

Written evidence submitted by Big Brother Watch

DCMS Sub-committee on Online Harms and Disinformation – Call for Evidence; Online safety and online harms

September 2021

About Big Brother Watch

Big Brother Watch is a civil liberties and privacy campaigning organisation, fighting for a free future. We're determined to reclaim our privacy and defend freedoms at this time of enormous technological change.

We're a fiercely independent, non-partisan and non-profit group who work to roll back the surveillance state and protect rights in parliament, the media or the courts if we have to. We publish unique investigations and pursue powerful public campaigns. We work relentlessly to inform, amplify and empower the public voice so we can collectively reclaim our privacy, defend our civil liberties and protect freedoms for the future.

Introduction

The Online Safety Bill, published in May of this year by the Department for Digital, Culture Media and Sport (DCMS) is a fundamentally flawed piece of legislation, destined to negatively impact the sanctity of the fundamental rights to privacy and freedom of expression in the UK. The proposed model, centred on imposing "duties of care" on all companies that enable people to interact with others online, to protect users from "harm", will force these companies to act as online police. Under the threat of penalties, this will compel online intermediaries to over-remove content.

We believe the Online Safety Bill in its current form is not fit to become law in a liberal democracy like the UK. In order to protect citizens' free expression and the free flow of information, the Bill must be materially altered.

The legislation engages the fundamental rights to freedom of speech and privacy, protected by Article 10 and Article 8 of the European Convention on Human Rights (ECHR) respectively. The

European Convention on Human Rights is clear that interference with these rights will only be lawful where they are provided by law, necessary and proportionate.¹ The presumption must rest in favour of protecting these rights and interference with them should come as a last resort.

The Bill has been widely criticised across the human rights sector. The international freedom of expression organisation, Article 19, has stated that if passed, the Bill would be “a chokehold on freedom of expression” and that it is “wary of legal frameworks that would give either private companies or regulators broad powers to control or censor what people get to see or say online”². Gavin Millar QC, of Matrix Chambers, has also been highly critical of the legislation. Talking about the impact the Bill could have on rights around the world he said,

“As someone who has undertaken many free speech missions for international organisations to countries with repressive free speech regimes such as China, Turkey, Azerbaijan there is a real risk that this legislation, if passed, will be used to justify repressive measures aimed at closing down free speech on the internet in these countries.”³

As well as our profound concerns regarding the proposed duty of care, we believe that the Government’s approach to this legislation will effectively mean that the legal standard for permissible speech online will be set by platforms’ terms of use rather than being clearly set out in statute. It is also our view that the broad definition of harm given in the legislation will result in a malleable, censorious online environment. Additionally, we believe that the regulatory model will give legal backing to a system often described as “surveillance capitalism”, demanding that online intermediaries fortify their terms of use and uphold them, compelling increased surveillance of users online.

We believe that as a minimum, provisions relating to so called “legal but harmful” content should be removed from the face of the Bill (clause 11), that the scope of the legislation should not include private messaging services (clause 39) and that the legislation should not attempt to introduce age-verification via the back door.

In responding to this call for evidence, we will respond to the following questions:

Is it necessary to have an explicit definition and process for determining harm to children and adults in the Online Safety Bill, and what should it be?

¹The Human Rights Act, EHRC, <https://www.equalityhumanrights.com/en/human-rights/human-rights-act>

²UK: Draft Online Safety Bill poses serious risk to free expression, Article 19, 26 July 2021, <https://www.article19.org/resources/uk-draft-online-safety-bill-poses-serious-risk-to-free-expression/>

³Government’s Online Safety Bill will be “catastrophic for ordinary people’s freedom of speech” says David Davis MP, Index on Censorship, 23 June 2021, <https://www.indexoncensorship.org/2021/06/governments-online-safety-bill-will-be-catastrophic-for-ordinary-peoples-freedom-of-speech-says-david-davis-mp/>

What are the key omissions to the draft Bill, such as a general safety duty or powers to deal with urgent security threats, and (how) could they be practically included without compromising rights such as freedom of expression?

Are there any contested inclusions, tensions or contradictions in the draft Bill that need to be more carefully considered before the final Bill is put to Parliament?

What are the lessons that the Government should learn when directly comparing the draft Bill to existing and proposed legislation around the world?

