ORG OPEN RIGHTS GROUP

BAD ADS
TARGETED DISINFORMATION, DIVISION AND FRAUD ON META’S PLATFORMS

April 2025

ABOUT ORG

Open Rights Group (ORG) is a UK based digital campaigning organisation working to protect fundamental rights to privacy and free speech online. We are a grassroots organisation with supporters and local groups across the UK.

Our work on data protection and privacy includes challenging the immigration exemption to UK data protection law, defending the General Data Protection Regulation (GDPR) from attempts to water down its provisions, and challenging uncontrolled and unlawful data sharing by online advertisers.

openrightsgroup.org

Report by **Dr James Riley**

EXECUTIVE SUMMARY

This report lays out clear evidence of how Meta enables bad actors to use its targeted advertising system to manipulate elections, spread disinformation, fuel division, and facilitate fraud.

Meta's social media platforms, Facebook and Instagram, sit at the intersection of the attention economy and surveillance capitalism. Meta's business model is built on maximising user attention while tracking behaviours, interests and harvesting personal information. This surveillance is used to categorise people into 'types'. Meta uses this profiling to sell the attention of these 'types' to would-be advertisers – a practice known as surveillance advertising.

This report brings together existing and new evidence of how bad actors can and have used Meta's targeted advertising system to access the attention of certain types of users with harmful adverts. These 'bad ads' seek to mislead, to divide, and to undermine democracy. Through a series of case studies, it shows how bad actors — from political campaigns to financial scammers — have used Meta's profiling and ad-targeting tools to cause societal harm.

The case studies in this report examine how bad actors use Meta's advertising systems to spread bad ads across five areas:

- **Democracy – voter suppression, the targeting of minorities, electoral disinformation, and political manipulation by the Trump campaign, Musk-backed dark money groups, and Kremlin-linked actors.**
- **Science – the COVID infodemic, vaccine disinformation and climate crisis obfuscation.**
- **Hate – sectarian division, far-right propaganda, antisemitism, and Islamophobia.**
- **Fear – targeting of vulnerable communities, UK Home Office migrant deterrence, and the reinforcement of trauma.**
- **Fraud – deepfake scams, financial fraud, and the use of targeted adverts to facilitate black market activities on Meta's platforms.**

This report evidences the individual and collective harms enabled by Meta's advertising model. Three major issues emerge, each requiring urgent action:

1. The Transparency Problem

Meta's ad system is insufficiently transparent about the profiled targeting categories advertisers choose. This opacity facilitates harmful advertising and prevents public scrutiny of disinformation, fraud, and manipulation.

Recommendation: Meta must be required to publish full ad targeting details in its public Ad Library. This should include all demographic, interest-based, and behavioural categories used by each advertiser for each advert. Greater transparency would deter some forms of harmful targeting and enable greater public scrutiny of harmful ad targeting on Meta's platforms.

2. The Moderation Problem

Meta's moderation heavily relies on user reporting of bad ads once they are circulating rather than preventing them from appearing in the first place. Meta's largely automated approval process and lax approach to ad moderation enable targeted disinformation and harm.

Recommendation: Meta must significantly expand both human and technological resources allocated to pre-publication ad moderation to tackle obvious disinformation, fraud and harmful ads upstream of publication rather than downstream of harm. A useful starting ground could be provided by provisions, in the Digital Services Act, which already establish transparency obligations covering both content moderation decisions of online platforms and the criteria which advertisers use to target advertisement. The DSA also introduces so-called anti dark-patterns provisions, that prohibit online service providers from misleading, tricking or otherwise forcing users into accepting targeted advertising against their best interest.[1]

3. The Profiling Problem

Meta's business model is built on profiling users by harvesting vast amounts of personal and behavioural data, yet it offers users no effective opt-out of surveillance and targeting for users to protect themselves from the harms evidenced in this report. Additionally, there is little awareness of how users can opt out of their data being used to train generative AI.[2]

Recommendation: Users should be presented with a clear and explicit opt-in option for profiling and targeting, and be warned that this means they can be targeted by bad actors seeking to mislead or defraud them. Given the invasive nature of the data collection and Meta's inability to demonstrate it can protect citizens from harm, informed consent must be required above and beyond the acceptance of lengthy terms and conditions. For users who do not opt-in to surveillance advertising, Meta should adopt contextual advertising within broad geographies – targeting ads based on the content users are presently engaging with rather than based on the surveillance and profiling of citizens.

Despite Meta's public assurances that it does not allow disinformation, voter suppression, hate and division, or fraudulent adverts on its platforms, the case studies in this report demonstrate it consistently enables these harms. This is not solely a problem of policy enforcement, but an issue with the fundamental architecture of Meta's opaque and poorly moderated advertising model built on surveillance, profiling, and microtargeting. Combined with Meta's unwillingness to mitigate the harms it enables, Meta's surveillance advertising continues to facilitate societal harm. Without legal or regulatory intervention, these threats to democratic integrity, public safety, and social stability will persist, and citizens globally will continue to be deprived of the right to a reality that is not shaped by opaque, targeted advertising systems available for hire by bad actors.

[1] See The Digital Services Act, available at: https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act_en

[2] ICO fails UK Meta users and allows social media giant to resume data scraping for AI https://www.openrightsgroup.org/press-releases/ico-fails-uk-meta-users-and-allows-social-media-giant-to-resume-data-scraping-for-ai/

INTRODUCTION

"Democracy is at risk from the malicious and relentless targeting of citizens with disinformation and personalised 'dark adverts' from unidentifiable sources, delivered through the major social media platforms we use every day."

- Damian Collins MP, former Chair of the Digital, Culture, Media and Sport Committee, 18 February 2019[3]

"SENATOR, WE RUN ADS"

"Senator, we run ads," Mark Zuckerberg smirked in response to US Senator Hatch's questioning of how Facebook sustains a business model when users do not pay for the service.[4] The exchange happened in 2018 during a joint hearing of the Senate's Judiciary and Commerce committees on data privacy violations and Russian disinformation on Facebook. It was widely seen as emblematic of the disconnect between the new, disruptive world of digital technologies and the old institutions and out-of-touch policymakers tasked with regulating them.[5] [6]

Facebook, and Meta's[7] other main social media platform, Instagram, sit at the intersection of what has been called the 'attention economy' and 'surveillance capitalism.'[8] [9] [10] The model consists of keeping users' attention on the platform, generating detailed profiles of each user based on demographics, location, interests and behaviours, and then selling the opportunity to target specific 'types' of users to advertisers.[11] The more people use Facebook, the more valuable its advertising system becomes, allowing for greater ad reach and increasingly precise targeting. This business model relies on collecting, processing, and analysing vast amounts of user data, which is then monetised for targeted advertising—a practice Amnesty International has described as "invasive." Because advertising revenue is tied to engagement, Facebook's recommender algorithm is designed to maximise attention. This means it prioritises content that keeps users interacting, regardless of whether that content is informative, misleading, or harmful. Inflammatory, polarising, and harmful content — including disinformation — tends to generate high engagement and is, therefore, often boosted by the algorithm.[12]

The algorithm is not, however, a fully autonomous entity. It is actively shaped, refined, and overseen by Facebook employees who adjust its priorities and weightings based on company objectives. Users themselves also play a role by liking, sharing, and reacting to content, which further informs how the system operates.[13] We may view it as a cybernetic organism — a cyborg, part machine, part human — but the responsibility for its harms lies with real decision-makers at Facebook, who determine how content is ranked, spread, and monetised.

[3] Digital, Culture, Media and Sport Committee. (2019). Disinformation and 'fake news': Final Report published. *UK Parliament.*

[4] NBC News. (2018). Senator Asks How Facebook Remains Free, Mark Zuckerberg Smirks: 'We Run Ads'. *YouTube: NBC News.*

[5] Burch, S. (2018). 'Senator, We Run Ads': Hatch Mocked for Basic Facebook Question to Zuckerberg. *The Wrap.*

[6] Stewart, E. (2018). Lawmakers seem confused about what Facebook does — and how to fix it. *Vox.*

[7] Facebook changed its parent company name to Meta in 2021 after a series of negative stories engulfed the company. In this report, which covers years before and after the name change, we use Meta and Facebook largely interchangeably. Now, however, Facebook represents one of Meta's platforms, along with others, including Instagram and WhatsApp.

[8] Naughton, J. (2018). Anti-Social Media: How Facebook Disconnects Us and Undermines Democracy by Siva Vaidhyanathan – review. *The Guardian.*

[9] Vaidhyanathan, S. (2018). Antisocial media: How Facebook disconnects us and undermines democracy. *Oxford University Press.*

[10] Zuboff, S. (2015). Big other: surveillance capitalism and the prospects of an information civilization. *Journal of information technology, 30(1), 75-89.*

[11] Collier, B. et al. (2024). Influence government, platform power and the patchwork profile: Exploring the appropriation of targeted advertising infrastructures for government behaviour change campaigns. *First Monday, 29(2).*

[12] Amnesty International. (2022). The Social Atrocity: Meta and the Right to Remedy for the Rohingya. *Amnesty International Ltd.*

[13] Oremus, W. et al. (2021). How Facebook Shapes Your Feed. *The Washington Post.*

BAD ADS AND BAD ACTORS

The goal of maintaining the attention of users is to sell parts of that attention to advertisers, but not all advertising on Meta's platforms comes from legitimate businesses selling products or services. Targeted adverts are one of the tactics in the modern disinformation playbook.[14] This report details case studies of societally harmful content spread using Meta's targeted advertising system. In this report, we use the term "bad ads" to refer to malicious adverts which spread disinformation (false information deliberately intended to mislead) and misinformation (the spreading of inaccurate information, without the intention to mislead).[15] We also use it to refer to broader categories of harm, including undemocratic interference in elections, the stoking of fear, hatred and division, and outright fraud. "Bad actors" are individuals, groups, or countries who behave in ways harmful or detrimental to society, and in the context of this report, those who purchase bad ads.

The Cambridge Analytica scandal represents an archetypical example of this. It was revealed in 2018 that Cambridge Analytica, a British consulting firm, gained access to 87 million Facebook users' personal information through an underhanded data-gathering operation, with the intention to sell their service of influencing democratic processes using psychometric profiles and targeted ads seeking to influence voter behaviour.[16] In 2022, Meta settled a class action lawsuit regarding the third-party processing of data, including the Cambridge Analytica scandal, by agreeing to pay $725 million.[17] This is not just a third-party data processing issue, however, Facebook offers advertisers the opportunity to use its own profiling of users to target specific groups. In the Myanmar genocide, which began in 2017, in addition to Facebook's algorithm boosting the hate that fuelled the violence, the Myanmar military used Facebook's advertising system to spread disinformation about the Rohingya people.[18]

Facebook implemented some changes in response to the Cambridge Analytica scandal and ongoing worries around political interference, including launching its Ad Library in a bid for more transparency. Previously, it was unknown what ads different groups had seen, but with the creation of the Ad Library, some information about advertising on Facebook was made public.[19]

HOW META'S PROFILING AND ADVERTISING WORKS

On Facebook, a user's feed can be disaggregated into three broad classes of content: (1) organic content, (2) advertising, and (3) suggested content.[20] Each of these is driven by Facebook's recommender systems, or algorithms, which profile people into types and suggest content believed to be most relevant to them. However, as has been discussed, even so-called 'organic' content, such as the posts you see from friends, is algorithmically curated based on prescribed preferences and priorities to maximise user attention.

Meta allows advertisers to target users based on a number of characteristics which users have either directly shared with Meta, or which Meta has profiled users as having. These can be broadly grouped into location, demographics, interests and behaviours:

- **Location: Target people in continents and regions, countries, states, districts, and even down to the postcode or ZIP code level.**

[14] CISA. (2022). Tactics of Disinformation. *Cybersecurity and Infrastructure Security Agency.*

[15] American Psychological Association. (2021). Misinformation and disinformation.

[16] ICO. (2020). Letter to the DCMS Select Committee. *Information Commissioner's Office.*

[17] Raymond, N. (2022). Facebook parent Meta to settle Cambridge Analytica scandal case for $725 million. *Reuters.*

[18] Amnesty International. (2022). The Social Atrocity: Meta and the Right to Remedy for the Rohingya. *Amnesty International Ltd.*

[19] Shukla, S. (2019). A Better Way to Learn About Ads on Facebook. *Meta.*

[20] Obem, A. & Wróblewska, M. (2023). Anxious about your health? Facebook won't let you forget. *Panoptykon Foundation.*

- **Demographics: Target people based on age (down to as young as 13, with some restrictions), gender and language. Demographics can be further specified with detailed options including relationship status, educational attainment, the age of a parent's children, employment industry, income and job title, among many others.**

- **Interests and behaviours: Target people based on their hobbies and interests, their activities on Facebook, across the web, ads they click on, and details about their device and internet connection.**[21] [22]

Meta also allows advertisers to upload their own "Custom Audiences" based on external data such as email lists, data from an app or Pixel tracker, as well as using Meta's own data.[23] These can be augmented using Meta's Lookalike Audience system, which "leverages information such as demographics, interests and behaviours from your source audience to find new people who share similar qualities."[24]

Meta is always updating its offerings for audience access. Recently, it launched Advantage+ which "lets you use Meta's advanced AI to find your campaign audience." Advantage+ uses numerous pieces of information from places such as Meta Pixel, which tracks people's activities across the web, to show ads to people Meta's AI predicts are most likely to respond.[25]

There are 'black boxes' littered throughout all these mechanisms. We have no real understanding of the totality of Meta's data files, how its algorithms work, and all the ways it profiles us. What we do know is that the selling of the use of these black-boxed systems fuels the overwhelming majority of its business model. Of Meta's $164.5 billion reported revenue in 2024, $160.6 billion flowed through advertising.[26]

ABOUT THIS REPORT

In this report, we detail the use of Meta's targeted advertising system to spread bad ads. It is divided into five thematic sections: Democracy, Science, Hate, Fear, and Fraud.

We draw evidence from three broad sources:

Real-world case studies. Investigative reporting, academic articles, and reports by NGOs, among others, that have unearthed real-world evidence for bad ads being targeted on Meta's platforms.

Experimental tests. Where organisations have tested Meta's moderation and approval process by designing and uploading ads which violate Meta's purported policies against spreading disinformation, fraud and hate.

