


ANTISEMITISM POLICY TRUST

Special Briefing



THE ONLINE
SAFETY BILL:
HOUSE OF LORDS
STAGES

The text and illustrations may only be reproduced with prior permission of the Antisemitism Policy Trust.

Published by the Antisemitism Policy Trust, copyright © 2022

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

Contents

Overview	4
Duties of Care	4
Duties of Care: User-To-User Services	4
Category 1 Companies: Small but high harm, high risk platforms	6
Category 1 Companies: Countervailing Duties	7
Newspaper Comment Boards	7
Duties of Care: Search Services	8
User Identity Verification and Anonymity	8
Ofcoms Powers and Duties	9
Penalties	9
Consulting experts and relevant groups	9
Super Complaints	10
Entirely Absent: Supply Chains	10
Restrictive Covenant	10
Media Literacy	10

Overview

The Online Safety Bill (OSB) introduces a British regime for managing online harms. It proposes a form of systems-based regulation, focussed on frameworks, business models, structures, risk management and user-empowerment. The Bill provides the overall framework for a raft of measures, a number of which will be set out in secondary legislation, and codes which will be introduced by Ofcom, the new online safety regulator. The Bill sets out how the Secretary of State and Ofcom will need to act in order to administer these framework duties. The services considered in scope for this regulation are those that have a significant number of users in the UK, if the UK is a target market, or if the service can be used in the UK by individuals and there are reasonable grounds to believe that there is a material risk of significant harm to individuals in the UK. This will of course

have implications for Ofcom's technical capabilities and capacity and there will likely be some shortfall requiring expert third sector bodies to monitor the UK's digital sphere. The proposed law treats user-to-user services (like Facebook) and Search engines (like Google) differently, applying stronger duties to some user-to-user services (placed into what is called a Category 1 group), for example, providing certain user-empowerment duties.

There have been numerous changes to the Bill both from its draft form and in the Commons, most notably the removal of risk assessments for designated so-called 'legal but harmful' content. Some of these changes have consequences which will be addressed in this briefing.

Duties of Care

The Bill introduces numerous 'duties of care' which differ according to the nature of the service being considered, and do not apply across the whole digital sphere. These include duties to tackle illegal content, to prevent fraudulent advertising, ensure user verification, report Child Sexual Exploitation and Abuse, report transparently, protect freedom of expression and more. 'Online Safety Objectives' will guide Ofcom in producing Codes through which companies will deliver on the duties (unless they have comparative measures in place of their own). The Safety Objectives include principles relating to safety by design.

Duties of Care: User-To-User Services

Part 3 of the Bill considers the duties on User-To-User Services, in Chapter 2. When it comes to addressing illegal content, all platforms will be required to develop 'suitable and sufficient' risk assessments which must be renewed before design changes are applied and are linked to safety duties (what the platform must do about the risk). Consideration will need to be given to who is using a platform, and how it works – a systems-based analysis and response. Fundamentally, platforms will need to ensure that illegal content is

not online and where it is, to get it offline swiftly. The Bill lists certain forms of illegal content the tackling of which platforms must prioritise, but all offences (in addition to those specified as priority) 'of which the victim or intended victim is an individual (or individuals)' are theoretically in scope.

The next section of the Bill replaces what was a similar approach for 'legal but harmful' content and in its place requires Category 1 platforms to have a 'user empowerment duty'. This has been presented as the third part of the so-called 'triple shield' or as the 'toggle' for users to opt-out of seeing specified content which includes, in relation to antisemitism, abuse targeting a person's race or religion. The Trust's view of this requirement is layered. In order to identify and to have an option to hide the content, platforms will need to know where it is and the volume of it. That is information which Ofcom might theoretically wish to probe. However, this is not as strong a protection as having the adult risk duties that were previously in the Bill (and which did not require content take-down), nor is it likely to be fit for purpose unless the toggle is automatically set to 'on'.

