Online Harms White Paper: Consultation Response [BILETA Response to the UK Government Consultation 'Online Harms White Paper']

Harbinja, E.; Leiser, M.R.; Barker, K.; Mangan, D.; Romero Moreno, F.; Dushi, D.

Online Harms White Paper: Consultation Response

Prepared on behalf of the British Irish Law Education and Technology Association (BILETA) by: Dr Edina Harbinja, Dr M.R. Leiser, Dr Kimberley Barker, Dr David Mangan, Dr Felipe Romero-Moreno, and Dr Desara Dushi

By email to onlineharmsconsultation@culture.gov.uk

Introduction

The British Irish Law Education and Technology Association (BILETA) has concerns about the broad scope of the proposals in the White Paper and how the proposals will be applied to platforms. The White Paper proposes co-regulation by a new regulator called OfWeb. Previous attempts to regulate broadcast and press (Ofcom and IPSO) might provide insights on what its scope and application might look like, but there are different principles, issues, and regulatory designs needed for platforms. If establishing a new regulator proves necessary (and we are sceptical in this regard), the key requirement is its independence. The White Paper proposes that OfWeb will be granted a delegated power to define an online harm any way it wants[1]**.** This is not only ripe for abuse, it does not meet commonly accepted 'quality of law' and 'reasonable foreseeability' standards. Furthermore, it could also subject a regulator to the whims of political and industry influence. This is potentially undemocratic and does not meet rule of law standards required in a democratic society. Despite parity between the offline and online world listed as a specific objective, the scope of powers goes far beyond parity to what is permitted by UK substantive law in the offline world. It regulates users and tech companies through the imposition of a "duty of care" applicable to content that is not necessarily *unlawful*, but regarded as *harmful*.

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?

There are a few elements to unpack in this question: a common understanding of terms stated as Government's objectives; the regulatory framework set out by the European Union; and the challenges therein. As well, this question presumes a standard notion of key terms such as transparency, trust, accountability, and 'online harm'. The capacity for public understanding is necessary to building a culture of transparency, trust and accountability. The potential threat is a 'transparency fallacy' - a point found in the General Data Protection Regulation (GDPR).

Given the prominence of European Union in regulating data protection, the EU framework could be seen as setting out a minimum standard. And so, any steps toward transparency may exceed the GDPR's framework. However, falling below these measures (or appearing to do so) will bring into question the commitment to transparency, trust and accountability objectives. The GDPR's principles are overarching and this topic is a matter of developing understanding. Queries stemming from the framework of the GDPR (such as the separation between ancillary function and regular systematic processing) may need to be addressed in any regulatory outcome.

Any measures undertaken by platforms that restrict free expression/privacy/data protection should be independently verified through both a human rights audit and an impact assessment. The audit should examine compliance top-down and horizontally. These documents should be transparent, filed with the regulator, and routinely assessed for compliance. Platforms should undertake child protection impact assessments and the regulator should encourage human rights and ethical/societal impact assessments for new and innovative products and services.

[1] See Section 2.2 of the White Paper: "This list is, by design, neither exhaustive nor fixed. A static list could prevent swift regulatory action to address new forms of online harm, new technologies, content and new online activities."

Advertising transparency and effective data protection remedies are only one part of the solution. The fact actors were able to manipulate users for the sole purpose of commercial gain, points to a vacuum of proper regulation among not only platforms, but political advertisers and data brokers. For example, targeted political campaigns have only deepened the debate on how to attach accountability mechanisms to actors engaged in high-risk political advertising. Regulating these actors should be the starting point before any co-regulatory instruments.