Primary Ad Library Evidence. New evidence from Meta's ad library for bad ads being targeted based on inferred evidence that ads were differentially delivered to different audience segments. While the Ad Library is a move towards transparency, it is also limited in that Meta chooses not to make public the targeting categories advertisers have used.[27]

This report contains discussions of **targeted disinformation, voter suppression, racial profiling, hate speech, violent extremism, refugee deterrence, fraudulent scams, and algorithmic harm.** It includes examples of **misleading political advertising, financial fraud, vaccine disinformation, and fear-based campaigns**, some of which reference **child mortality, racism, antisemitism,** and **Islamophobia.** Readers affected by these issues should be aware that some of the report's content may be distressing.

[21] Meta. About reaching new audiences. *Business Help Centre.*

[22] Meta. About detailed targeting. *Business Help Centre.*

[23] Meta. About custom audiences. *Business Help Centre.*

[24] Meta. About Lookalike Audiences. *Business Help Centre.*

[25] Meta. About Advantage+ audience. *Business Help Centre.*

[26] Meta. (2025). Meta Reports Fourth Quarter and Full Year 2024 Results. *Meta Investor Relations.*

[27] Sankaranarayanan, A. et al. (2024). The Facebook Algorithm's Active Role in Climate Advertisement Delivery. *Research Square.*

DEMOCRACY

(DON'T) GET THE VOTE OUT

Barack Obama's 2008 and 2012 election campaigns were celebrated as the state-of-the-art in modern electioneering. While data-led campaigns and microtargeted messaging had a longer history, Obama's team pioneered new big data techniques, the modelling of voter characteristics, and a combination of offline and online efforts to mobilise voters.[28] The goal of all this was not only to "get the vote out" from his Democratic base but to expand the electorate by mobilising young adults and minority groups to vote for the first time.[29] Facebook was central to this new world of political messaging.[30]

In 2016, however, similar data-driven techniques were used by the Trump campaign but for a different purpose. This time the goal was not solely to mobilise the vote, but, for certain groups, to deter it. A Channel 4 News investigation uncovered that Trump's digital campaign, which included a team from Cambridge Analytica, used an algorithm to profile Americans in key battleground states – predicting their political beliefs and likelihood to vote.[31] These profiles included detailed information on voters' "income, race, ethnicity, place of origin, religion, language, marital status, gun ownership and more".[32] One of the eight audience segments generated was labelled "deterrence" and consisted of people the Trump campaign did not want at the polls on voting day. That is, if these people did vote, they would more likely vote for Hillary Clinton. It was found that **African Americans were disproportionately assigned to the "deterrence" category, and the campaign used Facebook to target these individuals down to the level of districts and wards with ads and disinformation to dissuade them from turning out to vote for Clinton.**[33] Jamal Watkins, vice president of the National Association for the Advancement of Colored People, called the tactic a modern-day suppression campaign.[34] [35]

Trump 2016 Campaign Segments

- Core Clinton (committed Clinton supporters)
- Core Trump (committed Trump supporters)
- Get Out the Vote (Trump supporters who needed to be rallied to the polls)
- Persuasion (swing voters who could be convinced to vote for Trump)
- Deterrence (Clinton supporters who could be demotivated from voting)
- Disengaged Clinton (Clinton supporters unlikely to vote)
- Disengaged Trump (Trump supporters unlikely to vote)
- Deadbeats (apathetic voters with no clear candidate preference)

Figure 1: Trump 2016 digital campaign segments, including the "deterrence" category which disproportionately contained African American voters. (Source: Miami Herald)

[28] Trish, B. (2018). Big data under Obama and Trump: The data-fueled US presidency. *Politics and Governance, 6(4), 29-39.*

[29] Nisbet, M.C. (2012). Obama 2012: The Most Micro-Targeted Campaign in History? *Big Think.*

[30] Pilkington, E. & Michel, M. (2012). Obama, Facebook and the power of friendship: the 2012 data election. *The Guardian.*

[31] Channel 4 News. (2020). Revealed: Trump campaign strategy to deter millions of Black Americans from voting in 2016. *Channel 4 News Investigations Team.*

[32] Blaskey, S. et al. (2020). How the Trump campaign used big data to deter Miami-Dade's Black communities from voting. *Miami Herald.*

[33] Channel 4 News. (2020). Revealed: Trump campaign strategy to deter millions of Black Americans from voting in 2016. *YouTube: Channel 4 News.*

[34] Channel 4 News. (2020). Revealed: Trump campaign strategy to deter millions of Black Americans from voting in 2016. *Channel 4 News Investigations Team.*

[35] Channel 4 News. (2020). Revealed: Trump campaign strategy to deter millions of Black Americans from voting in 2016. *YouTube: Channel 4 News.*

The official Trump campaign was not the only group in 2016 using Facebook's advertising apparatus to target African American voters. The Internet Research Agency (IRA), a St. Petersburg company with explicit links to both the Kremlin and the Russian paramilitary organisation, Wagner Group, were also engaged in a parallel suppression campaign. A 2019 report by the US Senate Select Committee on Intelligence determined that "the IRA sought to influence the 2016 U.S. presidential election by harming Hillary Clinton's chances of success and supporting Donald Trump at the direction of the Kremlin".[36]

Paid adverts on Facebook and Instagram formed a part of the IRA's strategy.[37] **The IRA used both interest-based and location-based targeting to flood African Americans with messages seeking to divert their political energy away from political institutions and to boycott the election.** [38] The ads began by promoting African American racial identity early in the campaign to charge societal cleavages, then later pivoted to attempting to suppress votes when the election drew closer.[39] **The IRA also specifically targeted polarising messages to LGBTQ+, liberal, conservative, and Muslim American voters, in an attempt to increase identity-based tensions**. According to researchers at the University of Oxford, different messages were sent to each of these groups and were "designed to push and pull them in different ways" that were ultimately to the "benefit the Republican Party—and specifically, Donald Trump".[40] Professor Young Mie Kim at the University of Wisconsin, Madison, concluded the IRA intimately understood the political cleavages in the USA and designed ads, "to divide the public and interfere in the U.S. elections by utilizing data-driven, algorithm-based, microtargeting capacity that is readily available on Facebook/Instagram".[41]

In 2020, a Facebook spokesperson said: "Since 2016, elections have changed and so has Facebook – what happened with Cambridge Analytica couldn't happen today. We have 35,000 people working to ensure the integrity of our platform, created a political ads library… and have protected more than 200 elections worldwide. We also have rules prohibiting voter suppression and are running the largest voter information campaign in American history."[42] However, the 2020 presidential election brought more evidence of targeted fear-based disinformation. These included **Facebook ads describing Joe Biden as a communist targeted at Latino American voters in Texas and ads in Florida, where many Venezuelan Americans reside, that compared Biden to Venezuela's socialist President Nicolás Maduro.**[43]

[36] US Gov. (2019). Report of the Select Committee on Intelligence. Russian Active Measure S Campaigns And Interference In The 2016 U.S. Election. Volume 2: Russia's Use Of Social Media with Additional Views. *United States Senate.*

[37] DiResta, R. et al. (2019). The Tactics & Tropes of the Internet Research Agency. *New Knowledge.*

[38] Howard, P. et al. (2019) The IRA, Social Media and Political Polarization in the United States, 2012-2018. *Computational Propaganda Research Project. University of Oxford.*

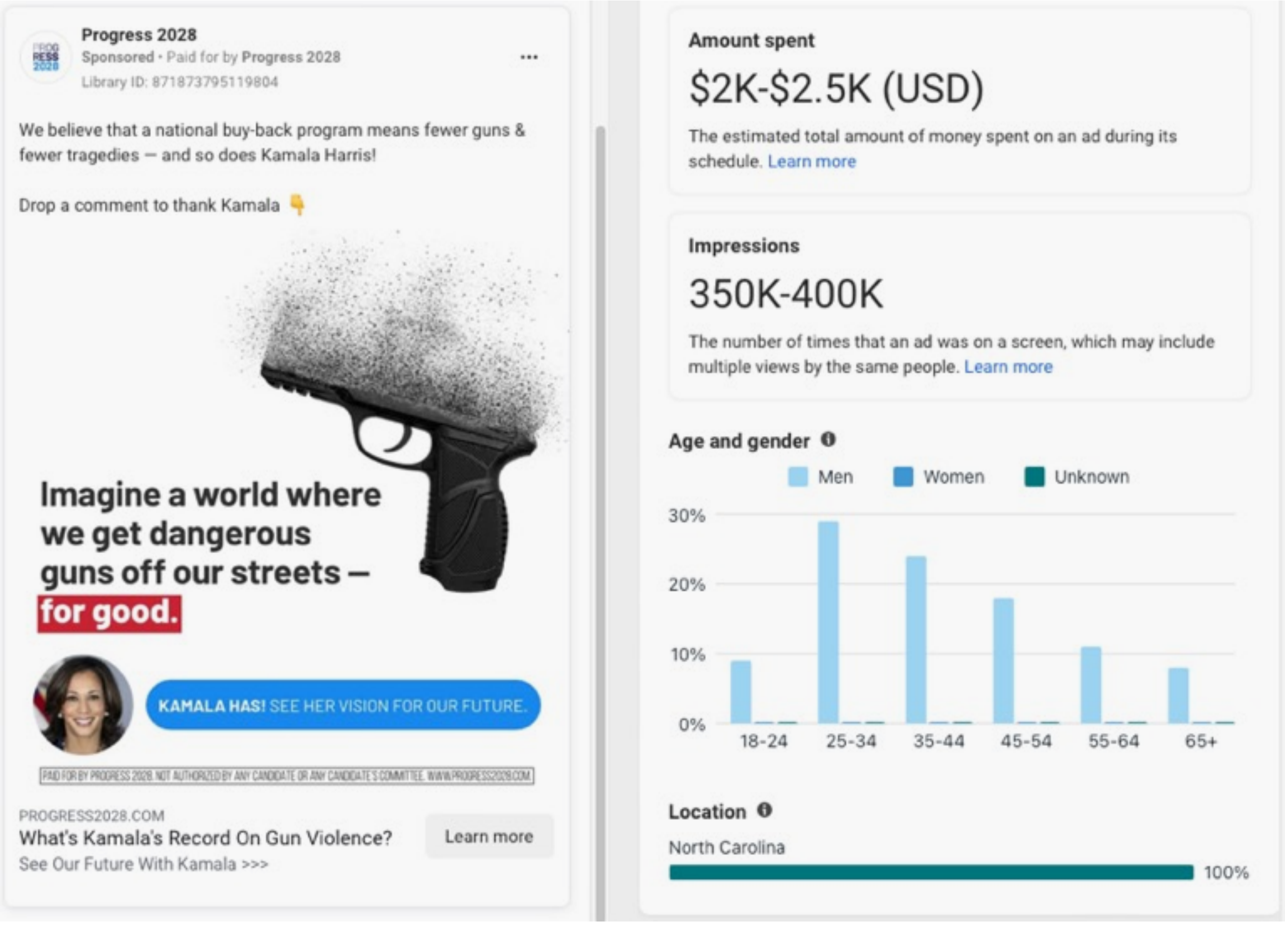
[39] Kim, Y.K. (2018). Uncover: Strategies and Tactics of Russian Interference in US Elections. Russian Groups Interfered in Elections with Sophisticated Digital Campaign Strategies. Project DATA. *The University of Wisconsin, Madison.*

[40] Howard, P. et al. (2019) The IRA, Social Media and Political Polarization in the United States, 2012-2018. *Computational Propaganda Research Project. University of Oxford.*

[41] Kim, Y.K. (2018). Uncover: Strategies and Tactics of Russian Interference in US Elections. Russian Groups Interfered in Elections with Sophisticated Digital Campaign Strategies. *Project DATA. The University of Wisconsin, Madison.*

[42] Channel 4 News. (2020). Revealed: Trump campaign strategy to deter millions of Black Americans from voting in 2016. *Channel 4 News Investigations Team.*

[43] Seitz, A. & Weisset, W. (2021). Inside the 'big wave' of misinformation targeted at Latinos. *Associated Press News.*



Progress 2028
Sponsored · Paid for by Progress 2028
Library ID: 871873795119804

We believe that a national buy-back program means fewer guns & fewer tragedies — and so does Kamala Harris!

Drop a comment to thank Kamala 👇

Imagine a world where we get dangerous guns off our streets — for good.
KAMALA HAS! SEE HER VISION FOR OUR FUTURE.
PAID FOR BY PROGRESS 2028 NOT AUTHORIZED BY ANY CANDIDATE OR ANY CANDIDATE'S COMMITTEE WWW.PROGRESS2028.COM

PROGRESS2028.COM
What's Kamala's Record On Gun Violence?
See Our Future With Kamala >>>
Learn more

Amount spent
$2K-$2.5K (USD)
The estimated total amount of money spent on an ad during its schedule. Learn more

Impressions
350K-400K
The number of times that an ad was on a screen, which may include multiple views by the same people. Learn more

Age and gender

Location
North Carolina 100%

Figure 2: Fake 'Progress 2028' initiative ads, funded by Musk-backed dark money group, purportedly backing Harris but linking to a disinformation website spreading exaggerated and false policy positions. (Source: Ad Library)

MUSK, DARK MONEY, AND DISINFORMATION

In the most recent US presidential election in 2024, human rights groups expressed concerns about the continued threat of targeted disinformation towards communities of colour.[44] During the election, Elon Musk's America PAC spent more money on advertisements targeting voters with conservative interests on Meta platforms than on X, Musk's own social media platform.[45] Musk has also been linked to various opaquely funded 'dark money' groups, such as Building America's Future, which funded numerous disinformation campaigns, including a fake initiative called Progress 2028. The name suggested it was a Democratic counterpart to the Heritage Foundation's Project 2025 – an outline for the reformulation of the American state and society — but in fact, it was a pro-Trump deception. The group targeted adverts on Facebook linking to a website containing false Harris policy positions such as 'mandatory' gun buy-back and driving licenses for undocumented migrants. The ads were targeted to over 800 audience segments, but due to the transparency limitations of the Ad Library, we cannot be sure which targeting categories were used beyond simple demographics and geographies.[46] The **mandatory gun buyback disinformation ads**

[44] Jain, Y. & Walk, T. (2024). Disinformation About US Elections Targets Communities of Color. *Human Rights Watch.*

[45] Montgomery, B. (2024). Elon Musk's pro-Trump Pac pouring millions into Facebook ads instead of X. *The Guardian.*