We are concerned that the terms of reference set out by the DCMS Sub-Committee on Online Harms and Disinformation make no reference to the rights impact of the Bill. The Bill is likely to have the greatest impact on freedom of expression of any piece of legislation passed in the UK in living memory, yet the Committee's questions make no reference to the impact on free speech. Nor do the terms of reference acknowledge the right to privacy and how this legislation may impact this right. As such, in the course of our response, we will attempt to highlight a number of rights concerns as well as answer the questions previously referred to.

Is it necessary to have an explicit definition and process for determining harm to children and adults in the Online Safety Bill, and what should it be?

We have fundamental concerns regarding the duty of care model at the heart of the Online Safety Bill, the obligations placed on platforms and the vague definitions of harm, all of which we believe would do damage to free speech in the UK.

At the heart of the Online Safety Bill is a shift towards increased liability on social media companies, who, under obligations placed on them through the legislation, must take responsibility for the speech and even private messages of members of the public on their sites. Such a move would have serious ramifications for freedom of expression and privacy online. Part 2 of the Bill sets out the new "duties of care" that the legislation places on all in-scope services.

It is extremely unusual for a private company to be under a "duty of care" to prevent harm resulting from the conduct of others.⁴ A duty of care ordinarily refers to a company's duty to ensure its own risk-creating actions do not cause physical injury to others (for example, a stockroom manager has a duty to ensure employees use safe lifting equipment to avoid physical harm). Clearly, these

⁴See also, UK Supreme Court, *Robinson v Chief Constable of West Yorkshire Police*, 2018

conditions do not apply to internet intermediaries – to provide platforms for people to interact is not a risk-creating action, and there is no risk of physical harm. We believe that this liability model, effectively developed in tort law, is highly inappropriate when deployed for the purposes of regulating free speech.

Introducing obligations of this nature marks a clear departure from a traditional regulatory approach towards online platforms, held in both the EU and US, which gives platforms immunity from liability for the content on their sites. This principle has been applied in regulatory frameworks with the specific intention of protecting the free expression and privacy of users online. A standard that directly applies is Article 15 of the EU’s E-Commerce Directive (this technically still applies to the UK as “EU retained law”), which prohibits member states from imposing general monitoring obligations on social media companies operating within their jurisdictions.⁵

In addition to their duties relating to potentially illegal content, “Category 1 services” (large social media companies) are obliged to fulfil additional duties “to protect adult online safety” including a duty to tackle “content that is harmful to adults”. This is set out in the legislation as follows:

11 (2) A duty to specify in the terms of service—

(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.

(3) A duty to ensure that—

(a) the terms of service referred to in subsection (2) are clear and accessible, and

(b) those terms of service are applied consistently⁶

These provisions are deeply problematic and pose a serious threat to free speech. The notion of a state-backed system of effectively forcing the removal or suppression of online expression that is lawful contravenes accepted human rights standards when it comes to limiting expression. The state should not endorse the censorship of expression that is lawful and limitations on free speech should

⁵Directive 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 ('Directive on electronic commerce') <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32000L0031>

⁶Draft Online Safety Bill, 2021, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/985033/Draft_Online_Safety_Bill_Bookmarked.pdf

be exercised only where necessary, where they are proportionate and where they are clearly prescribed in law.

The legislation sets out a definition of harmful content (and therefore harm) which category 1 platforms must endeavour to tackle through their terms of use. The definitions set out in Cl. 46 are:

(3) Content is within this subsection if the provider of the service 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 an adult of ordinary sensibilities

Or

(5) Content is within this subsection if the provider of the service has reasonable grounds to believe that there is a material risk of the fact of the content's dissemination having a significant adverse physical or psychological impact on an adult of ordinary sensibilities, taking into account (in particular)—

(a) how many users may be assumed to encounter the content by means of the service, and

(b) how easily, quickly and widely content may be disseminated by means of the service.⁷

In order to protect freedom of expression, restrictions on permissible speech should always be clearly defined in law, to safeguard rights and limit the possibility of overzealous or censorious enforcement. However, a “risk of... having a significant adverse physical or psychological impact”, even “indirectly”,⁸ is an overly broad definition and threatens application that would be damaging to free speech. An adverse psychological impact could, for example, refer to an offensive joke or footage of an emergency situation. It could even constitute the documentation of social injustice, such as the video of George Floyd’s murder which changed debates internationally about race, justice and authority. Such content might cause distress but is important to see and share for the benefit of society.

