ANTISEMITISM POLICY TRUST

Policy Briefing

August 2020

PUBLISHER LIABILITY AND ONLINE HARMS

Contents

Introduction

For over a decade, the Antisemitism Policy Trust has been working to counter antisemitism and related harms online. Over the past 25 years, legislation has been passed, in both the United Kingdom and abroad, which can be used to tackle online harms. However, with the advent and explosion of social media over the past decade or so, much of this legislation is outdated or ineffective.

The Antisemitism Policy Trust has submitted evidence to numerous Select Committees and to both the Government's Online Harms White Paper consultation in 2019, and the Petitions Committee Inquiry into Online Abuse in July 2020. However, many of these consultations failed to examine the fact that, at present, social media companies are not legally positioned as publishers, meaning they are exempt from many of the rules and regulations imposed on magazines, newspapers, television stations and other publishing outlets. This briefing paper will examine the legislation today in the United Kingdom, the Europe Union, America, Australia and in Germany, it will look at the European Code of Conduct for social media companies and will set out the rationale for whether or not to designate social media companies as publishers.

The American Precedent

The United States, often lauded as the bastion for Freedom of Speech and Expression, has legislation which limits access to indecent material online. The First Amendment does not provide individuals absolute free speech, as advocacy of illegal action, fighting words, commercial speech and obscenity are not protected.[1]

Section 230 of the Communications Decency Act, 1996: *"No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider"*

Following court cases with contradictory rulings, the Communications Decency Act (CDA)[2] was passed by the United States Congress as Title V of the Telecommunications Act of 1996.[3] Other than seeking to regulate minors' access to indecent material,[4] the Act has been described as conferring immunity upon the operators of Internet services, which are not to be deemed publishers of, and therefore not legally liable for, the words of third parties using their services.[5] [6] The CDA in effect improved service providers' ability to remove or supervise content according to their Terms of Service without fear of being considered a publisher.[7] This legislation can additionally act as a legal safeguard to bloggers and websites promoting controversial speech.[8] Bloggers would not be held accountable or liable for comments or content published by their users or by guest bloggers, because the blog owners themselves would not be considered as information content providers, nor publishers.[9] Immunity therefore holds, irrespective of editorial decisions or the knowing publication of offensive content.[10] Though some court cases have queried this general rule in respect of defamatory content.

The immunity for operators of internet services under the CDA Section 230, however, has no effect on laws such as criminal law, intellectual property law, state law and communications privacy law[11,12,13] and the operators therefore remain accountable to laws, including child pornography laws and obscenity laws.[14] Therefore, social media companies can be held liable if such content remains on their platforms without removal.

In recent years, Bills in congress have challenged CDA Section 230. The 'Allow States and Victims to Fight Online Sex Trafficking Act (FOSTA) and Stop Enabling Sex-Traffickers Act (SESTA) were introduced and FOSTA-SESTA was considered a major challenge to the act, exempting providers from immunity when knowingly supporting or facilitating Sex Trafficking.
Both before, during and since the FOSTA-SESTA Bills, issues of political neutrality and hate speech have been presented as reasons that platforms should have immunity revoked. President Donald Trump signed an Executive Order , currently subject to legal challenge, which would both remove platform immunity for lack of neutrality and remove the 'good faith' protection provided for in the Bill, from companies operating in a manner judged to be politically biased.

1 https://www.law.cornell.edu/wex/first_amendment

2 https://www.defamationremovallaw.com/legal-resource-center/what-is-section-230-of-the-communication-decency-act-cda/

3 https://uk.practicallaw.thomsonreuters.com/4-506-7494?transitionType=Default&contextData=(sc.Default)&firstPage=true&bhcp=1

4 https://www.law.cornell.edu/uscode/text/47/230

5 https://web.archive.org/web/20170313031105/https://www.eff.org/issues/cda230/infographic

6 https://web.archive.org/web/20170404022505/https://www.eff.org/issues/cda230

7 https://web.archive.org/web/20170521144245/https://www.eff.org/issues/bloggers/legal/liability/230

8 https://www.eff.org/issues/cda230

9 https://web.archive.org/web/20170521144245/https://www.eff.org/issues/bloggers/legal/liability/230

10 Ibid

11 https://www.law.cornell.edu/uscode/text/47/230

12 https://www.huffingtonpost.com/mary-anne-franks/section-230-the-lawless-internet_b_4455090.html

13 https://web.archive.org/web/20170521144245/https://www.eff.org/issues/bloggers/legal/liability/230

14 http://www.huffingtonpost.com/mary-anne-franks/section-230-the-lawless-internet_b_4455090.html

Unity in Europe: E-Commerce and Other Directives

Prior to the introduction of the E-Commerce Directive (ECD), the question of publisher liability was dealt with in many different ways across European Union member states. Even since the introduction of the Directive, the liability regime has been applied in different ways, including in diverging case law. That is to say, the matter is complicated, and Facebook's founder, to give but one example, has himself suggested the company no longer fits a simple technology, media or publisher definition.

