Amendments to the Online Safety Bill

At Index on Censorship, we work with people across the globe who are being censored by oppressive regimes. The UK Government's Online Safety Bill will be catastrophic for freedom of speech of British citizens online in its current form. Plans to force tech platforms to delete "harmful" content or face big fines will lead to many legal posts being deleted and we fear it will not make people actually safer. This paper sets out further policy ideas for amending the Online Safety Bill (the Bill') to protect freedom of expression and keep people safe. Amendments are in draft form only and based on wider stakeholder debate and the evidence submitted by us and many others to the Parliamentary Select Committees,

The proposed amendments to the Bill are broadly categorised in the following sections:

1. Ensuring the Duty of Care is not a duty to censor
2. Retaining Parliamentary Sovereignty
3. Digital evidence

[Text within the example amendment boxes are underlined to denote additional text and struck through to denote proposed text deletion]

BILL AMENDMENT PROPOSALS

1. Ensuring the Duty of Care is not a duty to censor

What is the problem?

The "Duty of Care" includes regulatory notions best applied to health and safety law in the workplace rather than freedom of speech online. The Duty of Care model's 'precautionary principle' shifts the balance from punishing individuals for breaking the law to state-sanctioned private censorship of content, for which there has been no judgement whether that content is legal or illegal. This shift, if enacted through the Online Safety Bill, would mark the most significant change in the role of the state over free speech in the UK since 1695.

Furthermore, the Bill proposes creating two tiers of free speech: free speech for journalists and politicians, and censorship for ordinary British citizens. The exemptions in the bill are both unfair and unworkable. The "journalistic content" exemption, will censor ordinary citizens and allow journalists to speak freely online. This is another outsourcing of our rights to Silicon Valley, who - given the lack of clear definition for "journalism" - will decide what is journalism

and what is not. This sets a dangerous precedent and will censor countless citizen journalists and online news sources.

What is the solution?

- **Removal of the Duty of Care and all reference to its structure in sections 5 and 6 of the Bill** - this would ensure that speech online was treated the same as speech offline.
- **If the Duty of Care is not removed, it should at the very least be amended to apply to effect illegal content only** (as set out by the Lords Committee in their report) and 'legal but harmful' content should be excluded from the duties that require codes of practice.

Amendments Proposed

Duty of Care: <u>Removal</u>

1. Duty of Care stripped out of Bill

   ~~Providers of user-to-user services: duties of care~~

   ~~(1) Subsections (2) to (6) apply to determine which of the duties set out in this Chapter and Chapter 3 apply in relation to a particular regulated user-to-user service...~~

Duty of Care: <u>Focuses only on illegal content</u>

1. *<u>11 Safety duties protecting adults: Category 1 services</u>*

   (1) The "duties to protect adults' online safety" in relation to user-to-user services are the duties set out in this section.

   (2) A duty to ~~specify in the terms of service~~ report bi-annually to the UK Parliament Online Safety Committee on —

   (a) how priority content that is harmful to adults is ~~to be~~ dealt with by the service (with each such kind of priority content separately covered) and

   (b) how other content that is harmful to adults, of a kind that has been identified in the most recent adults' risk assessment (if any kind of such content has been identified), is to be dealt with by the service.

2. *<u>Chapter 6 — Interpretation of Part 2</u>*

   48 Meaning of "Chapter 2 safety duty" and "Chapter 3 safety duty"

   (1) In this Part "Chapter 2 safety duty" means any of the duties set out in—

   (a) section 9 (safety duties about illegal content),

   (b) section 10 (safety duties for services likely to be accessed by children), ~~or~~

   ~~(c) section 11 (safety duties protecting adults).~~

   (2) In this Part "Chapter 3 safety duty" means any of the duties set out in—

   (a) section 21 (safety duties about illegal content), or

   (b) section 22 (safety duties for services likely to be accessed by children)

Standard of proving illegality

What is the problem?