[46] Koebler, J. (2024). Elon Musk-Funded PAC Supercharges 'Progress 2028' Democrat Impersonation Ad Campaign. *404 Media.*

were more likely to be delivered to males in key battleground states, including Michigan, Pennsylvania and North Carolina, presumably to convince conservative voters that Harris's views were more extreme than they in fact were.[47] [48] This was part of broader efforts by this coalition of Musk-backed PACs, which included the **targeting of disinformation at minorities and communities of colour on other platforms using ZIP codes and interests.**[49]

After the presidential election, Musk stated he would turn his campaign architecture to contesting midterm and judicial elections.[50] In February 2025, the Progress 2028 website pivoted to the Wisconsin Supreme Court race, purportedly supporting Democratic-endorsed candidate Susan Crawford. The website now contains fake Crawford quotes, likely designed to inspire an angry response through a contact form seemingly designed to collect information from voters, including ZIP codes and email addresses, which can be used for further targeting on Facebook or through direct mailings.[51] [52] [53]

DOPPELGANGER

In September 2024, the US Justice Department disrupted a covert Kremlin-directed influence operation codenamed Doppelganger. Doppelganger sought to reduce international support for Ukraine, bolster pro-Russian policies and interests, and influence elections, including the 2024 presidential election.[54] **Meta records confirmed the Russian Doppelganger operation used Facebook's advertising system to spread targeted pro-Russian propaganda**, in addition to tracking user reactions in real-time and adjusting disinformation campaigns based on engagement.[55]

One internal Doppelganger strategy document for a project called 'Good Old USA' detailed their tactics for the 2024 presidential election. Good Old USA had several goals including securing the victory of Donald Trump and reducing the approval rating of the Democratic candidate, in addition to reducing US public support for Ukraine and increasing the number of Americans who think the war should be ended soon with Ukrainian territorial concessions. The strategy included the use of targeted advertising on Facebook and Instagram. **Several audiences were marked as 'targets', including residents of swing states, residents of conservative states, Jewish and Hispanic Americans, along with communities of gamers and users of** 4chan (deemed as the "backbone" of US right-wing trends).[56] [57]

[47] Massoglia, A. (2024). Pro-Trump dark money network tied to Elon Musk behind fake pro-Harris campaign scheme. *Open Secrets.*

[48] Haberman, M. & Schleifer, T. (2024). Republican Operatives Function as Hidden Hand Behind Pro-Trump Efforts. *The New York Times.*

[49] Koebler, J. (2024). This Is Exactly How an Elon Musk-Funded PAC Is Microtargeting Muslims and Jews With Opposing Messages. *404 Media.*

[50] MSNBC. (2025). Elon Musk is trying to 'buy off' my opponent, Wisconsin Supreme Court candidate says. *MSNBC: Inside with Jen Psaki.*

[51] https://web.archive.org/web/20250224142249/https://progress2028.com/

[52] Swenson, A., & Bauer, S. (2025) A group funded by Elon Musk is behind deceptive ads in crucial Wisconsin Supreme Court race. *Associated Press.*

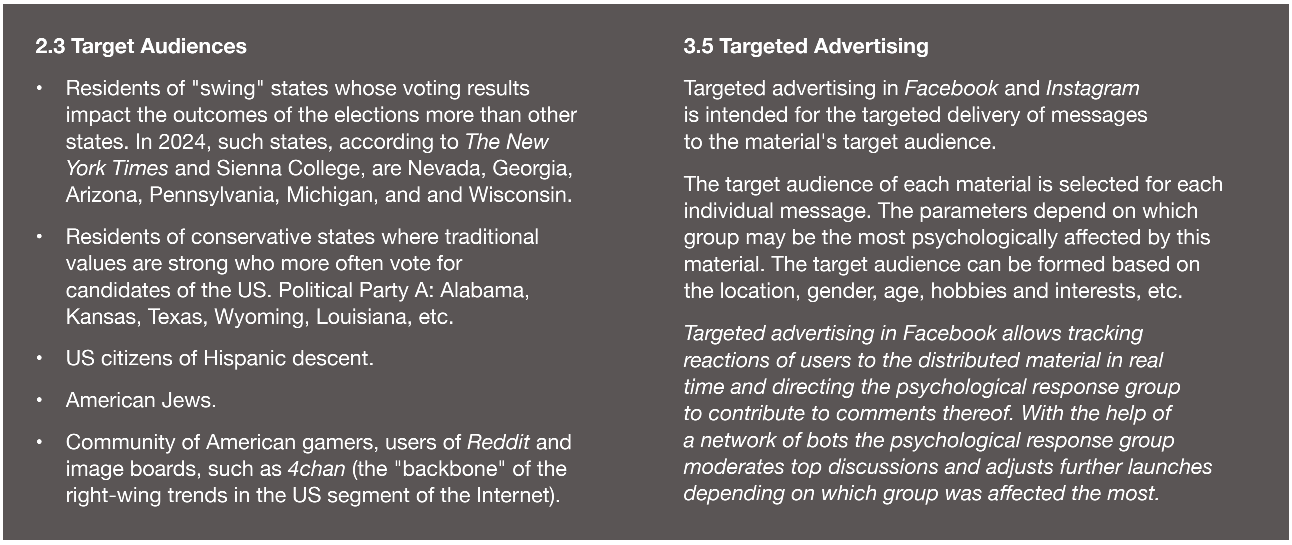
[53] In the process of compiling this report, we observed the Progress 2028 website pivot to focus on the Crawford campaign ahead of the release of its new wave of disinformation adverts. We shared our findings with the Crawford campaign.

[54] US DOJ. (2024). Justice Department disrupts covert Russian government-sponsored foreign malign influence operations. *U.S. Department of Justice.*

[55] US DOJ. (2024). United States of America v. Certain domains. Affidavit in Support of Seizure Warrants. Case No.: 24-mj-1395.

[56] US DOJ. (2024). United States of America v. Certain domains. Affidavit in Support of Seizure Warrants. Case No.: 24-mj-1395. Exhibit 8A. *The Good Old USA Project.*

[57] Gilbert, D. (2024). DOJ: Russia Aimed Propaganda at Gamers, Minorities to Swing 2024 Election. *Wired.*



2.3 Target Audiences

- Residents of "swing" states whose voting results impact the outcomes of the elections more than other states. In 2024, such states, according to *The New York Times* and Sienna College, are Nevada, Georgia, Arizona, Pennsylvania, Michigan, and and Wisconsin.
- Residents of conservative states where traditional values are strong who more often vote for candidates of the US. Political Party A: Alabama, Kansas, Texas, Wyoming, Louisiana, etc.
- US citizens of Hispanic descent.
- American Jews.
- Community of American gamers, users of *Reddit* and image boards, such as *4chan* (the "backbone" of the right-wing trends in the US segment of the Internet).

3.5 Targeted Advertising

Targeted advertising in *Facebook* and *Instagram* is intended for the targeted delivery of messages to the material's target audience.

The target audience of each material is selected for each individual message. The parameters depend on which group may be the most psychologically affected by this material. The target audience can be formed based on the location, gender, age, hobbies and interests, etc.

Targeted advertising in Facebook allows tracking reactions of users to the distributed material in real time and directing the psychological response group to contribute to comments thereof. With the help of a network of bots the psychological response group moderates top discussions and adjusts further launches depending on which group was affected the most.

Figure 3: Targeted advertising and target audience sections of ‘Exhibit 8A’, a translated Doppelganger Good Old USA project strategy document which aimed to secure the 2024 presidential election victory for Trump, and reduce American support for Ukraine (Source: US DOJ.)

The issue is not, however, contained to the USA. According to EU DisinfoLab, which has compiled a timeline of Doppelganger activities, the use of Meta’s advertisement platform is a “constant tactic used by Doppelganger operators”.[58] This has included using Facebook’s targeted advertising system to propagate anti-Ukraine messaging in Germany, France, Italy, Latvia, the UK, and Ukraine itself.[59] In 2024, not-for-profit Reset Tech, uncovered a pro-Kremlin campaign meddling in Moldova’s presidential elections and EU referendum. The researchers found the propaganda campaign was Facebook’s most profitable ad campaign for the entire country (up to 300,000 euros), representing almost a quarter of the total political ad spend on Facebook in Moldova since August 2020.[60] Meta is aware of Doppelganger’s malicious use of its platforms, yet has thus far been unable to stop it.[61]

Several reports have highlighted the privacy concerns regarding profiling, data use, and political microtargeting in Europe.[62] [63] [64] Open Rights Groups has recently warned about the potential for the centralisation of private company, political party, and state power in the UK through credit reference agencies and private companies hosting and creating political parties’ datasets and apps used for political canvassing.[65] Open Rights Group has also developed a tool for UK voters to opt out of political parties compiling and processing sensitive personal data.[66] However, as has been discussed in this section, beyond official political parties profiling and targeting voters with messages, bad actors use parallel techniques and targeted advertising to sow disinformation and disrupt democracy globally.

[58] https://www.disinfo.eu/doppelganger-operation

[59] Aleksejeva, N. et al. (2022). Russia-based Facebook operation targeted Europe with anti-Ukraine messaging. *Medium: DFRLab.*

[60] Atanasova, A. & Rusu, A. (2024) How Meta Benefits from Pro-Kremlin Election Meddling Ads in Moldova. *Reset Tech.*

[61] Meta. (2023) Draft: Adversarial Threat Report. *Second Quarter.*

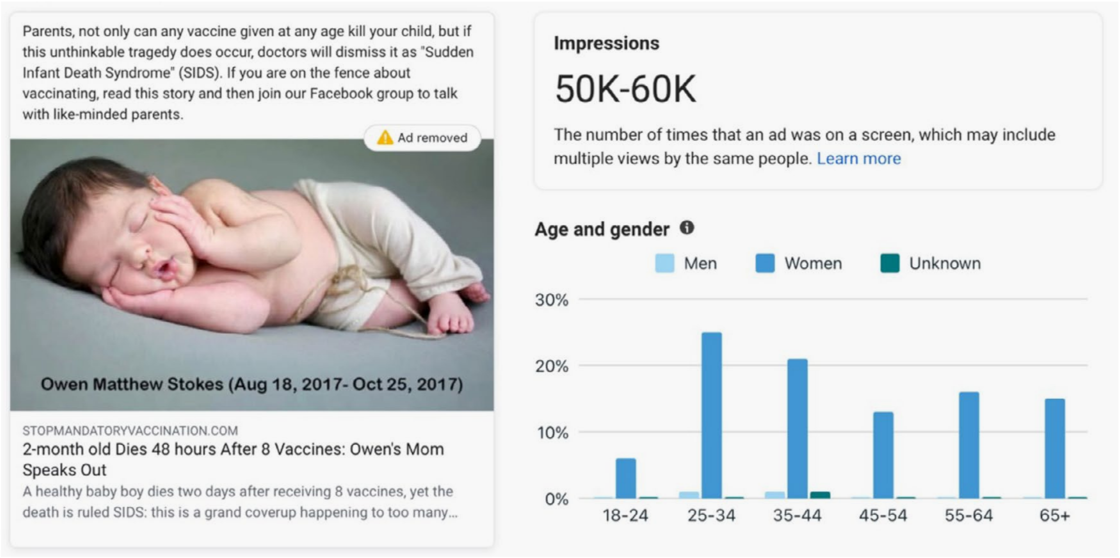
[62] ICO. (2018). Investigation into the use of data analytics in political campaigns: Investigation update. *Information Commissioner’s Office.*

[63] Iwańska, K. et al. (2020). Who (really) targets you? Facebook in Polish election campaigns. *Panoptykon Foundation.*

[64] Crowe, P. et al. (2020). Political Parties and Data Profiling: Who Do They Think We Are? *Open Rights Group.*

[65] Ohrvik-Stott, J. et al. (2025). Moral Hazard: Voter Data Privacy and Politics in Election Canvassing Apps. *Open Rights Group.*

[66] ORG. (2024). Opt-out of Political Parties Processing Your Data. *Open Rights Group.*



Parents, not only can any vaccine given at any age kill your child, but if this unthinkable tragedy does occur, doctors will dismiss it as "Sudden Infant Death Syndrome" (SIDS). If you are on the fence about vaccinating, read this story and then join our Facebook group to talk with like-minded parents.

⚠ Ad removed

Owen Matthew Stokes (Aug 18, 2017- Oct 25, 2017)

STOPMANDATORYVACCINATION.COM

2-month old Dies 48 hours After 8 Vaccines: Owen's Mom Speaks Out

A healthy baby boy dies two days after receiving 8 vaccines, yet the death is ruled SIDS: this is a grand coverup happening to too many...

