Summary:

- Industry should become more transparent about its moderation training processes, opening them up to experts and independent scrutiny
- Social Media companies should be funding expert third sector organisations to help improve their reporting and review processes if they cannot do so themselves
- There is a strong case for reviewing and consolidating legislation in the UK and binding social media companies to their own terms of service
- The “Code of practice for commercial social media platform providers on online abuse” section of the Digital Economy Bill should be retained by the Government during consideration of amendments
- More effective application of relevant judicial orders in sentencing including Banning and Criminal Behavior Orders would be welcome
- Effective single points of contact for the police and the Crown Prosecution Service with the social media companies are essential
- Police require increased digital capacity, better training specifically on criminal thresholds, better tools and partnerships with relevant expert organisations
- For the industry, investment in artificial intelligence systems should be increased. In the meantime, dual-factor blocking should be utilised. There is also a requirement for greater consistency of approach in moderation and better training
- Where industry firms contract out services, such as the provision of GIFS for direct messages, there should be strict standards in place
- Social media companies must abide by their own terms of service and should be judged against them. Introduction of and rating by an independent ombudsman should be considered
- Further consideration should be given to the wider implications of the results of the Society of Editors Online Moderation Guide as regards anonymity of accounts
- Reporting and tracking abuse is still far too cumbersome a process and must be improved
- Government should review laws at home and abroad that confer exemptions from liability upon social media companies for the content they host.
1. **Background: Work by the APPG Against Antisemitism to Address Cyber Hate**

1.1: For over ten years the APPG Against Antisemitism, with the support of the PCAA Foundation, has sought to find appropriate means to combat antisemitism online. Obviously social media can be a force for good but if determined and calculating racists are able to act with impunity that positive potential diminishes. The rationale for action is clear. According to the Community Security Trust, in 2011 only 12 of 609 recorded antisemitic incidents were from social media. In 2015 the number had risen to 185 and by 2016 the number was 287 incidents comprising 22 per cent of the overall total of 1,309. This number falls far short of the total number of antisemitic tweets, pictures posts and messages that exist or are accessible across the various social media platforms.

1.2 The 2006 All-Party Parliamentary Inquiry into Antisemitism report highlighted cyber hate as a key concern. It referenced the potential for “globalising hate” through increased access to the internet and highlighted the existence of antisemitic websites, platforms enabling racist commentary and the commercial distribution of antisemitic and other hate materials online. Former Attorney General Lord Goldsmith told the all-party inquiry that officials seeking to address racism online faced numerous jurisdictional challenges. Then Home Office Minister Paul Goggins however believed that either action taken to address child pornography or the 2006 Terrorism Act might have potential application for efforts to combat cyber hate. The all-party report highlighted the importance of international governmental co-operation for monitoring and regulation of the internet. The two formal recommendations of the 2006 report were that:

“…the Foreign and Commonwealth Office examines ways of convincing the governments of countries where antisemitic internet sites originate to take action to close them down. The United States in particular has been slow to take action in this area. We conclude that a new approach is needed in terms of freedom of expression that allows some limit on the public dissemination on the internet of material aimed at stirring up race hate and antisemitism”.

“... all providers of online payments systems adopt Offensive Material Policies which they undertake to actively police and that these organisations have clear mechanisms for members of the public to report any breaches of the policy. In addition we also recommend these providers strengthen their links with organisations such as Searchlight, which monitor the presence of racist, including antisemitic, material online, and respond quickly to any reports that their systems are being used to disseminate this material”

Following publication of the inquiry, a number of efforts were made to pursue successful implementation of the recommendations. PayPal and subsequently, major credit card companies took action to impede far-right groups using their systems to fund their activities as a result.

1.3: As part of ongoing efforts to improve policy pertaining to cyber hate and at the urging of the APPG against Antisemitism, in 2009 Ministers of the then Labour Government held a cross-party seminar at which they met with key agencies, Jewish community representatives and officials to discuss antisemitism on the Internet and actions to address it. Amongst the topics covered at this meeting included community cohesion, anonymity, jurisdiction and the disparity of global legislative frameworks. A series of actions were agreed including a recommendation that the Foreign and Commonwealth Office (FCO) work with the Cross-Government Hate Crime Programme to seek international co-operation to reduce the harm caused by hate crime on the Internet. The UK eventually secured an agreement in such regard within the OSCE Ministerial Council Decision 9/2009. Following that decision, a series of experts’ meetings were convened, but state responses to internet harm were inconsistent. The call to discourage state control of free speech through censorship, the promotion of freedom of expression and protection from harm have been seen as competing rather than complimentary aims in this regard.
1.4: Coalition Government Ministers hosted a second UK Ministerial seminar in July 2011, again at the urging of the APPG, which included industry and cross-departmental Government representatives. The meeting was framed as a chance for those working to combat internet hate to learn the lessons of campaigns against online child exploitation. The resultant discussion and policy recommendations fed into Home Office deliberations about broader Internet policy decisions being considered across government. The 2016 Government Hate Crime Action Plan continues this trend, with cyber hate concerns addressed directly. A third Ministerial conference is being planned for summer 2017 and the APPG has again been meeting Ministers and pursuing progress for its planning.

