Response to the public consultation on the Online Harms White Paper

BONAVERO REPORT NO. 3/2019

1 JULY 2019

Stefan Theil (Oxford), Oliver Butler (Oxford), Kate Jones (Oxford), Harriet Moynihan (Chatham House), Catherine O'Regan (Oxford), Jacob Rowbottom (Oxford)

EXECUTIVE SUMMARY

Our overarching recommendation is that any regulatory approach to online harms should be expressly founded on human rights law. Human rights law provides both a suitable normative framework as well as crucial guidelines to assist regulatory decision-making, especially in balancing competing rights and interests in the online sphere.

While the White Paper frames its approach as involving a duty of care, we believe that this terminology may be misleading: the codes of practice and the penalties for transgressions are better understood as conventional instruments of statutory regulation. We see a risk of knowledge asymmetry between the regulator and companies, especially with respect to determining 'reasonable steps' in the context of addressing online harms. This asymmetry may render it difficult to challenge a company's assertion of technical capabilities and feasibility, even where such claims appear dubious (see II below). These concerns may be partially alleviated through enhanced transparency (see Question 1).

We have rule of law concerns due to the broad scope of platforms and services covered. The regulator must be able to provide meaningful oversight and companies require clarity on what enforcement measures they can expect. Comparable legislation, like the Network Enforcement Law in Germany, is limited to larger companies and focused on a narrower set of platforms and services. If the broad scope outlined in the White Paper is retained, we suggest the regulator considers exempting certain companies partially or entirely from regulation (see Question 5).

Legislation should consider a two-tier approach to regulation which differentiates between: (a) harms with a strong evidence basis and a reasonably clear definition ('definite harms') and (b) harms with a weaker evidence basis or with a less clear and context-specific definition ('contextual harms'). While a prescriptive regulatory approach to definite harms seems appropriate, legislation may provide a more flexible oversight model for contextual harms. This would permit a degree of variation as to the standards applied by companies, increasing the choices available to users (see Question 8).

The regulator should ensure that the internal complaints management systems of companies are consistent, effective and efficient, as well as consider requiring them to provide for an external oversight board (see Question 3). It would be inappropriate for the Home Secretary to sign off on codes of practice without parliamentary oversight and a full public consultation by the regulator. The legislation should contain an express obligation on the regulator to protect and promote human rights in all its decisions and the codes of practice (see Question 4). On balance, we believe that a new regulator for online harms should be created (see Question 10).

Whether companies and individuals should be permitted to challenge decisions before tribunals depends on what decisions the regulator is empowered to make (see Question 14) and super-complaints may be appropriate given certain safeguards (see Question 2).

A pyramid of escalating enforcement powers seems appropriate but must be exercised in a manner compatible with human rights. The regulator should carefully consider the implications of any sanctions, especially a decision to block a website: evasion mechanisms are often available to users and there is considerable collateral harm when websites with overwhelmingly legitimate content are targeted. It is worth noting that the blocking of websites in other jurisdictions has led to successful human rights challenges (see Question 12).

RESPONSE TO THE PUBLIC CONSULTATION ON THE ONLINE HARMS WHITE PAPER

I. INTRODUCTION

The internet has revolutionised our ability to communicate and connect across historic social, political and geographic divides. This development bears great opportunities for the democratisation of free expression and the diversification of public discourse but has likewise broadened the potential for and impact of harm.

We welcome the vision of a free, open and secure internet outlined in the White Paper as well as the general objective of regulating online harms. However, care must be taken to ensure that such regulation is compatible with the rule of law and incorporates strong human rights protections. Companies require clarity as to what standards they must comply with and what oversight measures they can reasonably expect. This may entail partial or complete exemptions of some companies from regulation and oversight.

Regulations should be based on human rights law, particularly the Human Rights Act 1998, the UN Guiding Principles on Business and Human Rights, and international treaties to which the United Kingdom is a party. Human rights law provides a crucial framework and guideline for regulation, especially in balancing the myriad of competing rights and interests in the online sphere. Great care should be taken to ensure that codes of practice and regulatory interventions are compatible with human rights safeguards.