Impressions

50K-60K

The number of times that an ad was on a screen, which may include multiple views by the same people. Learn more

Figure 4: A 2018 advert by the anti-vax group Stop Mandatory Vaccination, with demographic details of ad impressions showing it targeting women in addition to new parents (Source: Ad Library)

SCIENCE

NEW PARENTS: VACCINES CAN "KILL YOUR CHILD"

In 2018, a US-based anti-vaccination group called Stop Mandatory Vaccination launched a flurry of **anti-vaccine adverts targeting new parents in the UK, the USA, and Canada**. Later that year, a mother of a young baby in the UK referred one of Stop Mandatory Vaccination's adverts to the UK's advertising regulator, the Advertising Standards Authority.[67] The advert stated: "Parents, not only can any vaccine given at any age kill your child, but if this unthinkable tragedy does occur, doctors will dismiss it as 'Sudden Infant Death Syndrome' (SIDS)." The advert also contained an image of a baby with his eyes closed with the accompanying text: "Owen Matthew Stokes (Aug 18, 2017 - Oct 25, 2017). Beneath the image read "stopmandatoryvaccination.com - 2-month old Dies 48 hours After 8 Vaccines: Owen's Mom."[68]

Stop Mandatory Vaccination told the ASA they had: "**targeted users with an interest in parenting because they intended to cause parents some concern before choosing to vaccinate their children.**"[69] However, as can be seen from the Facebook Ad Library statistics above, it is clear that **the advert was also specifically targeted at women**. A 2020 study covering this period found that **anti-vaccination adverts on Facebook are more likely to target women.**[70]

The ASA ruled in November 2018 that as the ad featured an image of a baby with its birth and death dates alongside a claim that "any vaccine given at any age kills your child," the ad implied that all vaccinations are proven to cause death. The ad also suggested that doctors misattribute vaccine-related deaths to Sudden Infant Death Syndrome, a statement likely to cause fear and distress, particularly among parents seeking factual information. As no evidence supported these claims, **the ad was deemed to breach the UK advertising Code by spreading unjustified fear or distress, and the ASA ruled it should not appear again.**[71]

[67] Boseley, S. (2019). Half of new parents shown anti-vaccine misinformation on social media – report. *The Guardian.*

[68] ASA. (2018). ASA Ruling on Larry Cook t/a Stop Mandatory Vaccination. *Advertising Standards Authority.*

[69] ASA. (2018). ASA Ruling on Larry Cook t/a Stop Mandatory Vaccination. *Advertising Standards Authority.*

[70] Jamison, A. M. et al. (2020). Vaccine-related advertising in the Facebook Ad Archive. *Vaccine, 38(3), 512-520.*

[71] ASA. (2018). ASA Ruling on Larry Cook t/a Stop Mandatory Vaccination. *Advertising Standards Authority.*

After the ruling, Cook, who had paid for the advert, said, “After many weeks of investigation, [the ASA] determined I violated their advertising policies and demanded I not run similar ads in the future and sign documentation agreeing to such. I refused. I will not sign any such documents agreeing to their terms and I will continue to promote my messaging to the parents of the United Kingdom.” Cook concluded, “The ASA does not have jurisdiction over Facebook or me.”[72] Indeed, Cook did not comply with the ruling, and the ASA instead worked with Facebook to take the advert down.[73]

One study found that during this period, just two groups were responsible for 54% of all anti-vaccine adverts posted on Facebook: Stop Mandatory Vaccination, led by Larry Cook, and the World Mercury Project, then chaired by the now United States Secretary of Health and Human Services, Robert F. Kennedy Jr..[74] [75]

THE GROUNDWORK FOR THE COVID-19 INFODEMIC

In November 2018, it was revealed that Facebook was not only allowing the promotion of anti-vaccination adverts to specific demographic audience segments of its billions of users, but **Facebook was also allowing advertisers to target users profiled as having an interest in “vaccine controversies.”**[76] This targeting, based on anti-vaccine interests, allowed anti-vaccine groups to target individuals seemingly more receptive to vaccine disinformation.[77] [78]

An investigation by Spotlight found that **Facebook not only offered the ability for advertisers to target people interested in “vaccine controversies”, but the company’s Ads Manager tool automatically suggested complementary targeting categories such as “new parents” and “parents with toddlers (01-02 years)”**.[79] This occurred at the same time as a study by the Royal Society for Public Health found that 50% of UK parents with children under five years old reported being exposed to negative messages about vaccines on social media.[80]

After years of organic and paid-for vaccine mis- and disinformation circulating on the platform, it took until March 2019 for Facebook to take a firm policy stance after receiving a letter from US Congressman Adam Schiff expressing his concerns.[81] Facebook announced that when they found ads containing vaccine misinformation, they would remove them, and they would also be removing targeting options such as “vaccine controversies.”[82]

[72] https://www.bmj.com/content/363/bmj.k4720

[73] Boseley, S. (2019). Half of new parents shown anti-vaccine misinformation on social media – report. *The Guardian.*

[74] Jamison, A. M. et al. (2020). Vaccine-related advertising in the Facebook Ad Archive. *Vaccine, 38(3), 512-520.*

[75] Foran, C. et al. (2025). Senate confirms RFK Jr. as Health and Human Services secretary. *CNN.*

[76] Dunn, W. (2018). Anti-vaccination advert banned – but Facebook still offers targeting of people susceptible to “vaccine controversies”. *The New Statesman.*

[77] Glenza, J. (2019). Majority of anti-vaxx ads on Facebook are funded by just two organizations. *The Guardian.*

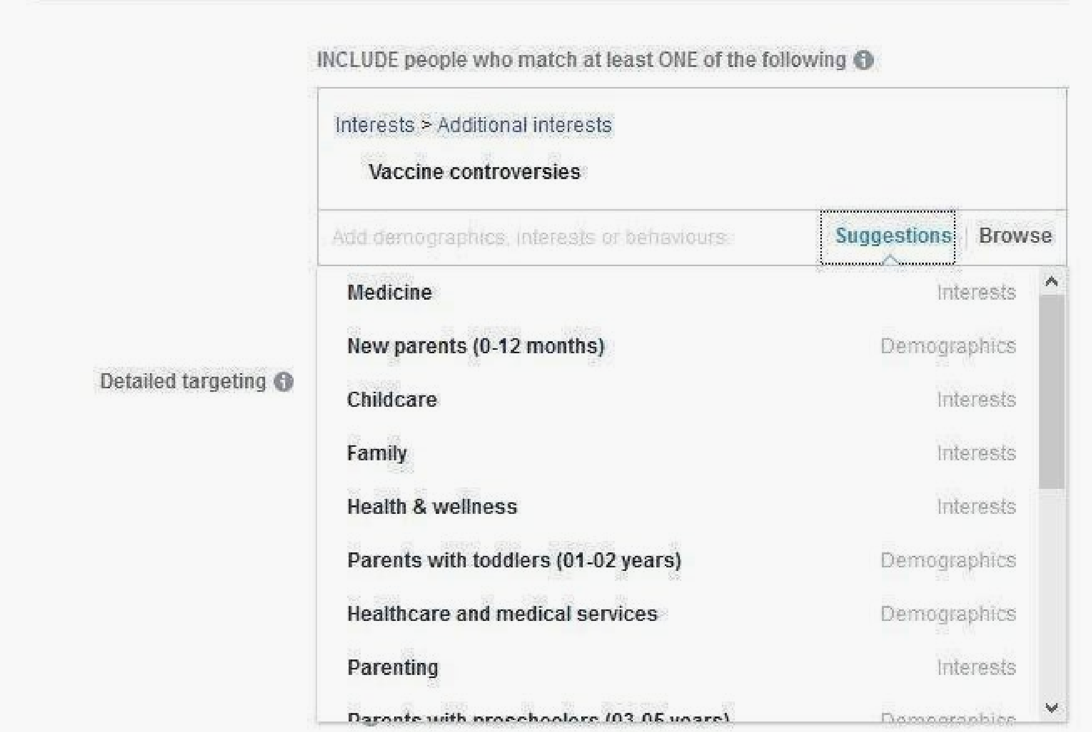
[78] Wong, J.C. (2019). Revealed: Facebook enables ads to target users interested in ‘vaccine controversies’. *The Guardian.*

[79] Dunn, W. (2018). Anti-vaccination advert banned – but Facebook still offers targeting of people susceptible to “vaccine controversies”. *The New Statesman.*

[80] RSPH. (2019). Moving the Needle: Promoting vaccination uptake across the life course. *Royal Society for Public Health.*

[81] https://thehill.com/policy/healthcare/430041-schiff-calls-out-facebook-google-over-anti-vaccination-information/

[82] https://about.fb.com/news/2019/03/combatting-vaccine-misinformation/



INCLUDE people who match at least ONE of the following

Interests > Additional interests

Vaccine controversies

Medicine	Interests
New parents (0-12 months)	Demographics
Childcare	Interests
Family	Interests
Health & wellness	Interests
Parents with toddlers (01-02 years)	Demographics
Healthcare and medical services	Demographics
Parenting	Interests
Parents with preschoolers (03-05 years)	Demographics

Figure 5: Facebook not only offered the option to target users categorised as having an interest in "vaccine controversies", but its algorithm suggested additional categories such as "Parents with toddlers" (Source: New Statesman)

The coronavirus pandemic, which began in late 2019, was accompanied by what the World Health Organisation (WHO) called the "infodemic" – an overabundance of information, including deliberate attempts to disseminate wrong information to undermine the public health response and advance alternative agendas. The WHO called on social media companies to combat the infodemic, and "further strengthen their actions to disseminate accurate information and prevent the spread of mis- and disinformation."[83]

In March 2020, as Mark Zuckerberg was publicly sharing updates about the work Facebook was doing to "limit the spread of misinformation"[84], investigative reporters at The Markup found that **Facebook was allowing advertisers to target people profiled as having an interest in "pseudoscience"**. The pseudoscience interest category contained 78 million people. A journalist at The Markup had received an advert selling a "radiation-blocking" hat at a time when 5G-related conspiracy theories were circulating concerning the origins of the virus. The advertiser stated that they had not selected the "pseudoscience" category and that Facebook had applied this interest category independently.[85] Also at this time, former Deputy Prime Minister Nick Clegg, then serving as Facebook's Vice President of Global Affairs and Communications, was publicly stating that Facebook would not allow posts saying bleach is in some way effective against coronavirus or that social distancing would not work.[86] Following these comments, investigative reporters at Consumer Reports uploaded adverts containing these exact messages. Facebook approved them all.[87]

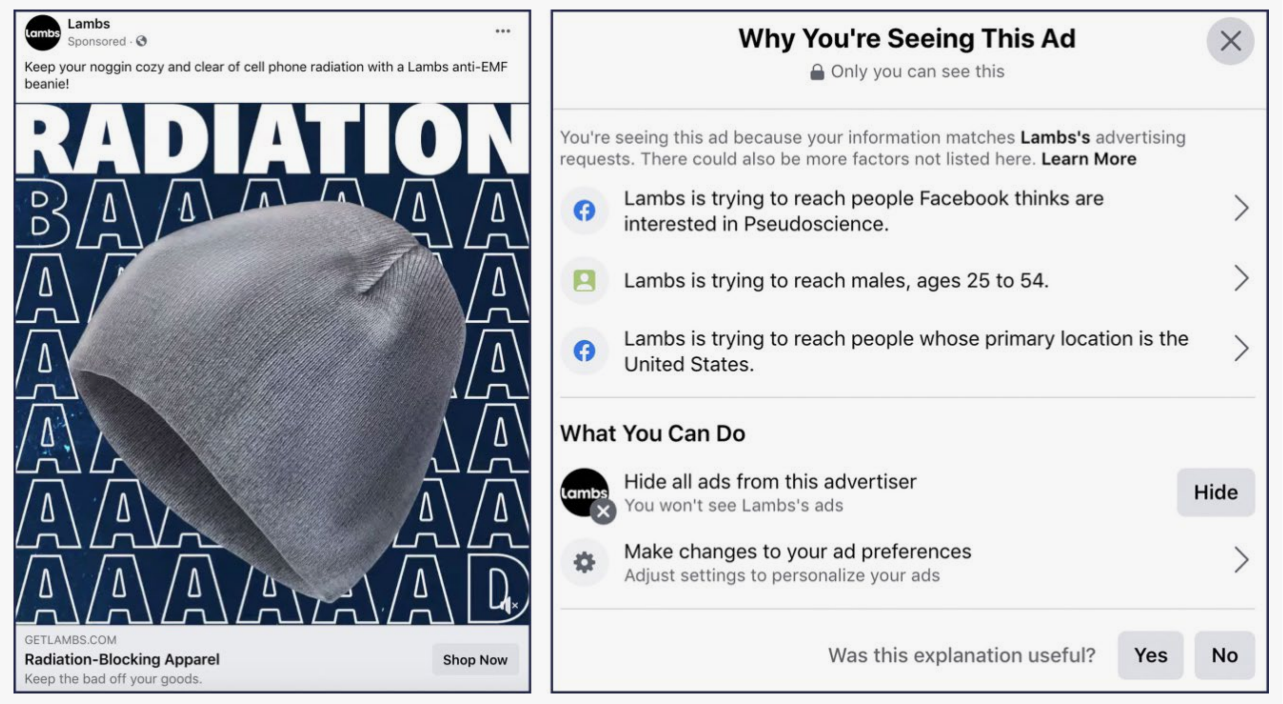
[83] WHO. (2020). Managing the COVID-19 infodemic: Promoting healthy behaviours and mitigating the harm from misinformation and disinformation. *World Health Organization.*

[84] Source: https://www.facebook.com/zuck/posts/10111806366438811

[85] Sankin, A. (2020). Want to Find a Misinformed Public? Facebook's Already Done It. *The Markup.*

[86] NPR. (2020). How Facebook Wants To Handle Misinformation Around The Coronavirus Epidemic. *All Things Considered: NPR.*

[87] Waddell, K. (2020). Facebook Approved Ads with Coronavirus Misinformation. *Consumer Reports.*



Lambs
Sponsored ·

Keep your noggin cozy and clear of cell phone radiation with a Lambs anti-EMF beanie!

GETLAMBS.COM
Radiation-Blocking Apparel
Keep the bad off your goods.

Why You're Seeing This Ad

Only you can see this

You're seeing this ad because your information matches **Lambs's** advertising requests. There could also be more factors not listed here. **Learn More**

Lambs is trying to reach people Facebook thinks are interested in Pseudoscience.

Lambs is trying to reach males, ages 25 to 54.

Lambs is trying to reach people whose primary location is the United States.

What You Can Do

Hide all ads from this advertiser
You won't see Lambs's ads

Make changes to your ad preferences
Adjust settings to personalize your ads

Was this explanation useful?