Unlike the Article 35 GDPR requirement that data controllers undertake DPIAs when processing personal data poses a high risk to the rights and freedoms of others, Ethical and Societal Impact Assessments (ESIAs) are rarely backed by legislative mandates. Most are undertaken by companies motivated by corporate social responsibility; however, there is a regulatory case for public and transparent ESIAs. A platform could benefit via social credit for undertaking. Mandatory ESIAs can be used to identify, understand, and mitigate the impact and effects of platforms on users at the planning stage, facilitates better decision-making at implementation stages, helps to avoid costly subsequent improvements, and, in some cases, mitigate liability.

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

Only if those designated bodies adhere to transparency requirements themselves, in particular, regarding their funding. Provisions for alternative views on the complaint should be permitted and encouraged. For example, European and national regulators exist to protect the fundamental right to data protection, yet there is no equivalent protection for free expression or digital rights. Complaints from business and enterprise should be adequately investigated, but views should also be sought from consumer protection organizations and civil society. The Regulator should also encourage views from users or help to establish user collectives to allow their opinions to be given standing.

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

For example, the grounds for 'super complaints' could include breaches and large-scale misconduct by platforms that affect users' human rights. However, we express concern about the consequences for users' Article 10 rights to free expression. There are no national authorities responsible for promoting free expression. Article 10 rights could be adversely affected by under-representation when investigating super-complaints.

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?

There should be no duty of care without guidance from the court. It should be courts that inform the duty, not platforms informing the courts. A statutory duty of care offers many opportunities to undercut or imperil the Government's stated aim of transparency, trust and accountability. The UK Supreme Court found it necessary in *Robinson v Chief Constable of West Yorkshire*[2] to reiterate the duty of care analysis from *Caparo Industries plc v. Dickman*[3], reminding the legal community that the test from this decision involved a response to an initial question of whether the duty of care under consideration was already established or established through an analogous duty of care that the courts had already considered. Due to the difficulty in drawing an analogy with current examples of duty of care (e.g. owners of physical property, where the potential harm is specific, well defined and owed to identifiable individuals), the duty of care as envisaged in the White Paper would struggle to satisfy this test. There

[2] *Robinson v Chief Constable of West Yorkshire* [2018] UKSC 4.
[3] *Caparo Industries plc v. Dickman* [1990] 2 A.C. 605.

is also a question as to how this duty of care would interact with other legislation, such as the Protection from Harassment Act 1997.

Furthermore, this does not address whether the duty of care is the right approach to addressing some of the government's concerns. Users are unsatisfied with police responses to genuine threats and online harassment. Several high-profile incidents of intimidation reveal the incompetence of Police forces to deal with misogynistic language and threats of violence against women and vulnerable users. Placing a duty of care on platforms to stop bad behaviour is not the same as ensuring victims have access to justice.[4] Trolls can simply open another account to continue sustained targeting of their victims. Mandating a duty of care to prevent online harms privatizes the public obligations society places on police to protect women, children, the vulnerable, and other members of society. The White Paper does not provide access to any methods to resolve disputes, nor does carefully balance rights to ensure oversight and accountability. We need more techno-legal solutions given the digital ecosystem. The duty of care requires speedier and cheaper forums of resolution and more mechanisms for law enforcement to tackle those who would use the platform to threaten and abuse. With the appropriate safeguards, police forces should be encouraged and empowered to take more action against platforms to locate and identify the users behind these accounts. More reporting mechanisms and graduated blocking schemes should be encouraged.

Government has rightly identified the threat to democracy that unregulated platforms pose. However, the plan puts the responsibility on platforms is ill-informed. Proper advertising and electoral regulation, as well as licensing plans to during electoral events, will do more to control nefarious actors than a duty of care. BILETA strongly urges stronger direct regulation over these actors before imposing a duty of care on platforms.

Under the White Paper's proposals, commercial entities seeking to build up goodwill with consumers through platforms interfaces may incur unnecessary risks. The White Paper proposals will hinder innovation and amalgamate employment protections with platform regulation. Any business with an interactive online presence must also be a safe workplace. The implications beyond platforms and any knock-on effects of additional regulation must be considered and scrutinized.

