

ANTISEMITISM POLICY TRUST

Special Briefing

THE DRAFT
ONLINE SAFETY BILL

Front cover: 'Child and binary codes'. Licensed by Jiris/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 © 2021

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

Overview:

The Draft 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 design.

The draft Bill provides the overall framework for a raft of measures to 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.

Duties of Care

Get Search In

The Bill comprises three thematic duties of care, accompanied by risk assessment duties and three countervailing factors that must be considered, for certain categories of service hosting user-generated content. The duties are not all-encompassing, and some are stronger than others, with a focus on proportionality. There is still room to impose a broader duty of Care, encompassing an obligation to take reasonable steps to address reasonably foreseeable harms, similar to the Health and Safety at Work Act or the proposals made by the Carnegie Trust.¹

The proposed duties in relation to user-generated content services cover:

- i. Protecting individuals from illegal material, notably terrorism, child sexual abuse and crimes that are directed at individuals
- ii. Protecting individual children from harm

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), such as addressing legal but harmful content.

- iii. Protecting individual adults from harm (on the largest² services with particular 'functionality' only)

In relation to search engines, there are parallel duties for illegal materials and children but not for adults.

Antisemitism Policy Trust³ research has found that changes to Google's algorithm reduced antisemitic searches. 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. Clause 30 (3)(C) of the Bill indicates there is a requirement for safety by design in search algorithm and indexing, the draft Bill could be significantly simplified by matching the duties on search with those on user-to-user services.

Recommendation 1: The Bill's current structure is cumbersome and complex. Introducing an overarching duty of Care, as Government previously indicated it would, perhaps with different application according to risk and other factors would be welcome.

Recommendation 2: If the current system is to persist, search services should have the same duties as user-generated content services.

What are the Terms?

The applicability of duties of care are set out in the draft Bill, with enhanced duties for services accessed by children and those in 'category 1' (the largest services and those with particular functionality). In summary, companies must risk assess for illegal content, and re-assess before making changes to their systems. They must keep a children's risk assessment if their services lend themselves to use by children and carry out a wider risk assessment for content harmful to adults, updating Ofcom as appropriate when within the category one group. The basic risk assessment must include methods to 'minimise' the presence, length of time, availability of illegal content (or 'mitigate' and 'prevent' such content in relation to children)—which must be removed 'swiftly' once made known to companies. These protections are to be in Terms of Service which must be applied 'consistently'.

Regrettably, the Bill provides very little detail on the duty to address legal but harmful content, leaving it up to platforms themselves to manage the risks to adults through Terms and Conditions⁴ which must ensure and specify how content is 'dealt with' (including a choice to do nothing about a particular type of content, so long as this is outlined clearly) and be consistently applied.

The current assessment criteria for categorising into group 1 or group 2 (size and functionality) are insufficient. There are countless so-called 'alternative' platforms which house extremist content and risk inspiring extremist behaviour.⁵ These should be subject to the additional duties set out in the Act.

Recommendation 3: There are no minimum standards set out for Terms and Conditions, expected to address harm to adults, which have proven to be hugely inconsistent across platforms. Terms and Conditions for addressing harmful content should be required to meet a minimum standard and the wording of the Bill should be amended to recognise this, including by defining what 'dealt with' means. Furthermore, risk assessments of harmful content performed by companies in scope should meet a requirement to be 'reasonable' to prevent gaming of the system.

Recommendation 4: The Categorisation of companies must be amended so that size and functionality alone are not the determinant factors for classification, but risk. The relevant schedule (4) needs amending, for example, to reference the risk register developed and maintained by Ofcom in Clauses 61 and 62.

Getting The Balance Right

There are several counterbalancing 'protective' duties, these are in relation to protecting:

- rights, such as free speech and privacy (including special provisions for category one companies)
- content of democratic importance (for category one companies)
- journalistic content (for category one companies).

Whilst perhaps well intentioned, the drafting of these duties appears problematic. There are those that actively undermine the democratic process, or cause harm using it as cover. For example, might standing in an election allow a far-right activist to claim their content should effectively be re-platformed once barred from a service? What about those standing on a neo-nazi platform more broadly? Might an individual or party argue misogynist content be of democratic important in the context of a campaign by the Women's Equality Party?

The explanations offered by officials, that platforms will have to perform a balancing act in respect of harms and content, have not provided reassurance.

The protective duty in relation to journalistic content presents similar and additional concerns. Journalistic content is poorly defined and might be read, as Hope Not Hate has suggested, as content “generated for the purposes of journalism”⁶. 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. Again, officials suggest this will require a balancing act by platforms, but we remain unconvinced. There are examples of far-right activists self-identifying as journalists, and news companies like InfoWars 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.

Furthermore, there is an explicit exemption⁷ for content present on the website of a recognised news publisher. This is deeply problematic. The Antisemitism Policy Trust has worked with Government and others for many 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 MHCLG) towards a guide delivered by the Society of Editors in 2014^{8,9}, which was inspired by discussion of this form of harm.

Recommendation 5: The wording of the counter-balancing duties should be amended to add specificity to the relevant categories of content, or to reference a set of standards to be outlined by the regulator.

Recommendation 6: The exemption in Clause 18 relating to newspaper comments boards should be removed or, at worst, amended to ensure publications have measures in place to address harm on relevant boards.