While the White Paper focuses its attention on online harms, it could additionally promote desirable social goods. We offer only a few tentative suggestions to illustrate how regulation might pursue other goals in addition to tackling online harms. For instance, as with traditional media in the context of elections, regulation might seek to offer all candidates and parties a fair chance to promote their message to a wider audience. This could include prioritising the messages of political parties and giving candidates an opportunity to prominently rebuke and correct false and misleading claims.

The global and cross-jurisdictional nature of the internet naturally lends itself poorly to a patchwork of national regulations. Therefore, achieving a greater degree of collaborative action, founded on a shared commitment to international human rights norms, ought to be a high Government priority. We urge the Government to work with its international partners to develop a cohesive approach to regulation of online harms. A global approach to online harms would bring about improvements for all stakeholders: regulators could more easily set and enforce minimum standards, online companies would benefit from increased legal certainty across the jurisdictions in which they operate and most importantly, users could better understand the freedoms and limitations for sharing content online.

II. DUTY OF CARE

Making powerful digital platforms more responsible and accountable to users in the way in which they host online content is important, especially given the asymmetric relationship they

have with their users, and the current lack of transparency around data collection on users and the use of algorithms to control what users see. However, we believe that framing the issue through a 'duty of care' requires clarification as the terminology may be misleading.

The envisioned framework would not grant any individual an action or remedy in the event that a duty has been breached: this would be the conventional understanding of a duty of care in negligence law. Instead, the White Paper proposes codes of practice, which companies are expected to follow and any penalties for transgressions would be imposed by the regulator. Such an approach appears more closely related to a conventional model of statutory regulation.

Regardless of the terminology employed, it is unclear how the regulator will identify a breach of a duty of care. To illustrate this point, suppose that a provision requires the regulator to determine whether a company has taken 'reasonable steps' (p.68) to prevent the dissemination of terrorist content. A company at the cutting edge of technology has significant expertise and knowledge on which measures it could employ and develop in the future. The asymmetry of knowledge will make it difficult for the regulator to challenge assertions of technical capabilities and feasibility. A company's inability to bring about a desired regulatory outcome may simply reflect a lack of interest in developing the necessary technology. For instance, the micro-blogging and social network site Tumblr was – contrary to earlier assertions – quickly able to develop mechanisms that blocked adult content from its site when the business model was threatened with removal from a prominent app store. The transparency suggestions we outline in our response to Question 1 would alleviate some of these concerns. The regulator could also consider issuing guidance on how it proposes that each of the codes of conduct will be applied in practice, as Ofcom has done in relation to some of its codes of practice.

On the other hand, the proliferation of certain harmful content on a website may in and of itself not be evidence of a failure to properly address the problem. A company may indeed be going well beyond what can be expected and still run afoul of an exacting definition of 'reasonable steps'. This is not to suggest that a flexible contextual or an exacting approach to reasonableness is inherently preferable. Rather, we raise the example only to highlight that a balance needs to be struck between the capacity for flexible responses in emerging situations, and fairness to companies trying to ascertain by what standards they must comply. The regulator may therefore wish to vary the regulatory standard according to the type of content and harm being addressed under a two-tier approach, which we suggest in our response to Question 8.

III. RESPONSES

Question 1: This government has committed to annual transparency reporting. Beyond the measures set out in this White Paper, should the government do more to build a culture of transparency, trust and accountability across industry and, if so, what?

We welcome the Government's focus on transparency and the initiatives proposed in the White Paper. In addition, the Government should consider two areas: (a) transparency in data collection and use, and (b) real-time transparency towards users, embedded in technological design. There is a widespread lack of awareness that social media companies are essentially advertising companies, and that their curation of content is designed to maximise advertising revenue. For instance, the Government should promote:

- Transparency in the curation of personal data: not only through the procedures dictated by current data protection law but, for example, technological design such that at the touch of a button, users can see all the information held on them and everything deduced from that information.
- Transparency in the uses of personal data: especially to whom it has been sold, for what purpose and for which sum. This could be the first part of a cultural shift towards sharing the commercial value of data with the users from whom the data originates.
- Transparency with the use of third-party cookies, particularly moving away from current 'nominal' consent to cookies towards a real understanding of what cookies entail and how they operate, thereby permitting a genuine choice along with specific opt-outs.
- Transparency on algorithms in real time for users. For example, enabling users to see on what basis content has been recommended to them and what inferences have been drawn about their identity. This could be the first step towards addressing discrimination in algorithms.

