BIG BROTHER WATCH

Big Brother Watch's Briefing on the Online Safety Bill for House of Lords Committee Stage

April 2023

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.

Contact

Mark Johnson

Advocacy Manager

Email: mark.johnson@bigbrotherwatch.org.uk

SUMMARY

- **The Online Safety Bill is a piece of legislation regarding the regulation of online intermediaries, destined to negatively impact the fundamental rights to privacy and freedom of expression in the UK.**
- **The proposed regulatory model centres on imposing "duties of care" on all companies that enable people to interact with others online, to protect users from "harm". Compliance with the regulatory regime will be adjudged by the regulatory body Ofcom.**
- **The legislation enlists social media companies to act as private online police and adjudicate on the legality of online content. This is bad for due process and undermines the rule of law. While illegality may be clear and obvious in some circumstances, these companies are ill-suited to make determinations of this kind (not least regarding the legal limitations of speech) and will consequentially over-remove online expression.**
- **Whilst the Government's proposed removal of provisions regarding so called "legal but harmful" expression are welcome, new plans to compel online intermediaries to remove content which contravenes their terms of use are deeply problematic.**
- **Social media companies' terms and conditions are bad for privacy and restrict expression significantly more than UK law. The Government recently amended the legislation to ensure that platforms uphold their terms of service consistently. Whilst this may sound benign, these terms of service are not neutral and have complex implications. We find it highly questionable as to whether the UK can uphold its obligation to protect freedom of expression under Article 10 if state regulators are legally compelled to add enforcement powers to foreign terms of service that generally censor speech far more than domestic laws.**
- **The Bill will do serious damage to the right to privacy in the UK and compels online intermediaries to use "proactive technologies" to scan and surveil the content of all users on their sites. Other measures in the Bill would force companies to surveil private messages and seek to introduce new ID requirements for internet access by compelling the use of age-verification.**
- **Unless the Bill is materially altered, it will do serious damage to both free speech and privacy in the UK.**

INTRODUCTION

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

2. We believe that 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.

3. 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 interferences with these rights are only 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.

4. The Bill has been widely criticised across the human rights sector and has rapidly become known as the "censor's charter".[2] The international freedom of expression organisation, Article 19, has stated that if passed, the Online Safety 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".[3] 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."*[4]

1 The Human Rights Act, EHRC, https://www.equalityhumanrights.com/en/human-rights/human-rights-act
2 Davis, D. These new laws to police the internet are a censor's charter that will have a chilling effect on free speech, Daily Mail Online, 23 June 2021, https://www.dailymail.co.uk/debate/article-9718867/DAVID-DAVIS-new-laws-police-internet-censors-charter.html
3 UK: 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/
4 Government'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-

5. As well as our profound concerns regarding the proposed duties placed on platforms, 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 regulator Ofcom's codes of practice and platforms' terms of use, rather than being clearly set out in primary legislation. It is also our view that the broad notions of harm established by the subsequent regulatory system 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 monitor millions of users in order to enforce increasingly fortified terms of service.

6. In a special committee stage considering only a small number of "re-committed clauses and schedules" the Government removed a key clause regarding lawful online expression considered to be "harmful" to adults. We welcome this move. However, a return to a policy of state enforcement of platforms' terms and conditions is a retrograde step which will also result in the censorship of lawful expression. The Government have also indicated their intention to make further changes to the legislation themselves in the House of Lords.

7. We believe that in the course of its passage through Parliament, this Bill must be materially altered in order to prevent serious damage being done to our rights to freedom of expression and privacy. We believe that **as a minimum, that the revived policy of state enforcement of terms and conditions must be dropped and in this regard Peers should support the amendment in the name of Lord Moylan engaging clause 65. We also believe that that the legislation should not include measures which require or encourage general monitoring and surveillance of people's communications, especially private messages and accordingly, peers should support the amendment in the name of Lord Clement-Jones to clause 110.**

8. This document is not a complete line-by-line analysis of the Bill. However, as Peers commence committee stage scrutiny, this briefing signposts the key threats to human rights which feature throughout the Online Safety Bill.

PART 2 – KEY DEFINITIONS

Clause 3 – Meaning of "regulated service"

9. Part 2, clause 3 sets out the scope of the legislation and describes what constitutes a "regulated service" for the purposes of the regulatory framework. This includes "user to user" and "search" services that have "links to the United Kingdom". Clause 49 sets out those services excluded

online-safety-bill-will-be-catastrophic-for-ordinary-peoples-freedom-of-speech-says-david-davis-mp/

from the scope of the new regulatory system, which include emails, SMS messages, MMS messages, comments and reviews on provider content, one-to-one live aural communications and news publisher content. However, under clause 192 the Government has also reserved the right to extend the duty of care to comments and reviews on provider content as well as one-to-one live aural communications if it is deemed "appropriate" based on the "risk of harm". This could mean, for example, that Zoom could have a duty to surveil calls to impose anti-harm rules.

10. Additionally, clause 192 (3) of the legislation awards the Secretary of State the power, through regulations, to exempt services of a particular description if they deem that the threat of "harm" on such services is low. As is the case throughout the Online Safety Bill, this provision gives the Secretary of State undue power to influence the regulatory framework and if exercised, this power could have serious implications from a markets and competition perspective.

11. The way in which the Bill covers any services with "links to the UK" also brings into question the extent to which the legislation could apply to communications that are sent from overseas but encountered by users in the UK. Given, the international free flow of information and conversation online, this has the potential to bring provisions within the regulatory framework into direct conflict with the laws of those states from which communications (viewed by users in the UK) are sent.