Figure 6: A 2020 advert selling "radiation-blocking" hat to Facebook users profiled as having an interest in "pseudoscience" (Source: The Markup)

It took until October 2020 for Facebook to explicitly ban adverts which discouraged people from getting vaccinated.[88] But warnings of the underlying cause of the so-called 'infodemic' came much earlier, with researchers at Ranking Digital Rights arguing in May that **"targeted advertising business models, and the opaque algorithmic systems that support them, are the root cause of their failure to staunch the flow of misinformation".**[89]

Facebook themselves seemed to implicitly acknowledge this, as in February 2021, they announced they had given "$120 million in ad credits to help health ministries, NGOs and UN agencies reach billions of people around the world with COVID-19 vaccine and preventive health information." Further, they provided training and marketing support to help governments and health organisations "move quickly and reach the right people with the latest vaccine information".[90] This move by Facebook was seen as "paradoxical" by some public health researchers, as Facebook's **"business decisions, and algorithms are in part responsible for the spread of misinformation about COVID-19 – including its failure to ban anti-vaccination advertisements until approximately seven months after the pandemic declaration".**[91] The Washington Post also revealed in the same year that Facebook had withheld key information from policymakers regarding how vaccine misinformation was spreading on its platforms.[92]

Ultimately, Facebook was a breeding ground for vaccine mis- and disinformation in the years preceding the COVID pandemic, which included the enabling of targeted vaccine disinformation using its advertising system. Further, Facebook's policies before and during the pandemic only seem to have had marginal effects on the spread of antivaccine content on

[88] Jin, K. et al. (2020). Supporting Public Health Experts' Vaccine Efforts. *Meta.*

[89] Maréchal, N. et al. (2020). Getting to the Source of Infodemics: It's the Business Model. *New America.*

[90] Jin, K. (2021). Reaching Billions of People With COVID-19 Vaccine Information. *Meta.*

[91] Zenone, M. et al. (2023). The Social Media Industry as a Commercial Determinant of Health. *Int J Health Policy Manag. 12:6840.*

[92] Lima-Strong, C. (2021). Facebook told the White House to focus on the 'facts' about vaccine misinformation. Internal documents show it wasn't sharing key data. *The Washington Post.*

its platform. A 2023 study in Science Advances concluded that Facebook's policies "may have reduced the number of posts in antivaccine venues but did not induce a sustained reduction in engagement with antivaccine content."[93] Similarly, a study of its initial 2019 policy shift found it only had relatively small effects.[94]

CLIMATE CHANGE DISINFORMATION

There is a wealth of evidence that Facebook profits from the spread of climate change disinformation, misinformation and greenwashing campaigns on its platform.[95] [96] [97] [98] Bad actors seeking to spread climate disinformation have used Facebook's profiling and targeting systems to send specific messages to specific audience segments.[99] **These climate disinformation campaign tactics are a dark mirror of the tactics called for by climate change communications scholars, who seek effective ways to inspire pro-environmental behaviour by segmenting audiences and delivering tailored messages.**[100] [101]

In 2020, ads on Facebook denying the climate crisis and the need for action were viewed over 8 million times.[102] Non-profit think tank Influence Map uncovered **strong evidence that opaquely funded climate change disinformation groups were using Facebook's ad targeting features to undermine public trust in the science of climate change**. The climate change disinformation ads were targeted particularly at males in US rural states. Further, 18 to 34-year-olds were more likely to be shown ads contesting the future consequences of climate change, while those 55 and older were more likely to be shown ads contesting the causes of climate change.[103] These ads revealed the "evolving misinformation playbook" from climate obstruction organisations, as they were differentially targeted to "tap into existing beliefs on fossil fuel consumption and potentially shape how people respond to climate change and voter choices."[104] A former director of sustainability at Facebook commented that **"the company's limited attempts to deal with the problem are failing to keep pace with powerful tactics like micro-targeting."**[105]

In 2024, researchers at the Massachusetts Institute of Technology and Northeastern University found that Facebook's algorithm was differentially distributing climate change adverts, despite the adverts not being explicitly targeted by advertisers. When ads were not targeted by location or demographics, Facebook's algorithm influenced the audience distribution by age, gender, and location. This created a price advantage for contrarian advertising (climate communication with contrary goals to climate action), particularly when targeting males, middle-aged and older individuals, and audiences in specific U.S. states. These specific groups have been

[93] Broniatowski, D.A. et al. (2023). The efficacy of Facebook's vaccine misinformation policies and architecture during the COVID-19 pandemic. *Science Advances, 9(37).*

[94] Gu, J. et al. (2022). The impact of Facebook's vaccine misinformation policy on user endorsements of vaccine content: An interrupted time series analysis. *Vaccine, 40(14), 2209-2214.*

[95] King, J. (2023). Deny, Deceive, Delay Vol. 2: Exposing New Trends in Climate Mis- and Disinformation at COP27. *Institute for Strategic Dialogue.*

[96] CAAD. (2024). Extreme Weather, Extreme Content: How Big Tech Enables Climate Disinformation in a World on the Brink. *Climate Action Against Disinformation.*

[97] SFH. (2021). In Denial - Facebook's Growing Friendship with Climate Misinformation. *Stop Funding Heat.*

[98] Gilbert, D. (2021). Facebook Is Making Millions Off Lies About the Climate Crisis. *Vice.*

[99] IM. (2020). Climate Change and Digital Advertising: Climate Science Disinformation in Facebook Advertising. *Influence Map.*

[100] AI. (2023). Communicating about Climate: Knowing Your Audience to Inspire Greater Action. *Aspen Institute.*

[101] Roser-Renouf, C. et al. (2014). Engaging Diverse Audiences with Climate Change: Message Strategies for Global Warming's Six Americas. *Climate Communication: Yale.*

[102] Carrington, D. (2020). Climate denial ads on Facebook seen by millions, report finds. *The Guardian.*

[103] IM. (2020). Climate Change and Digital Advertising: Climate Science Disinformation in Facebook Advertising. *Influence Map.*

[104] Holder, F. et al. (2023). Climate obstruction and Facebook advertising: how a sample of climate obstruction organizations use social media to disseminate discourses of delay. *Climatic Change, 176, 16.*

[105] Carrington, D. (2020). Climate denial ads on Facebook seen by millions, report finds. *The Guardian.*

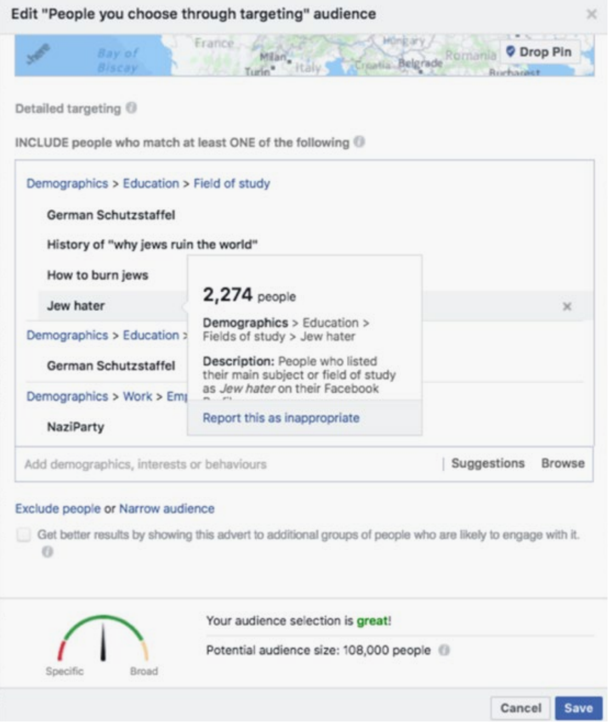
previously identified as more likely to have dismissive or doubtful attitudes about climate change. **This glimpse into the black box of Facebook's advertising delivery algorithm's biases shows that even in the absence of targeting on behalf of a bad actor, differential delivery can occur based on implicit profiling baked into the system.** Authors of the research warned that "without the inclusion of algorithmic bias in the framework of climate disinformation studies, the analysis of climate communication and disinformation research would be incomplete.[106]

HATE

ANTISEMITISM AND ISLAMOPHOBIA

In 2017, the Pulitzer Prize-winning nonprofit investigative journalism organisation, ProPublica, received a tip that **Facebook had listed several categories available for would-be advertisers to target to individuals with antisemitic interests.** These included the categories of "Jew hater," "How to burn jews," and "History of 'why jews ruin the world.'" These ad categories were likely automatically generated because users had listed these antisemitic themes on their profiles as an interest, an employer, or a field of study, with Facebook's algorithms transforming this data into advertising categories.[107]

To test if these categories were real, ProPublica paid to promote one of their news articles to individuals deemed to have these antisemitic interests. Facebook approved the ads for publication within 15 minutes. In response, a Facebook spokesperson told ProPublica: "There are times where content is surfaced on our platform that violates our standards… In this case, we've removed the associated targeting fields in question. We know we have more work to do, so we're also building new guardrails in our product and review processes to prevent other issues like this from happening in the future."[108]



Edit "People you choose through targeting" audience

Detailed targeting

INCLUDE people who match at least ONE of the following

Demographics > Education > Field of study

German Schutzstaffel

History of "why jews ruin the world"

How to burn jews

Jew hater

Demographics > Education >

German Schutzstaffel

Demographics > Work > Em

NaziParty

2,274 people

Demographics > Education > Fields of study > Jew hater

Description: People who listed their main subject or field of study as *Jew hater* on their Facebook

Report this as inappropriate

Add demographics, interests or behaviours | Suggestions Browse

Exclude people or Narrow audience

□ Get better results by showing this advert to additional groups of people who are likely to engage with it.

Your audience selection is great!

Potential audience size: 108,000 people

Figure 7: The 2017 ad buying process targeting antisemitic interests on Facebook's advertising portal (Source: ProPublica)

We are now in that future, yet the issues persist. Throughout 2023, Meta was criticised for approving adverts containing hate speech and incitements to violence against Palestinians, as well as for a lack of transparency regarding the funding of adverts attacking the pro-Palestine movement.[109] [110] Further, in 2024, corporate accountability group Ekō (formerly SumOfUs) uncovered a series of 23 ads – primarily paid for by Germany's far-right Alternative für Deutschland (AfD) party's regional groups and leading AfD elected representatives – containing hate speech, racist and anti-democratic narratives, and anti-immigrant disinformation. The ads, which had collectively received over 470,000 impressions, suggested

[106] Sankaranarayanan, A. et al. (2024). The Facebook Algorithm's Active Role in Climate Advertisement Delivery. *Research Square.*

[107] Angwin, J. et al. (2017). Facebook Enabled Advertisers to Reach 'Jew Haters'. *ProPublica.*

[108] Angwin, J. et al. (2017). Facebook Enabled Advertisers to Reach 'Jew Haters'. *ProPublica.*

[109] Biddle, S. (2023). Facebook allegedly approves paid ads containing hate speech & incitement against Palestinians. *The Intercept.*

[110] Visser, F. (2023). Facebook allegedly approves paid ads containing hate speech & incitement against Palestinians. *The Bureau of Investigative Journalism.*

Muslim immigrants were the main perpetrators of sexual assault, rape and other violent crimes, and called for mass deportations.[111] One of the ads Meta published was by the Saxony-Anhalt chapter of the AfD party, which was classified as a right-wing extremist group by state-level intelligence authorities in 2023. The evidence for this designation included Islamophobic, antisemitic and racist statements by elected officials of the chapter.[112]

Meta continues a lax approach to the moderation of paid-for hate on its platforms and a lack of contextual understanding of the environments in which it profits. In 2024, researchers uncovered a million-dollar hate network flooding social media ads in India.[113] Following this, in the lead-up to India's elections, Meta approved a series of inflammatory, violent adverts, including calls for Muslims to be burned.[114] The ads, submitted by Ekō to test Meta's approval process, directed hate towards Muslim minorities, disseminated disinformation and conspiracy theories prevalent in India, and incited violence through Hindu supremacist narratives. **As Facebook holds detailed location data on its users, it allowed the ads to be geographically targeted to highly contentious districts.** The ads were also placed as these districts entered the 'silence period', during which India's Election Commission prohibits any individual or group from posting advertisements or disseminating election materials from 48 hours before polling opens until after polling closes.[115]

In late July and early August 2024, after a tragic attack in Southport which left three children dead, a flood of anti-immigration demonstrations targeting mosques and hotels housing asylum seekers gripped the UK. A lack of information released from official sources created an information vacuum, leading to speculation and assertions that the attacker was a Muslim asylum seeker.[116] The UK's Science, Innovation and Technology Committee launched an inquiry in November 2024 to investigate the role social media and algorithms played in the violence following the attack.[117] How Meta adverts fed into the picture remains to be established, but on X, ads ran on accounts that were pushing misinformation and hate. This means not only did X financially benefit from the surge of falsehoods relating to the Southport attack, but so did high-profile accounts whose posts fanned the flames of violence.[118] The Cabinet Office spent over £160,000 on their own social media campaign attempting to quell the violence.[119] According to polling by YouGov in August 2024, 71% of Britons said social media networks did a bad job at tackling misinformation during the riots, and 66% of Britons said social media companies should be held responsible for posts by users inciting criminal behaviour during the riots. Further, 70% of Britons said social media companies are not regulated tightly enough.[120]

SECTARIAN HATE AND VIOLENCE IN NORTHERN IRELAND

In 2021, longstanding sectarian tensions across Loyalist and Unionist communities were again rising in Northern Ireland. This was in part due to anger regarding the Northern Ireland Protocol, which, in response to Brexit, instantiated a de facto customs border down the Irish Sea. This symbolic and legislative cleavage of Northern Ireland from the rest of

[111] Wyatt, V. (2024). New report: Meta profiting from far-right ads pushing hate speech and Islamophobia to voters in Germany. *Ekō.*

[112] Hülsemann, L. & Wilke, P. (2023). AfD in Saxony-Anhalt classified as right-wing extremist. *Politico.*

[113] Ekō. (2024). Slander, Lies, and Incitement: India's million dollar election meme network.

[114] Ellis-Petersen, H. (2024). Revealed: Meta approved political ads in India that incited violence. *The Guardian.*

[115] Ekō. (2024). As India election underway, Meta approves series of violent, inflammatory, Islamophobic AI-generated ads targeting voters.

[116] Coffey, J. & Moritz, J. (2025). Inadequate information released after Southport attack by authorities, says terror law reviewer. *BBC Panorama.*

[117] UK Parliament. (2024). Social media, misinformation and harmful algorithms. Science, *Innovation and Technology Committee.*

[118] CCDH. (2024). X Ran Ads on Five Accounts Pushing Lies and Hate During UK Riots. *Centre for Countering Digital Hate.*

[119] Milmo, C. (2024). Cabinet Office spent £160k on social media ads trying to quell far-right riots. *The i Paper.*

[120] Smith, M. (2024). Two thirds of Britons say social media companies should be held responsible for posts inciting riots. *YouGov.*

the United Kingdom ignited anger in Northern Ireland's Unionist communities.[121] In April, the riots that erupted in Belfast and further afield were, according to the police, at a scale that had not been seen for a number of years.[122]

During the increasing tensions leading to these riots, human rights campaign organisation Global Witness designed a series of adverts containing divisive, hate-filled messages and direct incitements of violence targeting Northern Irish Facebook users. This was done to experimentally test Facebook's advertising moderation systems in the context of increasing sectarian tensions. **The ads were targeted across the sectarian divide by using readily available proxies for Catholic and Protestant religious affiliation, themselves proxies for Unionism and Loyalism**. To achieve this, Global Witness targeted their adverts at users whom Facebook had profiled as having an interest in Protestantism and the Catholic Church. Additionally, they **geographically targeted communities using postcodes which fell across the Catholic Falls Road side and Protestant Shankill Road side of the peace wall in west Belfast**.[123]

Initially, two ads were designed and submitted to Facebook's moderation system. The first **targeted people who Facebook had profiled as having an interest in Protestantism**, saying "Northern Ireland is for the British - join the cause." The other **targeted people who Facebook had profiled as having an interest in the Catholic Church**, saying "They'll never leave the North of Ireland unless we make them." The next set of ads expressed the inferiority of and contempt for Protestants and Catholics, using offensive sectarian slurs, which violated Facebook's community standards regarding hate speech directed at the protected characteristic of religious affiliation.[124]

Finally, Global Witness submitted an advert that directly incited violence, containing the words "Voting hasn't worked, take to the streets."