Code It Be Magic

The Draft Bill empowers Ofcom to consult on, and to produce, numerous codes of practice. There are welcome elements of the Bill, for example Clause 29 includes a duty on Ofcom when designing the codes to liaise with “persons who appear to OFCOM to represent the interests of persons who have suffered harm as a result of encountering content online”. There are also elements that give some cause for concern, for example the Secretary of State’s power to direct Ofcom to modify its codes of practice to bring them in line with government policy, which should be removed.

Recommendation 7: Remove the power to direct Ofcom to modify its codes of practice to bring them in line with government policy.

The final sections of this initial chapter of the draft Bill include definitions for the earlier parts of the Bill, including a welcome reference to the impact of harm on “people with a certain characteristic (or combination of characteristics)”. This is welcome recognition of the case put forward by the Trust and key partners of ours including Glitch, that have been seeking recognition of intersectional harms. The different types of harm are however set to be defined in regulations, in consultation with Ofcom. The Trust was reassured to hear Secretary of State Dowden expressed his expectation that antisemitism would be a recognised harm¹⁰ and would similarly expect any respectable list of online harms to include anti-Jewish racism.

Ofcom's Powers and Duties of Care

On Ofcom

The Bill provides Ofcom with a number of powers to apply in the course of its regulatory work. Clauses 71 and 73 introduce senior management liability in respect of failing to provide information when presented with a notice by the regulator— this is welcome. The Government has signalled its intention to legislate following a two-year review of operations. We consider this an unnecessary delay which will strengthen the opposition and arguments against sanctions from technology companies.

At the other end of the scale, Clause 96 details provisions for publishing decisions by Ofcom but there is no provision to mandate publication of a breach notice by a service.

Recommendation 8: We recommend that Ofcom be granted full enforcement powers and criminal sanctions under chapter 6 of the Draft Bill in relation to senior management liability.

Recommendation 9: Amend the Bill to include powers to enforce publication of a breach notice by a regulated service.

In Clause 99, (2)(6b) Ofcom is compelled to consult on public opinion in respect of its activities but this section of the Bill, contrary to Clause 29, omits specific reference to consulting users that are experiencing harm.

Recommendation 10: Clause 99 should be amended to include users experiencing harm

Questions, Concerns and Matters Requiring Further Consideration

Anonymity

The gaping policy hole in the draft Bill is anonymity. The Trust has set out our position on this previously, detailing the impact of anonymity on Jewish communities online.¹¹ 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 be in line with existing measures in the e-Commerce Directive), and fines or other corrective measures could be put in place. We would suggest giving the platforms a year to become compliant. Companies should apply the 'Know Your Client/Customer' principle, familiar to those in the financial sector, with appropriate safeguards. Using some of the legal framework required by companies offline, such as Customer Due Diligence in The Money Laundering, Terrorist Financing and Transfer of Funds regulations 2017, online companies should verify users' identities before allowing use of their platform. This should be done even if use of the platform is free of charge and when users are not regarded as 'customers'. The Pan Canadian Trust Framework (PCTF) is but one example of a scheme that could be reviewed.

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.

Questions

There remain several outstanding questions about the Bill that we encourage parliamentarians to put to Ministers as the Bill progresses, for example:

- How will individual, as opposed to societal or cumulative harms be managed? The former are clearly captured in the Bill, the latter are not. Is there a point at which a particular number of individual harms become cumulative, and can these be addressed through the legislation?
- Can the Government provide more detail on the thresholds and description of harm which differs in criminal and tort law and has sometimes met such a high threshold as to be unrecognisable.
- What provisions for appeal will there be against Ofcom decisions? It is important that no loopholes exist which prevent effective enforcement of Ofcom rulings.
- To what extent will the Law Commission's work on Communications or Hate Crime legislation impact the Bill process?
- When will the Digital Media Literacy Strategy be published?
- Will the Government enable Ofcom to act by providing it with preparatory powers?

Footnotes

¹ https://www.carnegieuktrust.org.uk/blog/reducing-harm-social-media-duty-care/?gclid=CjwKCAjwieuGBhAsEiwA1Ly_nf1i4m_mACAvRtJmFtegB8KpdmDgxc2BK5BPIWryi_VBZpGb8dSihBoClaYQAvD_BwE

² The size threshold is yet to be determined; it may also have implications in respect of transparency duties.

³ <https://antisemitism.org.uk/wp-content/uploads/2020/06/APT-Google-Report-2019.1547210385.pdf>

⁴ See Clause 11

⁵ <https://cst.org.uk/news/blog/2020/06/11/hate-fuel-the-hidden-online-world-fuelling-far-right-terror>

⁶ <https://www.hopenothate.org.uk/2021/06/07/hope-not-hates-response-to-the-draft-online-safety-bill/>

⁷ See Clause 18

⁸ <https://www.societyofeditors.org/wp-content/uploads/2018/10/SOE-Moderation-Guide.pdf>

⁹ <https://www.bod.org.uk/bod-news/board-attends-society-of-editors-and-dclg-online-moderation-guide-launch/>

¹⁰ <https://committees.parliament.uk/oralevidence/2185/html/>

¹¹ <https://antisemitism.org.uk/wp-content/uploads/2020/12/Online-Anonymity-Briefing-2020-V10.pdf>

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



[@antiseppolicy](https://twitter.com/antiseppolicy)



Antisemitism Policy Trust



mail@antisemitism.org.uk

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