Question 4: What role should Parliament play in scrutinizing the role of the regulator, including the developments of codes of practice?

The European Data Protection Supervisor has noted that the Member States, the European Data Protection Board and the European Commission must support the drawing up of codes of practice considering the specific needs of micro, small and medium-sized enterprises[5]. Similarly, the Article 29 Working Party further observes that compliance with these codes also helps build transparency[6]. However, Page 7 of the White Paper states that "the government will have the power to direct the regulator regarding such codes".

This is extremely concerning.

The regulator should be completely independent of any Parliamentary and political influence, including when developing Codes of Practice. The Independent Reviewer of Terrorism Legislation is an illustrative example: in *Big Brother* the ECtHR stressed that the uniqueness of the Reviewer's role lied in "its complete independence from government"[7]. The UN Special Rapporteur's Report on Freedom of Expression stated that any law limiting the right to freedom of expression must be applied by a body

[4] Article 6. ECHR; Charter of Fundamental Rights of the EU (2009), Articles 41, 47, 48, 50, Available at http://www.europarl.europa.eu/charter/pdf/text_en.pdf.

[5] European Data Protection Supervisor (Opinion 3/2018) Opinion on Online Manipulation and Personal Data at Page 20.

[6] Article 29 Working Group Guidelines on Consent under Regulation 2016/679 at pg 19.

[7] *Big Brother Watch and Others v United Kingdom (*Applications nos. 58170/13, 62322/14 and 24960/15*)* at 160.

that is independent of any political power in a way, which is not arbitrary, including the possibility of remedy and challenge[8]. Moreover, case law from the ECtHR states that, in a field where abuse was highly likely, it was also in principle desirable to entrust supervisory oversight to a judge[9]**.** Parliamentary scrutiny should be limited to ensuring the regulator fulfils its duties appropriately.

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?

Following case law from the ECtHR and CJEU, if exceptional circumstances justified the adoption of technical measures such as the proposed disruption of search results, app stores, or links on social media, to satisfy the essential principle of proportionality, such measures must be specifically targeted, so that users were always able to legally access information.[10]

In view of ECtHR and CJEU case-law, the White Paper's proposal that ISPs should block non-compliant sites or apps after notification by the regulator is potentially disproportionate[11]. The paper's proposal requiring the regulator to issue a list of actors that have committed serious, repeated infringements, for ISPs to voluntarily block does not appear to be 'in accordance with the law'. Specifically, under Articles 8 and 10 of the Convention, any legislation should comply with the accessibility, foreseeability and rule of law principles[12]**.** Lastly, in terms of the White Paper's proposal to introduce both, civil and criminal senior management liability for non-compliant companies, this suggestion seems to be in line with Section 174 of the Companies Act 2006, as well as relevant case law[13].

In terms of proportionality, there is a serious concern about the effects the new regulation may have on human rights and freedom of expression. The key issue is the vague and undefined nature of many harms that the government proposes to regulate. While some of these harms are already regulated and clearly illegal (e.g. terrorist related content, content related to child abuse, extreme pornography etc.), a lot of the harms that the White Paper identifies as 'legal harms' (e.g. disinformation, trolling or intimidation) could potentially be within the remit of the protection awarded by Article 10 of the European Convention of Human Rights (the right to freedom of expression). Offensive content may well be harmful but not rise to the threshold of illegality and may even be protected speech. We do not advocate for this to change. The right to offend is *part* of free expression. The vague nature of *harms* as a group that are not *per se* illegal could be challenged under principles of the rule of law, proportionality and legal certainty.[14]

The second issue is the concept of 'duty of care', which Internet companies would owe to their users. Duty of care as proposed by the Government contradicts the established legal principles of the law of

---

[8] Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, (16 May 2011) at Page 8.

[9] *Klass and others v Germany (*Application no. 5029/71*)* at Para 56 *Big Brother Watch and Others v United Kingdom (*Applications nos. 58170/13, 62322/14 and 24960/15*)* at Para 58.