12. For example, the Polish Government have previously proposed a "social media free speech" law. The proposed law would prevent online intermediaries from removing content or banning users who do not break Polish laws.[5] In the event that this legislation and the Online Safety Bill were passed, the ability of social media users in Poland and the UK to communicate directly would be severely hampered by contradictory laws. In such a case, should a user in Poland issue a post on a large social media platform which, although lawful in Poland, could be viewed by users in the UK and deemed as "harmful to children" under the online safety framework, the intermediary in question would be posed with a complex legal dilemma. This could move us towards national digital silos and directly threaten the transnational interconnectedness of the internet as a whole.

13. The scope of the legislation has also been criticised by freedom of expression organisation Article 19, who have raised concerns about the breadth of the regulatory framework. In particular, the group have raised concerns about the

5 Poland proposes social media 'free speech' law, BBC News, 15 January 2021, https://www.bbc.co.uk/news/technology-55678502

extension of the regulatory framework to private messaging services, where it is likely to undermine the privacy guaranteed by end-to-end encryption.[6]

PART 3 – PROVIDERS OF REGULATED USER-TO-USER SERVICES AND REGULATED SEARCH SERVICES: DUTIES OF CARE

Clause 6 - Providers of user-to-user services: duties of care

14. 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 3 of the Bill sets out the new "duties of care" that the legislation places on all in-scope services.

15. Chapter 2 of Part 3 places duties of care on providers of regulated user-to-user services. According to the legislation, all regulated user-to-user services will be obliged to fulfil "illegal content risk assessment" duties, "illegal content" duties, duties regarding reporting and redress, duties regarding freedom of expression and privacy and record keeping and review duties. The legislation also places additional duties on services which "are likely to be accessed by children" and Category 1 (larger services).

16. The notion of duties of care was borne out of a proposal, put together by Professor Lorna Woods and Will Perrin of the Carnegie Trust in 2018/19, on tackling "internet harm". The model proposed a singular duty of care placed upon online intermediaries who would thus be liable for the welfare of online users in a similar vein to workplace health and safety regulations and the obligations an employer has to maintain the safety of employees. In doing so, the proposal cited the 1974 Health and Safety at Work Act.[7] The paper recommended that such a regime should be overseen by an independent regulator and proposed that Ofcom undertake this task.

17. This approach was criticised by civil society groups and members of the legal profession. Internet lawyer Graham Smith warned about the dangers of employing such an approach as a blanket measure for all internet regulation, pointing out that duties of care when it comes to risk of physical injury in public or semi-public spaces are often sector specific. He has also warned of

6 UK: 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/
7 Lorna Woods and William Perrin, "Internet Harm Reduction: a proposal", Carnegie UK Trust Blog, 30 January 2019, https://www.carnegieuktrust.org.uk/blog/internet-harm-reduction-a-proposal/

the unsuitability of this approach where a platform has to take responsibility for and govern the interactions of users.[8]

18. The freedom of expression NGO, Index on Censorship, has also been highly critical of the duty of care model. Arguing that it will put freedom of expression in "peril", the organisation set out its concerns in a paper on the duty of care, asserting that it "will reverse the famous maxim, 'published and be damned', to become, 'consider the consequences of all speech, or be damned'. It marks a reversal of the burden of proof for free speech that has been a concept in the common law of our country for centuries."[9] In a report published by the House of Lords Communications and Digital Committee, which was largely critical of the Government's draft Online Safety Bill, the Committee documented many of the problems with the duty of care model and acknowledged that there are many "legitimate concerns" regarding such an approach.[10]

19. The deployment of a liability model developed in tort law to mitigate risks of objective, physical harms, to regulate and likely curtail free speech, is highly inappropriate. We are deeply concerned by this model which would mark a significant step-change in how free expression is protected in the UK. A duty of care on the part of online intermediaries, which makes platforms liable for the interactions of individuals on the internet, is gravely threatening to free expression. This approach, which is preventative in its outlook, will prove to be excessively censorious as companies will over-zealously remove content in adherence with their obligations.

Clause 8 - Illegal content risk assessment duties

20. Clause 8 constitutes the first substantive duty placed on online intermediaries. The duty requires platforms to undertake "illegal content risk assessments" both in the immediate period following the establishment of the regulatory regime and when making notable changes to the service. Newly appointed regulator Ofcom will also maintain "risk profiles" of platforms and a change to this risk profile will also compel a platform to undertake a further illegal content risk assessment.

21. Whilst greater transparency of online intermediaries is very welcome, these measures may impact upon the extent to which individuals in the UK can access a free flow of information unimpeded. For example, given that the risk assessment duties will apply to any regulated service with "links to the UK",

8 Graham Smith, "Take care with that social media duty of care", Cyberlegal, 19 October 2018, https://www.cyberleagle.com/2018/10/take-care-with-that-social-media-duty.html

9 Right to type: How the "duty of care" model lacks evidence and will damage free speech, 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

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

this means that online services operating in foreign jurisdictions must be compliant with this duty in order to guarantee access to a UK audience. Under such a regulatory burden, it may be the case that online intermediaries based overseas instead opt for UK users not to be able to access their services.