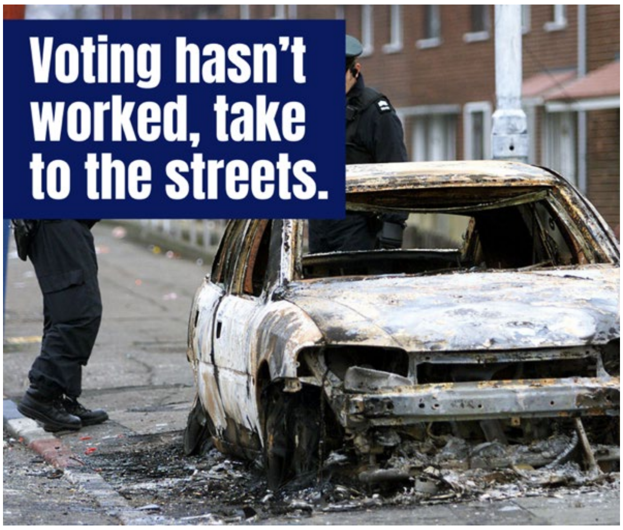


Figure 8: A 2021 ad targeted to people in Northern Ireland who Facebook had profiled as having an interest in Protestantism or the Catholic Church, in addition to those with postcodes on either side of the peace wall in west Belfast. (Source: Global Witness)

This ad was targeted in two ways. Firstly, as per the previous two examples, by using quasi-religious identifiers Facebook had profiled users as belonging to – Protestantism and the Catholic Church. Secondly, by **geographically targeting the postcodes of people who lived on either side of the peace wall in west Belfast, representing Catholic-majority and Protestant-majority communities**. It was across this very geographic divide that violence soon erupted.[125]

Every targeted ad Global Witness submitted was accepted for publication, often approved in just a few hours. None of the ads were ever seen by the targeted communities, however, as Global Witness withdrew them before they went live.[126]

When presented with this investigation, a Facebook spokesperson said several of the adverts violated their policies. They also said, "People's interests are based on their activity on Facebook -- such as the pages they like and the ads they click on -- not their personal

[121] Fitzpatrick, J. (2021). NI Protocol: Palpable anger but no return to violence, says Sheridan. *BBC NI Spotlight.*

[122] BBC. (2021). Belfast: Rioting 'was worst seen in Northern Ireland in years'. *BBC News.*

[123] GW. (2021). The Big Tech business model poses a threat to democracy. *Global Witness.*

[124] GW. (2021). The Big Tech business model poses a threat to democracy. *Global Witness.*

[125] Hirst, M. (2021). NI riots: What is behind the violence in Northern Ireland? *BBC News NI.*

[126] GW. (2021). The Big Tech business model poses a threat to democracy. *Global Witness.*

attributes." This represents a potential side-stepping of legal obligations, as there are protections in place regarding the processing of personal data around protected characteristics such as religious affiliation. Global Witness suggested: "Facebook is attempting to wriggle out of this obligation by claiming that people's interest in a topic such as Protestantism or the Catholic Church does not reveal anything about their religious views." Concluding, "**For as long as Facebook's business model is selling our profiles to advertisers, based on deeply personal predictions about us such as our religious views, the system will be open to abuse by those who wish to polarise us.**"[127]

From 2021, Global Witness has continued to experimentally test the ability to spread hate, division and disinformation through advertising on social media. Facebook has approved for publication test adverts inciting violence and genocide against the Rohingya in Myanmar in 2022,[128] adverts containing racist and far-right hate speech in Norway in 2022,[129] and adverts containing extreme violent hate directed at LGBTQ+ communities in Ireland in 2023.[130] Global Witness's investigations demonstrate Meta's and other social media platforms' consistent failure to implement their own policies on hate speech and disinformation in adverts on their platforms.[131] These tests confirm how easy it is for bad actors to use Facebook's advertising systems to sow social division, hate, and violence around the world. They also highlight how Facebook has little institutional understanding of the many geographies, cultures, and historical and political contexts in which it advertises and profits. It offers a targeted division on-demand service.

FEAR

DETERRING REFUGEES WITH PATCHWORK PROFILING

Between 2021 and 2022, **the UK's Home Office ran a series of targeted, fear-based ads aiming to deter refugees from crossing the English Channel in small boats.** The ads showed sinking boats, search dogs, and military-style drones, suggested smugglers would betray them, and they were likely to die in the Channel. The campaign, named Migrants on the Move, was delivered in collaboration with a 'migration behaviour change' agency called Seefar. Written in Arabic, Pashto, and Vietnamese, the adverts were designed to target refugees in Northern France and Belgium, aiming to 'nudge' them away from attempting the crossing.[132]

According to an investigation funded by the Scottish Institute for Policing Research, the targeting was highly invasive, splitting audiences into more than 600 different segments, some as small as a few hundred people—for example, Kurdish speakers in Brussels or Vietnamese travellers in Calais. Others had a reach of up to 100,000 people, such as all Arabic speakers over 18 in Brussels. **The ads also used 'patchwork profiles', which stitched together several interests, behavioural and language categories to reconstruct a targeted refugee**. For example, Pashto speakers, with an interest in the Afghan cricket team, who Facebook had flagged as "travelling away from family" or "away from home".[133] Some of these profiling categories were likely created by Facebook to market products to holidaymakers; the Home Office campaign hijacked them to target vulnerable refugees.[134]

[127] GW. (2021). The Big Tech business model poses a threat to democracy. *Global Witness.*

[128] GW. (2022). Facebook approves adverts containing hate speech inciting violence and genocide against the Rohingya. *Global Witness.*

[129] GW. (2022) An open door to hate: Meta approves ads containing far-right hate speech in Norwegian. *Global Witness.*

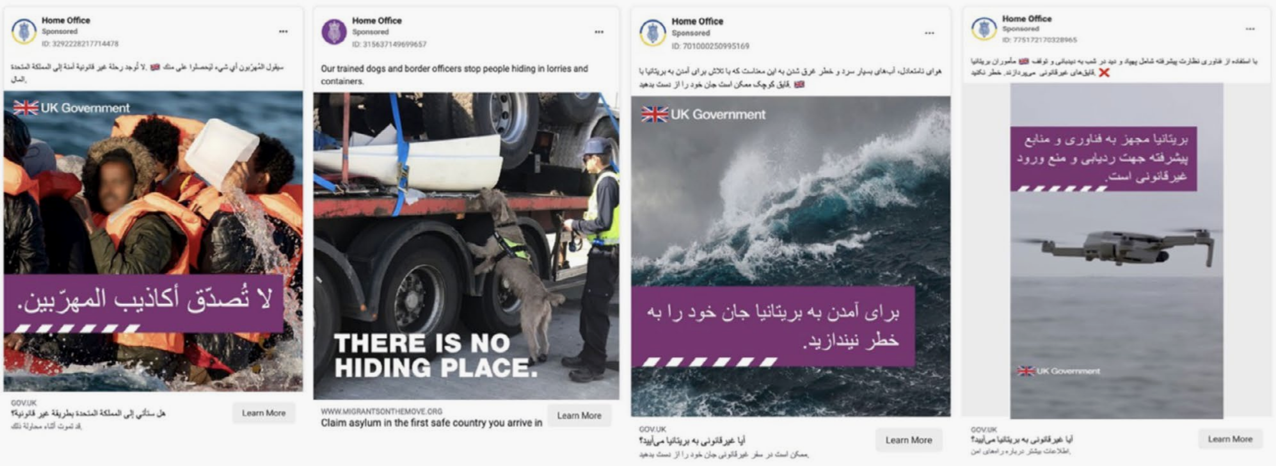
[130] GW. (2023). Extreme and violent anti-LGBTQ+ hate approved for publication by leading social media platforms. *Global Witness.*

[131] GW. (2023). A world of online hate and lies: Mapping our investigations into social media platforms' failure to tackle hate and disinformation. *Global Witness.*

[132] Collier, B. et al. (2023). Influence Policing: Strategic communications, digital nudges, and behaviour change marketing in Scottish and UK preventative policing. *Scottish Institute for Policing Research.*

[133] Collier, B. et al. (2024). Influence government, platform power and the patchwork profile: Exploring the appropriation of targeted advertising infrastructures for government behaviour change campaigns. *First Monday.*

[134] Collier, B. (2023). The UK Uses Targeted Facebook Ads To Deter Migrants. Now Meta Is Releasing the Data. *New Lines Magazine.*



Home Office Sponsored ID: 3292228217714478

Home Office Sponsored ID: 315637149699657 Our trained dogs and border officers stop people hiding in lorries and containers.

WWW.MIGRANTSONTHEMOVE.ORG Claim asylum in the first safe country you arrive in

Home Office Sponsored ID: 701000250995169

Home Office Sponsored ID: 775172170328965

UK Government
لا تُصدق أكاذيب المهرّبين.
THERE IS NO HIDING PLACE.
UK Government
برای آمدن به بریتانیا جان خود را به خطر نیندازید.
بریتانیا مجهز به فناوری و منابع پیشرفته جهت ردیابی و منع ورود غیرقانونی است.
UK Government
GOV.UK
GOV.UK
GOV.UK
Learn More
Learn More
Learn More
Learn More

Age: 18-65+

Gender: All

Language: Arabic

Bourseville, Fontaine-sur-Somme, Saint-Quentin-en-Tourmont

Interests: Afghan Premier League, Afghan Star, Afghan Wireless, Afghanistan, Afghanistan national cricket team, Afghanistan national football team, Aleppo, Baghdad, Cinema of Iran, Damascus, Eritrea, Football in Iraq, Homs, Iran, Iran national football team, Iraq, Iraq Football Association, Iraq national football team, Iraqi Kurdistan, Iraqi Premier League, Iraqi cuisine, Kabul, Kurdistan, Lebanon, MTN Syria, Music of Afghanistan, Music of Iran, South Sudan, South Sudan national football team, Sudan, Syria, Syria (region), Syria TV, Syria national football team, Syrian cuisine, Syrianska FC, The Voice of Vietnam, Vietnam national football team, Vietnamese language, mtn afghanistan

Location: TRAVELLING THROUGH: Blankenberge, Nazareth, Comines, Nord-Pas-de-Calais, Dunkirk, Grande-Synthe, Gravelines, Monchy-Breton, Saint-Martin-Boulogne, Picardie, Bourseville, Fontaine-sur-Somme, Saint-Quentin-en-Tourmont

Figure 9: Adverts and one "patchwork profile" example of the over 600 targeting segments from the Home Office's 2021 – 2022 Migrants on the Move campaign, which used fear to deter refugees from crossing the Channel. Images contain the text: "There are large ships in the ocean, which can be deadly for small ships. Do not take this risk." (Source: Scottish Institute for Policing Research)

The researchers who studied the campaign argued that in Brussels, Arabic speakers saw these ads while their French-speaking neighbours did not, creating a digital version of Theresa May's 'Go Home' vans, invisible to everyone except the targeted groups.[135] Despite its precision, the campaign also 'misfired', reaching (presumably perplexed) business travellers and holidaymakers in Mexico, India and Jordan. One of the study's authors, Dr Ben Collier, of the University of Edinburgh, suggested that "thousands of Arabic speakers around the world, including many visiting Brussels on holiday or for business, have been targeted by this campaign".[136] Concluding, "This is a really horrible fear-based campaign targeting refugees in France and in Belgium, and obviously, in France, it is absolutely illegal for you to target anything, or even collect data on ethnicity. And so I think the issues with the legality of using this stuff internationally is pretty questionable as well."[137]

This case raises serious ethical and legal concerns. The ads were explicitly designed to create fear, seeking to target already vulnerable people who had taken great risks to reach Europe. Given this, they were unlikely to deter crossings—instead, they inflicted additional harm and may have even deepened distrust in authorities, pushing asylum seekers further from support networks. This campaign is a stark example of how surveillance capitalism, profiling, and microtargeting can be weaponised by state actors.[138]

[135] White, M. (2024). Theresa May admits mistakes over migrant policies. *BBC News*.

[136] Collier, B. (2023). The UK Uses Targeted Facebook Ads To Deter Migrants. Now Meta Is Releasing the Data. *New Lines Magazine*.

[137] Benjamin, J. (2023). Government 'invasively' targeted digital users using 'offensive stereotypes'. *The Media Leader*.

[138] Collier, B. et al. (2023). Influence Policing: Strategic communications, digital nudges, and behaviour change marketing in Scottish and UK preventative policing. *Scottish Institute for Policing Research*.