[10] See Page 60 of the White Paper.

[11] This was confirmed in the ECtHR *Yildirim v Turkey* pg 29, as well as the CJEU *UPC Telekabel* [2013] [56].

[12] ECtHR *The Sunday Times v the United Kingdom* (No. 1), 6538/74, 26 April 1979, at Para 46. On the conditions of accessibility and foreseeability, see ECtHR *Kurić and Others v Slovenia*, (26828/06), 26 June 2012, at Para 341; ECtHr *Amann v. Switzerland* (27798/95) 16 February 2000 at Para 50; ECtHR *Slivenko v Latvia* (48321/99) 9 October 2003, at Para 100. The CJEU considers that the principles of legal certainty and legitimate expectations imply that "the effect of Community legislation must be clear and expectable to those who are subject to it": ECJ, 212 to 217/80, *Amministrazione delle finanze dello Stato v SRL Meridionale Industria Salumi and Others*, 12 November 1981 at Para 10; or "that legislation be clear and precise and that its application be foreseeable for all interested parties": CJEU, C-585/13, *Europäisch-Iranische Handelsbank AG v. Council of the European Union*, 5 March 2015, § 93; cf. ECJ, C325/91, *France v Commission*, 16 June 1993, § 26.

[13] *Donoghue v Stevenson*, *Norman v Theodore*, *Re D'Jan* and *Greson v HAE*

[14] UN, Declaration of the High-level Meeting of the General Assembly on the Rule of Law at the National and International Levels (2012), para 8 http://www.unrol.org/article.aspx?article_id=192;; EU, Charter of Fundamental Rights of the EU (2009), Article 49 (concerning the principles of legality and proportionality of criminal offences and penalties); European Convention on Human Rights, in particular 6(1), 7, 8(2), 9(2), 10(2) and 11(2).

negligence and liability. It broadens their scope to a wide range of potentially non-identifiable users, which would not necessarily suffer a physical or psychiatric injury as law normally requires. Rather the focus is on vague 'societal harms', which are outside the scope of the duty of care as normally conceived.

The third issue is the confusing concept of 'intermediary liability', established in 2000 in the EU E-commerce Directive and relevant CJEU case law. The Directive provides a safe harbour for internet 'hosts' (most of the companies the Government aims to regulate would fit into this category) and the protection from liability for illegal content stored on their platforms, provided that they do not have the actual knowledge about this content, and that they act expeditiously upon obtaining this knowledge. Importantly, the Directive prohibits the general monitoring of Internet users for the purpose of detecting such content. There is extensive CJEU case law on the matter[15] as well as the related ECtHR jurisprudence on Articles 8 and 10 of the ECHR and the liability of Internet platform**s.** Interestingly, the Government claims that the new regime would be compatible with the Directive[16]. We fully support compliance with the e-Commerce Directive. However, it would be difficult to imagine the duty of care without the general monitoring obligation that enables compliance with this duty. Therefore, we therefore strongly advise retaining the liability regime as it is, and especially, refrain from imposing any form of a general monitoring obligation. As explained in the next section, this would adversely impact users' privacy rights and their freedom of expression.

Effects

In application, users will have restricted ability to express themselves freely, will exercise less personal autonomy, media pluralism could be reduced, violating obligations under the EU Charter of Fundamental Rights and legitimate, online periodicals could be flagged for review and removed for not complying with the "status quo".[17] Alternative media and subculture content could trigger unjustified user complaints. Platforms are tasked with having to make case-by-case assessments of the quality of user-generated-content on the basis of opaque and criteria determined by a national regulator. For example, "trolling", one of the initial forms of content and behaviour in scope, is almost impossible to define and can be seen by some as a legitimate form of deliberative communication, especially when directed to politicians and celebrities. Terms like "violent content" are too opaque and could even capture a standard film-preview posted on Facebook. The White Paper's provision that platforms provide a reporting function is also problematic and could create unrealistic expectations among users about what speech should stay and what should be taken down or removed. Speech should never be judged on its subjective effects on a user. However, none of our laws regulating speech are applied via a subjective test. This also contravenes the longstanding principle from *Handyside v UK*[18]:

> "*Freedom of expression…is applicable not only to 'information' or 'ideas' that are favourably received or regarded as inoffensive or as a matter of indifference, but also to those that offend, shock or disturb the State or any sector of the population".*[19]

Any speech assessment will need to include qualitative questions whether content online should be treated differently to information offline for every individual user; the platform will need to understand the context of exchanges between every user on a platform and how people communicate offline with one another. The 'relevance' criterion of content moderation is particularly tricky, as it requires platforms to decide whether the content assailed *may* also harm at some point in the future. Platforms should not be placed in a position where they are forced to judge content on the effects that it may have on *some* users at *some* point.

---

[15] For example, *Scarlet Extended SA v Societe Belge des Auteurs, Compositeurs et Editeurs SCRL (SABAM)* (C-70/10) [2011] E.C.R. I-11959 (24 November 2011).

[16] For example, *Tamiz v the United Kingdom* (Application no. 3877/14) [2017] ECHR (12 October 2017).

[17] Article 11(2) EUChFR.

[18] ECtHR (1976) *Handyside v UK* (5493/72).

[19] At Para 49.

Discretionary regulatory designs for speech and expression are highly problematic. Due to the volume of user-generated content, it would be impossible to comply without deploying automation and technical measures. No filtering service is perfect; they will not catch all undesirable content and run the risk of over-blocking. When filtering is used under the threat of sanctions, platforms will err on the side of caution and could decide to block against content that could conceivably be harmful to users but also is in the 'public interest' to display; for example, real crime footage in order to generate leads on suspects or unpopular opinions. The decision to empower a new regulator to impose financial penalties on platforms confuses boundaries between public and private and extends the responsibilities of the latter to court-style decision-making to platforms who will be forced to apply an imprecise standard for evaluating whether user-generated-content is harmful to not only other users, but society-at-large. The broadening out of 'harms' cannot equate to the scaling up of financial penalties to encourage private policing.

At present, the White Paper lacks the clarity necessary in law. Any regulatory framework must be adequately accessible: Users "must be given an indication that is adequate in the circumstances of the legal rules applicable to a given case"[20]; Secondly, users must be able to moderate their behaviour in line with what is reasonably foreseeable[21]: Users "must be able - if need be with appropriate advice - to foresee, to a degree that is reasonable in the circumstances, the consequences which a given action may entail". The White Paper's proposed framework is insufficient to "give individuals an adequate indication as to the circumstances in which and the conditions on which the authorities are empowered to resort to any such measures."[22]

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

Case law from the ECtHR states that 'private life' is a wide concept without offering any exhaustive definitions. Therefore, Article 8 ECHR affords a right to 'private life' in its widest sense[23]. The ECtHR has explained that in order to strike a fair balance, both content and related communications data i.e. the who, when and where of communication must be treated equally. For instance, even though content can be encrypted it may disclose information about the sender or recipient. The related communications data (metadata) could portray an intimate picture of an individual through location tracking, social network and communication patterns mapping, Internet browsing tracking, and personal interaction analysis[24]. Moreover, pursuant to the European Data Protection Supervisor Opinion 6/2017, the definition of metadata should include any data which is processed 'in an electronic communications network' and any data processed by any other equipment to provide a service and that is not deemed content. Furthermore, the purpose or content of communication should not play a part in the treatment of its security and confidentiality. Additionally, any new legislation should not only protect the security and confidentiality of personal communications en route, but should safeguard the security and confidentiality of user communications data and equipment saved in the cloud[25]. The vast majority of communications data are stored in the cloud, following receipt[26]. Accordingly, these are some of the considerations we believe should guide the Government when defining personal communications.

