophobia, as well as many more that promote ‘traditional’ living (e.g. traditional gender roles, sexual and gender expressions) and countercultural philosophies ([Richards, 2019](#); [Valencia-García, 2020](#)).

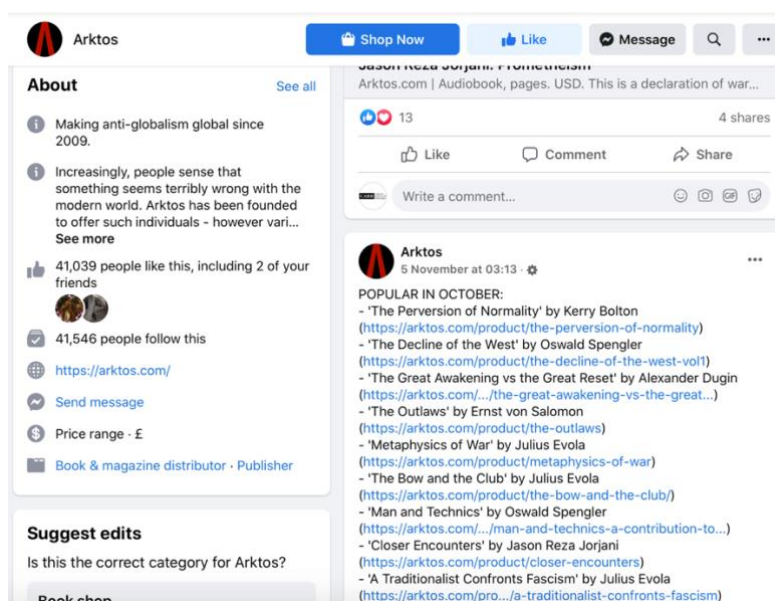
Arktos’ Facebook page has more than **41,000 likes and almost 42,000 followers**, as well as a relatively active, global ‘community page’ where other accounts can connect with each other. These figures, it should be noted, do not include the various pages Arktos manages for the individual texts it publishes, which each have as many as almost 3,300 likes and another 3,500 followers. As a small publishing house selling niche texts, Arktos has developed a significant social media following.

Arktos is extremely active on Facebook, posting several times a day. Outlinking represents a substantial enterprise for the group, as **72% of their top 50 posts in the last three years have included at least one outlink**. For Arktos, outlinking is becoming an increasingly popular tactic, as 27 of these top 50 (54%), and 17 of the top 25 (68%) posts both use outlinks *and* were posted within the last year. The popularity of these posts, it should be noted, is evidenced by hundreds of likes, shares, and comments.

Furthermore, patterns indicate that Arktos employs outlinking to push people towards their website, to simultaneously advertise specific publications, and to valorise the radical right. The preponderance of its outlinking indicates that some **88% of those posts in the top 50 that use outlinking are devoted to directing traffic to its official website and/or advertising individual books Arktos publishes**. Arktos rarely outlinks without this dual ideological-enterprising function. Several of the top posts discuss the lives and works of such figures as [Julius Evola](#) and [Guillaume Faye](#)—well-known pseudo-intellectuals promoting fascism, Islamophobia, and white supremacy—along with lesser-known figures including Dominique Venner and Pentti Linkola. Even in these posts not directly linked to a bookselling webpage, Arktos is nonetheless supporting its business by attempting to popularise these authors. In short, these posts and Arktos’

presence on Facebook overall are used to legitimise and normalise the hate-based narratives of these and other thinkers.

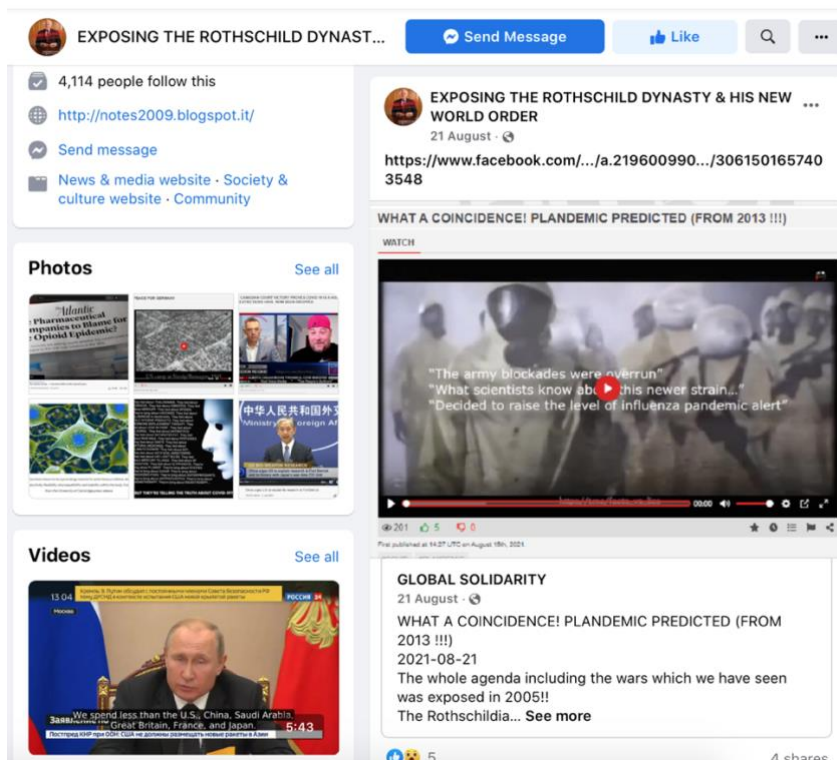
While it may be commonplace for vendors to use their social media accounts to drive readers to their websites, in the case of Arktos (and others like it), this action has an additional concerning effect. Outlinking amounts to monetisation:



Recently, many social media platforms and companies have attempted to distance themselves from violent and/or terrorist actors by demonetising them. Such efforts have included barring the use of ads on certain videos, refusing to host crowdfunding campaigns linked to extremists, and preventing terrorist-affiliated individuals or organisations from setting up mainstream bank accounts. Outlinking, in essence, is the opposite of these efforts. While not directly fundraising on Facebook, by using the platform to validate certain authors and ideas, and then selling books legitimating right-wing extremism, Arktos covertly develops a strong revenue stream. If even a fraction of the more than 40,000 followers or the several hundred users who engage with a given post go on to purchase an advertised book, Arktos has gained the capital it needs to stay afloat and [spread its hateful messages](#).