We do not need a new bureaucratic or regulatory framework. There is no doubt that racism, homophobia, and transphobia have no place in society, but they are already covered under existing laws. It is accepted that legislation governing online activity should be expanded to address the rapidly developing online space.

What is the solution?

The test for **illegal content should be 'actually' illegal, not that platforms have 'reasonable grounds to believe' content is illegal**. The reason why this change is needed is that platforms will interpret their 'reasonable belief' widely and depending on the political circumstances of the speech concerned, will not want to be blamed or fined for content remaining online. This will inevitably lead to over-censorship and there are currently insufficient free speech safeguards against such platform behaviour.

The test of 'actual knowledge of illegality' meets the criteria in the E-commerce regulations which is already applied to the law of libel. Aligning this criteria with pre-existing criteria already in use in other areas of law, which have proven to be adequate and efficient, would ensure the law is effective and certain.

There are a range of proof thresholds that we would propose in an amendment:

A. Believes there is an arguable case that content is illegal
B. Believes beyond reasonable doubt that content is illegal
C. Has sufficient evidence that content is illegal
D. Has a clear and convincing evidence that content is illegal

Thresholds A + B will require a lawyer from the online platform assessing the legality of speech that is to be removed based on existing case law and guidelines produced by the police and Crown Prosecution Service. Thresholds C+D will require an investigator to assess the sufficiency of evidence that the content is illegal - again based on ongoing dialogue with law enforcement authorities - in order to justify their removal of content.

This approach would better match the prosecutorial threshold in law of the Crown Prosecution Service and prevent a divergence in evidential standards offline and online, leading to over-censorship. This would also push the structure of the Online Harm framework closer to the model supported by the Centre for Policy Studies i.e. that platforms will take the lead from UK courts (adjudicating laws that Parliament has passed) on what is illegal rather than

figuring this out themselves. The amendment would also mean deferring to courts (whether criminal or libel) as the proper arbiter of an individual's intention underpinning their speech which is inevitably ignored by an algorithm.

We accept that the police, Crown Prosecution Service and courts do not always get cases right, as borne out by a previous campaign to amend the scope of section 5 of the Public Order Act 1986 or the attempted prosecution of a protester who insulted a horse or the Twitter Joke Trial. However, those examples were subject to a full judicial process and public scrutiny, a far cry from a picosecond process by an algorithm.

Amendments Proposed

Duty of Care: Defining 'illegal'

41 Meaning of "illegal content" etc

(1) This section applies for the purposes of this Part.

(2) "Illegal content" means—

(a) in relation to a regulated user-to-user service, content—

(i) that is regulated content in relation to that service, and

(ii) that amounts to a relevant offence;

(b) in relation to a regulated search service, content that amounts to a relevant offence.

(3) For the purposes of this section, content consisting of certain words, images, speech or sounds amounts to a relevant offence if the provider of the service has ~~reasonable grounds to believe~~ Threshold A or B or C or D —

(a) the use of the words, images, speech or sounds amounts to a relevant offence,

(b) the use of the words, images, speech or sounds, when taken together with other regulated content present on the service, amounts to a relevant offence, or

(c) the dissemination of the content constitutes a relevant offence…

44 Regulations under section 41

(1) When deciding whether to specify an offence, or a description of offence, in regulations under section 41, the Secretary of State must take into account—

(a) the prevalence on regulated services of content that amounts to that offence or (as the case may be) offences of that description,

(b) the level of risk of harm being caused to individuals in the United Kingdom by the presence of such content, and

(c) the severity of that harm.

(2) For the purposes of subsection (1), content consisting of certain words, images, speech or sounds amounts to an offence or to an offence of a particular description if there ~~are reasonable grounds to believe~~ Threshold A or B or C or D that—

(a) the use of the words, images, speech or sounds amounts to the offence or to an offence of that description,

(b) the use of the words, images, speech or sounds, when taken together with other regulated content present on regulated services, amounts to the offence or to an offence of that description, or
(c) the dissemination of the content constitutes the offence or an offence of that description...