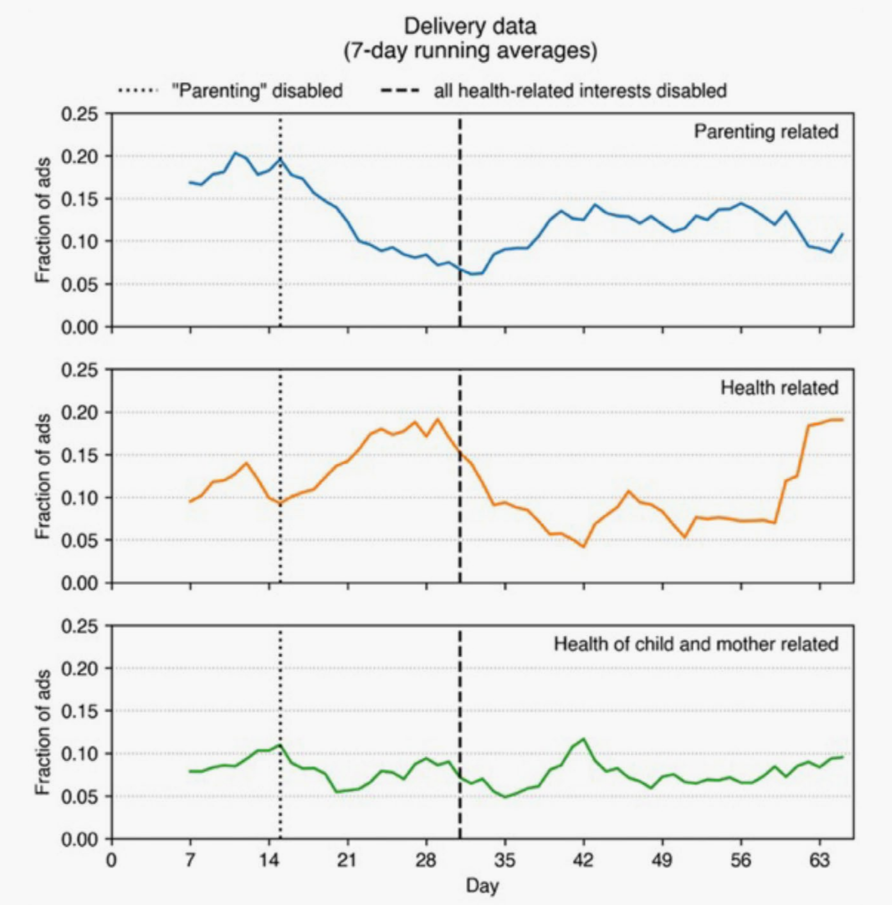


Figure 10: Graph displaying results of a study showing the disabling of specific ad targeting options is ineffective over time (Source: Panoptykon Foundation)

NO ESCAPE FROM ALGORITHMIC TRAUMA

In 2021, a young Polish mother became aware she was being bombarded with distressing content on Facebook. Her feed was flooded with health-related ads, particularly concerning cancer, genetic disorders, and other serious medical conditions. Some ads also contained crowdfunding campaigns for children or young adults suffering from these conditions. 'Joanna'[139] had recently suffered the loss of a loved one from cancer, and this disturbing content compounded her anxiety, being an unwanted reminder of her trauma.[140]

Panoptykon Foundation, a Polish human rights organisation, in collaboration with Dr Piotr Sapieżyński, a research scientist at Northeastern University, analysed over 2,000 ads in Joanna's Facebook feed over two months. One in five of these ads were health-related, often featuring terminally ill children or fertility issues. **Facebook had profiled her as having 21 sensitive health-related 'interests', including "oncology", "cancer awareness", "genetic disorder", and "spinal muscular atrophy"**. These profiling categories had been inferred by the platform and were likely based on her online activities both on and off Facebook.[141] "I wouldn't say Facebook caused my health-related anxiety," Joanna said, "but I feel it is exploited against me and it just fuels it and makes it worse."[142]

To test the level of control users have over the adverts they see, the researchers disabled a number of Joanna's profiled health-related interests. While the number of disturbing

[139] 'Joanna' is a pseudonym used to protect the woman's privacy.

[140] Głowacka, D. & Iwańska, K. (2021). Algorithms of trauma: new case study shows that Facebook doesn't give users real control over disturbing surveillance ads. *Panoptykon Foundation.*

[141] Głowacka, D. & Iwańska, K. (2021). Algorithms of trauma: new case study shows that Facebook doesn't give users real control over disturbing surveillance ads. *Panoptykon Foundation.*

[142] Murgia, M. (2021). Time to turn off Facebook's digital fire hose. *Financial Times.*

ads seen did change during the experiment, after two months, they had almost returned to the original level. The researchers claimed the experiment demonstrated that **"users are unable to get rid of disturbing content: disabling sensitive interests in ad settings limits targeting options for advertisers, but does not affect Facebook's own profiling and ad delivery practices"**.[143]

Over time, however, Joanna noticed an increased number of 'suggested posts' containing the same unwanted, unsettling themes she had tried to purge from the adverts in her timeline. In 2023, a follow-up study tasked Joanna with systematically clicking the option "Hide post – See fewer posts like this", which is supposed to retrain the algorithm away from its former profiling. Despite this, while the number of these unwanted suggested posts decreased in the first few days, subsequently, the number of posts exceeded the level before the intervention. The researchers concluded: "The buttons which should technically impact the content of the feed when clicked, and free the user of unwanted posts, do not work."[144]

Facebook's algorithm uses such a complex array of data points to pinpoint our interests that exactly how it works is unknowable, even to Facebook. There is also no way to fully turn this system off, even by "opting out". "We tell people . . . that removing interests or hiding topics will not stop every related ad, which is why we offer a range of ways to improve the ads experience," a Facebook spokesperson told the Financial Times.[145]

One woman, who had experienced a difficult pregnancy which led to a miscarriage, resorted to repeatedly Googling the word "miscarriage" in an attempt to shift Facebook's algorithm's profiling of her as pregnant and bombarding her with pregnancy-related ads. Facebook continued to show her adverts for baby products, prams, and parenting groups, compounding her distress. "I just didn't know what else to do," she said, "I felt really helpless throughout the pregnancy and now I feel even more helpless."[146]

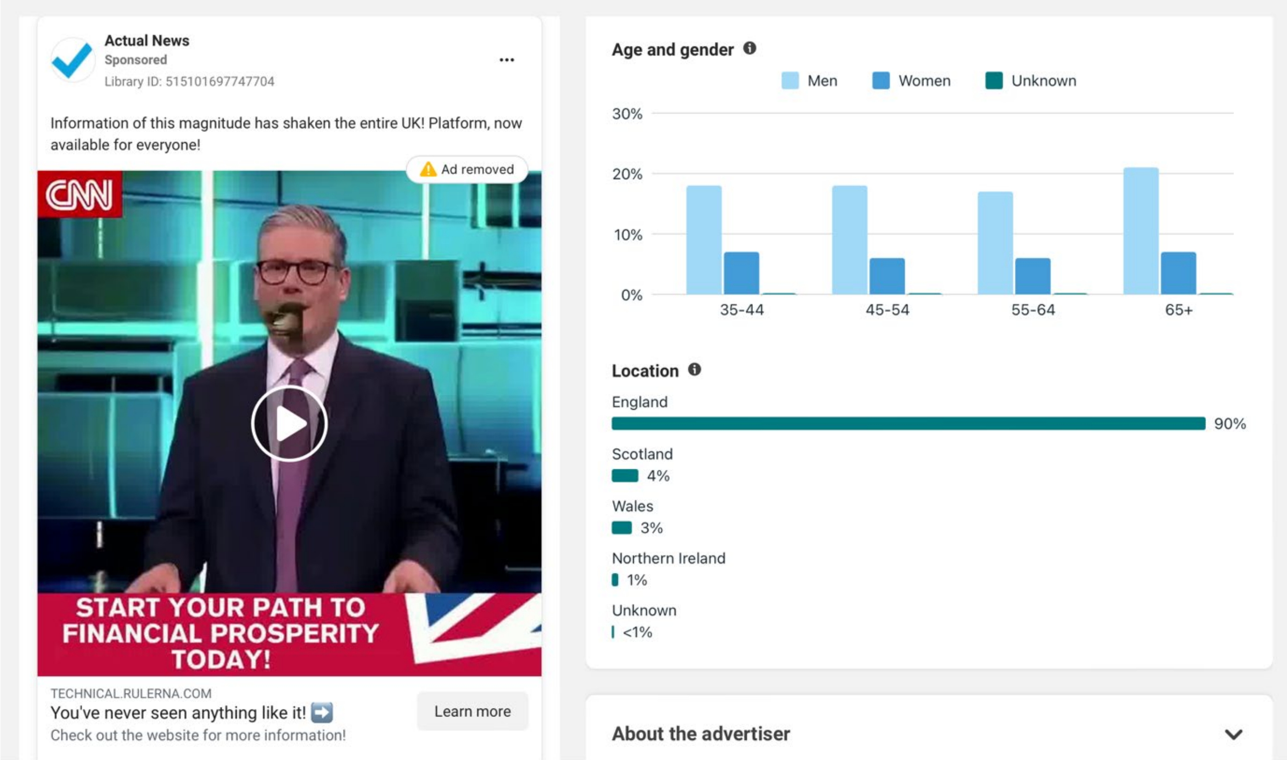
These examples demonstrate that bad ads don't have to be the result of bad actors, they can be a function of bad algorithms. In this woman's case, Facebook's algorithm was quick to learn she was pregnant but slow to learn she had a miscarriage. The detrimental effects of adverts are context-dependent and can occur when surveillance advertising is used for non-malicious intent by an advertiser. It is the nature of profiling itself, coupled with algorithmic recommendation, which is at the heart of this issue. **Inferring user interests and pushing content based on these profiled predictions can lead to the intensification of distress, trauma and anxiety for individuals online.** This is doubly concerning given that the purported options Facebook presents to users to reduce targeting based on profiling are seemingly ineffective buttons.

[143] Głowacka, D. & Iwańska, K. (2021). Algorithms of trauma: new case study shows that Facebook doesn't give users real control over disturbing surveillance ads. *Panoptykon Foundation.*

[144] Obem, A. & Wróblewska, M. (2023). Anxious about your health? Facebook won't let you forget. *Panoptykon Foundation.*

[145] Murgia, M. (2021). Time to turn off Facebook's digital fire hose. *Financial Times.*

[146] Moss, R. (2019). This Is What It's Like To Be Targeted By Baby Ads After Miscarriage Or IVF Struggles. *Huffington Post.*



Actual News
Sponsored
Library ID: 515101697747704

Information of this magnitude has shaken the entire UK! Platform, now available for everyone!

⚠ Ad removed

START YOUR PATH TO FINANCIAL PROSPERITY TODAY!

TECHNICAL.RULERNA.COM
You've never seen anything like it! ➡️
Check out the website for more information!

Learn more

About the advertiser

Figure 11: Ad Library data for one of the deepfake scam investment ads uncovered by Fenimore Harper. The deepfake Starmer invites people to "join us, and start your journey to financial wellbeing today" (Source: Ad Library)

FRAUD

DEEPFAKE FINANCIAL SCAMS

Recent advances in artificial intelligence have accelerated the production of deepfakes – a form of synthetic media characterised by plausible and realistic videos, pictures, text or audio of events which never happened.[147] [148] Deepfakes have become commonplace online, particularly imitations of high-profile figures, and are increasingly used in scam adverts. Ofcom data from 2024 found that 45% of adults in Great Britain reported seeing a deepfake scam advert, and 49% said they had seen a deepfake of a politician or political event. Worryingly, 32% of children between 8 and 15 years old had seen a deepfake scam advert, and 28% said they had seen a deepfake of a politician or political event. However, those who were retired were more likely to report seeing a deepfake scam advert (59%).[149] [150]

In 2024, media insight and research firm Fenimore Harper found that 43% of Meta ads about Starmer contained "harmful financial disinformation". These adverts used deepfakes of Kier Starmer and Prince William to promote a cryptocurrency scam website, suggesting it had been endorsed by the UK Government and Royal Family.[151] Up to £21,053 was spent, and the adverts reached over 890,000 people.

[147] US Gov. (2022). Phase 2: Deepfake Mitigation Measures. Department of Homeland Security.

[148] US Gov. (2022). Phase 2: Deepfake Mitigation Measures. Department of Homeland Security.

[149] Ofcom. (2024). A deep dive into deepfakes that demean, defraud and disinform.

[150] Ofcom. Open Data.

[151] FH. (2024). 43% of Meta Ads About Starmer Are Disinformation. Fenimore Harper.

Earlier in the year, Fenimore Harper, had uncovered a similar deepfake scam network impersonating Rishi Sunak, again promoting scam investment platforms.[152] Fenimore Harper argued the findings "indicate that AI-powered disinformation campaigns are a growing threat on social media platforms".[153] **We have inferred evidence of these ads being targeted, as Ad Library data shows demographic skews in delivery towards male Facebook users, and some adverts skewed in delivery towards older users**. However, again, we cannot tell from Meta's Ad Library if this differential delivery was due to the selection of interest categories which skew towards these categories, as Meta chooses not to make this information public.

The UK's National Crime Agency suggest the threat to the UK from fraud has grown in the last 10 years, with 89% of reported fraud cases now being cyber-enabled. [154] Fraud represents the most common type of crime experienced by individuals in England and Wales, with an estimated 3.6 million incidents in the year ending June 2024.[155] The NCA warned that the use of artificial intelligence and deepfakes will be adopted by more criminals to "increase the scale and sophistication of frauds."[156]

UK banks have also become increasingly concerned about the prevalence of such scams on social media. In 2023, Lloyds Banking Group warned of a surge in crypto investment scams and that a "growing number of British investors risk being defrauded by a wave of fake adverts posted on social media". Lloyds data suggested that 66% of all investment scams start on social media, with Facebook and Instagram being the most common sources, warning that over recent years criminals had "widened their net to target younger investors, who are often tempted by the supposed 'get rich quick' promise of cryptocurrency trading." According to Lloyds data, the most common age range for crypto scam victims is 25 to 34 years old.[157] In 2023, TSB Bank called on Meta to "face up to its responsibility" and do more to protect customers from "spiralling levels of fraud" on its platforms. The bank warned consumers could lose £250m from fraud on its platforms that year and called on Meta to filter out and block "obviously fraudulent" adverts.[158]

A Meta spokesperson said in 2023 "We don't want anyone to fall victim to these criminals, which is why our platforms have systems to block scams, financial services advertisers now have to be FCA (Financial Conduct Authority)-authorised and we run consumer awareness campaigns on how to spot fraudulent behaviour."[159] Given recent evidence, it is clear Meta continues enabling these frauds to be advertised to citizens, including the use of its targeting system, which allows the tailoring of messages to groups perceived to be more vulnerable to them.