The second most prolific user of outlinking among the selected case studies is **EXPOSING THE ROTHSCHILD DYNASTY & HIS NEW WORLD ORDER** ([@exposingrothschild](#)). Categorised as a “News & media website”, the page has more than 4,000 followers and almost 3,900 likes. First established in 2014, according to Facebook’s transparency information the group’s managers are mostly from Italy, the US and France, though there are additional managers in South Africa, Germany, Iceland, Puerto Rico, and

Turkey—all of which indicates just how international such conspiracist groups can be. The page hosts a lively, multilingual community replete with cross-posted photos and memes from other extreme accounts, and misinformation about such topics as COVID-19 and immigration. Here, page managers frequently use outlinks to communicate their underlying ideology.



As the name EXPOSING THE ROTHSCHILD DYNASTY & HIS NEW WORLD ORDER indicates, this ideology is fundamentally antisemitic and conspiratorial. Claiming that the Rothschilds are responsible for a ‘new world order’ is part of a larger, well-worn conspiracy theory (AJC), which claims the Rothschilds are a cabal aiming to exploit, harm, and/or eradicate white Europeans. The selection of the Rothschilds as the source of this ‘ZOG’—or ‘Zionist Occupied Government’—dates back decades. As a family of prominent Jewish bankers with close ties to European governments for centuries, the Rothschilds have been a longstanding metonym for Jewish greed and manipulation, significantly predating the arguably now more well-known [George Soros narrative](#). Although a historic family is presented as the main enemy, the page imparts new energy into this and other antisemitic conspiracies by promoting derivative outlinks claiming, for instance, that COVID-19 was planned by such elites, or that vaccines are part of a plot by ‘Zionists’. Even beneath the layers of clear antisemitic narratives, the most basic feature of EXPOSING THE ROTHSCHILD DYNASTY & HIS NEW WORLD ORDER is a belief that secret elites are controlling the masses.

Outlinking fits perfectly into this schema. For example, many of posts on this page relating to COVID-19 have been flagged in some way (typically with a banner discussing more accurate articles on the topic). It makes sense, then, that **34 (approximately 68%) of the top 50 posts** from the last three years have utilised outlinking. Further, the account relies heavily upon a specific outlinking strategy. While Facebook posts can include multiple links—and, as evidenced in other accounts surveyed as potential subjects, they often do—the links here almost exclusively include only one link per post. Collectively, these links guide readers to any array of websites promoting hate and misinformation, and only very infrequently are there links to articles by mainstream, respected news outlets. With respect to engagement, the audience tends to interact with the text via sharing rather than commenting or liking. The differences between likes and shares is stark—in one instance, an outlinking post that received only 23 likes was shared 422 times; in another, the rate was 25 to 364.

More specifically, the content of outlinking on EXPOSING THE ROTHSCHILD DYNASTY & HIS NEW WORLD ORDER, consistent with the account's stated interest, focuses upon alternative news and conspiracy theories. Although the group's title focuses on the Rothschilds, the family serves merely as one target amongst many 'others.' Among the most popular posts are those denouncing the United States as the most corrupt, aggressive, and amoral country in existence; those promoting conspiracy theories about politics in the Middle East; COVID-19; some utilising classic stereotypes, as well as articles about alleged Israeli atrocities in Palestine, all in order to paint a deeply antisemitic picture of society. At times the account marries several of these topics in a single article; thus, one outlink guides readers to an article entitled 'The Saudi Royal Family Are Crypto-Jews'. Still another outlink leads to a video discussing how the Rothschilds purportedly purchased Jerusalem and created Israel; still another reports that Israel is sterilising Ethiopian Jews. Although the *public* iteration of this account (there is a private group as well) has a smaller reach as compared with Arktos, it clearly demonstrates the extent to which outlinks can be used by groups purporting to be interested in news to drive audiences to read (and potentially embrace) racist conspiracy theories.



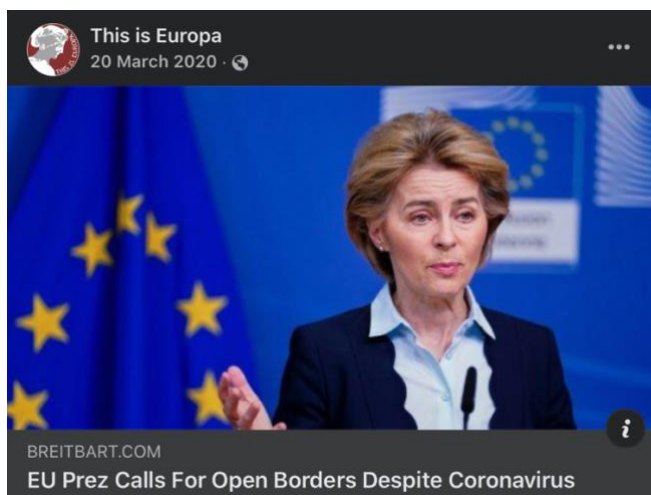
It warrants noting also that even within the most rigid readings of Facebook’s community standard policies, this account commits myriad violations. While the Rothschilds are often intended to serve as a metonym for Jewish people and/or a global elite, they are also individual people. The profile for the organisation features pictures of Jacob Rothschild, claiming that he is responsible for crimes against humanity. Meanwhile, what appears to be a **screenshot of video from the aftermath of the Christchurch attack and an outlink to the video** can be found on the page; notably, the screenshot is of a man who appears to be dead, with blood coming from his head onto the carpet as those around him scream and cry.