The E-Commerce Directive introduced a "special liability regime" (Section 4, Articles 12-15) and a "safe haven", providing three types of service providers exemption from liability under specific conditions. The three types of service are:

1. Mere Conduits (Article 12): This is the internet access providers (or sub-providers), or the passive pipes (in the UK this would be Sky, BT etc). So long as they are passive, they retain immunity.

2. Cached providers (Article 13): This is copies of a site maintained locally (so for example, a copy of a website page on one's computer) to ensure no loss of service. Liability applies in circumstances where, broadly, no change is made to the information provided by the source website.

3. Hosting providers (Article 14): these providers store data from their users, posted by the users for unlimited time. This was envisaged to be webhosting services. Hosts benefit from a liability exemption, or shield, when unaware of illegal activity (for civil claims) and, "do not have actual knowledge of illegal activity or information" (for other claims). Importantly, providers must "expeditiously remove, or block access to, such information once they are aware of their unlawful nature."

Whilst there is consideration of services that fit all three categories, there is also a developing recognition that some services do not fit well within one of the three categories predefined by the ECD.

Recital 42 of the Directive sets out the liability exemption should apply to passive service providers, but each category outlined above is understood to have different levels of passivity, with the precise application built up by complex case law. It appears, for example, that hosting providers are allowed to select or modify the data stored and select recipients.

There are important caveats in place in the Directive, for example, a service can be required to take measures to terminate or prevent infringements even if not held liable for them. There is also a prohibition on Member States imposing a general obligation to monitor the data stored or transmitted, nor to seek facts that indicate illegal activity (but specific monitoring is however considered). However, if data is modified in transmission, or access to data not blocked once the host is made aware of unlawfulness, the extra liability protection may be dropped, and they are open to member state laws.

There have been numerous legal cases which have resulted in a multitude of interpretations of the Directive's language and scope. For example, in Article 12, queries arose about the definitions of mere conduits and what 'interference' means. In Article 13, emerging technology has tested what qualifies as a cached provider and for Article 14, courts have had to determine if providers 'had knowledge' of illegality. With the development of Web 2.0 many aspects of the ECD have been challenged. For example, the special liability regime applies to information society services, these are "normally provided for remuneration" "by electronic means". Each part of these definitions has been called into question.

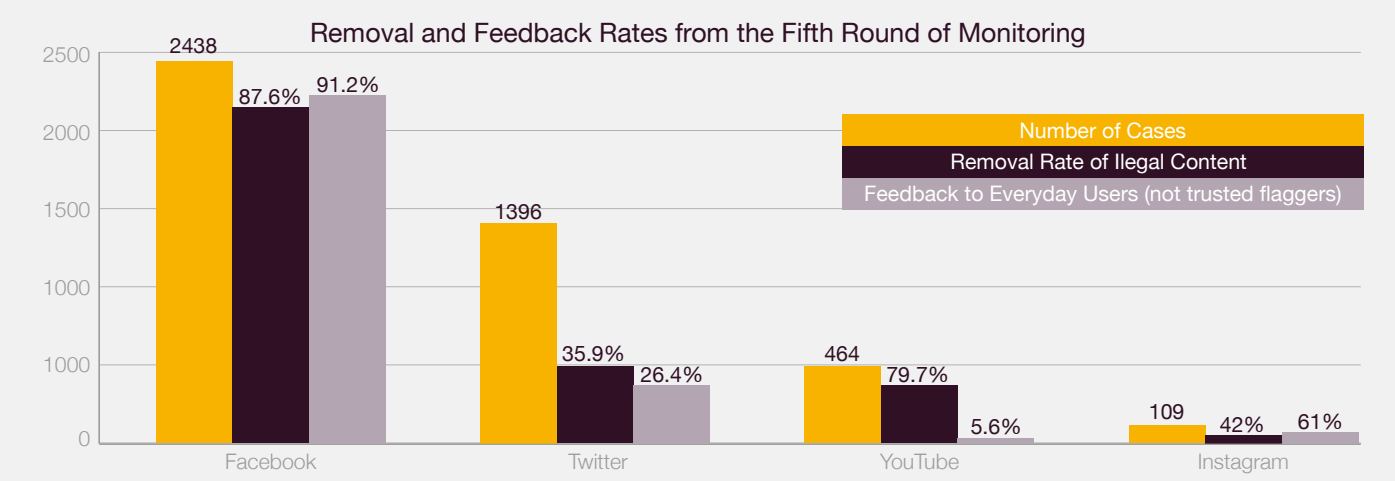
There are other relevant European Directives to consider, for example, the Audio-Visual Media Services Directive (AVMSD), due to be brought into law in Britain before Brexit, brings Video Sharing Platforms (VSPs) under scope for the rules in the audio-visual services single market. Specifically, VSPs will now be required to protect the general public from incitement to violence or hatred and from content constituting criminal offences (public provocation to commit terrorist offences, child sexual exploitation and abuse, and racism or xenophobia).