Clause 9 - Safety duties about illegal content

22. Clause 9 sets out a key operational user safety duty, which applies to all regulated services in scope and is central to the legislation. The most prominent subclauses read as follows:

> (*2) A duty, in relation to a service, to take or use proportionate measures to effectively mitigate and manage the risks of harm to individuals, as identified in the most recent illegal content risk assessment of the service.*

23. The clause continues:

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

> *(a) prevent individuals from encountering priority illegal content by means of the service;*
>
> *(b) minimise the length of time for which any priority illegal content is present;*
>
> *(c) where the provider is alerted by a person to the presence of any illegal content, or becomes aware of it in any other way, swiftly take down such content*

24. A list of categories of "priority illegal content" are set out in Schedule 7. The list is thematic and includes offences regarding assisting suicide, threats to kill, public order, drugs and psychoactive substances, firearms and other weapons, assisting illegal immigration, sexual exploitation, sexual images, proceeds of crime, fraud, financial services, other inchoate offences and a range of offences from the devolved nations. The Secretary of State has the power to add to the list through secondary legislation. According to Clause 53 (2) "'Illegal content' means content that amounts to a relevant offence." Apart from in the case of Child Sexual Exploitation and Abuse (CSEA) content, there is no requirement on the platforms to report potential criminal material to law enforcement bodies. This means that victims of crime will have no clear pathway to justice and instead will be reliant on the content in question simply being removed.

25. It is of particular note that the clause calls on platforms to "prevent" content of this nature. This will see the state compel online intermediaries to use

scanning and surveillance technologies in a way that has never been done before in a western liberal democracy. As internet lawyer Graham Smith has observed:

> *"This has a "predictive policing" element, since illegal content includes content that would be illegal content if it were, hypothetically, on the service."*[11]

26. Despite the definitions referenced in paragraph 24, removing so called "illegal content" for the purposes of complying with the regulatory system covers not only that which reaches conviction in a criminal court but anything that a platform determines could be illegal. This marks a significant departure from the rule of law as the provision constitutes the state asking private companies to make determinations on what constitutes illegality and the Bill gives little clarity on how a platform is to determine whether a piece of content is illegal or not. Whilst the identification of illegal material may be clear and obvious in some cases, in many others defining communications of this nature is a complex matter traditionally reserved for law enforcement bodies and the judicial system.

27. In order to remedy this lack of clarity regarding what constitutes potentially illegal content for the purposes of the regulatory framework, the Government additional provisions to the Bill which now feature in clause 170(6). The subclause states that content may be designated as "illegal" if a platform:

> *(a) has reasonable grounds to infer that all elements necessary for the commission of the offence, including mental elements, are present or satisfied, and*
>
> *(b) does not have reasonable grounds to infer that a defence to the offence may be successfully relied upon*

28. "Reasonable grounds to infer" that content could be illegal is significantly below the ordinary burden of proof required to determine that that a crime has been committed, meaning that under this definition, platforms will inevitably be forced to censor entirely lawful speech. What competency social media companies have to make determinations of this kind remains to be seen. Their lack of qualification to do so will cause this flawed regulatory framework to be rife with problems.

29. Further provisions in the Bill oblige Ofcom to issue guidance on how to make such determinations.

11 Smith, G. Mapping the Online Safety Bill, Cyberleagle blog, 27 March, 2022 https://www.cyberleagle.com/

30. The obligation for platforms to determine what constitutes illegality could become problematic around the limitations of free expression. Offences set out in the Public Order Act (1986) criminalise those who "stir up hatred" through their use of "words, behaviour or written material"[12]. These offences have been carefully developed through multiple rounds of rigorous Parliamentary scrutiny in order to protect minority groups. The full rigour of the criminal justice system and referral to established case law are necessary to make a conviction under offences of this nature. Social media companies are not capable of making such a determination.

31. The courts, Crown Prosecution Service (CPS) and the police are all bound by a duty under the Human Rights Act 1998 to act in accordance with the European Convention on Human Rights, including protecting the right to freedom of expression. No equivalent duty falls upon the platforms.

32. The risks to free expression are clear. Under rigorous obligations to protect people from "harm" on their sites, online intermediaries, who are not qualified to establish what constitutes illegal speech will over-remove content on their platforms under the threat of penalties. The consequential impact on free speech will be profound.

33. Writing about the legislation and in particular, clause 9, internet lawyer Graham Smith has said:

> *"It may seem like overwrought hyperbole to suggest that the Bill lays waste to several hundred years of fundamental procedural protections for speech. But consider that the presumption against prior restraint appeared in Blackstone's Commentaries (1769). It endures today in human rights law. That presumption is overturned by legal duties that require proactive monitoring and removal before an independent tribunal has made any determination of illegality."*[13]

34. Introducing obligations of the nature set out in clause 9 also marks a departure from traditional legal standards, held in both the EU and US when it comes to regulating online platforms, which give intermediaries immunity from liability for the user-generated content on their sites in order to protect users' freedom of expression and privacy. 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

12 Public Order Act, 1986, S.18-19 https://www.legislation.gov.uk/ukpga/1986/64
13 Smith, G. Mapping the Online Safety Bill, Cyberleagle blog, 27 March, 2022 https://www.cyberleagle.com/

general monitoring obligations on social media companies operating within their jurisdictions.[14]

Clause 11 - Safety duties protecting children

35. Clause 11 imposes new duties on platforms which are "likely to be accessed by children". This follows clause 10 which creates new risk assessment obligations on platforms which meet the same description and obligations on Ofcom to create platform "risk profiles". A site "likely to be accessed by children" is one which completes a "children's access assessment" and it is determined that children can access the service, where a service fails to perform this requirement or where Ofcom deems that the site in question can be accessed by children following a failure to comply with duties.

36. The provisions within the clause effectively demand that regulated services take responsibility for the safety of children who may access their site. Clause 11 states:

(2) A duty, in relation to a service, to take or use proportionate measures to effectively—

(a) mitigate and manage the risks of harm to children in different age groups, as identified in the most recent children's risk assessment of the service, and

(b) mitigate the impact of harm to children in different age groups presented by content that is harmful to children present on the service.

37. The clause continues:

(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 (for example, by using age verification, or another means of age assurance);

(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 (for example, by using age assurance).

