

ANTISEMITISM POLICY TRUST

Special Briefing



THE ONLINE
SAFETY BILL

Front cover: 'digital code number abstract'. Licensed by whiteMocca/Shutterstock.com

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]

Overview:

The Online Safety Bill (OSB) has the potential to introduce an effective regime for managing online harms. It proposes a form of systems-based regulation, focussed on frameworks, business models, structures, and risk management. 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, addressing harmful albeit legal content.

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, and legal but harmful 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. There are some improvements from the draft Bill in this section. 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) is theoretically in scope.

The next section repeats this approach albeit for adults, in relation to priority harms which are due to be set out in secondary legislation, and only for what are to be known as category 1 companies. These companies must not only risk assess and act on legal harms (though the level of required activity is much lower than for children or illegal content) but provide for user redress and empowerment and must protect journalistic content and that of democratic importance. There are higher protections for children, and one might argue content harmful to adults may be harmful for children too but we suspect a number of platforms will argue they are not designed for or aimed at children, or have minimum standards in place to avoid being classed as such.

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 somewhat further back, in Schedule 10.1 (1) (4) and (5) [page 192] and linked to Part 7 Chapter 2 (80 and 81). In summary, categories are determined by the number of users of a user-to-user platform (or sub-platform) and the ‘functionalities’ of that part of a service, which is later defined as relating to the mechanisms of a platform or service. 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. 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, 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 signaled 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.

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 (10) needs amending, to reflect this. We recommend adding a new section 10 1(1) (c). 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.

Category 1 Companies: Countervailing Duties

Section 15 of Chapter 2 includes a duty to protect content of democratic importance which is defined in (6b) (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 Section 16 there is a duty to protect 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.

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>

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 anti-racism 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 Sccop 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 the aforementioned joint committee simply ignores the actual working practices of 'bad actors'⁷ and this needs revisiting.⁸**

Duties of Care: Search Services

The duties pertaining to Search Services are contained in Chapter 3 of part 3 of the Bill. The exemption that applied to Search in the draft Bill from Category 1 status has been repeated in the Online Safety Bill. **This is not a realistic understanding of the harms caused by Search systems and requires address.**

Antisemitism Policy Trust research has found that changes to Google's algorithm reduced antisemitic searches.⁹ 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. **This requires amendment to ensure Search engine design features that lead to legal harms follow similar requirements for user-to-user platforms. We recommend amending the Bill to add a risk assessment, policy response and consistency of enforcement requirement for Category 2A search engines, with a power for Ofcom to command smaller high risk search engines to comply with Category 2A requirements.**

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

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

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

7 https://hopenoathate.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>

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

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 14) 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.

The Trust has set out our position on this issue previously, detailing the impact of anonymity on Jewish communities online.¹⁰ 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 magistrate’s 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.

Part 7 of the Bill sets out Ofcom’s various powers and responsibilities, including amending existing legislation to update its new status.

Ofcom’s Powers and Duties

Risk Assessments

Ofcom has some sensible requirements placed upon it in Chapter 3, including in relation to undertaking risk assessments and compiling a risk register. However, in clause 83 (6) it states in preparing the risk profiles which relate to the risk of harm mentioned in subsection (1) (c) [risk of harm to adults], “OFCOM must not take into account anything relating to non-designated content that is harmful to adults”. This leaves a lot to chance, given designated content is not due to be specified until secondary legislation is introduced. **To this end, we recommend that these categories be specified in the Bill, or at the very least, the Secretary of State or relevant departments give notice of their intentions of categories of harm.**

There should also be an exemption built in so that Ofcom, in agreement with the Secretary of State can include non-designated harms where these are considered to pose a significant risk of impact on the British populus.

Penalties

In its report on the Draft Online Safety Bill, the House of Commons Digital, Culture, Media and Sport Committee made the following finding and recommendation: “The Antisemitism Policy Trust has noted that while the Bill details how Ofcom should publish decisions for failures against the duty of care, it provides no such provisions for mandatory breach notices for service providers.

Such transparency would theoretically incentivise compliance while providing greater transparency for users of a service when a decision is made by Ofcom. We recommend that the Government include a provision in the Bill to mandate publication of a breach notice by a service. This should include details of their breaches against the duty of care and be available to view on the platform.¹¹

The Government's response to the committee explained that it considered the requirement on Ofcom to publish details of its enforcement activity sufficient transparency. We were sorry that such a provision was not included in the Bill, consider it a get out of jail free card for platforms and hope Government will accept a simple amendment to Chapter 6 of this part of the Bill, possibly in Section 128 or a new Section 129.

We are 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 154. This closes the loophole that allowed, for example, the grime artist Wiley to escape legal action in the UK for his antisemitic rant.

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 132 (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.

Super Complaints

We are pleased to see confirmation of the Super-Complaint 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 140 (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. There are examples in UK legislation, for example the Bribery Act, in which a company is liable if anyone performing services for or on the company's behalf, is found culpable of specific actions. Reference to supply chains, and similar culpability would be welcome.

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.

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

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.

There are other areas in which the Trust would like to see changes or clarifications, for example:

- **Whereas in the draft Bill, there was reference to “people with a certain characteristic (or combination of characteristics)”, which was welcome recognition of intersectional harms, no such text exists in the final Bill. The different types of harm are however set to be defined in regulations, in consultation with Ofcom, and we hope intersectionality will be considered.**
- **We are concerned that in Part 12, section 187, harm over time is not accounted for**
- We would like to see a more public facing aspect to risk assessments undertaken or enforced by Ofcom, and for Ofcom to take into account different experiences of different groups in relation to its guidance and regulatory effort
- 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 (section 40)

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

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]

Antisemitism Policy Trust Reports and Briefings