Question 2: Should designated bodies be able to bring 'super complaints' to the regulator in specific and clearly evidenced circumstances?

Question 2a: If your answer to question 2 is 'yes', in what circumstances should this happen?

We recommend that if a super complaints procedure is introduced the regulator should be given significant flexibility in the implementation and conditions of access. The role of the regulator should be focused on systemic compliance with codes of conduct, as opposed to resolving individual disputes. Therefore, the purpose of any super-complaint process should likewise be limited to raising awareness of systemic problems that the regulator is empowered to consider. Consistent with our view that human rights law should provide the normative underpinning of any regulation, we recommend giving the power to bring super complaints to the Equality and Human Rights Commission and similar broad-based human rights bodies.

Question 3: What, if any, other measures should the government consider for users who wish to raise concerns about specific pieces of harmful content or activity, and/or breaches of the duty of care?

Where companies have internal complaints procedures, the regulator should ensure that they are consistent, fair, effective and efficient at all stages of the process. The regulatory body should ensure that companies' internal complaints management systems are accessible for the individual user. Making a complaint should be easy, including for those individuals with various forms of disadvantage, and should permit the specification of the complaint with sufficient precision and detail. Regulation should ensure that reports can be issued without the need to subscribe and login to the services, which is particularly important for non-accountholders.

Once a complaint has been received, the company should process and evaluate it in a consistent, fair, and timely manner. This includes having clear standards of assessment and providing adequate staffing and training, as well as ensuring that content reviewers have the requisite cultural and language skills. The decision-making process should be overseen by management and have an internal appeal procedure, subject to clear deadlines for a final decision.

Given the current complaints management infrastructure (or lack thereof) of many online companies, there are legitimate concerns whether they are equipped to reach consistent and fair decisions in applying future codes of practice. The regulator may therefore consider requiring that larger companies set up or provide users with access to an external and independent oversight board to assist in promoting best practices. Such a board may adopt any or all of the following roles: it could review individual decisions on complaints (appeal model), it may monitor outcomes and address systemic concerns through a general audit (audit model), or review and oversee the development of the companies' terms and conditions and associated policies and guidelines based on human rights norms (normative review model). Such an external oversight board could help to ensure that companies' policies and activities comply with human rights law, particularly if the board can effect changes to company policy and terms and conditions.

Question 4: What role should Parliament play in scrutinising the work of the regulator, including the development of codes of practice?

It would be inappropriate for the Home Secretary to sign off on the codes of practice without parliamentary oversight and public consultation. There should be clear statutory requirements and a full public consultation process for any codes of practice drafted by the regulator. In particular, legislation should be drafted in more specific terms than the current reference to 'harms' in the White Paper suggests: legislation should identify specific categories of harm that have a sufficient evidence base and prescribe clear and proportionate regulatory sanctions.

The regulation of all harms requires a delicate balancing of the rights of individuals to express their views freely with the legitimate interests in tackling various online harms. The legislation should therefore contain an express obligation on the regulator to protect and promote human rights in all its decisions and specifically the drafting of the codes of practice.

Parliament may also wish to provide a two-tier approach to regulation as outlined in the response to Question 8 below. The two-tier approach avoids a broad brush and instead focuses the most prescriptive rules on the most serious harms for which there is a clear evidence base.

Question 5: Are proposals for the online platforms and services in scope of the regulatory framework a suitable basis for an effective and proportionate approach?

As it stands, the White Paper covers a wide range of platforms and services that all to a greater or lesser extent host user generated content and enable 'interaction with others.' We have concerns with this extremely broad scope and would encourage careful consideration of the effects of legislation passed in other jurisdictions.

From a rule of law perspective, companies require clarity on the codes of practice they must comply with and what oversight and enforcement measures they can reasonably expect. The regulator will need to take steps to ensure that it does not resort either to excessively vague standards or to highly selective oversight and enforcement actions. Great care needs to be taken in the calibration of regulatory requirements to the size and structure of companies, so as not to impose disproportionate requirements that push smaller enterprises out of the market, or unduly burden those for whom user-generated content is not core to their primary business model. Moreover, companies need to understand how the regulator intends to effect that

calibration. As outlined in our response to Question 4 there must be clear statutory requirement and a public consultation process for any codes of practice drafted by the regulator.