38. Content that is harmful to children is defined broadly as "of a kind which presents a material risk of significant harm to an appreciable number of children in the United Kingdom". According to the legislation "'Priority

14 Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market ('Directive on electronic commerce') https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32000L0031

content that is harmful to children' means content of a description designated in regulations made by the Secretary of State". How children may be prevented from encountering content of this kind must be set out in platforms' terms of use and "applied consistently". Once again the Bill compels platforms to "prevent" children from encountering content of this kind which will lead to greater levels of monitoring and surveilling users' activity in order to fulfil these preventative policing-style obligations.

39. The Online Safety Bill suffers throughout 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 manifestly illegal content online, this Bill seeks to eradicate broad-brush concepts of harm, which would result in a more restricted online experience for everyone.

40. Elaborating on the application of these duties on online intermediaries, the Bill states:

> *a provider is only entitled to conclude that it is not possible for children to access a service, or a part of it, if there are systems or processes in place (for example, age verification, or another means of age assurance) that achieve the result that children are not normally able to access the service or that part of it.*

41. 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. The Government have pledged to make this provision more explicit.[15]

42. 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.

43. Such a measure would also mean that internet users would have to volunteer even more personal information to the platforms themselves, which could 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.

15 New protections for children and free speech added to internet laws, DCMS, 28 November 2022, https://www.gov.uk/government/news/new-protections-for-children-and-free-speech-added-to-internet-laws

The Bill should not force online intermediaries to demand ID for internet access.

Clause 12 – User empowerment duties

44. Clause 12 sets out a number of new so-called "user empowerment duties". These measures are an attempt at creating quick, tech-based fixes to deeper problems embedded in large social media companies' business models and society at large. While granting users greater autonomy over their online experiences is a good thing, aspects of this clause actively undermine fundamental rights.

45. Clause 12 (2) states that platforms must comply with:

(2) A duty to include in a service, to the extent that it is proportionate to do so, features which adult users may use or apply if they wish to increase their control over content to which this subsection applies.

46. In an announcement regarding changes to the Bill late last year, the Government committed to developing the Bill's user empowerment duties further[16] as a substitute for removing provisions regarding so called "legal but harmful" speech online. Following amendments made in a special committee particular, category 1 platforms must now give users the option to block specified types of content where:

it encourages, promotes or provides instructions for—

(a) suicide or an act of deliberate self-injury, or

(b) an eating disorder or behaviours associated with an eating disorder.

47. Or:

if it is abusive and the abuse targets any of the following characteristics—

(a) race,

(b) religion,

(c) sex,

(d) sexual orientation,

(e) disability, or

(f) gender reassignment.

16 New protections for children and free speech added to internet laws, DCMS, 28 November 2022, https://www.gov.uk/government/news/new-protections-for-children-and-free-speech-added-to-internet-laws

48. Or finally:

if it incites hatred against people—

(a) of a particular race, religion, sex or sexual orientation,

(b) who have a disability, or

(c) who have the characteristic of gender reassignment.

49. What is particularly curious about the descriptions set out above is that many of these categories include forms of expression which are already illegal. For example encouraging suicide is a criminal offence under the Suicide Act 1961 and the Government have also stated their intention to make encouraging self-harm a criminal offence too[17]. Further, similar to the Bill's description of content which "incites hatred against people", the Public Order Act 1986 criminalises expression which stirs up hatred against groups of people with protected characteristics. It is unclear why users should simply be able to given options to block expression of this nature when it may in fact be manifestly illegal.

50. Clause 12 (6) also states that platforms must comply with:

A duty to include in a service features which adult users may use or apply if they wish to filter out non-verified users.

51. This is a flawed approach and treats online anonymity as inherently "unsafe". There is little evidence to suggest that anonymity itself makes online discourse more febrile. It is clear that MPs receive unacceptable abuse online. However, according to analysis conducted by the New Statesman Magazine which involved tweets sent to MPs since January 2021, there was little discernible difference in the nature or tone of the tweets MPs receive from anonymous or non-anonymous accounts.

52. While 32 per cent of tweets from anonymous accounts were classed as angry according to the metric used by the New Statesman, so too were 30 per cent of tweets from accounts with full names attached.[18] Similarly, 5.6 per cent of tweets from anonymous accounts included swear words, only slightly higher than the figure of 5.3 per cent for named accounts.[19] While there is no doubt that people communicate differently online from how they do in person, there is little evidence to suggest that behaviour differs substantially based on whether an individual is anonymous or not.

17 New protections for children and free speech added to internet laws, UK Government press release, 28 November 2022, https://www.gov.uk/government/news/new-protections-for-children-and-free-speech-added-to-internet-laws

18 van der Merwe, B. Are anonymous accounts responsible for most online abuse?, New Statesman, 21 October 2021, https://www.newstatesman.com/social-media/2021/10/are-anonymous-accounts-responsible-for-most-online-abuse

19 Ibid.

53. Neither is online anonymity or pseudo-anonymity a barrier to tracking down and prosecuting those who commit criminal activity on the internet. Police reporting shows that in 2017/18, 96% of attempts by public authorities to identify the anonymous user of a social media account, email address or telephone, resulted in successful identification of the suspect of their investigation.[20]

54. The police already have a range of intrusive powers to track down individuals online. The Investigatory Powers Act 2016 allows police to acquire communications data such as an email address and location of the device from which alleged illegal anonymous activity is conducted and use this data as evidence in court.

55. Despite the lack of necessity, this provision will force all companies in scope to create two-tiers of users, those who are willing to be verified and those who are not.