The European Commission Code of Conduct

In 2016, the European Commission, alongside the major social media companies (Facebook, Twitter, YouTube and Microsoft), announced, and committed to, a voluntary code of conduct in an effort to combat criminal hate speech on the internet. As part of this voluntary code, dozens of NGOs across Europe, including the Community Security Trust, Tell Mama and Galop in the United Kingdom, engaged in periodic monitoring to ensure the social media platforms adherence to the terms of the code, including on times of removal, whether users received notifications following their reports, the types of illegal hate being removed, and what sanctions were being handed down to users for posting illegal hate speech content.

The first results were released in December 2016[15] and on 1st June 2017,[16] the European Commission released the second evaluation. It has since released the third, fourth and fifth evaluations.[17] The second evaluation[18] found that a number of improvements had been made: social media companies were , on average, responding to more reports of illegal hate speech by removing content from their platforms; and companies were reviewing more complaints within 24-hours. Although the removal rate of content was higher when reports came from trusted organisations, companies were responding better to their users reports. There are still some improvements to be made, in particular the evaluation found that the quality of feedback and the decision-making process is noticeably different between social media companies.[19]

The most recent evaluation, released in June 2020, found that 90% of the notifications passed onto the social media platforms were reviewed within 24 hours, with 71% of content reported being removed.[20] 39 organisations participated in this round of monitoring, with Facebook, YouTube, Twitter, Instagram and Jeuxvideo being monitored. Although the rates of removals from the social media platforms has drastically increased since the first and second round of monitoring, the rates based on the form of hate being removed is vastly varied. For example, content calling for the murder of specific groups was removed in 83.5% of cases. However, content using defamatory words, which can also be deemed illegal in various locales, was only removed in 57.8% of cases. In the United Kingdom however, the removal rate across the board is only 42.5%, meaning self-regulation is not working.[21]



15 https://ec.europa.eu/newsroom/just/item-detail.cfm?&item_id=50840

16 http://ec.europa.eu/newsroom/just/item-detail.cfm?item_id=71674

17 http://ec.europa.eu/newsroom/just/item-detail.cfm?&item_id=50840

18 http://ec.europa.eu/newsroom/just/item-detail.cfm?item_id=71674

19 Ibid

20 https://ec.europa.eu/info/sites/info/files/codeofconduct_2020_factsheet_12.pdf

21 Ibid

Made in Britain

In the United Kingdom, laws exist to protect the victims of online crime, including certain acts of online abuse. Section 1 of the Malicious Communications Act 1988, Sections 4A and 5 of the Public Order Act 1986, Section 127 of the Communications Act 2003, and Sections 2 and 4 of the Protection from Harassment Act 1997, amongst others, all contain relevant clauses for taking action against crimes online and generally predate the widespread existence of social media.

None of the offences covered by the aforementioned Acts includes a specific defence, or exemption from liability, for an internet company that hosts material covered by one of these offences. A company might theoretically therefore find itself liable to criminal prosecution for encouraging or assisting one of these offences. In 2014, the House of Lords Communications Committee published a report into social media and criminal law. In reference to corporate liability for undesirable content, the Committee referenced Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 (the aforementioned E-Commerce Directive), harmonised into UK law by the Electronic Commerce (EC Directive) Regulations 2002 (SI 2002/2013). The Lords Committee explained that:

"Those regulations give immunity to websites from damages or criminal sanctions where they act merely as a conduit, cache or host, so long as they operate an expeditious "take down on notice" service. This acts as an incentive to website operators to remove illegal or actionable material. It is for the website itself to determine whether the material which they have been asked to remove is genuinely illegal or actionable."

The Lords Committee viewed that "Parliament has thus accepted the view that the liability of website operators should be limited in respect of content they host but which they have not originated." The Lords continued: "Website operators are *not necessarily* [emphasis added] accessories in liability to crimes. The law could be changed to clarify this." The Committee suggested an alternative approach might be the establishment of an ombudsman funded by website operators, to set policy and consider complaints. Subsequently, and most recently, the Home Affairs Select Committee recommended sanctions for companies failing to remove illegal content on request.