The broad scope of companies covered also raises concerns of principle, namely whether any regulator can cope with the vast number of platforms and services appropriately. If a regulator is practically only capable of providing oversight and enforcement for a few high-profile companies then that undermines its broader mission and legitimacy. The danger is that medium and lower profile companies, while technically covered by the statutory regime, might in practice not be regularly monitored and only infrequently subject to enforcement actions.

For good reasons, the Network Enforcement Law in Germany therefore limits regulation to larger companies and focuses on a narrow set of platforms and services. It would be advisable to study similar legislation from other democratic jurisdictions closely in drafting the legislation. Should the broad scope of platforms and services outlined in the White Paper be retained, the regulator might consider exempting certain companies partially or entirely from regulation, as well as introducing a two-tier approach to regulation as outlined in the response to Question 8.

Question 6: In developing a definition for private communications, what criteria should be considered?

The White Paper proposes that requirements to scan or monitor content should not apply to private channels. It correctly identifies the difficulty of identifying an appropriate and non-arbitrary cut off point between private one-to-one messaging and a 'whatsapp group of several hundred.' It is not, however, clear that the number of recipients is a suitable criterion for determining whether a communication is public or private. Instead, a better criterion for defining private communications is the degree of control enjoyed by the sender over the identity of the recipients of the communication. The effect is to draw private communications relatively broadly, excluding 'public' posts and potentially some 'newsfeed' posts.

If an individual sends hundreds of letters to members of a private association or group, the volume of private letters does not make these communications public. Alternatively, an announcement posed in a public place (i.e. to any individuals who in fact read the message) is not private merely because the number of members of the public who do read the notice is small. A similar point can be made even with regard to personal profile or 'timeline' posts, as users typically have a range of privacy settings that they may adopt, both for the profile itself and individual posts within it, and control over the number and status of connections (public, friend, acquaintance). Posts may be accessible to large numbers of individuals but only where the user chooses to permit access to their timeline and to particular posts within it. The metaphor of private spaces is not inappropriate given the levels of user control, even if some users in effect treat their profile page as an 'open house' and allow many 'friends' to freely enter that space. Different considerations might apply where that content is taken by a platform and displayed in a 'newsfeed' to others outside the control of the user who made the post.

Question 7: Which channels or forums that can be considered private should be in scope of the regulatory framework?

Question 7a: What specific requirements might be appropriate to apply to private channels and forums in order to tackle online harms?

Any regulation of the content of private communications requires great caution, both for reasons of privacy and freedom of expression, and should only address definite harms for which there is a strong evidence base and clear definition as outlined in our response to Question 8 below.

Question 8: What further steps could be taken to ensure the regulator will act in a targeted and proportionate manner?

The idea of extending regulation of online platforms and services beyond content which is prohibited by the criminal law is not wrong in principle, but great care must be taken not violate human rights as guaranteed by the Human Rights Act 1998, the European Convention on Human Rights and other international treaties. Legislation should consider a two-tier approach to regulation which differentiates between: (a) harms with a strong evidence basis and a reasonably clear definition ('definite harms') and (b) harms with a weaker evidence basis or with a less clear and contingent definition ('contextual harms'), for instance harms relating to online screen time and disinformation, respectively.

While a prescriptive regulatory approach to definite harms seems appropriate, legislation may provide a more flexible oversight model for contextual harms because their harmfulness is at times contested and context specific. Crucially, there are often a multitude of plausible interventions that might address contextual harms. The regulator may therefore wish to focus its attention on the consistent enforcement of existing company terms and conditions in such cases, as well as promoting transparency as outlined in our response to Question 1. These terms and conditions often already address many of the contextual harms flagged in the White Paper: the problem often lies with consistent, fair, effective and efficient enforcement as outlined in our response to Question 3.

As part of a flexible oversight model for contextual harms, the regulator could require companies to publish their community standards and associated policies, along with the details of their complaints management procedures and in-depth reports on outcomes. The regulator would then evaluate whether the complaints management system complies with human rights standards. This process may include scrutiny of the company's relationship with and performance of an external oversight board as suggested in our response to Question 3. Overall, a flexible oversight model for contextual harms would permit a degree of variation of standards depending on the online service and platform, increasing the choices available to the users.

