Your Response

Question 1: To assist us in categorising responses, please provide a description of your organisation, service or interest in protection of children online.
Is this a confidential response? (select as appropriate) No
The Antisemitism Policy Trust is a charity that works to educate and empower parliamentarians and policy makers to address antisemitism. For more than ten years, the Trust has provided the secretariat to the All-Party Parliamentary Group (APPG) Against Antisemitism. The Trust has advised the government, opposition parties, policy makers, civil servants, and regulators, including Ofcom, on policies relating to antisemitism, hate crime and online abuse. We have produced briefings and provided written and oral evidence about online safety, especially with regard to online antisemitism, both in relation to illegal and legal but harmful content. The Trust Chief Executive, Danny Stone, gave evidence to Parliamentary Committees scrutinising the Online Safety Bill, including the Joint Committee on the Draft Online Safety Bill, the Petitions Committee, and the Public Bill Committee. Our recommendations were adopted or advanced by some of these bodies, and by the DCMS Select Committee, and our work helped influence Government to add a breach notice provision to the Bill, enhancing Ofcom's powers

Question 2: Can you identify factors which might indicate that a service is likely to attract child users?
Is this a confidential response? (select as appropriate) [Please select]

Question 3: What information do services have about the age of users on different platforms (including children)?
Is this a confidential response? (select as appropriate) [Please select]

Question 3: What information do services have about the age of users on different platforms (including children)?

Question 4: How can services ensure that children cannot access a service, or a part of it?
Is this a confidential response? (select as appropriate) [Please select]

Question 5: What age assurance and age verification or related technologies are currently available to platforms to protect children from harmful content, and what is the impact and cost of using them?
Is this a confidential response? (select as appropriate) [Please select]

Question 6: Can you provide any evidence relating to the presence of content that is harmful to children on user-to-user and search services?
Is this a confidential response? (select as appropriate) No

Question 6: Can you provide any evidence relating to the presence of content that is harmful to children on user-to-user and search services?

The Trust has done extensive research into online antisemitic content, including on platforms and other services used by children. Search services are included in the Bill, but are regarded differently to user-to-user platforms. However, search engines are not passive indexes; they produce results based on algorithms and design features that can be programmed according to the service's interests. Google's autocomplete algorithm was been found to suggest antisemitic, racist and sexist content to users; typing the word 'are Jews' in the Google search bar, prompted an autocomplete suggestion '…. evil?'. This produced results that demonise or incite people to hate Jews. Microsoft Bing was found to direct users to hateful searches with the autocomplete "Jews are b*****ds" and Google's image carousel highlighted pictures of portable barbecues to those searching Jewish baby stroller.

Our research into Google found that its 'SafeSearch' option produced as many antisemitic results as its regular search. For example, when searching for the term 'Jew jokes' with the SafeSearch option disabled, 48% of the results produced by Google were found to be antisemitic – a high proportion in of itself. However, the same search phrase with the SafeSearch option enabled, produced an even greater proportion of antisemitic results – 57%. This places children and other vulnerable people at risk, because they wrongly assume that they are protected by Google's SafeSearch option. Many parents who allow their children to use Google with the SafeSearch option activated, do so under the assumption that it will limit exposure to harmful content. Unfortunately, this is not the case. While SafeSearch may be effective with regard to certain content, such as pornography, it is ineffective when it comes to antisemitic and potentially other racist materials.

Antisemitic results are not limited to regular search bars. Voice-activated services also produce potentially harmful content. As Andrew Percy MP evidenced in his speech during the Report Stage of the Online Safety Bill, Amazon's Alexa produced an antisemitic conspiracy theory in response to a search. It suggested – based on a single comment posted on Amazon's website – that the Jewish American-Hungarian philanthropist George Soros is responsible for all of the world's evils – a common trope based on antisemitic conspiracies. This is information that could reach millions of users of voice activated services around the world. Percy also pointed out that a Google search for the seemingly innocent words 'desk

Question 6: Can you provide any evidence relating to the presence of content that is harmful to children on user-to-user and search services?

ornaments' has produced top search results that included swastikas, SS bolts and other Nazi memorabilia.

Antisemitic search results come up in other languages too. The Trust found that asking Siri, in Spanish, "do the Jews control the media?" prompted a response that highlighting articles including details of "Jewish control international media" and an article arguing that "A world famous sociologist claims that the Jews control the media".