FAKE IDENTITY DOCUMENTS FOR SALE

In early 2025, a Sun on Sunday investigation revealed a network of accounts offering fraudulent British passports and driving licences. These were, according to the investigation, being advertised to migrants to help them evade police detection and secure jobs illegally after arriving in the country. A Meta spokesperson said, "Fraudulent activity is not allowed on our platforms and we remove ads and accounts which violate our policies."[160]

[152] FH. (2024). Over 100 Deep-Faked Rishi Sunak Ads Found on Meta's Platform. *Fenimore Harper.*

[153] FH. (2024). 43% of Meta Ads About Starmer Are Disinformation. *Fenimore Harper.*

[154] NCA. (2024). National Strategic Assessment: Fraud. *National Crime Agency.*

[155] ONS. (2024). Crime in England and Wales: year ending June 2024. *Office for National Statistics.*

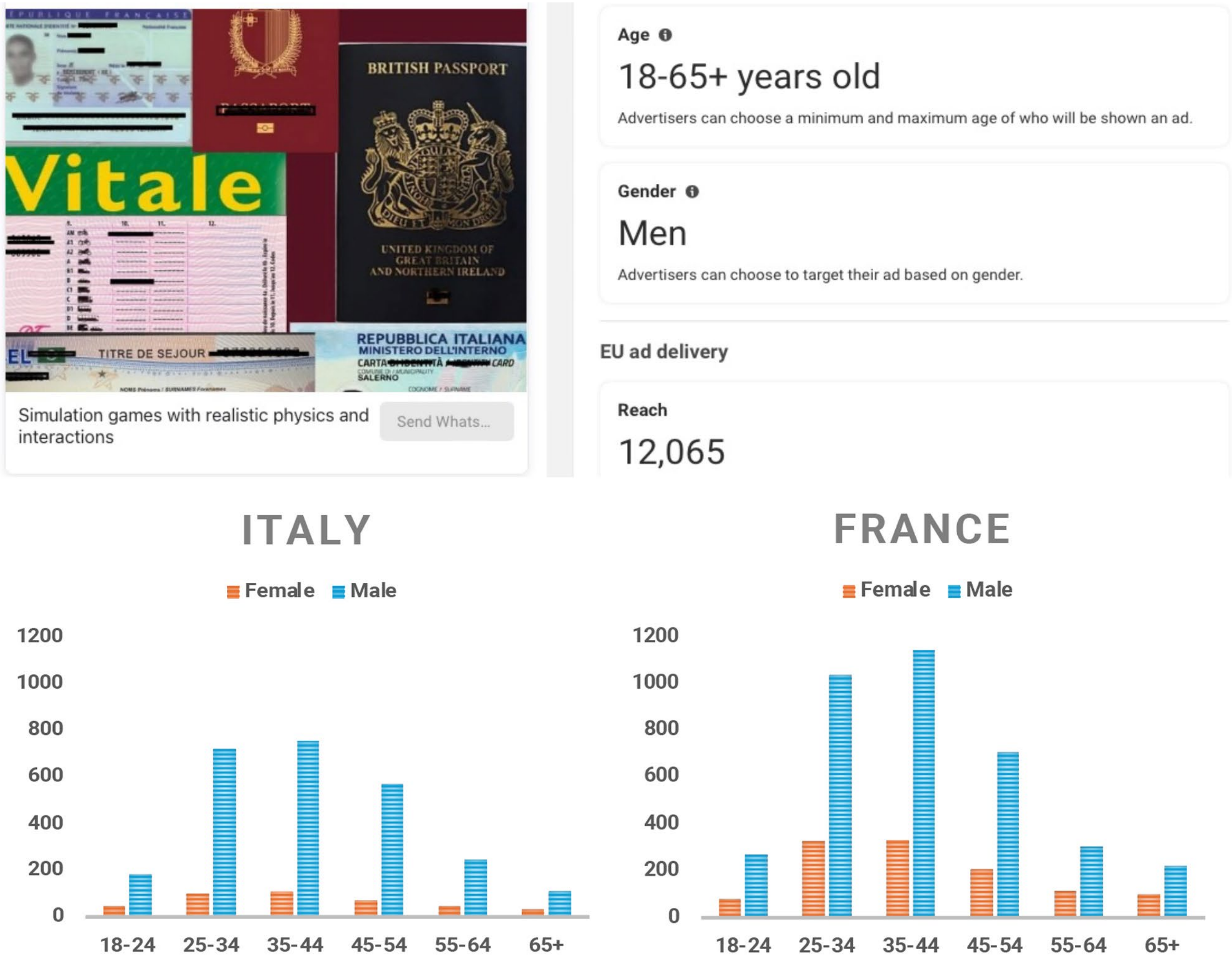
[156] NCA. (2024). National Strategic Assessment: Fraud. *National Crime Agency.*

[157] Lloyds. (2023). Lloyds Bank issues urgent warning over rising threat of crypto scams. *Lloyds Banking Group.*

[158] TSB. (2023). TSB calls for Meta to do more to protect customers from fraud – as consumers could lose £250m from fraud originating from Meta platforms in 2023. *TSB Bank.*

[159] Clark, J. & Hern, A. (2023). 'Mental anguish': how a crypto scam advertised on Facebook cost victim her life savings. *The Guardian.*

[160] Godfrey, T. (2025). 'CLICK & COLLECT' Illegal migrants are buying fake UK passports & driving licences for £5,000 before they travel to Britain. *The Sun.*



Simulation games with realistic physics and interactions

Age
18-65+ years old
Advertisers can choose a minimum and maximum age of who will be shown an ad.

Gender
Men
Advertisers can choose to target their ad based on gender.

EU ad delivery

Reach
12,065

Figure 12: Advertisement for a British passport and EU documents on Facebook in February 2025. The advert was targeted at to 18 – 65+ year-old men in Belgium, France, Germany, Italy, Malta, Netherlands, Poland, Portugal, and Spain. It was seen over 12,000 times, with 5,000 impressions in France. Men between 25 and 54 were more likely to see the ad. (Source: *Ad Library*)

This type of criminal document fraud has previously been identified on Meta's platforms.[161] When Which?, a not-for-profit that seeks to protect UK consumers, asked Meta why they allow such activity to proliferate on its platforms and why it does not proactively prevent it from appearing, Meta did not answer these questions. They did, however, respond by saying: "We do not allow fraudulent activity on our platforms, including the selling of forged documents."[162]

Given Meta had recently been warned about fake British passports and driving licences being sold on its platforms, **we conducted our own investigation into whether adverts for fraudulent documents were still running on Facebook, and if so, whether there was evidence that these criminal outfits were targeting specific groups.**

[161] Dillon, E. (2024). Criminal cyber gang targeting Revolut accounts in bogus driving licence scam. *Sunday World.*

[162] Lipson, F. (2023). Scamwatch: fraudsters lurking behind dodgy driving licence posts on Facebook. *Which?.*

To do this, we simply searched Meta's Ad Library for 'UK driving license' and 'driver's license'. We soon found numerous live adverts selling fake UK driving licenses, which were apparently "registered with DVLA" and could be obtained "without exams". We also found ads offering services such as "Licence Ban & points Remover." Whether these services are real or fake, they represent illegal activity.

We then searched for "British passports". Within minutes we found both live and completed ad campaigns that purported to sell British passports. One such ad, from a Facebook page disguised as a gaming page, advertised EU documents and British passports to men aged 18 and over in Belgium, France, Germany, Italy, Malta, Netherlands, Poland, Portugal, and Spain. This ad, which was live between February 11 and 13 February, reached over 12,000 people. Almost 5,000 of these impressions were in France, with males between 25 and 54 being more likely to see the ad. Due to the European Union's Digital Services Act, Facebook has been compelled to publish more open Ad Library information on demographic and location targeting information for ads appearing in the EU.[163] [164] However, Facebook still withholds data about more detailed demographics and any interest and behavioural targeting options chosen. What other specific targeting criteria the purchaser of the above advert used remains hidden.

Ads selling fake or forged identity documents – which one would normally associate with the dark web – are not hard to find on Facebook. This raises questions as to why Meta is allowing a black market in identity documents to continue to proliferate on its platforms. Different to the dark web, Facebook enables the purveyors of illegal goods to target their wares to audiences with specific demographics and geographies, with detailed interests and behaviours, who are more likely to be in need of buying them. By targeting refugees and asylum seekers, as some of these pages are clearly designed to do, these adverts are preying on already vulnerable populations, tempting them into courses of action which could cause further harm. Of course, there is no guarantee that the advertisers of these documents will ever deliver them, causing financial harm to those seeking to purchase them. If the fake documents are indeed delivered to the purchasers, this is clearly an illegal market both preying on and facilitating the movement of people outside of legal pathways and procedures of asylum-seeking. It raises the further question that if such adverts can be found by a simple manual search in minutes, why is Meta not allocating more of its substantial resources to employing people to monitor and prevent this illegal trade on its platforms?

[163] Meta. Beneficiary and payer requirements for ads targeting the European Union. *Business Help Centre.*

[164] Digital Services Act. *European Union*

CONCLUSION

Meta's business model relies on capturing users' attention while surveilling their behaviour, their interests, and their personal information to categorise people into 'types'. It uses this profiling to sell the attention of these 'types' to would-be advertisers.

Meta's advertising system is an opaque tool that has been repeatedly exploited by bad actors in attempts to suppress voters, spread propaganda, fuel division, and facilitate fraud. It has enabled targeted climate obfuscation, vaccine disinformation, financial scams, and the weaponisation of targeted messaging in ways that harm individuals, communities, and democracy.

Throughout the case studies in this report, Meta spokespeople have repeatedly claimed they do not allow these forms of content on its platforms. This is the expected legal and PR response. Its policy position may well be to not allow these types of adverts, but functionally, it consistently enables them. The long history of Meta's failure to uphold its own policies shows that self-regulation is an ineffective means to limit the societal harms being facilitated. There must be both domestic and transnational policy action to stem these harms and safeguard individuals, public safety, and democratic integrity.

This report lays out clear evidence of how Meta's profiling and targeted advertising systems can and have been weaponised for both individual and societal harm:

- **Democratic interference:** From the Trump campaign's Deterrence programme in 2016 to Russia's Doppelganger disinformation network seeking to undermine support for Ukraine, Meta's targeted advertising system is a tool used to undermine democracy around the world.

- **Vaccine and climate disinformation:** Meta's lax policies and implementation preceding and during the COVID-19 pandemic, as well as the business model itself, contributed to the infodemic. The system has enabled microtargeted messages of climate disinformation and greenwashing, obfuscating the urgent need for climate action.

- **Hate speech and division:** Advertisers have been able to target Northern Irish communities using proxies for protected characteristics, flaming sectarian divisions, while far-right networks across Europe and Asia have used Meta's ad platform to amplify hateful rhetoric and stoke violence.

- **Fear and trauma:** The UK Home Office ran fear-driven ads targeting refugees, warning them they would be arrested, betrayed, or drowned, while users with recent trauma are inescapably flooded with distressing adverts based on Facebook's profiling.

- **Financial fraud and identity scams:** Deepfake scams using Keir Starmer and Prince William's likeness pushed crypto fraud to hundreds of thousands, while fraudulent ID vendors used targeting to sell fake UK passports to men across Europe.

The case studies in this report evidence three major problems which require urgent action:

1 THE TRANSPARENCY PROBLEM

While Meta has shifted towards reducing the opacity of its advertising system by introducing the Ad Library, it is still insufficiently transparent regarding the profiled targeting categories advertisers choose. This opacity facilitates harmful advertising and prevents public scrutiny of disinformation, fraud, and manipulation. Public oversight should extend beyond aggregate audience, gender and location data – which are currently the only data that Meta chooses to make publicly available. Instead, advertisers' targeting categories must be open to public scrutiny, particularly when it is easy to target by proxy for protected characteristics or vulnerable populations on Meta's systems.

Recommendation: Meta must be required to publish full ad targeting details in its public Ad Library. This should include all demographic, interest-based, and behavioural categories used by each advertiser for each advert. Greater transparency would deter some forms of harmful targeting and enable greater public scrutiny of harmful ad targeting on Meta's platforms.

2 THE MODERATION PROBLEM

Meta must be proactive in preventing harmful ads, rather than relying heavily on user reporting once ads are circulating. The current model places a large burden on users to identify and flag harmful ads, allowing disinformation and harmful content to circulate before being removed. It is not always clear what is and is not disinformation, but clear frauds and scams, hate and division, circulating on the platforms could be easily spotted by human moderators with a contextual understanding of the geographies and communities in which Meta profits. Meta's overreliance on automated approval processes, and lack of institutional understanding of the geographic, political, and historical contexts in which it profits, results in harmful, misleading, and dangerous adverts being approved. Recent policy shifts demonstrate that Meta is moving further away from meaningful moderation, increasing the need for external regulatory and legal oversight.[165]

Recommendation: Meta must significantly expand both human and technological resources allocated to pre-publication ad moderation to tackle obvious disinformation, fraud and harmful ads upstream of publication rather than downstream of harm.

3 THE PROFILING PROBLEM

Meta's business model is built on profiling users by harvesting vast amounts of personal and behavioural data, yet it offers users no effective opt-out of surveillance and targeting for users to protect themselves from the harms evidenced in this report. Bad actors' ability to use Meta's ad targeting reinforces the need for users to have autonomy and agency over how they choose to be targeted. Meta's custom and lookalike audiences compound these problems, where bad actors can compile illegally obtained datasets and email lists and then use Meta's system to target citizens with similar characteristics. While Meta claims opt-out options exist, they also admit these are not fully effective. By not providing users with any meaningful right to object, Meta also gains an unfair advantage over other forms of media, such as TV and print, which rely more heavily on contextual models of advertising. There is, in the UK, very little public support for profiling and targeting. The UK communications regulator Ofcom has been tracking UK public attitudes towards so-called 'trade-offs' in data collection and use for some time and found in 2022 that just 17% of people were happy for companies to collect personal information to show them more relevant advertising.[166]

Recommendation: Users should be presented with a clear and explicit opt-in option for profiling and targeting, and be warned that this means they can be targeted by bad actors seeking to mislead or defraud them. The Court of Justice of the European Union has already found that, given the invasive nature of the data collection and Meta's inability to demonstrate it can protect citizens from harm, informed consent must be required above and beyond the acceptance of lengthy terms and conditions.[167] While this judgment does not constitute retained EU case-law in the UK, O'Carroll's settlement and the underpinning need to protect users from harmful advertising and other forms of invasive behavioural profiling suggest that the UK should follow the same lead. For users who do not opt-in to surveillance advertising, Meta should adopt contextual advertising within broad geographies – targeting ads based on the content users are presently engaging with rather than based on the surveillance and profiling of citizens.

Bad actors will continue to exploit Meta's advertising systems unless legal and regulatory efforts force change regarding transparency, moderation, and profiling. The case studies detailed in this report are about far more than technology or online life. They have implications for democracy, public safety, and the right to a reality that is not shaped by opaque advertising systems available for hire by bad actors.

[165] Kaplan, J. (2025). More Speech and Fewer Mistakes. *Meta.*

[166] Ofcom. (2022). Adults' Media Use and Attitudes Report.

[167] See, for instance, summary of CJEU Case C-252/21/ Judgement at: https://fra.europa.eu/en/caselaw-reference/cjeu-case-c-25221-judgement

ORG OPEN RIGHTS GROUP

openrightsgroup.org

Published and promoted by Open Rights Group, a non-profit company limited by Guarantee, registered in England and Wales no. 05581537. Registered address Space4 113-115 Fonthill Road, London, England, N4 3HH