We believe the toggle should be on by default. The SNP proposed an amendment to this effect in the Commons, seeking to amend what is now Clause 12 (4), replacing ‘made available to’ with ‘in operation by default for’. The Government did not provide its reasoning as to why this should not be possible. The Trust is of the view that those seeking pro-suicide, eating disorder and legal antisemitic content should have to request to see it, not be served it by default. We would welcome amendments to this effect, and would be pleased to assist with drafting.

Furthermore, whilst risk assessments for adult harms were removed from the Bill, it is illogical that risk assessments around the user-empowerment processes are not specified in the Bill – which itself is supposed to be predicated on risk. We would support amendments seeking to introduce risk-assessments tied to the user-empowerment duties, in order that they be well directed and effective.

There are other duties on Category 1 companies, including that they provide for user redress and must protect journalistic content and that of democratic importance. There are higher protections for children. One might argue content harmful to adults may be harmful for children too, but we suspect some platforms may seek to game the system to avoid such a classification. It is here that the Bill requires significant structural and definitional change.

Category 1 Companies: Small but high harm, high risk platforms

Though Category 1 is referenced early in the Bill, the determinants for categorisation are cited elsewhere in the Bill (Schedule 11). In summary, categories are determined by the number of users of a user-to-user platform (or sub-platform), the ‘functionalities’ of that part of a service (its mechanisms) and now, following the commons stages “any other characteristics of that part of the service or factors relating to that part of the service that the Secretary of State considers relevant”.

The Government’s response to the joint committee which scrutinised the draft Bill makes it clear that it considers reach is a key and proportional consideration when assigning categories and believes Secretary of State powers to amend those categories are sufficient to protect people. Unfortunately, this could leave many alternative platforms out of Category 1, even if they host large volumes of harmful material – or in some cases, are specifically designed with the intention of spreading harm. The duty of care approach is predicated on risk assessment. If size allows platforms to dodge the entry criteria for managing high risk there is a hole in the regime. Platforms including Bitchute, Odyssey, Minds, Gab and 4Chan, host extreme racist, misogynist, homophobic and other extremist content that radicalises and incites harm. The Community Security Trust has outlined in detail¹ some of the most shocking and violent materials on these sites and whilst illegal material has been present, much of that content is legal but harmful (and would be addressed in other environments, such as a football ground, cinema or on TV/radio). That lawful material can and has transferred to more mainstream platforms and has influenced real world events. The Antisemitism Policy Trust briefing on the connection between online and offline harms details how antisemitic terrorism, like the deadly attack in a synagogue in Pittsburgh, and deadly Islamophobic attacks, like the

Christchurch Mosque attacks, were carried out by men who were, at least in part, radicalised online and whom signalled their intent to attack online, and in some cases sought to livestream their attacks online.² Meanwhile, academic research underlines the threat from small “dark platforms” like 8kun³, including in relation to Covid conspiracy theories.

The existing user-empowerment duties for Category 1 services provide the most minimal amount of friction in the systems of user-to-user services. At the very least, some small, high-harm platforms should be required to put these in place. We believe that it is crucial that risk be a factor in the classification process determining which companies are placed in Category 1, otherwise the Bill itself risks failing to protect adults from substantial amounts of material that causes physical and psychological harm. The relevant schedule (11) needs amending, to reflect this. The proposed section would include a fallback option for Ofcom to add platforms presenting a (proven) significant risk of harm (irrespective of size) into the first Category. This would be in line with Ofcom’s powers in other broadcasting legislation, and means that not all companies are automatically dragged into the arrangement thus protecting start-ups and others. At the very least, the Government should confirm that the characteristics a Secretary of State might consider in setting Category thresholds would include those deemed high harm by Ofcom despite not meeting the minimum size or functionality requirement.