Similar to Arktos, **This Is Europa** ([@thisiseuropa](https://www.facebook.com/thisiseuropa)) is allegedly a ‘European based think tank’ that has garnered a significant following on Facebook. Although the account has been dormant with regards to *posting* new content for in much of 2021, in the 8 years it was active from 2012, there have been thousands of posts, including posts, videos, and photos to gain a following. This tactic has led to more than **82,000 likes as well as**

82,000 followers. Individual posts have received upwards of 1,200 likes and 195 comments. Facebook’s transparency data indicates that the account is run predominantly from Sweden, although one administrator is in Germany and another in Norway.

In its own words, This Is Europa is an [organisation](#) “highlighting the great demographic changes that have taken place in Europe over the last decades. We have tried to be a source of inspiration for people who see that our societies are moving in the wrong direction.” It claims that white Europeans who support immigration and multiculturalism, or who denounce nationalist efforts, actually secretly believe in This Is Europa’s anti-immigrant, ethno-nationalist ideas, but are too afraid to speak out. The group encourages Europeans (again, read: white Europeans) to endorse policies that align with identitarianism, racial segregation, and ethno-nationalism. Its content may be characterised as promoting racial hatred and pushing the Great Replacement narrative behind the Christchurch attack.

This Is Europa employs outlinking in 38% of its top 50 posts (19 of the 50). The Facebook page has maintained a steady usage of outlinking across the period covered by this report. Curiously, unlike Arktos, which employed outlinking to push traffic to its official website, This Is Europa did so in only 36.8% its top 50 posts using outlinking, and in only **14% of all of the top 50 posts as a whole.** In the majority of outlinking cases, much like EXPOSING THE ROTHSCHILD DYNASTY & HIS NEW WORLD ORDER, the outlinks lead to news articles or other Facebook pages – which we still consider outlinking, as it leads the reader to a different location for the purpose of spreading hate. Also, like EXPOSING THE ROTHSCHILD DYNASTY & HIS NEW WORLD ORDER, This Is Europa features a mixture of mainstream outlets, such as Reuters and CNN, though the preponderance is from more fringe content creators.





By contrast, the extremist account **The Warden Post** ([@TheWardensPost](#)) only makes limited use of outlinking, despite its influence as an intellectual vanguard of the radical right. With the motto “Think differently. Be dangerous.”, **The Warden Post** identifies its mission as uniting what they term ‘open thinkers’ to address “the converge [sic] of disasters towards forging a sustainable future.” These perceived disasters include demographic trends altering the racial and social hierarchies away from historic structures (patriarchal, Social Darwinist, white/European supremacist, heteronormative, illiberal) and towards inclusive, multicultural communities. Among the most troubling and potentially community-standard violating elements of their content is their fuelling of the Great Replacement theory that is behind so much real-world violence.

The top 50 posts in the last three years are largely drawn from ‘classic’ scholars to emphasise consistency and traditionalism. Many of The Warden Post's top posts feature quotations from figures such as Nietzsche and Plato, but also Guillaume Faye and Ernst Jünger. With an information section that links readers to The Warden Post’s Twitter, Instagram, and YouTube accounts, as well as their official website, readers are given an array of means through which to consume their racist messages. This arguably means that Facebook, by continuing to grant this and others like it a platform, cannot entirely keep track of the harm done by its existence. In short, the page can become a springboard into radical-right content on their other accounts as well as the suggested pages linked to the account.

That said, outlinking itself has not been a significant feature in the most popular of The Warden Post's posts, as between September 2018 and late August 2021 **only one of the top 50 posts included an outlink**. This 2019 post directed readers to *The Daily Mail* website and featured references to anti-establishment terrorist 'V', from the film *V for Vendetta*. The post received 133 likes and 31 shares, which is significantly smaller than the 348 likes and 919 shares of the most popular post (from 2020) from the same time period. Instead, The Warden Post appears to favour the use of hashtags on posts to generate attention, with **hashtags featuring in 28% of top posts**, sometimes with more than a dozen hashtags per post.

Whilst outlinking does not attract the greatest number of interactions with readers on any one post, it should not be discounted as having influence on subscribers to The Warden Post. For instance, when surveying posts from September 2021, all but two include outlinks, with posts sending readers to Patreon, Twitter, YouTube, and The Warden Post's official website. As The Warden Post puts out several messages per week, this leads followers to potentially engage with hundreds of outlinks over time. **The sustained usage of outlinking despite a seeming lack of popularity, suggests that those managing the account feel that it serves a useful purpose.** As The Warden Post publishes often, it may be that viewers feel obvious, direct engagement with any given post (expressed by way of like, comment, or share) to be superfluous; it is more about perpetuating an echo chamber.

In all the above-described case studies, outlinking serves as a means for an account to spread messages that Facebook might otherwise moderate – whether through banners with more accurate information, or through taking down the post – if they had been published directly on the site itself. Outlinks serve to circumvent Facebook policies, but also help generate a radical-right echo chamber, providing users opportunities to radicalise on and off Facebook.

Recommendations for Facebook

- Add authors **Guillaume Faye**, **Ernst Jünger**, and **Julius Evola** to trigger review.
- Cross-reference websites linked in a flagged group's 'About' section and observe the content on that website in light of Facebook policies.
- When a company attempts to set up a Facebook page linked to selling its product, trigger an automatic review of its website and self-professed services to gauge its propensity towards promoting content violating Facebook standards. If it appears to consistently profit on content Facebook would ban, then all outlinks on said Facebook page should be subject to automatic review. (An example of this is Arktos having a Facebook page that links almost invariably to the official arktos.com page).
- Reassess definition of hate speech in light of loopholes in distinctions between racism and conspiracy theories targeting individuals, institutions, and minority groups.