1.5: The Inter-Parliamentary Coalition for Combatting Antisemitism (ICCA) is a project of the PCAA Foundation. The organisation was founded in 2009 and sought to unite parliamentarians from around the world to lead the fight against resurgent global antisemitism. At its inaugural conference, the ICCA London Declaration\(^1\) was agreed and included recommendations for various bodies to take action against online hate. The declaration was subsequently signed by the leaders of all the UK’s major political parties.

1.6: In November 2010, the ICCA held its second conference in Ottawa at which it was agreed to establish an international task force on Internet Hate (not limited to antisemitism), to make policy recommendations to governments and international bodies. The task force met over the course of two years and published its findings in 2013\(^2\). It concluded that legislative responses to cyber hate would be inappropriate but resolved to create a permanent Anti-Cyber Hate Working Group to build best practices for understanding, reporting upon and responding to internet hate. In some instances, this group enabled Heads of Content Control from large companies to meet for the first time. The Group produced a statement of aspirational principles\(^3\) that all participating industry stakeholders committed to. Whilst aspirational, these principals for the first time provided a clear and jointly agreed framework through which companies could seek to address cyber hate.

1.7: In 2013 back in the UK, following a recommendation from the APPG and working in partnership with the Department for Communities and Local Government, the Society of Editors published a study into moderation of online comments boards\(^4\).

1.8: In 2014, a far-right activist named Garron Helm sent an antisemitic tweet to Luciana Berger MP. He was charged and successfully prosecuted and an orchestrated hate campaign followed. APPG chair John Mann initiated efforts to serve an Interpol warrant on a second activist, Andrew Anglin. Anglin is the editor of the Far Right ‘The Daily Stormer’ website and although US-based had been touring Europe including a spell in Germany where he made a film denying the Holocaust – an illegal act in Germany. Anglin was responsible for the online campaign against Luciana and John both contacted the relevant authorities and held an adjournment debate in the House of Commons about the campaign of hate online. As part of the efforts to address this far-right activity and another case relating to now jailed far-right troll Joshua Bonehill, John Mann MP and Matthew Offord MP visited Dublin under the auspices of the APPG to meet Twitter and Facebook at their European Headquarters. Facebook officials were open and engaging, Twitter officials however were very unhelpful.

1.9: In 2014 following an increase in antisemitism in the UK, related to an escalation of violence in the Middle East, the second All-Party Parliamentary Inquiry into Antisemitism was commissioned. Research carried out by Lancaster University to assist the MPs on the inquiry panel found that amongst the top 35 Twitter key words relating to Jews in summer 2014 were ‘Nazi’, ‘Hitler’, ‘Holocaust’ and ‘Nazis’. The report considered this and other matters. The panel recognised the reach and potential of social media for good but also to spread fear, the report reviewed a number of concerns about Twitter specifically which are discussed in the next section. The inquiry panel made a number of recommendations including:

\(^{1}\) http://www.antisem.org/london-declaration/


\(^{3}\) https://www.adl.org/cyberhate-response

“...that further research be carried out into the sources, patterns, nature and reach of the antisemitism on social media. Such learning can help to identify the most appropriate responses and effective deployment of resources to combat hate online.”

“...that the Crown Prosecution Service undertakes a review to examine the applicability of prevention orders to hate crime offences and if appropriate, take steps to implement them.”

“True Vision (www.report-it.org.uk) is a critically important framework for third-party reporting which deserves wider publicity and we recommend the government works with industry partners to organise a prominent awareness campaign about methods for reporting online racist and other abuse.”

“...that officials from the Foreign and Commonwealth Office and other relevant departments pursue a strategy of mainstreaming concerns about social media and cyber hate into existing discussion forums in Europe including the EU Fundamental Rights Agency working groups on hate crime, with a view to securing further Ministerial decisions.”

“...that the Crown Prosecution Service undertakes a review to examine the applicability of prevention orders to hate crime offences and if appropriate, take steps to implement them.”