While we agree with the White Paper that "The development of harmful activity online frequently involves a combination of activity taking place on both public and private communication channels"', it

[20] *ECtHR The Sunday Times v. the United Kingdom (No. 1), 6538/74, 26 April 1979* at Para. 47.
[21] *Rekvényi v Hungary*, 25390/94, 20 May 1999, At Para 34f.
[22] Note 20, *Sunday Times* judgement 1979, at Para. 49.
[23] ECtHR *Sidabras v Lithuania* (59330/00) at Para 43; *Bigaeva v Greece* (26713/05) at Paras 22–28.
[24] *Big Brother Watch and Others v United Kingdom* at Para 356.
[25] EDPS Opinion 6/2017 pg 12-13.
[26] WP Opinion 1/2017 para 40(c).

is essential that the notion of private communications is not conflated with publicly available content on the Internet, and that the protections granted by the e-Privacy Directive and Regulations are maintained. In particular, any requirements to scan or monitor illegal content should not apply to private channels, as this risks compromising the overall security of these communications (e.g. encryption), and leave a 'back door' for potential abuses by hackers and other adversaries.

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

Where there is an identifiable potential *illegal* harm (as listed in the white paper), those harms should be the focus rather than the forums on which the message/content is posted. An alternative model is to focus on all companies which encourage the sharing or discovery of user-generated content but if the latter option is chosen, where does this stop – for example, would this include only forums where content is not publicly posted, and does that include 'private groups' which could include upwards of 10 people?

Private messaging platforms – such as WhatsApp, Viber, or WeChat – are those which are potentially the most-harmful – as seen with cyberbullying and teen suicide for example[27]. Where the focus is on forums which contain private communications, there are free expression and privacy concerns, especially given the monitoring problems and current liability shield[28]. Private communications could refer to encrypted channels of communication – but this should not extend to 1:1 communications e.g. iMessage. Getting the balance wrong here could drive more users to resort to other privacy tools[29] causing third-party issues in the context of harm. Including all channels or forums in considerations of private could impose an unfair burden on small forums – something that previous legislation has been mindful of[30]. It is unclear from the consultation document how liberties will be protected with adequate safeguards if 'private' channels are to be included.

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

An anonymous contact point, whereby all users can make contact and report harmful content/messaging. The contact points should then rank the harm reported, and filter it accordingly to be addressed depending on the severity of the issue. There should be a differentiation between harmful content and illegal harmful content. This would mirror the approach used by New Zealand in its drafting of the Harmful Digital Communications Act 2015. We do not suggest replicating this legislation, but the approach adopted in the White Paper suggests other approaches to online harms have not been considered.

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

As laid out in the Internet Safety Strategy Green Paper (2017), a transparency report is the first step to accountability[31]. Beyond this, there should be separate accountability and reporting requirements imposed to report on human rights compliance and Article 8 and 10 ECHR protections. This could be achieved by human rights auditing by an independent body to ensure compliance with privacy and free

---

[27] NSPCC, 'On the edge: Childline spotlight report on suicide' (2014), Available at https://lfstest.nspxyz.net/services-and-resources/research-and-resources/2014/on-the-edge-childline-spotlight.

[28] Article 15, Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market.

[29] A Felt & D Evans, 'Privacy Protection for Social Networking Platforms' (2008) http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.143.5642&rep=rep1&type=pdf.

[30] For example, the Digital Economy Act 2010 imposed obligations on Internet Service Providers, but not all ISPS – only those with 400,000 or more subscribers.

[31] HM Government, 'Government response to the Internet Safety Strategy Green Paper' (May 2018) https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/708873/Government_Response_to_the_Internet_Safety_Strategy_Green_Paper_-_Final.pdf, 67.

expression rights. Additional measures include compulsory human rights impact assessments for platforms. A route of redress needs to be in place for users to challenge decisions regarding removal or takedown of content which has found not to be illegal or harmful.