Twitter

Last year, Twitter expanded its policy barring hateful speech to include “language that dehumanizes people on the basis of race, ethnicity and national origin.” In a [statement](#), Twitter added that it “will also continue to surface potentially violative content through proactive detection and automation.” A broader definition of hate speech allowed Twitter to include tweets that [incite fear or fearful stereotypes](#) about people due to a protected category, like race. As a consequence, the company **took action on 77% more accounts between July and December 2020**. Yet over the same period, legal demands from governments to take down content also increased. Actions taken ranged from removing a single tweet to banning an entire account. Having examined the Twitter accounts of three radical-right accounts, totalling more than 153K followers, it is clear there is still much that the social platform can do to fight harmful speech, including better addressing the issue of outlinking.

The most widely-recognised influencer among the online radical right is **Richard Spencer**, an American ‘identitarian’ known for being an antisemitic conspiracy theorist and of the alt-right movement. Spencer has 71.1K followers on Twitter and a reach of 71,289 (the latter number referring to the total amount of users on the platform who are likely to see his content). Through outlinking, between 2018 and 2021 Spencer was able to spread anti-government ideas, anti-Muslim and anti-black as well as antisemitic narratives, and sentiments that promote a binary language against freedom of sexuality. Among the top 50 tweets (shared and produced) on Spencer’s feed, between **43% and 88% of the outlinking was facilitated by Twitter**, with prominent outlink sites including Youtube.com (7%); the hosting service [favoured by the radical right](#) odysee.com (5%); amazon.com (4%); or centre-left news site axios (2%). Importantly, Twitter was the main domain that led Spencer’s followers to extremist material and to becoming stars in their own rights. Indeed, retweeting enabled a large increase in popularity for linked radical-right accounts such as @RichardHanania, @krystalball, and @OdyseeTeam (all of whom we recommend be monitored and/or banned). Given that most of Spencer’s feed is composed of retweeted hyperlinks rather than original content, this report suggests that focusing on outlinking practices rather than on account feeds may enhance anti-hate performance insofar as it would allow Twitter to detect echo chambers that serially violate terms of use.



Another radical-right influencer is **Chris Roberts**, a leading figure of the white supremacist far-right movement [American Renaissance](#). With a total of 567 Twitter followers and a reach of 583, Roberts’ use of outlinking is mostly characterised by hyperlinks to American Renaissance’s official website; that is, to content promoting ‘white power’ and bias against ethnic and sexual minorities and immigrant groups. Posts to this website comprise **at least 38% of Roberts’ feed**, and it is the top domain shared (79%) by his account to redirect followers to extremist content. That said, among Roberts’ top 50 tweets, or among those that produced the highest engagement among Twitter users, 54% of the outlinked content was enabled by Twitter itself via retweets. This was followed by 38% of outlinks leading to American Renaissance’s website; 6% to Gab; 2% to Babylonbee, a right-wing news page self-described as ‘Christian Satire’; and another 2% to UNZ, a radical-right website promoting antisemitism, Holocaust denial, conspiracy theories, and white supremacy. It follows that, among the top 50 posts by Roberts containing outlinks, **50% directly lead to extremist sites that either promote or enable white supremacy rhetoric**. More specifically, all these sites are characterised by; a sense of (white) victimhood; racism; the promotion of inequality (e.g., claims about that people are not equal); homophobia; and misogyny. All of these messages constituted part tweets incorporating outlinks that led to elaborated hateful texts. In other words, outlinking allows Twitter users to enter the radical-right stream of online hate speech by failing to recognise the impact of hyperlinks and the capacity for radical-right networks to use various tactics to circumvent terms of use.

Representing the official Twitter page for an organisation rather than an individual, the aforementioned **This Is Europa** (@[ThisisEuropa](#)) has pushed its narrative on the platform since April 2013 (although only posting sporadically between 2014 and 2018). The account has almost 3,000 followers. On Twitter, the organisation is more up-front about its racial leanings than it is on Facebook: “This is Europa is an independent pro-European online think tank that highlights anti-white policies.” Suggestive of outlook, among the accounts this account follows are Marine Le Pen, Nigel Farage, Tucker Carlson, Dr Jordan Peterson, and various members of fringe right-wing parties such as Ressamblent National and Alternativ för Sverige.



The account is, admittedly, sparse compared This Is Europa’s other social media platforms. Yet some of their tweets have considerable traction. For example, the top performing tweet has almost 3,400 retweets and more than 3,100 likes. Outlinking features heavily in the group’s tweets. Of those sent during the period in question, **85.7% featured outlinks, with each of the top five most popular having outlinks.** The outlinks frequently direct readers to other This Is Europa content, but also at times to other fringe sites. The most popular tweet by the group—with over 3,000 accounts engaging in both liking and retweeting—utilises an outlink to the site voiceofeurope.com, to a piece about UN migration policies. This Is Europa provides outlinks on topics including COVID-19 in Europe; the perceived desire for ‘remigration’ (or forced deportation of non-whites) in Europe, the ‘Great Replacement’ theory, and various alleged instances of anti-white bias in white-majority countries in Europe. This is Europa uses outlinks to present a declinist narrative about the state of white people in Europe, advancing a sense that if policies do not change shortly, white societies of Europe will be subdued by non-white peoples.



Recommendations for Twitter

- We recommend **monitoring the following accounts**, which tend to promote anti-government, misogynist, racist, and white supremacist narratives:
 - @TheBabylonBee
 - @alth0u
 - @ggreenwald
 - @ColumbiaBugle
 - @ThottonMather
 - @DarrenJBeattie
 - @VDAREJamesK
 - @RichardHanania
 - @krystalball
 - @OdyseeTeam
 - @ThisisEuropa
- **Ban Richard Spencer** from Twitter to reduce the impact of white supremacist content. Spencer's account reach is more than 70,000 people.
- **Ban Chris Roberts' account** from Twitter to reduce the impact of damaging content, especially that related to the American radical-right group American Renaissance.
- **Block outlinks and/or label tweets** leading to the following sites:
 - archive.nc-f.org (American Renaissance blog, distributing radical-right propaganda);
 - gab.com (chat service that facilitates the spread of radical-right propaganda);
 - unz.com (the website promotes antisemitism, Holocaust denial, conspiracy theories, and white supremacist material)
 - babylonbee.com (site that claims to distribute Christian 'satire' while promoting white supremacist, homophobic, and xenophobic narratives);
 - Odysee.com (hosting service for videos that has been favoured by the radical right) According to [The Guardian](#), the site "could do more to restrict extremists".
- Of the 100 sampled tweets, 88% of those produced by Spencer and 54% by Roberts were to content found elsewhere on Twitter itself, and so we recommend **strengthening Twitter's own Hate Conduct Policies**.
- **Enable image recognition** for Pepe the Frog and other radical-right symbols in all angles and perspectives (rather than only the front) to avoid their

manipulation by extremist users in order to spread the same messages. For example:



Reddit

In 2015, the Southern Poverty Law Center (SPLC) dubbed Reddit home to “[the most violently racist](#)” content on the internet. In 2018, when Reddit CEO Steve Huffman was asked about whether “obvious open racism” was against the company’s rules, he responded: “[It’s not](#)” (shortly thereafter he did say it was unwelcome). Last year, Reddit made a statement pledging that the company would “[take a stronger stance on fighting racism](#)”. Yet as of August 2021, the effectiveness of Reddit’s terms on online hatred and

content blocking remains questionable. While its [policy on hate speech](#) states that those who promote “hate based on identity or vulnerability” will be banned, it allows a remarkable amount of explicitly racist accounts to remain active. But even if banning measures were effective, they would still be limited. As a website aggregator, Reddit constitutes a platform whose content thrives on outlinking. Even if Reddit (like other sites) cannot be held accountable for harmful content distributed through outlinking to third-party sites, it is worth noting the role that Reddit’s own image and video hosting services have had in helping the radical right circumvent the company’s hate-speech policies.

In order to examine the impact (or lack thereof) of Reddit’s policies upon radical-right outlinking, this report selected the following three accounts: ‘**Behold, the Master Race**’ (“Supremacists of any share displaying their supremacy”, with 246,220 members), ‘**MetaCanada**’ (38K members), and ‘**Dark Enlightenment**’ (“The place to discuss Neoreaction and Mencius Moldbug”, with 23.8K members). The groups are located in the United States, Canada, and the United Kingdom, respectively; they share content in English, and have a total membership of half a million users. All of these groups are openly racist, and membership requires adherence to white-supremacist ideology. Among their messages, it is common to read (explicitly or implicitly) that non-white individuals are inferior and that the white ‘race’ is victim of a (liberal/democratic) conspiracy. Anti-Asian, anti-Black, and anti-Jewish speech abounds. In addition, all these groups promote homophobic language, in particular attacking the LGBTQI+ community and their demands for equality.

After analysing the most impactful posts produced by these radical-right ‘chat rooms’ on Reddit between the years 2018 and 2021, it appears that Reddit itself allows for up to **56% of outlinking activity that facilitates the spread of hate speech**. Indeed, Reddit’s own image hosting service, i.redd.it, appears to be a main impediment to banning extremist groups and accounts, and for hosting radical material with no consequences. Out of the three cases observed, **up to 100% of the most popular posts** – or those who received the highest level of engagement – **used outlinking to share extremist content**. Users did this much more frequently than uploading material directly from a browser, which can be more easily flagged by the site. In cases where outlinking represented at least 38% of distributed extremist content among the top 50 posts per group, between 10% and 56% of it was facilitated by Reddit’s own image hosting service. Other prominent outlinking sites include other image and text hosting services (16% to 58%), new media sites (from 2% to 46%), and right-wing and radical-right sites (2% to 26%).

Recommendations for Reddit

- **Apply hateful speech policy to Reddit's** image and video hosting service as established for Reddit's main feed.
- **Flag** the following hosting services:
 - <https://files.catbox.moe/>
 - <https://archive.fo/>
 - outline.com
- **Moderate the following accounts:**
 - <https://christopherrufo.com/>
 - [Unz.com](https://unz.com)
 - [Voiceofeurope.com](https://voiceofeurope.com)
 - [Oneangrygamer.net](https://oneangrygamer.net)
 - [Truepundit.com](https://truepundit.com)
 - [Thefederalist.com](https://thefederalist.com)
 - [Anonymousconservative.com](https://anonymousconservative.com)
 - <https://legalinsurrection.com/>
 - <https://www1.cbn.com/>
 - <https://thenewamerican.com/>
 - <https://www.eurocanadian.ca/>

YouTube

YouTube ranks among the more technically stringent of content moderators. The Google-owned platform's Terms of Service forbids [hate speech](#), medical [misinformation about COVID-19](#), and even [spam](#), with an entire moderation division devoted to regulating "violent or dangerous content." Specific regulations [bar content produced by terrorist](#) entities, those that utilise terrorist imagery, that praise or memorialise terrorist acts or individuals, or appear aimed at recruitment. YouTube's policies on hate speech are also expansive, barring not only content that attacks protected classes of citizens (based on gender, sexual orientation, race, nationality/ethnicity, religion, and disability), but also taking down content that dehumanises individuals, spreads conspiracy theories that deny instances of violence, or promotes any form of supremacism based on demographic categories. These and other policies led to YouTube removing [11.4 million videos](#) in Q2 2020.

At present, YouTube claims that every minute users upload 500 hours of content to its site. Given this scale, it has developed a method for moderation that combines AI and human moderators. According to its transparency statements, YouTube relies most heavily upon machine learning to flag content. The programme that YouTube uses to flag potential violating content is one that relies upon examples, with the programme *learning* to identify individual instances of problematic material by comparing it to a pattern of identified banned content. For example, after being shown a series of variations on the image of the symbol for designated terrorist group Feuerkrieg Division, the programme proactively searches through frames of content for similar images, flagging when a match is made.