The Bill also gives power to the Secretary of State to designate, through secondary legislation, specified categories of “harmful content” which Ofcom must incorporate into its codes of practice. While the Government have set out a definition of “harm” in the Bill, what specific “harms” will be set out in secondary legislation remain opaque, and subject to change.

⁷Ibid.

⁸Ibid.

This level of influence over the regulatory regime (alongside a number of other provisions throughout the legislation), will give the government of the day a huge amount of executive power to ultimately influence the permissibility of speech online. It is also unclear whether these specific “harms” will even have to meet the aforementioned standard of posing a risk of having an “adverse physical or psychological impact”. The Bill defines “priority content that is harmful to adults”, which the Secretary of State sets via secondary legislation, as “content of a description designated as such in [those] regulations” (cl. 46(9)) – that is, harmful content is whatever the Secretary of State says it is.

The risk of politicisation of free speech limitations is further exacerbated by the highly subjective threshold for so-called “harmful” speech that applies to platforms. In the definition of “content that is harmful to adults”, which applies to platforms (cl.46), the content in question must risk direct or indirect harm to an “an adult of ordinary sensibilities”⁹. The appropriateness of such a definition was brought into question by internet lawyer Graham Smith, who argues that such a term does not nullify subjectivity from the measure of harm.¹⁰ He also notes that such a test ordinarily refers to a “reasonable person of ordinary sensibilities” and thus the objective nature of this term has been further eroded.¹¹ Furthermore, where content “may reasonably be assumed to particularly affect people with a certain characteristic” or of a “certain group”, the platform must apply that characteristic or group to the “adult of ordinary sensibilities” (cl. 46(4)). These “characteristics” or “groups” can be any at all and are not restricted to protected characteristics, leaving an absurdly broad and subjective framework for “harm” that is skewed towards censorship and open to abuse.

The legislation goes even further in Cl. 46 (7), where it defines content which could have an “indirect” impact and could risk an adult acting in a way which may cause “harm” to another person¹². Obligations on platforms to tackle “harmful content” of this kind are deeply problematic and would pave the way for sweeping online censorship. These measures play on the idea that adults do not have individual agency and exposure to others’ speech, even where lawful and not incitement, could cause others to perpetuate “harm”. This could make online spaces sanitised spaces where it would be impossible to document violence or other societal ills.

⁹ Draft Online Safety Bill, 2021,

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/985033/Draft_Online_Safety_Bill_Bookmarked.pdf

¹⁰Smith G. On the trail of the Person of Ordinary Sensibilities, Cyberleagle, 28 June 2021

<https://www.cyberleagle.com/2021/06/on-trail-of-person-of-ordinary.html>

¹¹ibid.

¹²Draft Online Safety Bill, 2021,

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/985033/Draft_Online_Safety_Bill_Bookmarked.pdf

Clause 11, which creates this “safety” duty for platforms, has been roundly criticised by freedom of expression groups in the strongest of terms; a warning which must not go unheeded by policymakers. Index on Censorship criticised this provision in a paper, which stated:

“Legal but harmful’ has been defined in the draft Bill as causing “physical or psychological harm”, but how can this be proved? This definition opens up significant problems of subjectivity. The reason, in law, we do not use this definition for public order offences is that it is hard for citizens to understand how their words (written or spoken) could cause psychological harm in advance, especially on the internet where we do not know our audience in advance.”¹³

Concerns around the breadth of the definition of harm have also been expressed by lawyer Graham Smith:

“What is an adverse psychological impact? Does it have to be a medically recognised condition? If not, how wide is it meant to be? Is distress sufficient? The broader the meaning, the closer we come to a limitation that could mean little or nothing more than being upset or unhappy. The less clear the meaning, the more discretion would be vested in Ofcom to decide what counts as harm, and the more likely that providers would err on the side of caution in determining what kinds of content or activity are in scope of their duty of care.”¹⁴

In a blog on the Government’s Online Harms agenda, the lawyer Ashley Hurst previously wrote that the Government should “focus on what is illegal and defined, not legal and vague.”¹⁵ An article by digital rights organisation, Matrix, also referred to these provisions within the legislation as an attempt at “centralising and regulating relative morals”¹⁶.