Further immunity from prosecution was conferred on social media providers through the Defamation Act 2013, which reformed defamation law in relation to the right to freedom of expression. Section 5 of the Act includes defences for 'Operators of websites'. A website operator has a defence to charges by showing it was not they who directly 'posted' a statement on a website. The defence can be defeated if three conditions are met, including the operator failing to respond to a notice in accordance with any provisions contained in regulations.[22] Therefore, if a platform is alerted to illegal content on their platform, whether by an organisation, an individual or a trusted flagger, they have a legal requirement to act on that report and remove the illegal content. Where a successful defamation action has been taken, the courts can now order the platform to remove the material. How viable this is in practice, is questionable.

Section 103 of the Digital Economy Act 2017 meanwhile, includes provision for a non-statutory, or voluntary Code of Practice for providers of online social media platforms. This Code is broader than, for example, the Australian licencing system outlined below, and includes guidance on conduct which involves "bullying or insulting the individual, or behaviour likely to intimidate or humiliate the individual." After various consultation, in April 2019 Government published the Code, which can best be described as 'High Level'[23].

22 http://www.legislation.gov.uk/ukpga/2013/26/section/5/enacted

23 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/793324/Code_of_Practice_for_providers_of_online_social_media_platforms.d.pdf

On The Western Front

In Germany, the Network Enforcement Act (known widely as NetzDG) has been in force for two years, following criticism that the government was slow to act in tackling hate online and in response to the lack of self-regulation of social media platforms.[24] Germany has been leading the charge for action against social media platforms for their inaction on cyber hate. NetzDG was approved by the country's cabinet in draft,[25] and later passed by parliament. The act requires social media networks to assign a complaints representative, responsible for taking down or blocking content that is evidently criminal, 24 hours after receiving the initial report.[26] Where content is not immediately recognised as illegal, companies have seven days to act. Social media companies can receive fines of up to 50 million Euros if the deadlines set in the bill are not upheld.[27]

Social networks are also required to follow up with, and explain the outcome of, an appeal to complainants, as well as provide quarterly reports detailing both the number of complaints and the networks decision-making procedures.[28] Amendments to the Bill widened its scope and new categories of criminal content, including child pornography, were added, together with clauses enabling courts to order social networks to disclose the identity of any user who posted unlawful material online.[29] The Bill did not create new criminal offences but gave social media platforms the responsibility for removing unlawful content.[30]

Enforcement of the Bill is still queried[31] and NetzDG, as of early 2020, is being amended to respond to wide criticism and problems it has encountered since its inception two years prior. The legislation has not been shown to be effective in forcing platforms to remove more content that is deemed to be illegal and harmful.[32] The new amendments would oblige the service provider to give users an "easily recognisable, directly accessible, easy-to-use and permanently available procedure when perceiving the content for transmitting complaints about illegal content."[33] This would give ordinary users the ability to report illegal content through a separate system, often only reserved for law enforcement. The law would also critically widen illegal hate speech offences as it would criminalise "conduct far in advance of aggressive opinions and calls for violence."[34]

24 https://policyreview.info/articles/news/germany-amending-its-online-speech-act-netzdg-not-only/1464

25 http://www.bmjv.de/SharedDocs/Downloads/DE/Artikel/04052017_Faktenpapier_GesE_NetzDG.pdf?__blob=publicationFile&v=2

26 http://www.independent.co.uk/news/world/europe/germany-fake-news-social-networks-fine-facebook-50-million-euros-illegal-content-hate-speech-angela-a7668731.html

27 Ibid

28 https://www.ft.com/content/c10aa4f8-08a5-11e7-97d1-5e720a26771b?mhq5j=e5

29 http://www.independent.co.uk/news/world/europe/germany-fake-news-social-networks-fine-facebook-50-million-euros-illegal-content-hate-speech-angela-a7668731.html

30 https://policyreview.info/articles/news/germany-amending-its-online-speech-act-netzdg-not-only/1464

31 http://www.independent.co.uk/news/world/europe/germany-fake-news-social-networks-fine-facebook-50-million-euros-illegal-content-hate-speech-angela-a7668731.html

32 https://policyreview.info/articles/news/germany-amending-its-online-speech-act-netzdg-not-only/1464

33 Ibid

34 Ibid

In The Outback

Australia have adopted a similar model to those outlined by the UK House of Lords Communications Committee. The Australian Parliament legislated for a 'Children's e-safety Commissioner'. The office of the Commissioner administers a complaints system for cyber-bullying material targeted at Australian children and, amongst other roles, promotes online safety for children, publishes reports and encourages and conducts research.

The Australians also created a two-tiered scheme for the rapid removal of cyber-bullying and other material. The content in question, found on social media or a relevant electronic service, must be considered by "an ordinary or reasonable person" as likely to have an effect on a particular child, or in general be seriously threatening, intimidating, harassing or humiliating. The impact can be direct or indirect.