56. Pseudo-anonymity is vital to many minority groups and a tool in the armoury of those who want to hold the powerful to account. Many LGBT people, particularly those who are not "out", may choose to navigate the internet anonymously in order to give themselves the freedom to explore their identity without disclosing this to anyone. Online anonymity is also important to many survivors of sexual violence or domestic abuse, who might prefer to seek support without revealing their identity. Anonymity is also crucial to the work of journalists, human rights activists and whistleblowers in the UK and all around the world. Attempts to undermine online anonymity in the UK would set a terrible precedent, likely to be emulated by authoritarian governments in other jurisdictions. This measure, would treat all users as described above as second-class citizens online.

57. This provision is unlikely to do anything to keep those who are verified "safe" online, but it could prevent those who rely on anonymity from accessing others who do disclose their identity. For example, a high-profile LGBT role-model may inadvertently prevent anonymous LGBT people exploring their sexuality from accessing their account if this measure was widely-used across major platforms.

Clause 13 - Duties to protect content of democratic importance

58. 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", "new publisher content" and "journalistic content".

20 Original Government response to "Make verified ID a requirement for opening a social media account", Parliamentary Petition, https://petition.parliament.uk/petitions/575833

The Government claims 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.

59. The first of these provisions, clearly borne out of concern that platforms could reprimand politicians in a similar way to former President Trump, oblige intermediaries to take into account whether content is of "democratic importance" when moderating content. It calls on platforms to ensure that their systems and processes "apply in the same way to a wide diversity of political opinion".

60. According to the Bill, content of democratic importance is that where "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."

61. The vague nature of this categorisation will only create additional complications for the platforms as they are simultaneously told to deal with 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.

Clause 14 - Duties to protect news publisher content

62. Out of recognition for the impact that the legislation still threatens to have on the freedom of the press, the Government introduced clause 14 as an amendment to place further duties on platforms to protect "publisher content".

63. The clause obliges platforms to take prescribed steps when "taking action" against news publisher content including setting out why the platform had taken the action in question, "how the provider took the importance of the free expression of journalistic content into account when deciding on the proposed action" and detail regarding how the new publisher may appeal the decision. These steps are not necessary if the platform thinks they may incur "criminal or civil liability" for the content in question. However this extensive protection for news publisher content does beg the question why such protections cannot be afforded to ordinary users online.

Clause 15 - Duties to protect journalistic content

64. The carve-out in clause 15 requires online platforms to consider whether content is "journalistic" when enforcing their terms of use, and to create an expedited appeals process for the reinstatement of removed journalistic content. These appeals processes apply not only to the creator of the content in question but also those sharing it.

65. The legislation defines "journalistic content" in 15 (9) as follows:

> *(a) the content is—*
>
> *(i) news publisher content in relation to that service, or*
>
> *(ii) regulated user-generated content in relation to that service;*
>
> *(b) the content is generated for the purposes of journalism; and*
>
> *(c) the content is UK-linked*

66. Given the breadth of this definition it is clear that the platforms will retain a large degree of power to designate which types of content are "journalistic". It is not apparent how independent 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 journalism and more disintermediated news sharing. If carve-outs are only afforded to the journalists and media operators that social media companies choose, an unhealthy monopolisation will be quick to return.

Clause 16 - Duty about content reporting

67. Clause 16 of the Online Safety Bill creates new mandated reporting obligations which all in-scope platforms will have to adhere to in some form. The clause states that intermediaries will have:

> *(2) A duty to operate a service using systems and processes that allow users and affected persons to easily report content which they consider to be content of a kind specified below (with the duty extending to different kinds of content depending on the kind of service, as indicated by the headings).*

68. This covers mandatory reporting mechanisms for content which is deemed to be illegal on all services and reporting mechanisms content which could be harmful to children on platforms that are "likely to be accessed by children" .

Clause 17 - Duties about complaints procedures

69. Clause 17 creates a duty on intermediaries to ensure that complaints systems are integrated into their systems and processes in order that they fulfil the aforementioned duties. This includes the mandating of complaints procedures

for users who have had their content removed because the service provider in question believed that it may be illegal or harmful to children.

70. Whilst a duty on platforms to integrate appeals processes into their processes is a welcome step when it comes to protecting freedom of expression online, the reality is that many platforms already offer this function, which in many cases lacks transparency or rigour. There is nothing in the legislation about improving or setting minimum standards for these appeals processes. Further, this measure will make little difference if the bar for what is considered acceptable online is considerably lowered. The duty also requires platforms to create complaints processes for users who believe a platform is in breach of their operational safety duties.

Clause 18 - Duties about freedom of expression and privacy

71. Clause 18 constitutes an attempt on the part of the Government to in some way balance the damage to individuals' rights to freedom of expression and privacy as a result of the Bill. It is a weak duty which will do nothing to protect these rights. The duty set out in the clause is written as follows:

> *(2) When deciding on, and implementing, safety measures and policies, a duty to have regard to the importance of protecting users' right to freedom of expression within the law.*

72. Unlike the previously considered operational safety duties, which compel companies to "prevent" and "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.

73. 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 these fundamental rights.

Clause 19 - Record-keeping and review duties

74. Clause 19 obliges platforms to keep a record of all risk assessments conducted to comply with duties in the previous clauses and to keep a record of steps taken to comply with duties that are not described in the codes of practice. It is without doubt that greater levels of accountability and transparency from online intermediaries are needed.

Clause 20 - Providers of search services: duties of care

75. Clauses 20-29 replicate many of the provisions previously set out in clauses 8-19 but for search services as opposed to user-to-user services.

76. The right to freedom of expression in an online setting not only concerns the ability of individuals to impart information but also to receive it. In this regard, a free flow of information and the right to freedom of expression go hand in hand.