However, most recently and most prominently, the House of Lords Communications and Digital Committee Recommended that Clause 11 be removed from the Bill entirely. In the Committee’s report, which followed a Parliamentary inquiry into freedom of expression online, the Committee stated:

“We do not support the Government’s proposed duties on platforms in clause 11 of the draft Online Safety Bill relating to content which is legal but may be harmful to adults. We are not

¹³Right to Type, Index on Censorship, June 2021, <https://www.indexoncensorship.org/wp-content/uploads/2021/06/Index-on-Censorship-The-Problems-With-The-Duty-of-Care.pdf>

¹⁴Smith G. The Online Harms edifice takes shape, Cyberleagle, 17 December, 2020, <https://www.cyberleagle.com/2020/12/the-online-harms-edifice-takes-shape.html>

¹⁵Hurst, A. Tackling misinformation and disinformation online, Inforrm, <https://inforrm.org/2019/05/16/tackling-misinformation-and-disinformation-online-ashley-hurst/>

¹⁶Almeida, D. How the UK’s Online Safety Bill threatens, Matrix, 19 May 2021, <https://matrix.org/blog/2021/05/19/how-the-u-ks-online-safety-bill-threatens-matrix>

convinced that they are workable or could be implemented without unjustifiable and unprecedented interference in freedom of expression. If a type of content is seriously harmful, it should be defined and criminalised through primary legislation. It would be more effective—and more consistent with the value which has historically been attached to freedom of expression in the UK—to address content which is legal but some may find distressing through strong regulation of the design of platforms, digital citizenship education, and competition regulation.”¹⁷

The idea that social media platforms should be compelled to remove broad categories of so-called “harmful” content would lead to two distinctive tiers of permissible speech in the UK in the online and offline worlds. Not only would our online public squares be restricted and censored but free speech more broadly would be chilled as a result.

If we are to avert the Online Safety Bill doing permanent damage to the right to free speech in the UK, as a minimum, Clause 11 should be removed from the Bill.

Clause 10 of the Bill sets out safety duties that intermediaries must undertake with a focus on users who are children. The provisions within the clause effectively demand that regulated services take responsibility for the safety of children who may access their site. Clause 10 (3) states:

(3) A duty to operate a service using proportionate systems and processes designed to—

(a) prevent children of any age from encountering, by means of the service, primary priority content that is harmful to children;

(b) protect children in age groups judged to be at risk of harm from other content that is harmful to children (or from a particular kind of such content) from encountering it by means of the service.¹⁸

It is vital that children are protected online but as a result of these provisions, this legislation would result in internet-wide censorship at intolerably strict levels. There are many forms of content which could be considered “harmful” for children to view but should not be wiped from user-to-user platforms, for example adult humour or the documentation of crime or violence.

¹⁷Free for all? Freedom of expression in the digital age, House of Lords Communications and Digital Committee, 22 July 2021, <https://committees.parliament.uk/publications/6878/documents/72529/default/>

¹⁸Draft Online Safety Bill, 2021, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/985033/Draft_Online_Safety_Bill_Bookmarked.pdf

Throughout the Online Safety Bill, the legislation suffers from being overly broad in its aims. Rather than focus on upholding the rule of law and ensuring platforms take steps to work with law enforcement to protect children from genuinely illegal content online, this Bill seeks to eradicate nebulous concepts of harm which would result in a more restricted online experience for everyone.

Clause 10 (9) also states:

The duties in this section extend only to such parts of a service as it is possible for children to access.¹⁹

Given the huge popularity of social media and the vast number of users on each of the major platforms, the likelihood that a social media site may be accessed by children is high in any case. This means that unless a platform undertakes invasive age verification checks and then age-gates user-generated content at a granular level, content moderation on the site in question must be tailored for children.

This directly threatens both free expression and privacy rights online. The measures will force platforms to comply with higher thresholds for the acceptability of content unless they verify users' age using ID. This means mandating age verification and would be hugely damaging to privacy rights online. Online anonymity is crucially important to journalists, human rights activists and whistleblowers all over the world. Even tacit attempts to undermine online anonymity here in the UK would set a terrible precedent for authoritarian regimes to follow and would be damaging to human rights globally.

