

**Written evidence submitted by Dr Edina Harbinja Senior Lecturer in Law at Aston University, Aston Law School (OSB0145)**

I am grateful for the opportunity to give oral evidence at the Committee's session on 23 September 2021. I would also like to take the opportunity to respond to this call in writing. As a law academic working in this area for the last decade, I am very interested in the issues surrounding the regulation of technology and digital rights.

**Summary**

- a) There is a need to address the current discrepancies in the regulation of online content. I support the need to improve the protection of vulnerable users online, children in particular.
- b) However, I would like to express grave concerns over the proposed Draft Online Safety Bill (hereinafter: the Bill), concerning its substantive aim, clarity, certainty, possible practical and legal effects.
- c) The policy aim is clear, but it is unlikely that the proposed regulatory model, and the duty of care, in particular, will achieve this intention without implicating other issues and causing unintended consequences for UK users, the market and other stakeholders.
- d) The Bill potentially impacts free speech adversely, particularly in the areas of democratic and journalistic content, content moderation and concerning 'legal but harmful speech'. The Bill also has the potential to adversely affect other human rights, such as privacy, which it allegedly aims to promote.
- e) There are concerns over Ofcom's suitability to perform the role of the digital regulator.

**Objectives of the Bill**

***Will the proposed legislation effectively deliver the policy aim of making the UK the safest place to be online?***

1. I argue that the Bill in its current form will not deliver the policy aim.
2. As detailed below, there are some serious concerns related to the effects of the Bill on the critical aim of 'making the UK the safest place to be online'. The Government boldly expressed an ambitious plan to show 'global leadership with our groundbreaking laws to usher in a new age of accountability for tech and bring fairness and accountability to the online world'.<sup>1</sup> After Brexit, it was easier for the UK to claim 'regulatory agility' and

---

<sup>1</sup> Gov.uk Press Release, 'Landmark laws to keep children safe, stop racial hate and protect

commitment to innovation outside Brussels and its 'slow bureaucratic legislative and regulatory processes'. The UK can almost act as a 'regulatory sandpit', where mechanisms can be explored and tested more quickly, for better or for worse.

3. The Bill's key aims are as follows: to address illegal and harmful content online (terrorist content, racist abuse and fraud, in particular) by imposing the duty of care concerning this content; to protect children from child sexual exploitation and abuse (CSEA) content; to protect of users' rights to freedom of expression and privacy; to promote media literacy. Arguably, some of these are more important than others for the Government since the Bill does not include adequate mechanisms to protect freedom of expression and privacy, for example.
4. I agree with organisations such as the Electronic Frontier Foundation that these are important aims<sup>2</sup> considering how much power platforms have in controlling and moderating communications and other information. Nevertheless, these aims must be balanced against protecting freedom of speech, other human rights and avoiding censorship.<sup>3</sup>
5. The Bill only tangentially mentions human rights or digital rights. In section 12, it vaguely mandates 'duties about rights to freedom of expression and privacy'. This section in the Bill seems disjointed from the rest of the proposal and almost acts as just a necessary add-on, as I have argued earlier. Privacy invasions covered by the Bill are limited to 'unwarranted'.<sup>4</sup>

### ***Is the "duty of care" approach in the draft Bill effective?***

6. A duty of care as the model for regulating speech online is problematic in many ways. This is a duty derived from the health and safety law and imposed on certain service providers to moderate user-generated content in a way that prevents users from being exposed to illegal and harmful content online. Many commentators, including myself, have expressed their reservations about the concept during Online Harms consultations and elsewhere.<sup>5</sup> Dr Leiser and I have emphasised the difficulty of the duty of care to satisfy the test set by courts, *inter alia*, due to the difficulty in drawing an analogy with current examples of duty of care (e.g. owners of physical property, where harm is specific, well defined and owed to identifiable individuals).<sup>6</sup>

---

democracy online published' (12 May 2021), <https://www.gov.uk/government/news/landmark-laws-to-keep-children-safe-stop-racial-hate-and-protect-democracy-online-published>

<sup>2</sup> EFF and OTI Joint Comments in Response to UK Online Harms White Paper (2019), [https://newamericadotorg.s3.amazonaws.com/documents/UK\\_Online\\_Harms\\_White\\_Paper.pdf](https://newamericadotorg.s3.amazonaws.com/documents/UK_Online_Harms_White_Paper.pdf).