77. Clauses 20-29 transpose many of the duties set out in Part 3 for user-to-user services and apply them to search services. This includes illegal content risk assessment duties, risk assessment duties specifically for services "likely to be accessed by children", safety duties relating to potentially illegal content, safety duties where the service is "likely to be accessed by children" as well as content report and complaints duties. Unlike with user-to-user services, there is no stipulation of obligations based on the size of the service in question and as such, no additional duties for larger services.

78. Given that search services are frequently used for educational purposes, the idea that such a service is "likely to be accessed by children" is high. The duties set out in this section threaten to age-gate these services and exclude those unwilling or unable to verify themselves from using them.

79. The legislation imposes further duties upon search services to "have regard to" freedom of expression and privacy but these are weak checks on an otherwise deeply restrictive model. Reporting, redress and record-keeping duties also apply.

80. As with user-to-user services, we are deeply concerned that the broad definitions and the weight of the obligations placed upon these intermediaries will mean that search services feel obliged to censor heavily. As such, this runs the risk of stifling the free flow of information online. This is a retrograde step given the otherwise democratising power of the internet.

81. Of all of the major digital markets, the field of search engines is among the most monopolised, with Google overwhelmingly acting as the major market player. Given the expensive regulatory costs of the proposed online safety regime, far from taking power from online platforms like Google, this legislation will obstruct market entry to potential new services and entrench powerful actors such as Google.

Clause 36 - Codes of practice about duties

82. Building on the duties of care, clause 36 instructs the newly appointed regulator, Ofcom, to draft codes of practice setting out how social media companies can fulfil their obligations when it comes to regulating content

that is deemed to be illegal or " harmful". Compliance with the relevant duties is met if a platform takes the steps set out in the codes of practice, which they will have to integrate into their company's "systems and processes".

83. The effect of this step is to fortify social media companies' terms of use, ensuring that they are upheld, and to clearly identify companies that fail to comply, who risk sanction. It would seem a controversial position for a government-appointed regulator to oversee private companies in effectively upholding those terms and conditions – sets of rules that are not neutral, and which have complex implications.

84. In drafting the codes of practice, Ofcom must consult with the Secretary of State amongst others and Parliamentary approval comes in the form of a negative resolution of the House which affords minimal scrutiny. Once again this brings into question the independence of the regulator the significant influence the executive will have to influence the regulatory framework.

39 - Secretary of State's powers of direction

85. A running theme throughout the entirety of the Online Safety Bill is the way in which the Government awards itself a huge amount of executive power to shape this proposed system of online speech moderation and as a result, to influence discourse.

86. Clause 39 typifies this level of executive control as it gives the Secretary of State of the day the power to effectively influence Ofcom's codes of practice, which set out how intermediaries can reach compliance with the relevant duties. The key element of the clause is set out as follows:

> *(1) The Secretary of State may direct OFCOM to modify a draft of a code of practice submitted under section 38(1) if the Secretary of State believes that modifications are required—*
>
> *(a) for reasons of public policy, or*
>
> *(b) in the case of a terrorism or CSEA code of practice, for reasons of national security or public safety.*

87. Ofcom must comply with this direction.

88. This is incredibly dangerous and opens the entirety of this flawed system up to politicisation. The Secretary of State's power of direction would allow the Government to pressure Ofcom into writing codes of practice that would shape the permissibility of categories of online content based on the political mood.

89. It is wholly inappropriate for our right to free expression to be curtailed by secondary legislation which is unamendable and allows for little parliamentary oversight. In these circumstances, the power exercised by the online regulator and Secretary of State would bypass the full democratic process, creating a two-tier speech system whereby the increasingly ubiquitous online tier would be, for all intents and purposes, untethered from decades of existing law and highly susceptible to political swings of the day. This situation is precisely what Government should be seeking to prevent – not endorse.

90. This was a problem recognised by the joint committee of both Houses of Parliament who undertook pre-legislative scrutiny of the Bill. Addressing this clause in their report on the draft Online Safety Bill, the Committee said:

> *"The powers for the Secretary of State to a) modify Codes of Practice to reflect Government policy … give too much power to interfere in Ofcom's independence and should be removed."*[21]

91. Concerns about this provision were also raised during committee stage in the House of Commons when SNP DCMS spokesperson John Nicolson said:

> *"The regulator must not be politicised in this way. Regardless of the political complexion of the Government, when they have too much influence over what people can say online, the implications for freedom of speech are grave"*[22]

This recommendation was not adhered to by the Government in drafting the full Bill and as such this power must be removed by Parliament.

Clause 65 - Further duties about terms of service

92. The obligation to tackle expressly lawful expression deemed to be "harmful to adults" was without doubt the most controversial provision set out in the Online Safety Bill due to the damage that it threatened to do to free speech in the UK.

93. During last year's Conservative Party leadership contest, many of the candidates identified the problems with this provision leading the Government to review its inclusion in the Bill.[23] Following months of speculation, the Secretary of State, Michelle Donelan, announced her intention to remove the "legal but harmful" clause from the legislation and

21 Joint Committee on the Draft Online Safety Bill, Report of Session 2021–22, 10 December 2021, https://committees.parliament.uk/publications/8206/documents/84092/default/

22 Online Safety Bill, House of Commons Official Report, 14 June 2022, p. 354, https://publications.parliament.uk/pa/bills/cbill/58-03/0004/PBC004_OnlineSafety_1st17th_Compilation_29_06_2022.pdf

23 Hymas, C. Free speech online will be protected with amended duty of care law, say Sunak and Truss, Daily Telegraph, 19 August 2022, https://www.telegraph.co.uk/politics/2022/08/19/free-speech-online-will-protected-amended-duty-care-law-say/

this clause was then removed when the Bill entered a special committee in mid-December.