Such a measure would also mean that internet users would have to volunteer even more personal information to the platforms themselves, which would likely be stored in large centralised databases. Further, many people across the UK do not own a form of ID and would directly suffer from digital exclusion.

The Bill should not force online platforms to introduce mandatory age verification via the back door.

What are the key omissions to the draft Bill, such as a general safety duty or powers to deal with urgent security threats, and (how) could they be

¹⁹ Draft Online Safety Bill, 2021, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/985033/Draft_Online_Safety_Bill_Bookmarked.pdf

practically included without compromising rights such as freedom of expression?

The Online Safety Bill is not a law enforcement Bill. No obligations are placed upon platforms to collaborate with or provide evidence to law enforcement agencies. Rather, the legislation is effectively a content-takedown Bill, deputising online intermediaries to remove content that they believe to be illegal and content that meets a broad definition of “harm”.

As such, it would be highly inappropriate to integrate additional measures relating to urgent security threats, which should involve law enforcement agencies and security services.

The question set out by the Committee is based on the premise that the legislation does not already fundamentally compromise rights such as freedom of expression. As previously discussed, it does, both by increasing liability on online platforms and placing obligations upon regulated services to tackle open-ended concepts of harm.

Are there any contested inclusions, tensions or contradictions in the draft Bill that need to be more carefully considered before the final Bill is put to Parliament?

The legislation sets out a number of additional duties relating to freedom of expression and privacy, protecting “journalistic content” and content which is of “democratic importance”. However, far from creating effective protections in these areas, the provisions create points of conflict within the legislation and are, nevertheless, largely outweighed by safety duties that will encourage platforms to over-remove content.

The duties relating to freedom of expression and privacy, which engage all regulated services, are particularly weak and read as follows:

(2) A duty to have regard to the importance of—

(a) protecting users’ right to freedom of expression within the law, and

(b) protecting users from unwarranted infringements of privacy, when deciding on, and implementing, safety policies and procedures.²⁰

²⁰Draft Online Safety Bill, 2021, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/985033/Draft_Online_Safety_Bill_Bookmarked.pdf

Unlike the operational safety duties, which compel companies to “minimise” illegal or so-called harmful content on their sites, this duty only instructs tech companies to “have regard to the importance” of free expression and privacy.²¹

The duties specifically imposed upon category 1 services are no more conducive to effectively protecting freedom of expression on large social media platforms than the aforementioned requirements. They compel platforms to undertake impact assessments on the way in which their systems and processes effect freedom of expression and privacy on the platform and set out how they might remedy any threats to these rights on their platform.²² Once again, this duty is significantly weaker than the operational safety duties and will do little to materially protect free expression online.

The weakness of these provisions was fairly characterised by internet Lawyer Graham Smith when he said:

“No obligation to conduct a freedom of expression risk assessment could remove the risk of collateral damage by over-removal. That smacks of faith in the existence of a tech magic wand. Moreover, it does not reflect the uncertainty and subjective judgement inherent in evaluating user content, however great the resources thrown at it.”²³

The legislation addresses the point of conflict between the operational safety duties and duties to protect free expression and privacy in clause 36 (5):

36 (5) A provider of a regulated user-to-user service is to be treated as complying with the duty set out in section 12(2) (duty about freedom of expression and privacy) if the provider takes such of the steps described in a code of practice which are recommended for the purposes of compliance with a Chapter 2 safety duty (so far as the steps are relevant to the provider and the service in question) as incorporate safeguards for—

(b) the protection of users’ right to freedom of expression within the law,

or

(b) the protection of users from unwarranted infringements of privacy.²⁴

²¹ Ibid.

²² Draft Online Safety Bill, 2021, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/985033/Draft_Online_Safety_Bill_Bookmarked.pdf

²³ Smith, G. *Harm Version 3.0: the draft Online Safety Bill*, Cyberleagle Blog, May 2021, <https://www.cyberleagle.com/2021/05/harm-version-30-draft-online-safety-bill.html>

²⁴ Draft Online Safety Bill, 2021, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/985033/Draft_Online_Safety_Bill_Bookmarked.pdf

This suggests that so long as platforms follow the code of practice on operational safety duties, this will be sufficient to comply with their free expression and privacy duties. This demonstrates the inherent weakness of the duties to have regard to freedom of expression and privacy, which pay lip service to these fundamental rights in a Bill which otherwise damages them.