<sup>3</sup> Big Brother Watch, Response to the Government's proposed Online Safety Bill, 12 May 2021, <https://bigbrotherwatch.org.uk/2021/05/big-brother-watch-response-to-the-governments-online-safety-bill/>

<sup>4</sup> Edina Harbinja, 'U.K.'s Online Safety Bill: Not That Safe, After All?', Jul 8, 2021, <https://www.lawfareblog.com/uks-online-safety-bill-not-safe-after-all>

<sup>5</sup> Mark Leiser and Edina Harbinja. 2020. "CONTENT NOT AVAILABLE". Technology and Regulation 2020 (September), 78-90. <https://techreg.org/index.php/techreg/article/view/53>

7. There is little detail in the Bill on how the duty will operate, which raises concerns over censorship, over-removing, and ‘collateral damage.’<sup>7</sup> The Bill leaves decisions related to the legality of content to private, profit-driven platforms, the task that has been traditionally dealt with by courts. The model proposed places platforms in a position where they are directly controlling the speech of an individual.<sup>8</sup>

***Does the Bill deliver the intention to focus on systems and processes rather than content, and is this an effective approach in moderating content? What role do you see e.g., safety by design, algorithmic recommendations, minimum standards, default settings?***

8. The proposed Bill mainly focuses on content, with systems and processes left to be regulated by Ofcom’s codes. It is difficult to say that the Bill delivers the intention to focus on systems and processes.
9. The Bill should emphasise safety by design, human rights impact assessments and setting minimum standards rather than emphasising content moderation, which can lead to unwanted filtering and blocking of legitimate speech. Other examples of helpful impact assessments include child protection impact assessments, ethical/societal impact assessments for all products and services, including retroactively.

***How does the draft Bill differ to online safety legislation in other countries (e.g. Australia, Canada, Germany, Ireland, and the EU Digital Services Act) and what lessons can be learnt?***

10. The Bill differs from the EU proposal in various significant ways. From a digital rights perspective, the EU proposal is considered to be more suitable. Other examples mentioned in the question are not exemplary, and I do not advise using them as a blueprint for the new UK law.<sup>9</sup>
11. A good example is the EU’s Digital Service Act Proposal definition of very large platforms as those having ‘average monthly active recipients of the service in the Union equal to or higher than 45 million’.<sup>10</sup> New obligations for these platforms include mandatory risk

---

<sup>6</sup>Ibid.

<sup>7</sup> Graham Smith, ‘Harm Version 3.0: the draft Online Safety Bill’ Cyberlegale Blog, 16 May 2021 <https://www.cyberleagle.com/2021/05/harm-version-30-draft-online-safety-bill.html>

<sup>8</sup> Kim Barker, ‘Taming our Digital Overlords: Tackling Tech Through ‘Self-Regulation’?’ in Ignas Kolpkas and Julija Kalpokiene (eds) *Intelligent and Autonomous* (Brill, forthcoming 2022)

<sup>9</sup> For a commentary on Germany, see e.g. Heidi Tworek and Paddy Leerssen, ‘An Analysis of Germany’s NetzDG Law’, April 15, 2019,

[https://www.ivir.nl/publicaties/download/NetzDG\\_Tworek\\_Leerssen\\_April\\_2019.pdf](https://www.ivir.nl/publicaties/download/NetzDG_Tworek_Leerssen_April_2019.pdf) ; For the Australian context, see Zahra Zsuzsanna Stardust, ‘A new online safety bill could allow censorship of anyone who engages with sexual content on the internet’, *The Conversation*, February 18, 2021, <https://theconversation.com/a-new-online-safety-bill-could-allow-censorship-of-anyone-who-engages-with-sexual-content-on-the-internet-154739>

assessment, mitigation of risk, independent audit, online advertising transparency, obligations related to recommender systems, data access by the Independent Coordinator. With some refinement, however, the Proposal retains principles of the existing intermediary liability, including the prohibition of general monitoring, which could violate Article 8. of the ECHR. All these are some valuable lessons to be learned.

12. The EU's Proposal for the Digital Services Act<sup>11</sup> includes establishing a Digital Services Coordinator in each member state and the European Board for Digital Services, tasked to advise the Digital Services Coordinators and the Commission. In my view, the Digital Authority proposed by the House of Lords earlier<sup>12</sup>, not Ofcom, would be a more suitable counterpart to these European partners.

