INDEX ON CENSORSHIP

A VOICE FOR THE PERSECUTED

# THE WAY FORWARD:

# HOW TO MAKE THE ONLINE SAFETY BILL WORK FOR THE BRITISH PEOPLE

SEPTEMBER 2022

# The Way Forward: How to make the Online Safety Bill work for the British people

## INTRODUCTION

Index on Censorship has been tracking the Online Safety Bill and the precursor debate on 'online harms' for many years. It seems that finally, under the premiership of Liz Truss, there may be the prospect of sensible and workable regulation on how content is moderated online. It has been encouraging to hear the Prime Minister confirming at PMQs on 7 September 2022 that the Bill will be amended to better balance the trade offs on safety online and individual freedoms.

So how does the government move forward with one of its signature pieces of legislation in this Parliament?

Firstly, the Bill has benefitted from high level scrutiny in Parliament and in particular by the DCMS, Joint Select Committee and Lords' Communications and Digital Committees. These bodies have carefully heard evidence from the tech industry, civil society, victims of abuse, academics, legal experts and others and produced detailed analyses that were ignored by the previous Secretary of State for DCMS, the fourth to be charged with the task of regulating online harms. It is time to listen to the detailed scrutiny that Parliament has undertaken.

Secondly, it is vitally important that the government legislates in a way that observes the principle of legality, namely that people, business and courts can apply the rules that Parliament sets with certainty. The common law has a long tradition of presuming the legislative intent of Parliament as contributing to existing bodies of law, including fundamental constitutional principles and rights[1]. So far however, the Online Safety Bill has sought to invert the historic right of free speech online so that perceptions of subjective harm trump objective legality. Our

[1] https://ukconstitutionallaw.org/2022/09/12/michael-foran-interpretation-after-the-human-rights-act-the-principle-of-legality-and-the-rule-of-law/

legal analysis[2] indicates that if the Bill is not amended with practicability in mind, free speech would be threatened by over removal of content and tech companies would likely leave the UK - diminishing our thriving digital economy/society. Efficacy and workability should be at the heart of this new legislation, not vague commitments to cure the ills of the internet.

Thirdly, the new Secretary of State has a promising track record on this issue given her work on the Higher Education (Freedom of Speech) Bill. Indeed her comments on the problems the Higher Education Bill was designed to resolve could equally be applied to the problems the Online Safety Bill will create: "*the damage that the erosion of free speech causes goes well beyond the classroom. It hits our communities, where ingenuity and diversity of ideas have flowed from throughout our country's history. It stifles creativity, where some of our greatest artists and composers have made their name challenging the accepted wisdom of the day. The implications for our economy and our public life are catastrophic*."[3] The new Secretary of State for DCMS should apply the dictum that what is legal to say should be legal to type.

It is never too late for the government to listen and heed the extensive warnings to prevent bad regulation. In this briefing, we have highlighted the five key areas for decision-makers to consider in order to protect free speech for adults. These changes are not an attempt to 'water down' the Bill but to purposefully demarcate protections that are only appropriate for children as distinguished from adults:

1. **Remove the 'legal but harmful' provision completely**
2. **Prevent the algorithmic censorship of lawful speech**
3. **Illegality should be decided by courts not corporates**
4. **Private communication and encryption need protection**
5. **Parliament should be sovereign**

---

[2] https://www.indexoncensorship.org/wp-content/uploads/2022/05/Legal-analysis-of-the-impact-of-the-Online-Safety-Bill.pdf

[3] https://www.gov.uk/government/speeches/minister-donelan-addresses-policy-exchange

# REFRESHING THE ONLINE SAFETY BILL

Throughout the last few years, civil society has consistently raised the alarm about the impacts the Online Safety Bill will have on ordinary people's free speech - yet hardly any substantial changes have been made. Now is the time for a refresh. We recognise that codification of online content moderation can be good for society and ultimately benefit the growth of the digital economy. However, this will only be possible if the Bill is limited in scope and specific in its ambit to the particular areas online that people need most protection from. The borderless nature of the online environment poses compliance challenges and governments around the world, such as in New Zealand[4], have recognised that there will need to be many different types of regulation as well as better public education and law enforcement to truly make us all safer online. This won't all happen under the current flabby, incoherent, Online Safety Bill.

## 1. Remove 'legal but harmful' provision completely

In the Bill, the duty to censor content with a 'material risk of significant harm to an appreciable number of adults', means that platforms will need to assess not whether content is illegal or not, before restricting access, but will have a duty to proactively censor content with the potential to cause harm. This wide-reaching nature of the 'legal but harmful' provision risks weaponising censorship of perfectly legal content.

A 'Health and Safety' approach to speech contradicts our historic common law on freedom of expression. Many lawyers agree that the Bill is not compatible with free speech rights under UK and international law[5]. The government should not invent a new category of lawful but 'harmful speech' for adults in a free society.