The very nature of the legislation, which compels social media companies to take liability for content on their sites, means that platforms of this kind will be forced to monitor and surveil users more than ever before. This approach is a serious threat to online privacy and cannot be remedied by asking platforms to simply give “regard” to this fundamental right.

In fact, by mandating so called “technology notices” (see clause 64)²⁵ the Bill will compel social media companies to read the messages of their users to scan for potential “harm”. Far from “reigning in” big tech companies, this legislation gives foreign companies license to spy on the communications of British citizens, supporting an exploitative business model that erodes privacy rights.

In addition to the aforementioned duties, the Online Safety Bill also places an obligation upon category 1 regulated services to “protect” content of “democratic importance” and “journalistic content”.²⁶ The Government claim that this legislation will not threaten free expression online - however, if this is the case it begs the question of why these carve-outs are necessary.

These provisions, clearly borne out of concern that platforms could reprimand politicians in a similar way to former President Trump, include an obligation on platforms to apply the safety duties in a politically neutral manner:

13 (3) A duty to ensure that the systems and processes mentioned in subsection (2) apply in the same way to a diversity of political opinion.²⁷

This demonstrates a recognition on the part of the Government that the fortification of and mandated adherence to platforms’ terms of use will create a more politicised, censorious environment online. However, these provisions effectively exempt politicians themselves from this new system of regulation.

In describing what content of “democratic importance” would constitute, the Bill states:

_Safety_Bill_Bookmarked.pdf

²⁵ Draft Online Safety Bill, 2021,

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/985033/Draft_Online_Safety_Bill_Bookmarked.pdf

²⁶Ibid.

²⁷Ibid.

*6 (b) the content is or appears to be specifically intended to contribute to democratic political debate in the United Kingdom or a part or area of the United Kingdom.*²⁸

The vague nature of this categorisation will only create additional complications for the platforms as they are simultaneously told to remove content which could subjectively be considered “harmful”, but not that which is considered a part of “democratic political debate”. Given the sweeping nature of this description and the regulatory burden dealt to them, it is likely that intermediaries will take a narrow interpretation of this provision and give additional protection to the expression of elected officials. As a result, these exemptions present as one rule for politicians, who will have greater privileges to speak freely online, and another rule for the population at large.

When setting out a duty upon category 1 services to protect “journalistic content”, Clause 14 states that platforms have:

*A duty ... to make a dedicated and expedited complaints procedure available to a person who considers the content to be journalistic content*²⁹ (Where the complainant is the person who shared or created the content in question.)

Services are also obligated to create such a dedicated and expedited complaints process for *all* users where action has been taken by the platform on “journalistic content”.³⁰ However, the legislation provides only a loose definition of what “journalistic content” should constitute and states that platforms are to set out a means of identifying journalistic content. The definition given in the Bill is as follows:

14 (8) (a) the content is—

(i) news publisher content in relation to that service, or

(ii) regulated content in relation to that service;

(b) the content is generated for the purposes of journalism; and

*(c) the content is UK-linked.*³¹

It is unclear how freelance or citizen journalism would fit within this description. A democratising effect of the internet has been the opening of spaces for marginalised voices, blogs, campaign

²⁸Ibid.

²⁹Draft Online Safety Bill, 2021, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/985033/Draft_Online_Safety_Bill_Bookmarked.pdf

³⁰Ibid.

³¹Ibid.

journalism and more disintermediated news sharing. Citizen journalism online has made a significant contribution to media as a whole, offering new and diverse perspectives, rapid story-telling, inclusive media and audience participation. Citizen journalism has played a major role in 21st Century political events,³² including the Occupy movement and the Arab Spring, and this has relied on the more equal playing field online for individuals to gain exposure and generate revenue. If carve-outs are only afforded to the journalists and media operators that the social media companies choose, this unhealthy monopolisation will only be exacerbated.

What are the lessons that the Government should learn when directly comparing the draft Bill to existing and proposed legislation around the world?

Policymakers must heed the warning signs from similar legislation passed overseas, which has often seriously repressed speech, had major unintended consequences and been emulated by more authoritarian actors around the world.

In 2017 the German Government passed the Network Enforcement Act, also known as 'NetzDG'. The UK Government have been quick to draw comparisons with the German law in modelling their own Online Safety Bill, without acknowledging the serious impact it has had on rights.