94. Whilst the removal of the "legal but harmful" provisions are welcome, the Government have replaced this measure with a new clause which has serious ramifications for freedom of expression.

95. Clause 65, "Further duties about terms of service", imposes a duty on Category 1 services to have proportionate systems and processes in place to take down or restrict access to content, and to ban or suspend users, in accordance with their terms of service (subclause 3); and that the terms regarding the take down or restriction of access to content, and the banning or suspension of users, must be clear, accessible and sufficiently detailed (subclause 4, paragraph a) and applied consistently (paragraph b).[24] As such, cl.65 imposes a duty on widely used services to consistently apply the terms of service that deal with suppression and censorship of speech and users.

96. Whilst it is reasonable that members of the public expect that private companies uphold their terms of service on speech, that does not justify the transformation of these private agreements into statutory duties. Likewise, we would not expect the Big Tech companies' terms of service on privacy that deal with their exploitative data collection and targeted advertising practices to be transformed into statutory duties. Rather, we look to UK regulators to protect the individual rights users have in UK domestic law on these platforms, such as the Data Protection Act, Human Rights Act and Equality Act. In reinforcing censorship policies on the digital public square that do not exist in UK law, clause 65 creates serious legal friction, unintended consequences and human rights issues.

97. In practice, major platforms' terms of service are extensive and enable companies to suppress and censor vast categories of lawful speech, as well as suspend and ban users for expressing such lawful speech.[25] As such, clause 65 may be considered to have some similar features to the now-removed "legal but harmful" powers, giving state regulators the role of ensuring types of lawful speech are suppressed online. However, clause 65 goes further to task OFCOM with ensuring that individuals who express lawful speech are suspended or banned from platforms where terms of service allow, thereby limiting those individuals in expressing themselves more widely beyond the speech in question, incurring a far wider interference with those individuals' right to freedom of expression.

24 Online Safety Bill Amendment Paper, 8th December 2022: https://publications.parliament.uk/pa/bills/cbill/58-03/0209/amend/onlinesafety_rm_pbc_1208.pdf

25 For examples, see The State of Free Speech Online by Big Brother Watch, September 2021: https://bigbrotherwatch.org.uk/wp-content/uploads/2021/09/The-State-of-Free-Speech-Online-1.pdf

98. The draft Online Safety Bill contained a clause (cl. 11(3)(b)) requiring that platforms' terms of service regarding "harmful" content were applied consistently.[26] The revised Bill slightly narrowed the duty, requiring that all terms of service relating to "priority content that is harmful to adults" would be applied consistently.[27] It is important to note that this "harmful" content was to be defined by the Secretary of State, whilst the terms dealing with it were up to the platform. As previously discussed, the Government have now removed this provision, thereby removing the legal invention of speech that is "legal but harmful for adults",[28] which we welcome, due to the clear human rights conflicts that such a construction gave rise to.

99. However, clause 65 introduces a wider duty about terms of service on platforms, with a wide interference on freedom of expression, as the duty applies to all terms of service whatever they may be regarding the platform's policies on speech suppression and censorship, and user suspensions and bans. This duty is not restricted to so-called "harmful" content as per the previous Bill, but whatever content the platform wishes to suppress or censor, and whatever people the platform wishes to suspend or ban.

100. Platforms have no obligation to align their terms of service with freedom of expression criteria as per Article 10.2 of the European Convention on Human Rights, but must only have "regard' to "the importance of" users' freedom of expression (cl.18). Putting the word "particular" before the word "regard", as the Government proposes adding, makes little material difference to this position. As noted by leading free expression barrister, Gavin Millar KC:

"It is not clear what right is being referred to in these clauses [cl.18]. But they appear to reflect the old common law maxim that a citizen can say anything provided it is not prohibited by law. It is certainly not a reference to the Article 10 right which is fundamental, supranational human right that is not dependent on compliance with domestic laws. The value of these provisions in protecting free speech online is again extremely limited."[29]

101. Platforms' terms of service can, and frequently do, change according to changing management and external political trends. In general, corporate terms of service are designed to protect platforms' business interests and

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

27 Online Safety Bill as of December 2022, cl. 13(6)(b): https://publications.parliament.uk/pa/bills/cbill/58-03/0209/220209.pdf

28 Online Safety Bill Amendment Paper, 8th December 2022, Amendment 7: https://publications.parliament.uk/pa/bills/cbill/58-03/0209/amend/onlinesafety_rm_pbc_1208.pdf

29 A legal analysis of the impact of the Online Safety Bill on freedom of expression – Gavin Millar KC for Index on Censorship, May 2022 - https://www.indexoncensorship.org/wp-content/uploads/2022/05/Legal-analysis-of-the-impact-of-the-Online-Safety-Bill.pdf

legal protection, and as such give platforms absolute power over their content policies – in the case of services hosting user-generated content, that means absolute power over what users can and cannot say.