Importantly, to a certain extent, YouTube outsources moderation. '[We rely](#) on the YouTube community as well as experts in our Trusted Flagger program to help us spot potentially problematic content by reporting it directly to us.' The Trusted Flagger program sees individual users as well as employees of governments and NGOs trained in YouTube guidelines and then trusted to flag potential violations, flags that are then reviewed by YouTube content moderators. Utilising these mechanisms, YouTube asserts that its "violative view rate" (i.e. the number of views a piece of content has before it is removed) declined by [70%](#) in 2020.

However, YouTube's machine learning still has many flaws, and the number of successful appeals (when content creators believed their videos had been deleted incorrectly) [grew fourfold](#), to 160,000 videos, in 2020. Another criticism levelled against the

company is that the platform is [still quite slow](#) at deciding what content it should ban, leaving harmful, standard-violating content up for long periods of time. For example, in 2019 YouTube [faced an outcry](#) after it left content from Steven Crowder untouched, which included racist and homophobic harassment of then-Vox journalist Carlos Maza. Only after a period of public outrage did YouTube ban Crowder's ability to run ads (again, this not the same as banning the account altogether). A similar example would be the case of EA Koetting, whose Satanist content has been linked to a double murder in London, and which was only [taken down](#) from YouTube follow outside pressure.

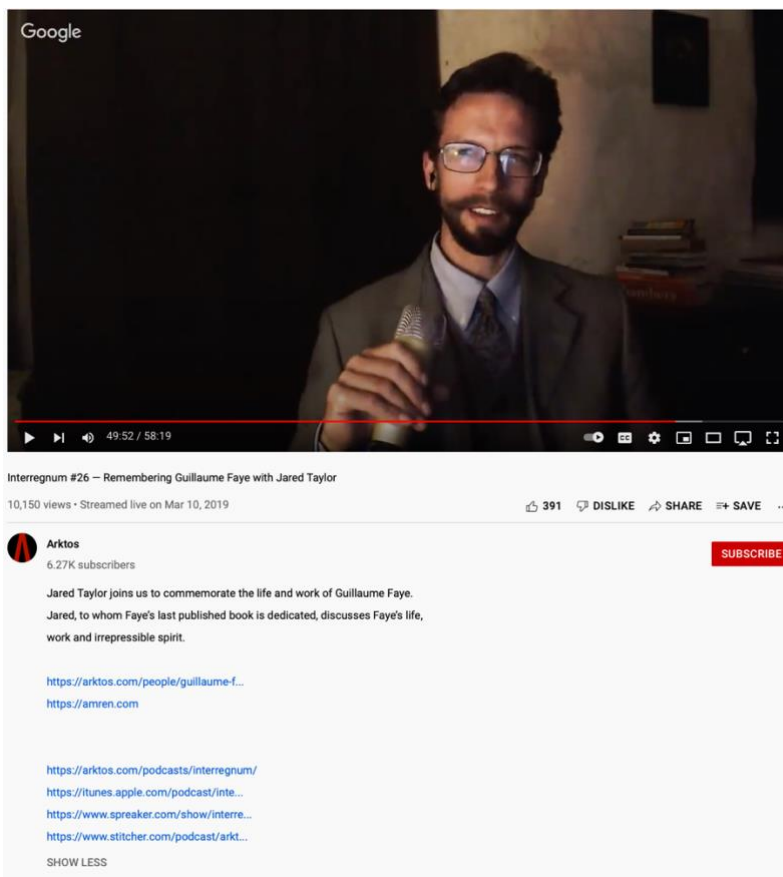
These examples speak to a larger issue with YouTube content moderation: [the strike system](#). According to YouTube regulations, when a video violates YouTube standards, the uploading account is notified via email, with an option to challenge YouTube's categorisation. After review, if the video remains classified as violating YouTube standards, but is the first of its kind to be released by the channel, it is merely given a warning. If the channel has already received a warning, then the video serves as a 'strike'. Should the channel then commit an offense warranting a second strike, there is a two-week ban on new uploads. Only after a third strike within 90 days can an account then be terminated. However, mechanisms do exist to terminate channels for a single, egregious offense. Although decisions appear extremely subjective and lacking transparency, it nevertheless appears that the majority of terminated channels come from violating the three-strike policy. Notably, then, if a channel uploads barred content outside the 90-day window, it can, at least for some time, avoid being entirely deplatformed. From a counter-terrorism perspective, however, the uploading of hateful video content mimics Pandora's box: once the video is uploaded, it is nearly impossible to erase its presence from the internet altogether (see, [for example](#); and [also](#)).

While the video format may appear as a barrier to outlinking, this is actually far from the case. For example, the radical right utilises outlinks with significant frequency on YouTube channels, thus making radicalisation a multimedia process. For instance, Arktos also has a YouTube channel, which was set up more than 11 years ago to support its activities. Today, the channel has approximately **6,250 subscribers** and its top video has received almost **50,000 views**. With respect to content, the Arktos channel focuses on book talks or book reviews; material aimed at selling the viewer both on the book and on its ideological conclusions. Those videos not explicitly linked to a specific book almost invariably feature an interview between Arktos staff and a radical-right ideologue. Topics include a supposedly impending ethnic apocalypse; identitarianism; and spirituality as related to racial and ethnic geopolitics. These beliefs appear to have captured a relatively attentive audience. Whereas Arktos' most popular video is a book talk with an author from 2015, which drew more than 49,000 views, the most popular

within this report’s timeframe (and second most popular overall) has some 19,000 views – after some two years online. Of the 41 total posts released in this three-year period, the viewing range stretches from 575 to nearly 20,000. Moreover, engagement can be seen through the likes and comments, which range from 29 to 650 in terms of likes, and from 29 to 175 in terms of comments.



Instructively, the **Arktos** ‘About’ page features five outlinks (to Facebook, Instagram, Twitter, Telegram, and the Arktos website), while individual videos can also include numerous outlinks. Indeed, **100% of its videos from the selected time period feature outlinks**. Much like its Facebook presence, Arktos’ use of YouTube is aimed at directing viewers to its website, where they can purchase its books. This is made obvious by the fact that popular videos from this period feature at least one link to its website (or another online purveyor of its publications). Arktos thereby promotes its racist and conspiratorial ideology through the pseudo-intellectual interviews it records and disseminates, allowing the outlinks to monetise ‘identitarian’ white supremacy.