### **Content in Scope**

***The draft Bill specifically includes CSEA and terrorism content and activity as priority illegal content. Are there other types of illegal content that could or should be prioritised in the Bill?***

13. No. This content is usually very clearly illegal, and it is easier for companies to distinguish this type of content from other types that may include legitimate speech, e.g. trolling.

***The draft Bill specifically places a duty on providers to protect democratic content, and content of journalistic importance. What is your view of these measures and their likely effectiveness?***

14. Considering numerous and continuing issues with content moderation,<sup>13</sup> this duty will just introduce an additional burden to the problematic area of private policing of free speech.

15. Content included in the duty to protect content that is "of democratic importance" is broadly defined as content intended to contribute to democratic political debate in the UK or a part of it. Essentially, this is a duty not to remove this particular type of speech. The definition is very broad and overlaps with 'journalistic content', so more effort is

---

<sup>10</sup> Article 25, EU, Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC, COM/2020/825 final

<sup>11</sup> Ibid, sections 2 and 3.

<sup>12</sup> House of Lords Communications and Digital Select Committee, 'Regulating in a Digital World', 9 March 2019, Chapter 6, <https://publications.parliament.uk/pa/ld201719/ldselect/ldcomuni/299/299.pdf>

<sup>13</sup> Kate Klonick, 'The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression', Yale Law Journal, Vol. 129, No. 2418, 2020, Available at SSRN: <https://ssrn.com/abstract=3639234>

needed to distinguish between the two categories. Arguably, these should be merged, as it could be difficult for a service provider to ascertain what type of speech belongs to each of them (e.g. election-related content, Covid-19 related content). Also, there is a concerning question of whether political speech should be distinguished from other important forms of free speech and when health-related speech becomes political and vice versa.

16. In terms of journalistic content, the Bill introduces a very wide definition of 'journalistic content'. The provider will be required to 'make a dedicated and expedited complaints procedure available to a person who considers the content to be journalistic content'. This procedure should be available for both democratic and journalistic content, given the overlap between the two types of content.

17. The definition of journalistic content seems to cover user content 'generated for the purpose of journalism', as long as there is a UK link. The Government's press release noted that 'citizen journalists' content will have the same protections as professional journalists' content'.<sup>14</sup> In practice, it will be difficult to implement and ascertain if a given user post should be deemed journalistic and a take-down challenged.

***Are the definitions in the draft Bill suitable for service providers to accurately identify and reduce the presence of legal but harmful content, whilst preserving the presence of legitimate content?***

18. Harm is defined very vaguely 'The provider...has reasonable grounds to believe that the nature of the content is such that there is a material risk of the content having, or indirectly having, a significant adverse physical or psychological impact on a child (adult)'. It is unclear what indirect harm includes, and it could be a very wide category that impacts human rights. The Bill will be supplemented by codes of practice developed by Ofcom and secondary legislation, which will further define what constitutes harmful content.

19. The definition of 'legal but harmful content' can encompass legitimate speech. Section 46(3) of the Draft Bill defines legal but harmful content as 'content having, or indirectly having, a significant adverse physical or psychological impact on an adult of ordinary sensibilities.'

20. First, the meaning of 'indirect harm' requires further explanation as platforms can use this to remove user content by applying a very low threshold.<sup>15</sup> This creates the danger of censoring a vast amount of content that is neither illegal nor harmful. Due to the fines and other enforcement mechanisms introduced in the Bill, platforms and service

---

<sup>14</sup> Note 20.

<sup>15</sup> Edina Harbinja, 'The UK's Online Safety Bill: Safe Harmful, Unworkable?' (18 May 2021) <https://verfassungsblog.de/uk-osb/>.

providers are likely to take down content without thoroughly investigating whether it is harmful.