2. Retaining Parliamentary Sovereignty

What is the problem?

The Bill in its current form gives extensive powers to Ofcom turning it into a free speech super regulator. This would be the first time since the 1600s that written speech will be overseen by the state in the UK. OfCom's new powers and the proposed influence of the Minister for DCMS are not adequately checked to avoid the politicisation of online content moderation. OfCom also has a duty to prepare codes of practice which help regulated services comply with their various duties. These codes of practice will be advisory only, but it is likely that compliance with them will be equated to compliance with the wider duties under the Bill. As such, they are likely to become a very significant feature of the regulatory landscape.

What is the solution?

We would propose that supervision of emerging problematic arenas of legal speech online should fall under a new Online Safety Committee[1] who could make recommendations on criminal laws that require updating. This approach would have far more democratic legitimacy than codes of practice by a regulator. The powers of the new Committee should involve scoping evidence and conducting consultations as well as balancing what applicable legislation may be needed beyond the Online Safety Bill. There should be no shortcuts to law-making in the UK, especially when it concerns censoring people's speech online.

Amendment Proposed

1. *11 Safety duties protecting adults: Category 1 services*

   (1) The "duties to protect adults' online safety" in relation to user-to-user services are the duties set out in this section.

   (2) A duty to ~~specify in the terms of service~~ report bi-annually to the UK Parliament Online Safety Committee on —

   (a) how priority content that is harmful to adults is ~~to be~~ dealt with by the service (with each such kind of priority content separately covered) and

   (b) how other content that is harmful to adults, of a kind that has been identified in the most recent adults' risk assessment (if any kind of such content has been identified), is to be dealt with by the service.

[1] See the question by the chair on page 46 https://committees.parliament.uk/oralevidence/2794/pdf/

2. *Chapter 6 — Interpretation of Part 2*

48 Meaning of "Chapter 2 safety duty" and "Chapter 3 safety duty"

(1) In this Part "Chapter 2 safety duty" means any of the duties set out in—

(a) section 9 (safety duties about illegal content),

(b) section 10 (safety duties for services likely to be accessed by children), ~~or~~

~~(c) section 11 (safety duties protecting adults).~~

(2) In this Part "Chapter 3 safety duty" means any of the duties set out in—

(a) section 21 (safety duties about illegal content), or

(b) section 22 (safety duties for services likely to be accessed by children)

3. Digital evidence

What is the problem?

There is a large emphasis on taking down content in the Bill but a distinct lack of legislative safeguarding of digital evidence that may be needed to punish offending online or even to prove innocence. If content deletion is not coordinated with investigatory requirements and limitation periods on criminal offences, it may hamper authorities ability to prosecute. Deleting online abuse can also leave individuals unaware they are in danger or impact victims' right to access justice. Some of the worst examples in recent times include evidence of war crimes in Syria being removed by YouTube as part of its work to remove terrorist content and propaganda from the platform.

What is the solution?

Category 1 platforms should be required to archive and securely store all evidence of removed content from online publication for a set period of time akin to HMRC requirements on public companies to keep financial records for 6 years from the end of the last company financial year they relate to.

Amendment Proposed

PART 8 - Duty to keep records of removed content

(1) All providers of Category 1 services must keep adequate records of removed content.

(2) Adequate removal records means records that are sufficient—

(a) to show and explain the company's reason for content removal,

(b) to disclose with reasonable accuracy, at any time, the content removal decision of the company at that time, and

(c) to enable the Senior Manager to ensure that any content removed comply with the requirements of this Act

(3) Records of removed content—

(a) *must be kept at its registered office or such other place as the directors think fit, and*

(b) *must at all times be open to inspection by the company's officers.*

(4) All records must be kept at a place in the United Kingdom, and must at all times be open to such inspection authorised by judicial warrant .

(5) Content removal records that are required by this section X must be preserved for a period of six years from the date on which they were made.