“...that the government offers additional resources to the police to enhance and develop policing and investigation of online hate crime.”

1.10: Since the report was published, True Vision has received additional support from DCLG and in 2016 as a direct response to the inquiry recommendation, the CPS renewed its social media guidance.

1.11: In recent months, members of the APPG and the PCAA Foundation have: continued to encourage social media companies to act, engaged with the reclaim the internet campaign, joined European Commission efforts to ensure voluntary takedown of illegal hate material within a 24 hour period, learned from German Ministers during the third ICCA conference about their plans to legislate to ensure social media took action against hate online, met companies seeking to tackle cyber hate in Israel, met with social media companies bilaterally and most recently, worked to support amendments to the Digital Economy Bill. Members of the APPG and the PCAA Foundation will continue to focus on efforts to address cyber hate in the coming years.

2. Reflections on Experience of the Industry Position on hate speech

2.1: Over a number of years engaging with the industry bilaterally, through conferences, governmental and other initiatives, certain themes, messages and policy positions have recurred. These industry positions in some cases might assist efforts to combat hate online if reversed.

2.2: The industry is keen to promote counter-speech and occasionally restorative justice as central planks of its response to hate speech. Money has been directed to research, for example with Demos, on collecting empirical evidence about what works in counter-narrative exercises. Facebook has certainly developed ‘in-line’ reporting which pre-drafts messages for users to send to others sharing questionable content. This is very important work but of course not the only way in which cyber hate should be addressed.

2.3: Industry organisations will usually cite constantly evolving and adapting reporting processes. Twitter in particular, has been making significant improvements to its systems with efforts to control the visibility of Tweets and for example, to address so-called ‘Dog Piling’ and other online trolling strategies. In addition, we have been informed that both Twitter and Facebook use case-studies to improve training and to update reporting systems. However, industry organisations whilst offering briefings about their systems appear totally averse to showing MPs or others how their moderation training is carried out in situ or to share, for example, the training modules so that experts might help improve them. This is something that could and should change.

2.4: Most of the internet companies rely on NGOs or expert groups like (in the case of antisemitism) the Community Security Trust for assistance in reactively regulating their platforms. For example, Facebook, Twitter and some others have a whitelisting arrangement whereby reports of hate material are prioritised for review if submitted by said NGOs. There is some outreach too and these companies and Google have certainly offered promotional support to relevant anti-racism charities. However, the social media platforms have generally claimed insufficient resources are to blame for their reactive approach. There appears to be muted enthusiasm to enter into formal financial partnerships and specifically to fund third party organisations that might filter reports and support efforts to better regulate content on these platforms. Certainly organisations like Tell Mama and the Red Button have made such offers before. Should all social media companies offer funds to, for example, True Vision there is no doubt that such an initiative could help collate and prioritise reports and provide a single point of contact for social media companies.

2.5 Generally, where there has been a reticence to take action, the stated reasons have been technical or bureaucratic. This ranges from refusal to discuss individual cases when antisemitism has occurred to suggestions that racist pictures cannot be removed. The companies tended to claim they are mere hosts, not content controllers but this is no longer a viable position. Ultimately industry is willing to listen and engage but action has tended to occur through PR crises or the threat or actuality of funds being removed.
3: Action to Address Online Hate:

3.1 In Law, fortunately, Britain has a good national record as regards convictions for internet-related incitement, malicious communications and other online abuse. In the R v Sheppard and Whittle case two men appealed against convictions for possessing, publishing and distributing racially inflammatory material contrary to the Public Order Act 1986. They had between them composed, edited and uploaded racist material online to a website hosted by a server in the USA. The judge ruled that the UK courts had jurisdiction to try the case as a substantial measure of their activities had taken place in the UK. This sets a very important precedent for prosecutions of online hate and is not sufficiently well known or publicised. The prosecution of individuals who sent antisemitic twitter communications to Luciana Berger MP and others also set important precedents. However, as John Mann MP has previously advised the Home Affairs Select Committee, consideration of consolidating legislation in the UK and adding a concept that social media companies should be required to keep their own terms of service might be advisable. The Public Order Act, The Communication Act, the Malicious Communications Act and tens of other acts offer means to combat cyber hate but there is a strong case for reviewing and consolidating legislation and bringing it into the social media age.