Any social media service may volunteer to participate in Tier 1, including small social media services. The Commissioner can recommend large social media services be declared Tier 2. These second-tier services are subject to legally binding notices and penalties. In addition, there is a system of end-user notices, requiring users that post offending material to remove, refrain from posting or apologise for material that has been posted.

Complaints can be made to the Commissioner by, or on behalf of, a child, and an investigation follows on agreed cases.

Fundamentally, the expectation exists in the Australian system that social media and internet companies responsibly apply their own Terms of Service, which should prevent cyber bullying.

The Rationale For And Issues With Full Liability

Social media companies commission, edit and curate content for broadcast or publishing and as such, are benefiting from an absence of liability. The companies at present pick and choose their status depending on how it suits them to be defined at a given time. The legal case between app Six4Three and Facebook saw the latter company argue in court that it was protected as a publisher under the first amendment for making editorial decisions not to publish content, whilst claiming protection under the Communications Decency Act because it was not a publisher, something it repeatedly claims in public.[35] The platforms are not simply hosts, nor neutral or mere conduits as they do apply community guidelines, albeit inconsistently.

However, the current definitions of publishers and language used in legal framing does not match the development of the technology we now use. Platforms like Facebook and Twitter have so much content streaming through their service that to introduce liability for individual posts might not be possible. Platforms systems and moderation cannot keep pace with the rate of upload, and despite their claims often only act after the event. In addition, the platforms have been proven to have liability in certain circumstances, where they have failed to act expeditiously, and with prior knowledge and so full immunity is not a given. If one assigns the companies as publishers, it might make it an impossibility for them to operate their model in the UK. This is aside form jurisdictional concerns, whereby cases might need to be filed, or action taken, in Ireland (the EU/UK base) or America, which operate under a different ruleset. To this end, maintaining the status quo whereby courts have deemed liability under certain circumstances may be a preferable option to redefinition in narrow regional parameters. This would also go some way to preventing social media preference for their companies to be entirely envisaged as 'good faith operators/good Samaritans' in acknowledging that sometimes they do act in bad faith.

In fact, opening up the debate on liability, for example to undertake a fundamental rewrite of what constitutes an online publisher or media company, could be beneficial for the social media companies which might like a 'good Samaritan' exception, similar to the United States, which would make it harder to prove wrongdoing than is perhaps the case in Europe.

There is an obvious lack of consistency in regulation. Horror was expressed by parents that Peppa Pig cartoons were spliced with gruesome images on YouTube. Had the same material been broadcast on television, Ofcom would have acted, and outsourcing responsibility to children and parents to use television more wisely would not have been the result. OFCOM, the potential online harms regulator as proposed by the Government's Online Harms White Paper, can fine telecoms companies for failure to meet licence conditions or for failure to respond properly to information requests, or fine broadcasters for failing to meet agreed standards but even parliamentary committees struggle to solicit simple answers and data from social media companies. People have free speech when they go on television, radio, write in the press but they are bound by regulatory limits and in the case of newspapers, they are held responsible for publishing libellous material. Social media companies are not.

The Government previously discussed and undertook a review of liability[3637], ultimately considering the matter too complex to resolve. In response to the Committee on Standards in Public Life it said "The Government has been clear that social media platforms are no longer just passive hosts, and we need a new approach. We need to think carefully about what level of legal liability social media companies should have for content on their sites, and we need to fully understand the consequences of any changes."[38]

35 https://www.theguardian.com/technology/2018/jul/02/facebook-mark-zuckerberg-platform-publisher-lawsuit

36 https://www.gov.uk/government/speeches/pms-speech-at-davos-2018-25-january

37 https://www.gov.uk/government/speeches/pm-speech-on-standards-in-public-life-6-february-2018

38 https://www.gov.uk/government/publications/government-response-to-the-committee-on-standards-in-public-life-review-of-intimidation-in-public-life

Following the former Prime Minister's commitment to a review, the Internet Safety Strategy Green Paper outlined that the Government was looking at legal liability for illegal content shared on social media sites: "The status quo is increasingly unsustainable as it becomes clear many platforms are no longer just passive hosts. Whilst the case for change is clear, we also recognise that applying publisher standards of liability to all online platforms could risk real damage to the digital economy, which would be to the detriment of the public who benefit from them. That is why we are working with our European and international partners, as well as the businesses themselves, to understand how we can make the existing frameworks and definitions work better, and what a liability regime of the future should look like. This will play an important role in helping to protect users from illegal content online and will supplement our Strategy".[39]

By the time the Online Harms White Paper was published, the Government had changed its tone:

"The new regulatory framework will increase the responsibility of online services in a way that is compatible with the EU's e-Commerce Directive, which limits their liability for illegal content until they have knowledge of its existence, and have failed to remove it from their services in good time."

