CarnegieUK TRUST

CHANGING MINDS • CHANGING LIVES

Response to Joint Committee on Human Rights Inquiry on Democracy, Free Speech and Freedom of Association

May 2019

Background

1. Carnegie UK Trust (CUKT) is a not-for-profit organisation focused on improving wellbeing through a range of research, advocacy and community programmes. Since early 2018, it has supported work on new proposals for internet harm reduction instigated by William Perrin (a former UK Civil Servant, who is now a Carnegie UK Trustee) and Professor Lorna Woods (Professor of Internet Law, University of Essex, and an EU national expert on free speech and data regulation). Their work has focussed on the development of a statutory duty of care to reduce reasonably foreseeable harms on social media enforced by a regulator. Full details on the proposals can be found on the CUKT website. [1]

2. Our work has garnered support from, among others, the Commons Science and Technology Select Committee[2], the Lords Communication Committee[3], Chief Medical Officer[4] and the NSPCC[5]. The Government, in its recent Online Harms White Paper[6], has also proposed introducing a statutory duty of care, drawing heavily on our thinking. We are currently considering the detail of the Government proposals and will respond to its consultation shortly.

3. This consultation response does not address all the questions posted in the Committee's consultation terms of reference but focuses primarily on the third: "*what is the role of social media in relation to free speech and threats to MPs? How, if at all, should it be regulated?*" We particularly look at this question through the lens of the initial responses to the Government's White Paper from groups campaigning for to protect free speech.

4. This is an area which is of particular importance to us: we worked with Anna Turley MP on her Private Members Bill 'Malicious Communications (Social Media)'[7], which was motivated by the evident impact abuse from social media was having on women in public life.

1 https://www.carnegieuktrust.org.uk/project/harm-reduction-in-social-media/

2 https://www.parliament.uk/business/committees/committees-a-z/commons-select/science-and-technology-committee/news-parliament-2017/impact-of-social-media-young-people-report-published-17-19/

3 https://www.parliament.uk/business/committees/committees-a-z/lords-select/ communications-committee/inquiries/parliament-2017/the-internet-to-regulate-or-not-toregulate/

4 https://www.gov.uk/government/publications/uk-cmo-commentary-on-screen-time-and-social-media-map-of-reviews

5 https://www.nspcc.org.uk/globalassets/documents/news/taming-the-wild-west-web-regulate-social-networks.pdf

6 https://www.gov.uk/government/consultations/online-harms-white-paper

7 Malicious Communications (Social Media) Bill 2016-2017: bill documents available at: https:// services.parliament.uk/bills/2016-17/maliciouscommunicationssocialmedia.html

5. Indeed, in December 2017 the Committee on Standards in Public Life reported that: '*A significant proportion of candidates at the 2017 general election experienced harassment, abuse and intimidation … The widespread use of social media platforms is the most significant factor driving the behaviour we are seeing. Intimidatory behaviour is already affecting the way in which MPs are relating to their constituents, has put off candidates who want to serve their communities from standing for public offices, and threatens to damage the vibrancy and diversity of our public life.*'[8]

Balancing freedom of expression with protecting against harm

6. The publication of the Online Harms White Paper has prompted many free speech campaigners to argue that the introduction of a statutory duty of care will lead to overly cautious censorship by tech platforms who, in seeking to avoid regulatory sanctions, will remove or prohibit online content that is otherwise legal and permissible offline. We believe that some of the vagueness in the Government's current description of their proposed duty of care allows for such an interpretation: a focus on reactive or censorious content moderation and take-down speeds (which admittedly might be necessary in relation to some serious crimes) does not fully align with the system-level, by-design principles of our work.

The CUKT duty of care proposals

7. Under our proposals, social media service providers should each be seen as responsible for a public space they have created, much as property owners in the physical world. Everything that happens on a social media service is a result of corporate decisions: about the terms of service, the software deployed and the resources put into enforcing the terms of service. As in the physical world, a statutory duty of care focuses on the objective – harm reduction – and leaves the detail of the means to those best placed to come up with solutions in context: the companies who are subject to the duty.

8. Parliament should guide the regulator with a non-exclusive list of harms for it to focus upon. We have that this should include: the stirring up offences including misogyny, harassment, economic harm, emotional harm, harms to national security, to the judicial process and to democracy. The government proposes introducing a new electoral offence of intimidation. While the government's thinking appears to be at an early stage we expect that the 'harms to democracy' elements of our proposal would include the gist of this offence the government has worked it up[9]. There is a long tradition in regulators arbitrating complex social issues and these harms should not be problematic.

9. In our proposals, we argue that the scope of regulation would cover services that:

   - Have a strong two-way or multiway communications component;
   - Display user-generated content publicly or to a large member/user audience or group.

8 Intimidation in Public Life: A Review by the Committee on Standards in Public Life December 2017 Ref: ISBN 978-1-5286-0096-5, Cm 9543, HC 1017273996 2017-18 https://www.gov.uk/ government/publications/intimidation-in-public-life-a-review-by-the-committee-on-standards-inpublic-life

9 "Protecting the Debate: Intimidation, Influence and Information' government response to consultation" Cabinet Office (May 2019) https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/799873/Protecting-the-Debate-Government-Response-2019.05.01.pdf

The regime would cover reasonably foreseeable harm that occurs to people who are users of a service as well as those who are not users of a service.