Considering user-to user platforms, research into online antisemitism by the Antisemitism Policy Trust, in collaboration with the Community Security Trust (CST) and the Woolf Institute, found that antisemitism flourishes on Instagram, with antisemitic hashtags often associated with conspiracy theories. Such hashtags are sometimes attached to posts that have no direct relationships to the content of the post, meaning that they are displayed to a large pool of users, who in most cases are not actively looking for antisemitic content, but are exposed to it and to its harmful influence. A case of antisemitic supply rather than demand. Our research concluded that Instagram requires improved algorithmic filtering to address conspiracies and antisemitic hashtags, and that it needs to improve its community standards against hateful content. Worryingly, a recent report found that 45% of children under 13 in the US are using Instagram, despite the platform's age restrictions. Of these, 36% reported that they had been exposed to harmful experiences on the platform. Users aged 13-17 make up some 9% of users.

From our own experience, antisemitic material that includes Holocaust denial and revisionism and conspiracy theories against Jewish people, are very easily found on both YouTube and Amazon UK, with both platforms' automated systems suggesting more harmful content to viewers. These are services that can, and are, being used by children.

Considering the harmful content produced by search engines, we have called for search engines to be treated in a similar to user-to user services as regards risk assessments when it comes to legal but harmful content that children are likely to be exposed to, applying stronger duties of care to larger search engines. Additionally, the risk of exposure by children to

Question 6: Can you provide any evidence relating to the presence of content that is harmful to children on user-to-user and search services?

harmful but legal content on voice search assistants, as demonstrated above, can easily lead users to increasingly extreme and even illegal content. Greater clarity on the responsibility of voice activated search systems should be provided by Ministers in relation to the Online Safety Bill's requirements, but so too Ofcom should be mindful of the potential harm caused by these services.

Question 7: Can you provide any evidence relating to the impact on children from accessing content that is harmful to them?

Is this a confidential response? (select as appropriate)

No

Racist content online has consequences offline. There have been numerous violent attacks on Jewish targets involving people who were radicalised online by being exposed to antisemitic conspiracy theories – the vast majority of which would be deemed legal content. When it comes to incidents of antisemitism that involve children, schools in the UK have seen a rise in antisemitic incidents which includes verbal attacks on pupils and staff. CST recorded 50 antisemitic incidents against Jewish pupils or staff at schools in the 1st half of 2022. Of the 2,255 antisemitic incidents recorded by CST in 2021, 99 occurred in schools. Whilst it is not yet clear that there is a direct causal link between these statistics and online activity, anecdotally, during the last Israel-Gaza conflict, TikTok was used widely to stir up activism in schools, and there were examples of antisemitism stemming from those demonstrations, and so we are confident to suggest that there is some connection between the two.

Question 8: How do services currently assess the risk of harm to children in the UK from content that is harmful to them?

Is this a confidential response? (select as appropriate)

[Please select]

Question 8: How do services currently assess the risk of harm to children in the UK from content that is harmful to them?

Question 9: What are the exacerbating risk factors services do or should consider which may have an impact on the risk of harm to children in the UK?

Is this a confidential response? (select as appropriate)

No

Referring back to the results of our study about antisemitism on Instagram, the platform's algorithm actively suggested antisemitic conspiracist content to users even without any prompts. The study shows that one of the risk factors, is the way Instagram designs its algorithm and deploys it. For search services, autocomplete functions and indexing are key. In algorithmic design, indexing, autodirect or other related prompt systems, companies should be considering and consulting, where they have them, lists of key harmful words or phrases, and particularly those covered under Equalities legislation as having protected characteristics.

This requires a systems based approach to how platforms are designed and operate, rather than content-based regulation. This means expecting companies to employ safety-by-design and other features that help minimise the presence of illegal and legal but harmful content, and automatically limit users' exposure to such content where it is present.

Question 10: What are the governance, accountability and decision-making structures for child user and platform safety?

Is this a confidential response? (select as appropriate)

[Please select]

Question 10: What are the governance, accountability and decision-making structures for child user and platform safety?

Question 11: What can providers of online services do to enhance the clarity and accessibility of terms of service and public policy statements for children (including children of different ages)?

Is this a confidential response? (select as appropriate)

No

See our reply to Ofcom's consultation on the First Phase of Online Safety Regulation, Q5.

Question 12: How do terms of service or public policy statements treat 'primary priority' and 'priority' harmful content?[1]

Is this a confidential response? (select as appropriate)

[Please select]

Question 13: What can providers of online services do to enhance children's accessibility and awareness of reporting and complaints mechanisms?

Is this a confidential response? (select as appropriate)

[Please select]

[1] See A1.2 to A1.3 of the call for evidence for more information on the indicative list of harms to children.

Question 13: What can providers of online services do to enhance children's accessibility and awareness of reporting and complaints mechanisms?