"Our review found that, while it is important to ensure that companies have the right level of liability for illegal content, this is not the most effective mechanism for driving behavioural change by companies. The existing liability regime only forces companies to take action against illegal content once they have been notified of its existence. It therefore does not provide a mechanism to ensure proactive action to identify and remove content. In addition, even if reforms to the liability regime successfully addressed the problem of illegal content, they would not address the full range of harmful activity or harmful behaviour in scope. More fundamentally, the focus on liability for the presence of illegal content does not incentivise the systemic improvements in governance and risk management processes that we think are necessary. We concluded that standalone changes to the liability regime would be insufficient. Instead, the new regulatory framework takes a more thorough approach. It will increase the responsibility that services have in relation to online harms, in line with the existing law that enables platforms to operate. In particular, companies will be required to ensure that they have effective and proportionate processes and governance in place to reduce the risk of illegal and harmful activity on their platforms, as well as to take appropriate and proportionate action when issues arise. The new regulatory regime will also ensure effective oversight of the take-down of illegal content and will introduce specific monitoring requirements for tightly defined categories of illegal content.[40]".

The case for accountability to a regulator, rather than liability as a publisher is therefore strong, at least from the Government's perspective.

39 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/708873/Government_Response_to_the_Internet_Safety_Strategy_Green_Paper_-_Final.pdf

40 https://www.gov.uk/government/consultations/online-harms-white-paper/online-harms-white-paper

Legal But Harmful Materials

It is important to separate out liability for illegal content from legal but harmful materials. The role of antisemitic conspiracy theories and Holocaust denial, which often do not cross the threshold to illegal online harms, are important concerns. Facebook judges nudity as unacceptable, but Holocaust Denial is deemed unproblematic. Holocaust denial is grossly offensive, can be a conduit for incitement to racial hatred against Jews and a steppingstone to radicalisation, particularly on the far right.[41] Legal but harmful material can drive terrorism.[42] Atrocities in Pittsburgh, San Diego, Christchurch and multiple others, and the deification of the murderers, have proven this to be true. Cases like the 'Pizzagate' affair, which saw a gunman in America attack a family pizzeria after imbibing legal but harmful material online, such as false conspiracy theories about a paedophile ring being run in the pizzeria, underline why companies must have plans in place to address these issues. Freedom of speech is, of course, critical but there are potential impingements on people's right to life under the current legislative arrangement. The broadcast of hate should be regulated and a company's failure to act should be viewed as a failure of responsibility as a publisher.

Legal but harmful speech is already regulated in other areas. The BBFC, Britain's film certifier, uses "discrimination" as a category that it considers when classifying potentially harmful content. This can result in a higher age classification where the viewers are judged too young to be able to critically understand the racist or discriminatory commentary. The BBFC also refuses to classify content which is likely to cause "harm risks to potential viewers and, through their behaviour, to society". For instance, the BBFC refused to classify the online film 'Hate Crime' in 2015 because it consisted of nothing but an extended scene in which a Jewish family is subjected to racist abuse, violence and sexual violence in their own home. The BBFC concluded there was a risk that some viewers may be entertained by the abuse, and associate with the attackers. This is a sensible limit on the freedom of speech and expression the filmmaker sought to exploit

Similarly, the Audio-visual Media Services Directive sets out requirements including for certain legal but harmful content for linear broadcasters, on demand services, and in the latest revisions, video sharing platforms.

Facebook, Twitter and others have previously received funds through advertising which have been used to promote information leading to real world harms, such as harmful content on Coronavirus or theories about 5G. Thankfully, the immediate offline impact of such content was clear to politicians and subsequently social media companies, as 5G masts were being burnt and people were being peddled dangerous Coronavirus cures, so the companies acted to stop this. The same harmful impact is not immediate in hateful legal speech, so the social media companies have been less compelled to act. When they have done so it has often been through pressure from advertisers, seeking to manage reputational risks. In recent years, in addition to the 5G and coronavirus content, some fake news has been challenged. Facebook took action after Stoneman Douglas school shooting in Florida in 2018 to remove what it branded "abhorrent" posts claiming shooting survivors were actors or part of a conspiracy. This again, demonstrated publisher oversight editorship and curation.

The platforms approach to misinformation, or fake news in general, is inconsistent. Sometimes it is demoted or taken down, sometimes accompanied by 'fact checked' counter information, but hate speech is subject to less intervention. Social media companies are not freedom of speech havens. Alternative platforms that have seemingly fewer rules and appear to be totally unwilling to moderate content on their platforms, like Gab,[43] Bitchute,[44] 4Chan and 8Chan,[45] are proven extremist and terrorist-inciting online sewers, sometimes only acting on notice from police to remove materials from proscribed organisations but in a piecemeal way and with much illegal content remaining on the platform.