3.2 An amendment was accepted recently to the Digital Economy Bill. The Lords amendment moved by Baroness Jones and supported by other cross-party and independent peers will ensure a “Code of practice for commercial social media platform providers on online abuse”, this is something that MPs of all parties should be actively supporting when consideration of amendments takes place. In her speech to the amendment, Baroness Jones referred to the Australian system of licencing social media companies. The Australians have established a tier scheme with two levels focussed on the removal of cyberbullying material from social media services. Based on an Act of parliament, companies are required to have a complaints management system, terms of use prohibiting cyberbullying and referral to a relevant ombudsman for complaints deemed inadequately addressed. Tier 1 is voluntary, Tier 2 mandatory with legally binding notices and penalties. The clause in the Digital Economy Bill would put us on a more equal footing with other countries whilst being worthy in its own right. Other elements of the Bill, for example, the role of the BBFC may also have future application for efforts to address hate speech and we would urge parliamentarians to support such measures.

3.3. Whilst the law provides a framework for action to combat hate online, industry feedback is that action by judges still does not go far enough in practical terms to deal with the peddlers of cyber hate. Judges should be issuing sentences where relevant, that incorporate orders to ban individuals from holding multiple electronic devices, ensure they are required to forfeit passwords, retain internet browsing history, delete the offending social media accounts etc. This happened in the Joshua Bonehill case but more effective direction from judges and application of relevant judicial orders including Banning and Criminal Behaviour Orders would be welcome.

3.4 At present the police lack sufficient expertise—or resource—in dealing with hate speech and hate crime on the internet. Whilst priority for funding must be determined by the forces themselves, having a single point of contact for the police and the Crown Prosecution Service with the social media companies is essential. This exists in theory but not always in practice. In one case of which we have knowledge, in reference to a tweet that quite obviously broke the law by someone with a previous criminal record, police were told that details of the Twitter account holder could only be passed to them if passed at the same time to the account holder. This approach must be reassessed as the tweet, aside from falling foul of the law will have broken twitter’s own terms of service. Ultimately, the Police need increased digital capacity, better training specifically on criminal thresholds, better tools and partnerships with relevant expert organisations. If successful, the MPS online hate crime hub is one potential model. Internet companies meanwhile must find better working relationships with police and more appropriate modes of action where their terms of service have already been contravened.

---

6 [http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/home-affairs-committee/antisemitism/oral/35121.html#Panel2](http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/home-affairs-committee/antisemitism/oral/35121.html#Panel2)
3.5 For the industry, investment in artificial intelligence systems should be increased. Algorithms must be created that more readily link abusive words, accounts and pictures to filter and readily identify problem users and remove them. In the meantime, dual-factor blocking which seeks out and blocks accounts and tweets which have repeated and multiple use of pejorative words, perhaps as supplied by the Internet Watch Foundation would be a good starting point. Social media acts as much as a search engine as a communication tool and what is left up can be found by children and others. The implications are significant and there is, as we have already established, a requirement for greater consistency of approach and better training of moderators.

3.6 Reporting and tracking is far too cumbersome a process for those reporting abuse. Social Media companies must do better. In addition, there should be greater analysis of whether social media companies are abiding by their own terms of service and delivering consistency in moderation. The European Commission is presently working with NGO’s to monitor the success of the companies in this regard and early indications are that the results will not be good. A ratings system through an independent ombudsman should be considered by Government.

3.7 Where industry firms contract out services, such as the provision of GIFS for direct messages, there should be strict standards in place to ensure racist motifs are not available when searching for simple words like Muslim, Jew or Christian. Only recently – and possibly even now – vile antisemitic GIFs were easily discoverable searching the word Jew.

3.88: Further consideration can and should be given to the wider implications of the results of the Society of Editors Online Moderation Guide. It is clear that tackling online anonymity – which does in some cases serve a purpose – can address abuse online. Being forced to give a real name or to prioritise accredited account holders on discussion boards should be considered more carefully by the digital industry.

3.99: What happens on social media has serious real life consequences. Lack of resource is no defence from social media companies. People co-ordinate attacks and can alert cells to action. The legal ramifications of action taken after co-ordination through social media should be reviewed and companies must be held more accountable. Government should review laws at home and abroad (for example, the 1996 American Communications Decency Act) that confer exemptions from liability upon social media companies for the content they host.

The principle that we as a country screen or otherwise take action against inappropriate content in other media but not on social media needs to change. Social media should not be afforded special status, we must ensure consistency of approach.

About Us:

The Parliamentary Committee Against Antisemitism Foundation (PCAAF) is a registered charity centred on the principle that the struggle against prejudice and discrimination is not just the responsibility of the victims.

The APPG against Antisemitism was established to combat antisemitism and help develop and seek implementation of effective public policy to combat antisemitism. With committed members and an enviable reputation within Westminster and Whitehall for its impartiality, integrity and measured approach, the Group enjoys unique standing.

Prepared by Danny Stone MBE