1 <https://cst.org.uk/news/blog/2020/06/11/hate-fuel-the-hidden-online-world-fuelling-far-right-terror>

2 <https://antisemitism.org.uk/wp-content/uploads/2020/08/Online-Harms-Offline-Harms-August-2020-V4.pdf>

3 <https://www.tandfonline.com/doi/full/10.1080/21670811.2021.1938165>

Category 1 Companies: Countervailing Duties

Section 13 of Chapter 2 includes a duty to protect content of democratic importance which is defined in (6) (b) as “content that is or appears to be specifically intended to contribute to democratic political debate in the United Kingdom or a part or area of the United Kingdom”. In Sections 14/15 there is a duty to protect news published and journalistic content. We agree that there is an imperative to protect democratically important and journalistic content. However, the way in which these duties are set out in the draft Bill, mean that extremists, who actively undermine the democratic process by disseminating hateful and racist material, disinformation and other harmful content, will be protected under the law. For example, a racist activist standing for election might be able to demand their harmful material be re-platformed once removed, claiming bias or discrimination against the platform.

What would happen in the case of someone spreading misogyny in an electoral race against a candidate from the Women’s Equality Party? We know, as the UN’s Special Rapporteurs on free speech have made clear, that political speech can and does cause harm.⁴ In other words, if democratically important speech causes harm, what guidance will be offered beyond leaving the matter to the platforms to decide? Furthermore, without a duty to promote such content, how will platforms ensure spaces are not closed to those often left out of democratic debate? **Whilst perhaps well intentioned, the current drafting of this duty is not workable in practice and should be reconsidered.**

The protective duty in relation to journalistic content presents similar and additional concerns. Journalistic content is poorly defined and might be read, as the antiracism NGO Hope Not Hate has suggested, as content “generated for the purposes of journalism”.⁵ To this end, citizen journalists’ content achieves the same protections (through Ofcom) and enhanced routes to appeal (through platforms) as content from any major national publication. There are examples of far-right activists self-identifying as journalists, and news companies like InfoWars, Rebel Media or Urban Scoop which spread hateful and dangerous conspiracy theories. There is little assurance on the face of the Bill that content produced by such individuals and companies might not be offered special protections, or otherwise benefit from these duties as currently drafted. We could end up in an awful situation where far-right citizen journalists are legitimised by winning a complaint and thereby ‘proving’ they are indeed journalists. We are also not convinced by the explanations offered to the Trust and others by officials, that platforms will have to perform a balancing act in respect of harms and content (especially given some platforms are designed for harm).⁶ **Sadly, the Government’s response to questions about these concerns ignores the actual working practices of ‘bad actors’⁷ and this needs revisiting.⁸**

Newspaper Comment Boards

Under the current regime, newspaper comment boards are to be excluded from the regulatory remit of Ofcom and remain unregulated. The Antisemitism Policy Trust has worked with Government, civil service and

4 <https://www.article19.org/resources/free-speech-rapporteurs-a-blueprint-for-politicians-and-public-officials/>

5 <https://hopenothate.org.uk/wp-content/uploads/2021/09/Online-Safety-Bill-2021-09-v1-2.pdf>

6 <https://hopenothate.org.uk/2022/03/17/ill-be-back-the-rise-of-far-right-alt-tech/>

7 https://hopenothate.org.uk/wp-content/uploads/2022/03/state-of-hate-2022-v1_17-March-update.pdf

8 <https://www.gov.uk/government/publications/joint-committee-report-on-the-draft-online-safety-Bill-government-response/government-response-to-the-joint-committee-report-on-the-draft-online-safety-Bill>

other anti-racism organisations for years, to highlight the abuse on newspaper website comment forums. For example, as secretariat to the APPG, the Trust worked with the Department of Communities and Local Government (now DLUHC) towards a moderation guide delivered by the Society of Editors in 2014⁹ which was inspired by discussion on this form of harm.