41 https://www.splcenter.org/fighting-hate/extremist-files/ideology/holocaust-denial

42 https://www.counterterrorism.police.uk/neil-basu-welcomes-online-safety-measures/

43 https://www.adl.org/blog/when-twitter-bans-extremists-gab-puts-out-the-welcome-mat

44 https://www.hopenothate.org.uk/wp-content/uploads/2020/07/BitChute-Report_2020-07-v2.pdf

45 https://www.theguardian.com/technology/2019/aug/04/mass-shootings-el-paso-texas-dayton-ohio-8chan-far-right-website

How Would Change Occur and Who Would Support It?

Codifying the existing law emanating from Europe will be difficult. Platforms will already contest whether or not they 'have knowledge' of wrongdoing and will argue that it could be against natural justice to hold them accountable for something they had knowledge of and later addressed. However, as above, there is a gap at present whereby platforms shaping information are not widely responsible for harms occurring on their platforms.

Social media companies must be accountable to a regulator for failing to apply minimum standards, acting irresponsibly, or breaching a future statutory Duty of Care in relation to takedown according to terms of use. The regulator would potentially licence the firms or otherwise have them notify to a regulator they had users in a particular jurisdiction, have the ability to issue strong fines and individual senior management liability would be introduced. The latter point is perhaps the most important if wider liability is unachievable. The Terms would need to be clarified but this could bring good practice on an industry-wide basis, whereby notice for clearly stated harms and easily established violations, not actioned by algorithms or human reviewers, is actioned within a time period and failure to act leads to penalties. Required policies for repeat offenders, on anonymity, decency and discrimination could be required as part of this regulatory regime.

Codes of Practice on Terrorism and Child Abuse have already been announced, but a Code of Practice on (how to prevent) harm is, from the Antisemitism Policy Trust's perspective, a necessity. This should not be left to a later stage with all the uncertainty that brings.

Addressing legal harms is critical and a Code of Practice in this area would be an important anchor for shared learning and operational standards.

In respect of the law, after Brexit, there is the potential to amend the regulations that allowed for the E-Commerce Directive to be implemented. The question of liability is extremely complex and there could be unforeseen consequences in pushing for full liability. However, it might be possible to address the existing liability exemption by clarifying, to a degree, when platforms are considered to have 'actual knowledge'. The Government should be consulting legal experts on this issue as part of the drafting process

As Baroness Kidron has said, by introducing liability and having a better regulatory approach, we can revolutionise the way our online world works and better safeguard of future and prepare our children for the digital experiment we are all a part of.[46]

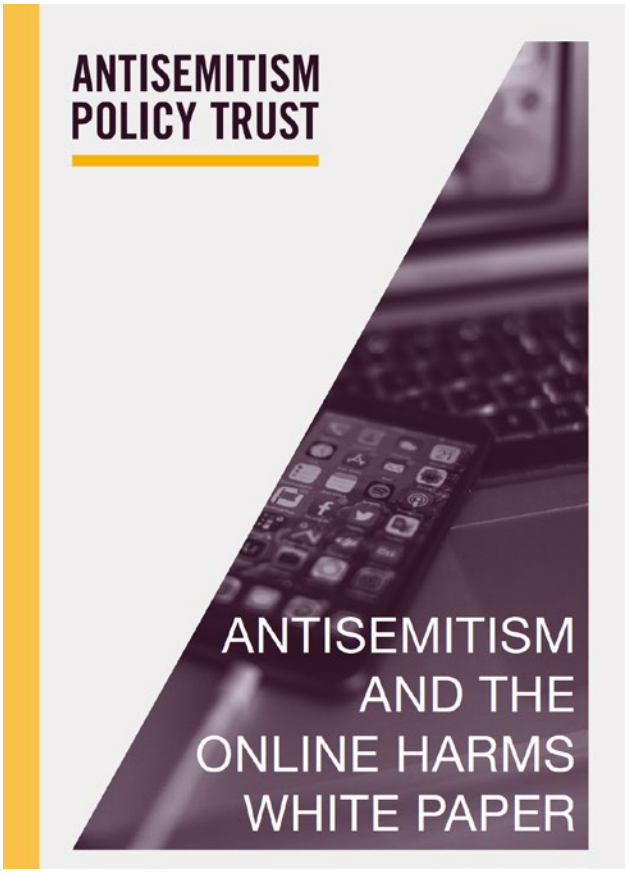
A number of parliamentary bodies, organisations and individuals support change, including: The Home Affairs Committee, The Committee on Standards in Public Life[47] and both Houses of Parliament (through Baroness Kidron and former MP, now Lord John Mann. There is wide support for change in civil society, in the media and in the third sector, including and not least, the Antisemitism Policy Trust.