The Warden Post has also established a YouTube page, although it only appeared in February 2020 and to date has just over 1,000 subscribers. Nonetheless, the channel has produced a considerable number of videos with a sizable rate of views per upload during this report’s reporting period. The most popular video within this time frame, for instance, accrued more than 7,300 views in less than one year. Likewise, all of **The Warden Post’s** top-ten most popular videos have over 1,000 views. Much like Arktos, the brunt of **The Warden Post’s** videos consist of interviews on topics such as ‘Dysgenicis and Eugenics’ and ‘The Reconquista in Western academia’. In keeping with its style on Facebook, The Warden Post’s YouTube channel positions itself as a space for debate about the apparently imminent decline of the West.

The Warden Post’s YouTube channel also relies upon outlinking with relative frequency, with **nearly half the top 50 videos (48%) in the last three years utilising outlinks** (including **3 of the top 5 posts**). The outlinks lead to an array of different websites. Several, for instance, lead viewers to Arktos or similar booksellers. Others point people towards the social media accounts of The Warden Post as well as of other radical-right groups. Still more outlinks connect to small websites focused on white supremacist ideas, or to sites where donors can support the group through alternative financial

means. The outlinks expressly lead to a book or article being discussed by the speakers of the video, but can also serve as a guide to further radicalising content.

Although seemingly dormant for years, finally, **This Is Europa** also has a YouTube channel with almost 24,000 subscribers “[that highlights anti-white policies](#)”. This is despite a series of conditions appearing to make the channel of little interest on the radical-right scene. **This Is Europa** has not posted new YouTube content since March 2018 and has only ever released nine videos, with six of these released more than seven years ago. The content may be characterised as nostalgic, equating an immigrant-free Europe with a utopia for which white people should struggle to return. While falling outside of the time period of this study, it is worth mentioning that the channel has a considerable viewership: the top-performing video has more than 200,000 views, and even the least popular has upwards of 15,000 views. **All of the videos utilise outlinking**. While one video features an outlink to **This Is Europa’s** Facebook page only in the video itself, all others include outlinks in the textual description of the video. Most frequently, viewers are guided to the official website for the group, though in the majority of cases they are shown several outlinks; to funding pages (PayPal, Bitcoin, Patreon); merchandising websites; listicles for further content; or outlinks to other social media platforms. As is the case with the other case studies, **This Is Europa** uses YouTube as a means not only to directly radicalise its audience, but to monetise content otherwise prohibited on the site due to its extreme, hateful nature.

Recommendations for YouTube

1. **Alter the strike system** to prevent channels from uploading content that circumvents ban punishments;
2. For **text-based content**, cross-reference book titles, topics, and authors on radical right sites to avoid platforming or monetising white supremacy.
3. Set up **detection metrics for fundraising** media within uploaded content, in order to ensure demonetisation of radical right content;
4. Set up the **algorithm to read for outlinks**, both in the descriptor for a video and in the video itself.

Reassessment of Stated Hypotheses

1. Outlinks feature URL patterns that may allow for algorithmic detection, with most outlinking concentrated on a few websites.
 - Was the hypothesis supported: Yes.
2. A large proportion of those producing outlinks are individuals with affiliations to banned terror groups.
 - Was the hypothesis supported: No. We did not find that there were direct links with banned terror groups (notwithstanding indirect ideological parallels between the two, such as ‘the Great Replacement’).
3. Outlinking is a strategy used to advance radical-right narratives, but also to instil a sense of community and belonging among those who find, read, agree with, and further disseminate the outlinked material.
 - Was the hypothesis supported: Yes.
4. Anonymity on social media platforms also contributes to outlinking.
 - Was the hypothesis supported: Yes (with the exception of certain Twitter profiles).

Conclusions and Recommendations

This report has shown that Facebook's moderation policies and lack of consideration for hidden hate-content has resulted in its exploitation by radical actors. As a consequence, Facebook can and has become a platform a) for authors and ideas that incite hateful speech and potentially offline violence, b) for the acquisition of merchandising that reinforces hatred, and c) for the dissemination of messages that can facilitate the normalisation of extremist ideologies, enabling online radicalisation. In response, Facebook can tackle outlinking by:

- Offering more accurate information concerning what constitutes hate speech online.
- Improving procedures allowing it to flag and take down harmful posts as well as ban accounts.
- Putting in place structures to identify users on the path of radicalisation in order to intervene and facilitate deradicalization.
- Undertaking the latter in cooperation with law enforcement and other counter-terrorism bodies.

Furthermore, Facebook should invest in further analysis on a series of related topics:

1. Facebook can learn more about the formation of echo-chambers on the platform via the outlinking-like activity of cross-posting between Facebook pages.
2. As this study has only examined public groups, Facebook should support additional research into outlinking in private groups. The latter seem crucial in the dissemination of online extremism (as they allow for the development of a particularly strong sense of community, effectively luring users experiencing a severe sense of isolation or offline social constraints).
3. Any enhancement of anti-hate action developed by Facebook should apply not only to public accounts, but private ones as well.

Albeit a smaller platform, a broader definition of hate speech has allowed Twitter to act against a larger (and growing) number of accounts, including tweets that incite fear or fearful stereotypes. Nevertheless, anti-government, misogynist, racist, and white supremacist narratives still have a noticeable presence on Twitter, which is further amplified through outlinking. Accordingly, focusing on outlinking practices rather than on account feed may enhance moderation of the radical right, allowing for the detection

of self-reinforced networks violating the terms of use rather than individual transgressions.