The Act threatens fines of up to €50 million for social media companies that fail to remove illegal content within 24 hours. It is extremely heavy-handed and the imposed threat of such a large fine incentivises profit-driven social media companies to err on the side of caution and over-censor content.

Around the time of its passing, Human Rights Watch called on German lawmakers to “promptly reverse” NetzDG and explained that it is “vague, overbroad, and turns private companies into overzealous censors to avoid steep fines, leaving users with no judicial oversight or right to appeal.”³³

Far from protecting the most vulnerable, the Bill has created hate speech martyrs and has seen the account suspensions of people from minority groups as a result of targeted mass-reporting. The Bill has also been cited and emulated in Russia, Venezuela, the Philippines and Turkey amongst other

³²Citizen Journalism, Encyclopaedia Britannica, <https://www.britannica.com/topic/citizen-journalism>

³³Germany: Flawed Social Media Law – Human Rights Watch, 14 Feb 2018, <https://www.hrw.org/news/2018/02/14/germany-flawed-social-media-law>

countries.³⁴ However, the legislation, which has been criticised by rights groups around the world, is arguably less stringent than the Online Safety Bill.

As the Online Safety Bill goes through Parliament, policymakers should be mindful of the fact that governments around the world will be closely watching the course of this legislation. If it remains inherently bad for human rights, malign actors and authoritarian states will seek to emulate it or legitimise their own practices by citing its example.

The legislation makes clear that private messaging services will be within scope and therefore, platforms will be obliged to uphold duties of care in these channels. This is a dangerous direction and will result in growing surveillance online, even in spaces intended for users to hold a private conversation.

There are important technical issues to consider when imposing the “duty of care” on companies’ private messaging channels. Some companies offer structural privacy to their services – for example, the end-to-end encryption offered by instant messaging/VoIP apps WhatsApp and Signal. It is concerning that the Government’s intentions appear to deliberately make privately designed channels of this kind incompatible with platforms’ obligations set out in the Bill.

The Bill gives Ofcom the power to mandate the use of technology to identify and remove certain types of illegal content. The legislation states:

64 (4) A use of technology notice under this section is a notice relating to a regulated user-to-user service requiring the provider of the service to do either or both of the following—

(a) use accredited technology to identify public terrorism content present on the service and to swiftly take down that content (either by means of the technology alone or by means of the technology together with the use of human moderators to review terrorism content identified by the technology);

(b) use accredited technology to identify CSEA content present on any part of the service (public or private), and to swiftly take down that content (either by means of the technology alone or by means of the technology together with the use of human moderators to review CSEA content identified by the technology).³⁵

³⁴Germany’s balancing act: Fighting online hate while protecting free speech, Politico, 1 October 2020, <https://www.politico.eu/article/germany-hate-speech-internet-netzdg-controversial-legislation/>

³⁵Draft Online Safety Bill, 2021,

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/985033/Draft_Online_Safety_Bill_Bookmarked.pdf

It is vital that terrorism and CSEA content are removed from the internet. However, the risk of such content being stored or shared does not justify breaking encrypted channels, sacrificing the security, safety and privacy of millions of users. Given that private messaging services are within the scope of the legislation, the provision above does imply that certain types of technology could be used to break, erode or undermine the privacy and security provided to messaging services by end-to-end encryption. This could involve the use of a technique known as client-side scanning, which would create vulnerabilities within messaging services for criminals to exploit or could open the door to a greater level of surveillance through use of this technology.³⁶

It is not unreasonable to think that such technology would be escalated in time, put to use in other areas and result in increased surveillance of individuals' private messages.

As with other areas of the Bill, one of the real risks when it comes to legitimising new surveillance technology is that it will be emulated and indeed, will embolden, authoritarian regimes around the world to undertake similar practices but for even more undemocratic means.

Private communications are fundamental for our safety and privacy – and are critical for protecting journalists, human rights activists and whistleblowers all around the world. If the Government use this Bill to attack private communications, this will impact upon safety online for all and will set an example for more authoritarian regimes to follow.