---

[4] https://www.dia.govt.nz/diawebsite.nsf/Files/Proactive-releases/$file/Cabinet-material-about-the-initiation-of-the-media-content-regulatory-review.pdf

[5] Gavin Millar QC legal analysis of the Online Safety Bill, Index on Censorship report (2022) https://www.indexoncensorship.org/wp-content/uploads/2022/05/Legal-analysis-of-the-impact-of-the-Online-Safety-Bill.pdf

As former Supreme Court Justice, Lord Jonathan Sumption QC stated:
*"As applied to adults, the whole concept of restricting material which is entirely legal is a patronising abuse of legislative power. If the law allows me to receive, retain or communicate some item of information in writing or by word of mouth, how can it rationally prevent me from doing the same thing through the internet? Why should adult internet users be infantilised by applying to them tests directed to the protection of the most sensitive minorities?"*[6]

**The concept of legal but harmful material for adults should be completely removed from how platforms conduct their content moderation obligations** because it will incentivise tech companies to over-censor legal content to avoid fines of up to 10% of global revenue from Ofcom. This measure will create a censorious UK internet, out of sync with British standards on freedom of speech and without actually making anyone safer from harm.

Suggested Amendment

- Page 13, lines 9-10, leave out Clause 13, subsection (4) (a) and (b) and delete Clause 54 of the Bill

## 2. Prevent the algorithmic censorship of lawful speech

The duty to take or use proportionate measures to manage the risk of harm to individuals means that platforms will use crude algorithms to assess whether content is illegal or not, before restricting access to it or removing it completely.

AI has no regard for nuance or context. Reliance on algorithmic content moderation will disproportionately impact marginalised communities, victims of crime, and journalists, and remove content on the basis that it falls under the 'legal but harmful' category. Crucial evidence of harm and abuse would be removed before it can be reported to the authorities, as well as masking potential threats to victims. Studies have also already

[6] Jonathan Sumption, The Spectator (August 2022)
https://www.spectator.co.uk/article/the-hidden-harms-in-the-online-safety-bill

found that algorithms disproportionately identify posts written in African American Vernacular English as “rude” or “toxic,” reflecting and amplifying racial bias in AI[7].

Censorship by algorithm is by definition not a transparent process meaning citizens will be deprived of the ability to understand how and why their speech online is being curtailed. This will happen despite the legal requirement for restrictions on free speech to be ‘sufficiently accessible to the individual who is affected by the restriction’, and will likely be challengeable in the courts.

In addition, UK journalists will be unable to report on complex world events as opposed to their counterparts in other countries as they will only be able to access a ‘sanitised’ version of social media online. This will impugn investigative journalism, as press protections under the Bill only addresses journalism when it is published, not within the process of journalism being produced.

Judging whether criminal offences have occurred is a difficult and technical exercise and one which police, the Crown Prosecution Service and courts often get wrong. The Bill’s illegal content duties however, will lead private companies to pre-emptively take down significant amounts of content that they ‘reasonably believe’ to be ‘illegal’ even if they are not, in order for tech companies to avoid being penalised themselves. This would represent a huge leap in the powers of technology companies - making tech executives some of the most powerful actors for the UK’s internet. Obvious examples of where this will be problematic are assessments of satirical content, content in which users discuss trauma or other challenging topics (which lends itself to being wrongly assessed by technology as either illegal or harmful).[8] This will inevitably lead to broad censorship and there are currently insufficient free speech safeguards against such platform behaviour.

[7] Index on Censorship, RIGHT TO TYPE (2021)
https://www.indexoncensorship.org/wp-content/uploads/2021/06/Index-on-Censorship-The-Problems-With-The-Duty-of-Care.pdf
[8] Gavin Millar QC legal analysis of the Online Safety Bill, Index on Censorship report (2022)
https://www.indexoncensorship.org/wp-content/uploads/2022/05/Legal-analysis-of-the-impact-of-the-Online-Safety-Bill.pdf

One of the best ways of ensuring that public interest materials are not inadvertently removed is to increase (not decrease) the thresholds for content moderation to narrow the reach of platform's new state-sanctioned powers. The Bill should set the threshold for defining illegal content which matches criminal investigations and prosecutions offline. This will ensure that there is a level of consistency between speech that is criminalised through the courts and the standards applied by service providers online. Otherwise the Bill is in danger of creating a de facto 'legal but harmful' speech through unaccountable and unlawful definitions of what speech is illegal. The new Bill should include a **requirement for human moderation and particularly legal expertise on crime which will ensure that only content that is actually illegal is removed from platforms.**

Suggested Amendment

- Page 8, line 9, Clause 9, sub-section (6) remove 'reasonably considers' and replace with 'has sufficient evidence that' with explanatory additional text that evidential sufficiency is a legal test that can not be undertaken by an algorithm.