21. Further, the service provider will need to determine if the content is harmful (directly or indirectly) to children or adults ('service has reasonable grounds to believe that the nature of the content is such that there is a material risk of harm...'). The standard used for this assessment is a risk of harm to adults or children of 'ordinary sensibilities'. The concept is taken from the tort of abuse of private information, but in that context, it operates in an entirely different manner as the person with ordinary sensibilities is the person whose information is disclosed, not the person who receives the information.<sup>16</sup> This is a vague legal standard, which does not correspond to the well-established standard of a 'reasonable person', for instance, and leaves many open questions, such as whether those 'easily offended' fall within this category.
22. Some of these harms are already regulated and illegal (e.g. terrorist-related content, content related to child abuse, extreme pornography etc.). However, a lot of the harms that the Government's White Paper<sup>17</sup> identified 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 be harmful but not rise to the threshold of illegality and may be protected speech.
23. 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. None of our laws regulating speech are applied via a subjective test. This also contravenes the longstanding principle from *Handyside v UK*<sup>18</sup>:

*"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".<sup>19</sup>*

24. Regarding harms that online speech may cause, any speech assessment will need to include qualitative questions about 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 parity principle*). For this to happen, the platform will need to understand the context of exchanges between every user on a platform and how people communicate offline with one another, and this is very difficult, if not virtually

---

<sup>16</sup> *Campbell v Mirror Group Newspapers Ltd* [2004] UKHL 22.

<sup>17</sup> Online Harms White Paper: Full Government Response to the Consultation, <https://www.gov.uk/government/consultations/online-harms-white-paper/outcome/online-harms-white-paper-full-government-response>

<sup>18</sup> ECtHR (1976) *Handyside v UK* (5493/72).

<sup>19</sup> At Para 49.

impossible.<sup>20</sup> Different platforms have different social norms and communication practices, and this should be respected (e.g. it is not realistic to expect the same language on 4Chan, Reddit, TikTok, Instagram and Mumsnet).

25. Therefore, I believe that the new legislation should apply only to illegal content and leave out the odd and vague concept of legal but harmful. If the content is really harmful, legislation should outlaw it. Otherwise, it should remain within the domain of free speech.

### **Services in Scope**

26. Schedule 1 of the Bill specifies that services and content excluded from the scope of this regime are: emails, SMS messages, MMS messages (for these three categories, only if they represent 'the only user-generated content enabled by the service', so for example, Messenger does not qualify and it would be regulated), comments and reviews on provider content, 'one-to-one live aural communications' (communications made in real-time between users and only if the communications consist solely of speech or other sounds and do not include any written message, video or other visual images, so Zoom does not qualify and it is within Bill's scope), internal business services, paid-for advertisements and news publisher content (needs to be a 'recognised news publisher'), and certain public bodies services. The list of exceptions is thus relatively narrow.

27. Moreover, the Bill gives a significant power to the Secretary of State for Digital, Culture, Media and Sport to amend the Schedule 1 (exempt services) and either add new services to the list of exemptions or remove some of those already exempt, based on the assessment of the risk of harm to the individuals. This power gives the Government minister quite a lot of discretion, which, if misused, could lead to policing private messaging and communications channels such as Messenger or Zoom. The rationale is to address illegal content such as terrorist and 'child sexual exploitation and abuse' (CSEA) content, which is disseminated through these channels, but questions around its effect on encryption, security and privacy remain open.<sup>21</sup>

### **The role of OFCOM**

28. The choice of OFCOM as the regulatory body for the proposed regime is concerning. The regulatory body for media and electronic communications should not be the same body tasked with addressing online content and speech regulation.

---

<sup>20</sup> S Banerjee, Chua, A. Y., & Kim, J. J.. 'Don't be deceived: Using linguistic analysis to learn how to discern online review authenticity. *Journal of the Association for Information Science and Technology*', (2017) 68(6), 1525-1538.

<sup>21</sup> Heather Burns, 'Encryption in the Online Safety Bill' (ORG, 20 July 2021) <https://www.openrightsgroup.org/blog/encryption-in-the-online-safety-bill/>

29. With Ofcom in charge, 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 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 etc.).
30. As I noted before, the House of Lords Communications and Digital Committee's suggestion to establish a new body, the Digital Authority, is a more suitable option. The regulator should have expertise in Internet regulation, cybercrime and online offences *and* human rights law. This would provide a balanced and proportionate oversight and protect the fundamental rights and freedoms of Internet users. Judicial oversight is crucial, and any regulator should be independent with clear pathways for judicial remedies and reviews.
31. The Bill creates a potent online regulator with potent enforcement mechanisms, which could have significant and lasting effects on businesses and digital rights. There is a real danger that Ofcom may not be able to undertake this role effectively, given all other areas within its regulatory remit, plus its lack of human and technical capacity.

*16 September 2021*