Even more than in the case of Twitter, Reddit constitutes a platform whose content thrives on outlinking. Most concerningly, Reddit's own hosting services are allowing for the larger sharing of extremist content. To better moderate extremist messaging, Reddit should apply its hateful speech policies to its own image and video hosting services. At present, Reddit only intervenes by flagging or blocking outlinks that lead to extremist content shared on third-party sides – in other words, dangerous outlinks that force the user to leave Reddit's page to keep browsing. However, if the outlinks lead to extremist material uploaded on Reddit's own hosting service, the platform does not intervene in any way.

YouTube's predominately video-sharing platform makes it a notably different case when compared to Facebook, Twitter, or Reddit. While the format may appear something of a barrier to outlinking, this is not the case. In fact, the radical right utilises outlinks with a frequency on YouTube channels that is comparable to that of Facebook. Oftentimes, the viewer is guided to official websites for an extremist group through links appearing either in the video description or within the video itself. In most cases, outlinks lead to various forms of revenue generation. The nature of outlinking on YouTube suggests that an overhaul of the platform's detection metrics and algorithms is needed, one especially focusing on hyperlinks shared within videos. Without making such changes, YouTube contributes to the fundraising and further growth of extremist groups, both online and offline.

In summary, outlinking on social platforms allows users to engage with radical-right content through hyperlinks intended to circumvent terms of use (particularly those focused on account surveillance rather than 'migrated feed'). This report therefore concludes that outlinking is indeed a strategy effectively used by the radical right to: help promote hate-filled narratives; facilitate processes of online radicalisation; and instil a sense of belonging among those sharing outlinked extremist material. That said, the report has shown that outlinking is, in fact, a process that tends to rely upon a limited number of external sites. It also has a number of other patterns, which indicate that the alteration of AI detection mechanisms should allow for more precise identification of extremist outlinking. Such modifications will be pivotal in reducing online extremist networks.

General Recommendations

Finally, this report makes the following general recommendations:

- Incorporate the above-listed radical-right websites into existing databases for the detection of hate speech, radical online networks, and extremist propaganda. This would include hosting sites such as archive.org, outline.com, archive.fo, files.catbox.moe, as well as websites like arktos.com, gab.com, freespeechlibrary.com, legalinsurrection.com, odysee.com, archive.nc-f.org, babylonbee.com, voiceofeuropa.com, thisiseuropa.net, or unz.com.
- Consider new requirements for users such as providing verified identities when creating an account, ideally as reflected in their national identification documents. This could be a valuable step toward moderating extremist outlinking as it might deter users who otherwise would not openly share extremist material.
- Social media platforms should more closely collaborate with law enforcement and other counter-terrorism institutions like [GIFCT](#), particularly in monitoring legal cases with a relevant online component. This could help establish an effective and responsible foundation for strategic engagement in litigation and research.

This report contends that implementing these recommendations can effectively stop social media platforms acting as a pipeline for radical-right extremism. This would leave space, instead, for them to either become powerful sites for deradicalization, aid counter-terrorism efforts, and contribute to promoting inclusion and support for liberal democratic values.

Appendix 1: On the Lookout for Outlinking

(A Practical Guide to Uncovering Hateful Hyperlinks)

WHAT IS OUTLINKING?

In essence, outlinking is the practice of using a hyperlink specifically for the purpose of disseminating hateful content.



How Does Outlinking Undermine Facebook’s Community Standards?

Outlinking aims to subvert Facebook’s methods for automatically flagging of hateful content and/or disinformation as expressed in key terms and images. In circumstances in which the outlink’s URL, HTML meta text, and/or URL thumbnail do not feature any of the prohibited items, the posting user can **freely promote content normally barred on Facebook**, as well as direct traffic from Facebook onto overtly extremist websites.

WHAT ARE CURRENT COMMON OUTLINKS ON FACEBOOK?

According to our research, the following sites were the **top five** most consistently used by the **sampled extremist Facebook users**:

1. arktos.com
2. thisiseuropa.net
3. rightoftheright.com
4. humansarefree.com
5. breitbart.com

WHAT ARE CURRENT COMMON OUTLINKS ON OTHER PLATFORMS?

According to our research, the following sites and accounts are the **top ten** most consistently used by the sampled extremist users on **Twitter, Reddit, and YouTube**:

1. archive.org
2. gab.com
3. thisiseuropa.net
4. i.reddit.it and v.reddit.it
5. i.imgur.com
6. breitbart.com
7. arktos.com
8. telegram.com
9. minds.com
10. sticher.com

PATTERNS FOR IDENTIFYING COMMON EXTREMIST OUTLINKS ON FACEBOOK

Although there are some variations in outlinking habits across the various social media platforms and even between individual accounts of the same platform, there are certain commonalities that can be observed and utilised to enable automatic flagging. Here are **five actions to be taken that would significantly reduce outgoing traffic to right-wing extremist websites**.

1. Aside from already flagged authors (i.e. Adolf Hitler, James Mason), the following authors should be flagged, and posts related to their works should be reviewed for potential violations of community standards given the high correlation between references to them and standard-violating content:

- a. Guillaume Faye
- b. Julius Evola
- c. Ernst Jünger
- d. Savitri Devi
- e. Alexandr Dugin

2. In addition to widely known terms and numbers linked to violating content (i.e. 'Heil Hitler,' 'ZOG,' racial slurs, 14/88), there are several others frequently utilised in outlinks that violate Facebook's community standards:

- a. European New Right
- b. New World Order
- c. Globalist
- d. Great Awakening
- e. Great Reset

3. With its clear 'heroes,' the radical right operating on Facebook also has clear enemies. Outlinks used in posts about the following individuals frequently lead to content that otherwise violates Facebook's community standards:

- a. Jacob Rothschild (in earnest, all Rothschilds and the collective Rothschild dynasty)
- b. Bill (& Melinda) Gates
- c. George Soros
- d. Dr Tedros Adhanom Ghebreyesus
- e. Mark Zuckerberg

Academic Consulting Services Ltd.
(www.academicconsulting.co.uk)

The Centre for Analysis of the Radical Right
(www.radicalrightanalysis.com)



Academic
Consulting
Services Ltd.