We hold serious concerns that the financial and criminal penalties set out in the legislation would result in overzealous application of companies' duties and thus compel widespread censorship. However, the Bill also goes much further and sets out penalties which are inherently draconian. The legislation would give Ofcom license to seek Service Restriction Orders (e.g. forced removal of services from the app store) or Access Restriction Orders (ISP blocking), either of which must be approved in court.³⁷ The proposal for search engine, intermediary and ISP blocking is severe and is a threat to free expression.

Concerns about service restriction orders and access restriction orders were also raised by free expression group Article 19 in their response to the Bill. Addressing what they described as "disproportionate sanctions" the group stated:

³⁶*Fact Sheet: Client-Side Scanning*, The Internet Society, March 2021,
<https://www.internetsociety.org/resources/doc/2020/fact-sheet-client-side-scanning/>

³⁷Draft Online Safety Bill, 2021,
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/985033/Draft_Online_Safety_Bill_Bookmarked.pdf

“Website (or service) blocking is almost always disproportionate under international human rights law because in most cases, websites would contain legitimate content. In practice, blocking is a sanction that would penalise users who would no longer be able to access the services that they like because a provider hasn’t removed enough content to the liking of Ofcom or the Minister. It is also the kind of measures that have been adopted in places such as Turkey. It is therefore regrettable that the UK is signalling that these types of draconian measures are acceptable.”³⁸

These are extremely serious sanctions with wide-ranging effects, including on third parties such as search engines and ISPs and the public more widely. The idea of the British Government appointing a regulator to enforce ISP blocks and search-engine controls over information is extraordinary. Such severe sanctions are chilling and reflect the extreme nature of this proposed legislation, which sets an awful precedent internationally and is at odds with fundamental liberal democratic values.

As such Clauses 91-95 should be removed from the Bill.

Conclusion

The Online Safety Bill poses a greater threat to freedom of speech in the UK than any other law in living memory. In the course of this consultation response, we have attempted to set out a number of our key concerns regarding the impact that this legislation would have on fundamental rights in the UK, whilst also attempting to answer the Committee’s questions.

It is vital that policymakers consider the impact on the right to free speech and privacy in the course of their scrutiny of this legislation.

Whilst we believe that the Bill is fundamentally flawed in its approach, the legislation suffers particularly from broad definitions, overbearing provisions and measures which grant the executive excessive power over the process.

There are a number of measures that policymakers could take to limit the detrimental impact of this legislation on free speech and privacy. Whilst we will also be responding to the pre-legislative scrutiny committee’s call for evidence with a full analysis of the Bill and full set of recommendations, some of the most important amendments that we recommend are set out below.

³⁸UK: Draft Online Safety Bill poses serious risk to free expression, Article 19, 26 July 2021, <https://www.article19.org/resources/uk-draft-online-safety-bill-poses-serious-risk-to-free-expression/>

Recommendations for policymakers

- **The legislation should not force platforms to take down or suppress speech that is lawful.**

In its current form the Online Safety Bill would create a two-tier system for freedom of expression, with extra restrictions for online speech which is considered subjectively harmful. Any further restrictions on our right to free speech must be in line with UK law and decided on in a democratic, parliamentary process.

If we are to avert the Online Safety Bill doing permanent damage to the right to free speech in the UK, as a minimum, Clause 11 (relating to “legal but harmful” content) should be removed from the Bill.

- **Private conversations should not fall within the scope of the Bill.** The legislation extends duties of care to private messaging services and threatens end-to-end encryption. Private communications are vital for our safety and privacy – and are critical to protect journalists, human rights activists and whistleblowers all around the world. Moves to erode privacy online undermine the fundamental right to privacy and would make us all less safe.

In order to protect the right to privacy, clauses 63-69 (technology notices) should be removed from the Bill. Further, the scope of the legislation (Clause 137) should be refined and private messaging services should be excluded from the Bill entirely.

- **The Bill should not force online platforms to introduce mandatory age verification.** At present, the legislation requires companies which are “likely to be accessed by children”, to moderate content accordingly unless they can prove otherwise. This will push platforms towards mandatory age verification which is invasive, would cause digital exclusion and sets a poor example for autocratic regimes around the world to follow.
- **The Bill should not give authorities the power to block or prevent UK citizens from using online intermediaries.** These measures are incompatible with liberal democratic ideals, could cause wholesale damage to both free expression and the free flow of information and would be cited and emulated by malign actors all around the world.

Clauses 91-95 should be removed from the Bill.