A study looking at comments relating to reports by news outlets in the UK, France and Germany in 2021, found that antisemitism was easily found in comment on social media profiles of articles that pose as a ‘trigger,’ such as reporting on events in Israel, primarily on violent clashes between Israel and the Palestinians.¹⁰ UK publications attracted roughly twice as many antisemitic comments as French and German publishers. Considering this, comments boards on publishers’ websites can also be fertile ground for promoting antisemitism. Currently, some publishers, such as The Times, moderate their comment boards, but are not legally obligated to do so. However, since comments are made by readers, not journalists, **we believe that comment boards should not be protected under the Bill, but adhere to the same rules as other user to user content.**

Duties of Care: Search Services

The duties pertaining to Search Services are contained in Chapter 3 of part 3 of the Bill. Whilst Search Services are bound by the same duties as user-to-

user services in respect of illegal content, the user empowerment duties – that is friction in the system – does not apply to these services, this despite work we have undertaken which demonstrates the harm caused by search. Antisemitism Policy Trust research has found that changes to Google’s algorithm reduced antisemitic searches.¹¹ We have also worked with Microsoft Bing on similar issues, including its search bar promoting users towards the phrase ‘Jews are b*****ds’. Exemptions to the duties on search systems might equally apply to Amazon Alexa or the Siri service, despite these facilities directing people to antisemitic content. The joint committee recognised that search engines are more than passive indexes. “They rely on algorithmic ranking and often include automatic design features like autocomplete and voice activated searches that can steer people in the direction of content that puts them or others at risk of harm”.... “It is reasonable to expect them to come under the Bill’s requirements and, in particular, for them to conduct risk assessments of their system design to ensure it mitigates rather exacerbates risks of harm”. Sadly, the Government failed to heed to committee’s warning and at present people will be left open to harm through search.

We would welcome further amendments to introduce more stringent requirements on Terms of Service for Search Systems, similar to those set out for Category 1 services in section 65 of the Bill.

User Identity Verification and Anonymity

Part 4 (Section 57) of the Bill introduces a new duty on Category 1 entities, in Chapter 1, to include a verification option for relevant services. In an earlier section of the Bill (Part 3, Section 12) Category 1 companies are required to give users the ability to

“filter out non-verified users”, as part of the “user empowerment duties”. This is a welcome step forward but we believe more can be done to address anonymous abuse.

⁹ <https://www.societyofeditors.org/wp-content/uploads/2018/10/SOE-Moderation-Guide.pdf>

¹⁰ <https://decoding-antisemitism.eu/publications/second-discourse-report/> p.9-10

¹¹ <https://antisemitism.org.uk/wp-content/uploads/2020/06/APT-Google-Report-2019.1547210385.pdf>

The Trust has set out our position on this issue previously, detailing the impact of anonymity on Jewish communities online.¹² We also recognise the importance of anonymity to certain individuals in different contexts. It is the Trust's position that it should be the responsibility of a given platform to determine the degree of anonymity it wishes to offer users, though there should be a risk-assessed approach, with incentives against hateful content, and severe penalties for anonymous abuse. In our view, if a crime or a libel has been committed in the UK on regulated technologies and companies in scope cannot or will not provide proof of identity, where a court order demands it (subject to an appropriate burden of proof), then a range of options should be considered.

The Trust believes that the civil or criminal liability

should pass to the platform itself (this would accord with existing measures linked with the UK's implementation of regulations derived from the e-Commerce Directive, for example), and fines or other corrective measures could be put in place. We would suggest giving the platforms a year to become compliant. We recommend the underlining of police powers to compel revelation of identity through production orders be added to the Bill, with appropriate civil liberties safeguards, and a suggested amendment to this effect is being produced.

Of course, the earlier reference to small, high harm platforms holds particular relevance here given the volume of anonymous accounts which tend to operate on such mediums.

Ofcoms Powers and Duties

Penalties

We were delighted that, as the Minister recognised publicly in the Commons, the Antisemitism Policy Trust's recommendation (endorsed by the DCMS Select Committee¹³) that Breach Notices should be included in the Bill were heeded. We were also very pleased to see the Law Commission's recommendation that those persons habitually resident in the United Kingdom can be considered to have committed an offence when temporarily overseas in Part 10, section 165. This closes the loophole that allowed, for example, the grime artist Wiley to escape legal action in the UK for his antisemitic rant. The Trust has called for a more stringent senior management liability regime from before the Bill was introduced to the House of Commons and look forward to seeing what proposals the Government brings forward to improve the penalty regime. **We would support an amendment, as Government indicated it would be bringing**

forward, which brought the UK in line with the 'Irish model'.