10. Central to the duty of care is the idea of risk. If a service provider targets or is used by a vulnerable group of users (e.g. children), the risk of harm is greater and service provider should have more safeguard mechanisms in place than a service which is, for example, aimed at adults and has community rules agreed by the users themselves to allow robust or even aggressive communications.

11. Regulation in the UK has traditionally been proportionate, mindful of the size of the company concerned and the risk its activities present. Small, low-risk companies should not face an undue burden from the proposed regulation. Baking in harm reduction to the design of services from the outset reduces uncertainty and minimises costs later in a company's growth.

12. Our work proposes that the appointed regulator (we recommend Ofcom) employ a harm reduction method similar to that used for reducing pollution: agree tests for harm with civil society, users and companies; run the tests; the company responsible for harm invests to reduce the tested level; test again to see if investment has worked and repeat if necessary. If the level of harm does not fall or if a company does not co-operate then the regulator will have sanctions.

13. The nature of fast-moving online services is such that we argue that the regulator adopt a formalised version of the precautionary principle[10] in relation to the question of whether social media is implicated in the existence of harm, acting on emerging evidence rather than waiting years for full scientific certainty about services that have long since stopped.

Protecting freedom of expression

14. We believe that, if properly applied and implemented, our proposals will reduce the harmful impact of threats and abuse directed at MPs, and others, while protecting freedom of expression and the functioning of democracy. As we argue in our most recent paper[11]: "while the impact of the right to freedom of expression is important, it is not the only right that might be relevant"; moreover, "there is no hierarchy between the rights – all are equal … freedom of expression does not have automatic priority." Further, the State may well have obligations to consider the impact of the speech of some on the ability of others to express themselves: there is a concern about the silencing effect of aggressive, violent and abusive communication.

15. We discuss the application of Article 10 (freedom of expression) of the European Convention in Human Rights (ECHR) in our report; we refer the Committee to that section for detailed consideration (pp 14-18). We note that the right protected is broad: it includes the right to hold opinions as well as to express them; the right to be silent as well as to speak; and the right to receive information as well as covering a wide range of topics, including the controversial or offensive.[12]

10 http://www.hse.gov.uk/aboutus/meetings/committees/ilgra/pppa.htm

11 "Online Harm Reduction: a statutory duty of care and a regulator" (April 2019): https://d1ssu070pg2v9i.cloudfront.net/pex/carnegie_uk_trust/2019/04/08091652/Online-harm-reduction-a-statutory-duty-of-care-and-regulator.pdf

12 Smajić v Bosnia and Herzegovina (App no 48657/16), decision 16 January 2018, para 49

16. The notion of a restriction on the exercise of the right is wide and any regulatory regime would have to satisfy the requirements of Article 10(2) ECHR according to a three-stage test:

    - the interference is prescribed by law
    - the interference is aimed at protecting one or more of a series of interests listed in Article 10(2
    - the interference is necessary in a democratic society – this, in essence, is a form of proportionality test but where 'necessary' has been understood to mean 'pressing social need'.[13]

17. Assessing our duty of care proposal in terms of Article 10, we note that it has the potential to constitute an interference with speech. But the right to freedom of expression is not absolute and the State may (in some instances must) take action to protect other interests/rights. The right to private life is an obvious example here, but also security (which might justify actions against terrorism) and the right to life, as well as more generally the prevention of crime (for example, child pornography). Significantly, the State should take action to safeguard pluralism, and to safeguard the rights of all to speak and to ensure that there is no discrimination in individuals' actual entitlement to rights. In balancing rights, the State may have a broader margin of appreciation.

18. It is highly likely that a regulatory scheme for the purposes of ensuring online safety would therefore be in the public interest as understood in relation to Article 10(2). This, of course, does not suggest that rights should be treated cavalierly but simply that the rights framework does not seem to rule out such regulatory action. The question would then be about the proportionality of the measure, as well as its implementation.

19. In terms of the proposal, in focussing on the structure and operation of the platform no particular viewpoint is necessarily targeted by the regulation; indeed, no types of speech are so targeted. It is content neutral.

20. We set out a range of mechanisms whereby a social media platform could satisfy the duty of care that do not involve take down or filtering of the content by the platform[14]. Underlying these mechanisms is the idea that the architecture of a platform influences content and the way that people communicate; central to the duty of care is the idea that providers consider the likely (adverse) impacts of their services – including 'rewarding' abusive and extreme content - before deployment. Furthermore, the proposal envisages that there will be a range of platforms (as there currently are) and that these will have different standards and approaches, so allowing space for different ways of communicating and potentially acting as a safeguard for diversity.

21. The aim of our proposal is to facilitate best practice and for a regulator only to take regulatory enforcement action as a last resort, again emphasising the proportionality of the structure. Further, we emphasise that the regulator must be independent, so that decisions about the acceptability of speech in particular instances are taken out of the political arena.

13 For approach to this question see e.g Cumpn and Mazre v. Romania (App no 33348/96), ECHR 2004-XI, para 90

14 See chapter 9 of our recent paper for more detail

22. Taking a systems approach could be part of a human rights respecting mechanism for tackling online harms, such as abuse and harassment of MPs.

23. We are happy to provide the Committee with further information on our work or to attend a hearing to discuss it in more detail.

Professor Lorna Woods
William Perrin
Maeve Walsh

May 2019