Question 14: Can you provide any evidence or information about the best practices for accurate reporting and/or complaints mechanisms in place for legal content that is harmful to children, or users who post this content, and how these processes are designed and maintained?
Is this a confidential response? (select as appropriate) *[Please select]*

Question 15: What actions do or should services take in response to reports or complaints about online content harmful to children (including complaints from children)?
Is this a confidential response? (select as appropriate) *[Please select]*

Question 16: What functionalities or features currently exist that are designed to prevent or mitigate the risk or impact of content that is harmful to children? A1.21 in the call for evidence provides some examples of functionalities.
Is this a confidential response? (select as appropriate) *[Please select]*

Question 16: What functionalities or features currently exist that are designed to prevent or mitigate the risk or impact of content that is harmful to children? A1.21 in the call for evidence provides some examples of functionalities.

Question 17: To what extent does or can a service adopt functionalities or features, designed to mitigate the risk or impact of content that is harmful to children on that service?

Is this a confidential response? (select as appropriate)

No

As mentioned, when it comes to antisemitic and other racist content, platforms should employ safety by design features and create friction in the system that can minimise illegal and legal but harmful content that children may be exposed to, from being uploaded.

Platforms and search services should also restrict exposure to harmful content by designing algorithms that direct children to counter-information, steering them away from harmful and racist conspiracy theories and disinformation and instead offering content that is reliable and moderate. An example of good practice is Facebook's suggestion of reliable sources about vaccines, when their system recognises content relating to vaccines on their platform. Our research on Instagram, mentioned earlier, revealed that the opposite is happening, with the platform suggesting antisemitic content to users. A recent engagement we had with Google produced disappointing results, in so far as Google was not able to appropriately target counter-speech on antisemitism to various audiences. Ensuring systems for counter-speech are appropriate and workable should be something Ofcom keeps in mind and asks providers about, including whether they have experienced any problems with re-direct efforts.

In addition, user-to-user platforms should improve their community standards against hateful content that can incite, radicalise and have negative impact on children's emotional and mental wellbeing. Platforms should design and consistently enforce their Terms and Conditions in a way that contributes to building a positive and safe online community for children.

Question 18: How can services support the safety and wellbeing of UK child users as regards to content that is harmful to them?

Is this a confidential response? (select as appropriate)

No

Services can offer support in a variety of ways. When it comes to antisemitic and other racist content, services should offer reliable sources of information that promote tolerance and inclusion and moves young users away from sources that promote hate, conspiracy theories and racism. In addition, systems can use algorithms that can evaluate and identify sign of mental health issues and suggest that users seek support offline. Ensuring reliable partner organisations, in the case of antisemitism, the CST, Antisemitism Policy Trust, Holocaust Educational Trust and others, have priority indexing can be helpful. It is important that in relation to harmful content, the appropriate reporting mechanisms (for example CSTs incident reporting facilities) are given priority on a page, platform or server.

Question 19: With reference to content that is harmful to children, how can a service mitigate any risks to children posed by the design of algorithms that support the function of the service (e.g. search engines, or social and content recommender systems)?

Is this a confidential response? (select as appropriate)

[Please select]

Question 20: Could improvements be made to content moderation to deliver greater protection for children, without unduly restricting user activity? If so, what?

Is this a confidential response? (select as appropriate)

No

See our reply to Ofcom's consultation on the First Phase of Online Safety Regulation, Q11.

Question 21: What automated, or partially automated, moderation systems are currently available (or in development) for content that is harmful to children?

Is this a confidential response? (select as appropriate)

[Please select]

Question 22: How are human moderators used to identify and assess content that is harmful to children?

Is this a confidential response? (select as appropriate)

[Please select]

Question 23: What training and support is or should be provided to moderators?

Is this a confidential response? (select as appropriate)

No

See our reply to Ofcom's consultation on the First Phase of Online Safety Regulation, Q11.

Question 24: How do human moderators and automated systems work together, and what is their relative scale? How should services guard against automation bias?
Is this a confidential response? (select as appropriate) *[Please select]*

Question 25: In what instances is content that is harmful to children, that is in contravention of terms and conditions, removed from a service or the part of a service that children can access?
Is this a confidential response? (select as appropriate) *[Please select]*

Question 26: What other mitigations do services currently have to protect children from harmful content?
Is this a confidential response? (select as appropriate) *[Please select]*

Question 27: Where children attempt to circumvent mitigations in place on a service, what further systems and processes can a service put in place to protect children?
Is this a confidential response? (select as appropriate) *[Please select]*

Question 28: Other than those covered above in this document (the call for evidence), are you aware of other measures available for mitigating the risk, and impact of, harm from content that is harmful to children?
Is this a confidential response? (select as appropriate) *[Please select]*