Human oversight should be a requirement wherever there are takedown procedures in place. This is essential to ensure that fundamental rights are upheld and are not unjustifiably interfered with. It is also essential to have human oversight where automated systems are not reliable enough to be used in isolation – something made abundantly clear through the YouTube Content ID system[32] and the likelihood of errors[33].

Question 9: What, if any, advice or support could the regulator provide to businesses, particularly start-ups and SMEs, comply with the regulatory framework?

This depends on the regulatory framework introduced. Independent oversight should exist aside from the regulatory body – especially where that body is Ofcom or is affiliated to industry leaders. A code of conduct will provide clarity and training for human moderators. It can also support mental health and well-being. SMEs should be given appropriate support to ensure that there is protection in place for their staff to ensure that they are able to meet the demands of any regulatory framework, and are not driven out of the market as a result of being unable to cope with the increased regulatory measures.

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?

It is evident that the government expresses a strong preference for a co-regulatory model for platforms. Social media platforms would be particularly affected by the new regulatory framework. While this model has its benefits (e.g. stronger legitimacy than self-regulation, based on powers given by the Parliament, expertise, principle-based regulation, flexibility, cooperation with the industry), there is a danger of uncritically replicating the model of broadcast regulation into the online environment. Broadcast regulation has a very different historical rationale and justification (i.e. regulating entities who have the access to scarce resources, i.e. spectrum, those who produce and distribute content at a large scale, and exercise editorial control with little or no freely user-created and generated content), whereas the need for the regulation of the Internet is largely different (i.e. there are not scarce resources of the same sort, but user-generated content, individual speech and privacy implications, open and free Internet). While it is evident that self-regulation has failed on various instances, given the scandals we have witnessed, companies have started improving their regulatory mechanisms (e.g. Facebook's Oversight Board). In our view, the key here is making sure that users have the right to redress in accordance with those procedures, as well as that there is not a general obligation to monitor users, and that the current liability regime is overseen more efficiently by the regulator. Thus, it is not so much about new powers or the duty of care, but enforcement powers and the necessary oversight. If the government wishes to introduce a regulator, ideally this should be a new public body, with expertise in Internet regulation, cybercrime and online offences *and* human rights law. This would provide a balanced and proportionate oversight and the protection of fundamental rights and freedoms of Internet users.

Question 11: A new or existing regulator is intended to be cost neutral: on what basis should any funding contributions from industry be determined?

At present, a cost-neutral means of engaging with this issue is not evident. Still, there are models that may be *cost-effective*. Lord Justice Briggs, as he then was, identified the Civil Resolution

[32] YouTube, 'How Content ID Works' https://support.google.com/youtube/answer/2797370?hl=en; J Bailey, 'YouTube Beta Testing Content ID For Everyone' Plagiarism Today (2 May 2018) https://www.plagiarismtoday.com/2018/05/02/youtube-beta-testing-content-id-for-everyone/.

[33] J M Urban, J Karaganis, and B Schofield, 'Notice and Takedown in Everyday Practice', UC Berkeley Public Law Research Paper No. 2755628. http://dx.doi.org/10.2139/ssrn.2755628

Tribunal (CRT)] (located in British Columbia, Canada) as an example of a user-friendly approach to the online-dispute-resolution framework. The CRT may offer such a cost-effective example. We implore decision-makers to not repeat the access problems that arose with the Employment Tribunal Fees scheme found to be unlawful.[34] Indeed, excessive fees or social media ring-fenced taxes, as supported by some civil society groups, could be found unlawful or they could disproportionately affect business and innovation. Therefore, rather than introducing new taxes, the government should make sure that current taxes are being collected appropriately. There are too many instances where companies have failed to pay their existing contributions.

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?