Consulting experts and relevant groups

Throughout the Bill, Ofcom is required to consult with various entities and individuals on different areas of its work. For example, in part 3 of the Bill, in Chapter 6, Ofcom is mandated to consult those experiencing harm when drawing up Codes of Practice, which is specified as "persons who appear to OFCOM to represent the interests of persons who have suffered harm as a result of content to which the code of practice is relevant".

In Part 7 of the Bill, in Chapter 7 (section 141 (2) (6B)), Ofcom is required to consult with consumers on their various experiences. **It would be helpful for Secretary of State to clarify that expert and representative groups should form part of that consultation exercise.**

¹² <https://antisemitism.org.uk/wp-content/uploads/2020/12/Online-Anonymity-Briefing-2020-V10.pdf>

¹³ <https://committees.parliament.uk/publications/8609/documents/86961/default/>

Super Complaints

We are pleased to see confirmation of the SuperComplaint function in Part 8, Chapter 2 of the Bill, through which substantial evidence of systematic issues affecting large numbers or specific groups of people can be heard by Ofcom.

We are however significantly concerned that section 150 (3) specifies 'eligible entities' must meet criteria specified in regulations made by the Secretary of State. This leaves another important area of specificity to secondary legislation. **We would at the very least expect Secretary of State to establish or indicate some basic principles for eligibility during the passage of the Bill so that organisations like the Antisemitism Policy Trust and Community Security Trust can be reassured we will be heard on matters pertaining to antisemitism, on which we are expert.**

Entirely Absent: Supply Chains

Platforms, particularly those supporting user-to-user generated content, employ services from third parties. In the past, this has included Twitter explaining that racist Gifs were not its own but provided by another service. YouTube found it difficult to give precise figures for its moderator team given a number of moderators were employed or operated under third parties. The UN guiding principles on business and human rights and OECD guidance on responsible business conduct both cover supply chains. There are examples in UK legislation, for example the Bribery Act 2010, in which a company is liable if anyone performing services for or on the company's behalf, is found culpable of specific actions. Specific reference to supply chains, and similar culpability would be welcome rather than, for example, more vague references to content being encountered 'by means of the service' in eg 8(5)(b) and taken to mean the entirety of a platform.

This matter was discussed at the Committee stages in the Commons, in relation to outsourced moderators and working practices, with the Government stating that existing employment law was sufficient to protect UK workers, and putting the view that the Bill was not

the place to address these concerns but there was limited discussion of the type of outsourcing that might see a platform state that a service is not within scope, and the service to suggest it is not user-to-user but a business-to-business facility. **We would welcome further clarification on this point.**

Restrictive Covenant

The Trust believes that the Bill should include a restrictive covenant on senior government and Ofcom officials with direct responsibility for any platform within the regulatory ambit of the Bill. The revolving door between Government and social media platforms can be sinister and certainly undermines public trust.

Media Literacy

There is next to no detailed direction for the UK's Digital Media Literacy strategy in the Bill. There is a special imperative that online media literacy be well-conceived and delivered from an early age in the UK. This will help address the growing tendency towards conspiracy theory material online and gets to the heart of educating about antisemitism. We would encourage greater detail in this regard.

- We support greater user powers, including the creation of an ombudsman to hear individual complaints, something absent from the current Bill.
- We would like to see qualification of the Secretary of State's powers to direct Ofcom in relation to public policy

We would be pleased to discuss these with anyone interested to engage further on these topics.

Antisemitism Policy Trust Reports and Briefings



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](https://twitter.com/antisempolicy)



Antisemitism Policy Trust



mail@antisemitism.org.uk

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