The legislation should include an express obligation for the regulator to protect and promote human rights in all its decisions and specifically the drafting of the codes of practice as outlined in our response to Question 4. The regulator should expect that the codes of practice may face human rights challenges in the courts from the affected companies and individual users.

Question 10: Should an online harms regulator be: (i) a new public body, or (ii) an existing public body?

Question 10a: If your answer to question 10 is (ii), which body or bodies should it be?

There are advantages and disadvantages to either creating a new regulator or choosing an existing regulator. However, on balance it is our view that there is a good case for establishing a new regulatory body. Ofcom and other regulators were created with specific tasks in mind and may therefore not be institutionally suitable for the functions outlined in the White Paper: regulating online harms is significantly different from telecommunications and broadcast. There would also be risks to freedom of expression if a single body, Ofcom, were to determine content standards both for broadcast media and social media. A new regulator would develop expertise in the specific area of online harms and adopt procedures and practices tailored to online harms.

Whichever option is followed, the regulator will have to engage with other bodies that have some responsibility for different aspects of online activity (such as the ICO, IWF). It will be important to ensure that users and companies have clarity about where to take complaints and other concerns. The regulator for online harms will therefore have some role in coordinating policies and procedures among the different regulators.

Question 12: Should the regulator be empowered to i) disrupt business activities, or ii) undertake ISP blocking, or iii) implement a regime for senior management liability? What, if any, further powers should be available to the regulator?

We consider a pyramid of escalating enforcement powers appropriate, but these powers must be exercised in a manner that is proportionate to the transgression of the codes of practice and overall compatible with human rights. The powers may range from publication of reports and compliance rankings, to administrative fines and on towards more serious disruptions of the business activities for persistently non-compliant companies. Powers to block access to websites entirely are already part of some enforcement regimes (see for instance the age verification rules for pornographic websites), but should be reserved for extreme cases and only as a last resort.

Even in extreme cases, the regulator should take account of the implications of its decision to block a website. There are many evasion mechanisms (for instance, virtual private networks, mirrors of banned websites, darknet websites) that are accessible for users with basic technological competence. Furthermore, any blocking of websites, especially those from popular and large service providers will cause inadvertent harm, especially to users engaged in entirely legitimate use of the blocked site. Given the potential for unintended consequences, it is difficult to see how blocking large and popular websites whose content is overwhelmingly legitimate could ever be a proportionate response to infractions of codes of practice. Public outrage may be considerable and lead to the proliferation of the above-mentioned evasion mechanisms. It is also worth noting that the blocking of websites in other jurisdictions has led to successful human rights challenges, for instance in the context of Turkey blocking access to a popular video sharing website.

Question 13: Should the regulator have the power to require a company based outside the UK and EEA to appoint a nominated representative in the UK or EEA in certain circumstances?

Yes, this would be a useful mechanism to ensure that companies are accountable for their service provision in the United Kingdom. As outlined in the introduction, we urge the

Government to work with its international partners to develop a cohesive approach to regulation of online harms.

Question 14: In addition to judicial review should there be a statutory mechanism for companies to appeal against a decision of the regulator, as exists in relation to Ofcom under sections 192-196 of the Communications Act 2003?

Question 14a: If your answer to question 14 is 'yes', in what circumstances should companies be able to use this statutory mechanism?

Question 14b: If your answer to question 14 is 'yes', should the appeal be decided on the basis of the principles that would be applied on an application for judicial review or on the merits of the case?

The answer to this question depends on what decisions the regulator is empowered to make. Judicial review may be adequate for infrequent and general challenges, while a specialised tribunal could deal more efficiently with frequent and specific challenges that benefit from institutional expertise and streamlined processes. The regulator could also be required to set up an internal appeal process in relation to certain decisions, in order to correct its errors before the matter is raised with courts and tribunals. Such a right of appeal will be particularly important where formal sanctions are imposed on a social media company. The question refers to a statutory mechanism for companies to make an appeal. Depending on the decision at stake, a right of appeal could also be extended to other parties, such as those bringing complaints to the regulator.