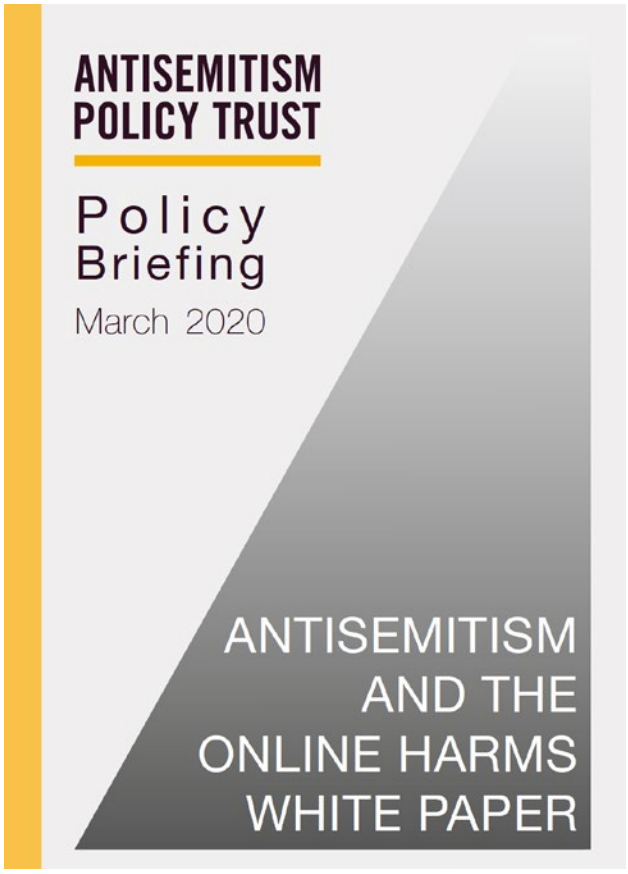
46 https://www.telegraph.co.uk/news/2020/05/21/hold-up-online-duty-care-laws-will-mean-government-capitulating/

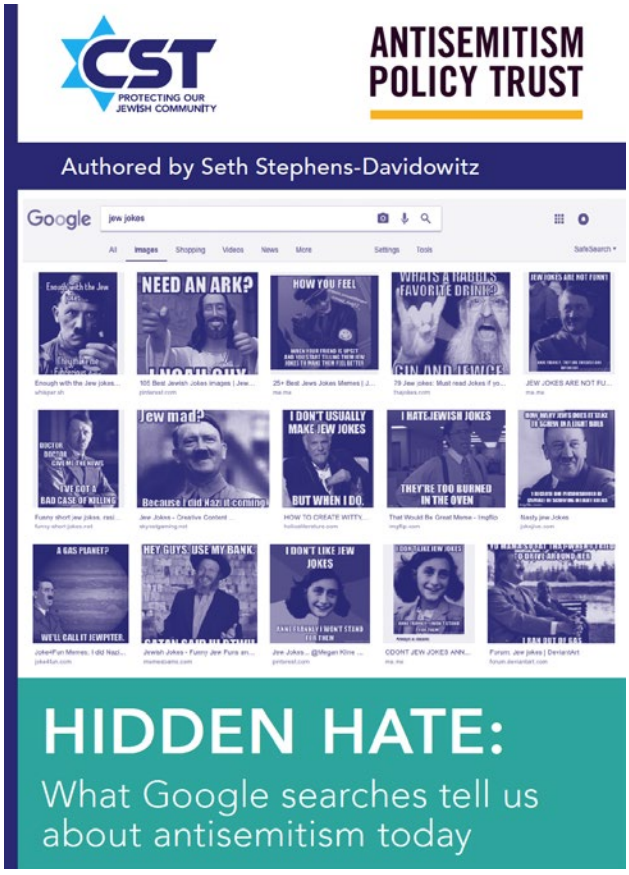
47 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/666927/6.3637_CO_v6_061217_Web3.1__2_.pdf

Antisemitism and Online Harms

Related Publications from APT







The Antisemitism Policy Trust's mission is to educate and empower parliamentarians, policy makers and opinion formers to address antisemitism. It provides the secretariat to the British All-Party Parliamentary Group Against Antisemitism and works internationally with parliamentarians and others to address antisemitism. The Antisemitism Policy Trust is focussed on educating and empowering decision makers in the UK and across the world to effectively address antisemitism.

Contact APT


www.antisemitism.org.uk


@antisempolicy


Antisemitism Policy Trust


mail@antisemitism.org.uk

The Antisemitism Policy Trust is a registered charity (1089736) [England] and company (04146486) [England and Wales]