The regulator should not have the power to disrupt the business, as this should be placed into the hands of judicial authorities, if absolutely necessary. ISP blocking is already being done on a large scale in the UK (IWF and ISPs for different types of content), and options for blocking clearly illegal harmful content exist. In terms of the vaguely defined 'harms', this option should not be used, as it has the potential to limit lawful speech on platforms. Generally, we would not recommend going beyond the current powers different regulators have (e.g. the ICO), but only for illegal content. Senior management liability could be implemented for intentional or negligent, large scale breaches and offences (e.g. manipulation and political advertising, hacking, data misuse scandals etc.), making sure that there are appropriate appeals procedures.

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?

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?

Question 15: What are the greatest opportunities and barriers for (i) innovation and (ii)adoption of safety technologies by UK organisations, and what role should government play in addressing these?

Question 16: What, if any, are the most significant areas in which organisations need practical guidance to build products that are safe by design?

Organizations need practical guidance on the detection of cases of live-streaming of child sexual exploitation and sexual abuse (CSEA). The government should foster the development of new technological tools for the detection and automatic flagging of live-streaming of CSEA in a way that protects the privacy of users. (i.e. the further development of Google API tool or Microsoft's PhotoDNA for detection of not only pre-existing content but also newly published live-streamed content through AI technology that prevents any breach in privacy rights of users). For the development of these products, organisations need assistance with research, as there is a lack of data on methods and tools

[34] (As it was constituted) by the Supreme Court's decision in *R (on the application of UNISON) v Lord Chancellor* [2017] UKSC 51.

used by perpetrators of live-streaming of CSEA. There is a need for further research so that organisations with have the necessary data needed to develop an effective safety by design product in this field. Any deployment of technology should be undertaken after a human rights impact assessment and any organizations should be subjected to human rights audits in the same manner as the Internet Watch Foundation.

Question 17: Should the government be doing more to help people manage their own and their children's online safety and, if so, what?

The government should increase awareness of the types of online risks faced by children of all ages, and increase the promotion of available methods of safety and protection, including technological ones, and increase the promotion of available reporting mechanisms. An example of this would be promoting the teaching resources, safeguarding advice, training and consultancy services for schools and colleges offered by NSPCC. The government should focus on education and changing mentalities to reporting online harms faced by children, ensuring compliance with international access to justice obligations for children.[35] This is one of the main reasons behind low reporting by children.

Ethical and societal impact assessments encourage reflection on any potential backlash associated with product development and the harms associated with targeting products at children. ESIAs evaluate risks to the rights and freedoms of children from marketing and advertising built around product descriptions. In evaluating their communication strategies towards children, platforms and developers should assume children have a greater susceptibility to manipulation; consider the effects of stereotypes; and advertising that might disproportionately affect this demographic. Because our understanding of how emotionally targeting affects children is evolving constantly, platforms and developers should develop Codes of Best Practices, consider joining voluntary marketing codes and standards, and support other relevant government and NGOs efforts.

Question 18: What, if any, role should the regulator have in relation to education and awareness activity?

The regulator should monitor institutions and organisational efforts to educate and raise awareness. The regulator should require periodic reports on the effectiveness of these activities and ensure organisations reach the general public each time a new technological tool, service or product is developed. This will help to ensure that the general public remains educated on how to safely use new products and services. We also support the introduction of Child Safety Impact Assessments to be filed with the regulator every time a company designs an online service or product marketed at children.

Submitted in the name of BILETA and the following individuals

Dr Aysem Diker Vanberg, Senior Lecturer, University of Greenwich, School of Law and Criminology
Dr Maureen Mapp, University of Birmingham
Gavin Sutter, Senior Lecturer in Media Law, Queen Mary University of London
Professor Abbe Brown, University of Aberdeen

[35] UN, Convention on the Rights of the Child (1989), Articles 12(2), 40, Available at http://www.ohchr.org/EN/ProfessionalInterest/Pages/CRC.aspx.