## 3. Illegality should be decided by courts not corporates

The Bill defines "Illegal Content" as "content that amounts to a relevant offence" and also includes "Priority Illegal Content" such as child sexual exploitation and abuse (CSEA), terrorist content, and other priority offences. Illegal Content includes an offence under any United Kingdom law wherein "the victim or intended victim is an individual or (individuals)[9]. There is also a category of priority offences which include assisting suicide, threats to kill, intentional harassment that causes fear of violence, racially or religiously aggravated public order offenses, offenses regarding firearms, sexual offenses, assisting unlawful immigration, proceeds of crime or fraud, and more.

These definitions are extremely broad and include criminality that no algorithm or tech employee has a hope of being able to properly substantiate on the basis of published

[9] https://www2.datainnovation.org/2022-uk-online-safety-bill.pdf

content such as tweets or instagram posts. The criminal law engaged is complex and requires specific scrutiny by people with professional expertise and sometimes lay members of a jury. It is wrong to pretend to the public that tech companies' algorithms removing content will make them safer than real investigative and prosecutorial authorities putting individual allegations of offending through the criminal justice system.

The resolution to this problem is for the scope of illegal content moderated under the Bill to be narrowed so that it only includes criminal offences that require a high level of *objective* evidence e.g. CSAM rather than offences that have a high level of *subjective* interpretation e.g. use of threatening words or disorderly behaviour within the hearing or sight of a person likely to be caused harassment, alarm or distress (s.5 Public Order Act 1986). This amendment to the Bill should assist in the protection of free speech and prevent the worst problems of censorship as encountered in offline cases such as the 'Twitter joke trial' concerning Paul Chambers[10].

**There should only be one set of duties on illegal content (i.e. Clause 9(3) in the Bill) that is applied to one narrowly defined and coherent schedule of priority illegal content.**

Suggested Amendment

- Delete Clauses 9(2) and 21(2) and Schedule 7

## 4. Private communication and encryption need protection

This Bill proposes that accredited technology can be used to scan both publicly and privately shared content for anything related to terrorist activities. The scanning of everyone's private messages is a huge violation of privacy and completely disproportionate given existing policing and investigatory powers.

Despite evidence from countless technology experts and civil society, the government's amendment to break end-to-end encryption and widen the scope of the proactive

[10] https://www.judiciary.uk/wp-content/uploads/JCO/Documents/Judgments/chambers-v-dpp.pdf

monitoring requirement will undermine security and leave users vulnerable to attack, doing exactly the opposite of what it is supposed to - keep users safe.

The Bill will push the UK away from Five Eyes cyber security standards and could open up UK citizens' private communications to foreign hostile states. The former head of GCHQ has said that weakening security for everyone is not the solution. The UK Information Commissioner's Office (ICO) also intervened in the encryption debate with an unequivocal endorsement of end-to-end encryption.

The Bill could inadvertently damage Trans-Atlantic trade of digital services. US tech companies offering encrypted services such as Whatsapp will be unable to access the UK's digital market due to risks of being exposed to proactive monitoring. Fresh data shows that London is the number one destination globally for US tech companies looking to expand outside of their home market[11] but the Bill would force many large US companies to withdraw from the booming UK tech ecosystem, now valued at $1trn.[12]

**Private communications must be removed from the scope of the obligations to moderate content under the Bill.**

Suggested Amendment

- Page 87, line 19, Clause 104 sub-section (2) (b), remove 'whether' and 'or privately'

## 5. Parliament should be sovereign

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 also has a duty to prepare codes of practice

[11] Data by London & Partners - https://media.londonandpartners.com/news/london-looks-to-strengthen-tech-trade-and-investment-links-with-us
[12] Figures by the Digital Economy Council and Dealroom - https://www.siliconrepublic.com/start-ups/uk-tech-sector-1-trillion-dollars-northern-ireland#:~:text=An%20analysis%20by%20Dealroom%20for,tech%20ecosystem%20at%20%241trn.

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.

The Secretary of State of the day is also given a plethora of new powers in this Bill, the most dangerous being the ability to designate what content should be considered 'harmful to adults' for up to 28 days without any approval from parliament, police, or the courts. This means a homophobic Secretary of State could have a statutory instrument censoring LGBTQ+ content passed without full parliamentary scrutiny in as little as a day.

There is no precedent in the Western World for such accumulation of executive power. Neither OfCom's new powers nor the proposed influence of the Secretary of State for DCMS are adequately checked to avoid the politicisation of online content moderation. Indeed such unbridled power would be more suited to an autocratic nation state. In an increasingly polarised political playing field, allowing the whims of a single politician to dictate what content should be removed is incredibly dangerous and undemocratic.

If red lines 1-4 were observed, there would be no need for Clauses 40 and 148 of the Bill as the narrowed scope would render them otiose. However, it is important as a matter of freedom of speech principle to ensure that limitations on citizens' speech should only be mandated through the legislative process in Parliament.

**If speech that is currently lawful needs to be curtailed for legitimate public interest reasons, this process should go through full Parliamentary scrutiny via primary legislation**.

<u>Suggested Amendment</u>

- Remove Clause 40 and 148 from the Bill