102. Category 1 services typically have terms of service that permit the suppression of speech far beyond the limitations on speech in UK law. As such, clause 65 shows a worrying lack of commitment to the UK's laws and case law on free speech that have evolved over many years. For example:

a) under Twitter's policy against "misgendering", gender critical feminists, trans people and any other commentators can be and have been censored, suspended or banned for using words such as "cis", "TERF", "guy" and "dude"[30] (Twitter's policy also defines the relevant protected characteristics as "gender" and "gender identity",[31] not "sex" and "gender reassignment" as per the Equality Act 2010).

b) under Facebook's community guidelines, social or political exclusion on the basis of what the company calls "protected characteristics"[32] is prohibited.[33] This gives Facebook latitude to prohibit black-only political groups, or a social group only for black women, or other single-sex social groups.

c) under Facebook's community guidelines, generalisations about groups inferring inferiority are prohibited.[34] Under this policy, women can be and have been censored and suspended for quips such as "men are so stupid," whilst a black activist was censored for describing white people as "fragile"[35] (at the time, the book White Fragility was a New York Times bestseller).

d) under YouTube's community guidelines during the pandemic, any content that contradicted "health authorities" was prohibited. Under this policy, a speech by David Davis MP at Conservative Party conference, in which he criticised the Government's Covid pass policy, was removed from the platform.[36] Major sections of the Labour Party, Liberal Democrats and Green Party also opposed mandatory Covid passes during the pandemic. The video

30 For examples, see The State of Free Speech Online by Big Brother Watch, September 2021, pp.53-60: https://bigbrotherwatch.org.uk/wp-content/uploads/2021/09/The-State-of-Free-Speech-Online-1.pdf
31 Hateful Conduct Policy, Twitter, last accessed 8th December 2022: https://help.twitter.com/en/rules-and-policies/hateful-conduct-policy
32 Facebook's definition of 'protected characteristics' is out of sync with the Equality Act 2010 and UK hate crime, including not only the five protected groups of disability, race (and national origin), religion, sexual orientation and gender reassignment as per the UK's hate crime definition, but also sex, gender, caste and serious disease.
33 Hate Speech Community Standards, Facebook, last accessed 8th December 2022: https://transparency.fb.com/en-gb/policies/community-standards/hate-speech/
34 Ibid.
35 For examples, see The State of Free Speech Online by Big Brother Watch, September 2021, p.18 and p.21: https://bigbrotherwatch.org.uk/wp-content/uploads/2021/09/The-State-of-Free-Speech-Online-1.pdf
36 YouTube U-turns over David Davis vaccination passports clip after protest, BBC News, 14 October 2021, https://www.bbc.co.uk/news/uk-politics-58915092

was reinstated after he complained, but his viewpoint did indeed contradict authorities at the time.

103. Under clause 65, content moderation of this nature would no longer be simply a contract between the company and the user but would be brought under statute; Big Tech companies would be exercising public law functions by consistently suppressing such lawful speech; and OFCOM would be tasked with ensuring that the policies above are enforced consistently.

104. This displaces the UK's legal standards for permissible speech online with platforms' terms of use.

105. The provisions above are an egregious threat to freedom of expression in the UK. The notion of state-backed censorship of lawful expression contravenes long-held human rights standards on protecting freedom of speech. The state should not curtail or endorse the censorship of expression which is lawful. Limitations on free speech should be exercised only where necessary, where they are proportionate and where they are clearly prescribed in law.

In order to protect freedom of expression and limit the possibility of overzealous or censorious enforcement, restrictions on permissible speech should always be clearly defined in law, not in platforms terms of use with the state acting as a guarantor. Accordingly, Peers should support the amendment to clause 65 in the name of Lord Moylan.

Clause 68 - Transparency reports about certain Part 3 services

106. Clause 68 (1) states:

> *(1) Once a year, OFCOM must give every provider of a relevant service a notice which requires the provider to produce a report about the service (a "transparency report").*

107. This is welcome and could empower users to gain a greater understanding of how large social media platforms operate.

PART 5 - DUTIES OF PROVIDERS OF REGULATED SERVICES: CERTAIN PORNOGRAPHIC CONTENT

Clause 70 - Duties about regulated provider pornographic content

108. The scope of the Online Safety Bill has been widened from its previous draft formation to include commercial pornography websites. This follows an attempt by the Government to create digital ID checks for those viewing adult

content of this nature through the Digital Economy Act 2017. Despite passing the Act into law, the measures set out in the legislation regarding the age-gating of pornographic material were abandoned due to inherent flaws with the legislation when it came to protecting privacy. Unless the Government addresses these issues, it will be confronted with the same problems once again.

109. Clause 70 (2) states that these platforms will have:

> *A duty to ensure that children are not normally able to encounter content that is regulated provider pornographic content in relation to the service (for example, by using age verification).*

110. According to the Bill's explanatory notes "provider pornographic content":

> *"is pornographic content which is published or displayed on a service by the service provider itself, or an individual acting on behalf of the service provider. It does not include user-generated content".*

111. While it can be argued that age-gating pornographic content is of secondary concern when it comes to safeguarding human rights, there are inherent privacy issues at hand. The collection of identity documents or biometric data for access to adult-content websites is a recipe for disaster, matching personal identifiers with adults' viewing habits. Not only does this risk compromising intimate elements of individuals' private lives but it poses a threat to members of the LGBT community who may not be "out" and openly willing to reveal their sexual preferences.

112. We recognise the need to regulate pornographic content and to do so in a way which prevents children from accessing material of this kind. However, as per the highlighted risks set out above, such a scheme cannot proceed without embedding serious privacy safeguards in its application.

113. For example, as Open Rights Group have observed, GDPR has provided a number of safeguards when it comes to data protection, but it does not, on its own, protect information that is as potentially revealing as a person's pornographic viewing history.[37] The organisation has set out other minimum standards for achieving a system which is safe and secure and argues that in order to safeguard individuals' privacy age verification systems should:

> *"process the minimum personal data necessary to verify your age; additional personal data should not be collected, irrespective of whether it is subsequently securely deleted. Personal data must not be kept for longer*

37 Age Verficiation Facts, Open Rights Group, https://www.ageverificationfacts.org.uk/over-18s/

than is necessary to achieve the purpose of age verification, and must not be used for other purposes, such as advertising."

114. Similar provisions set out in the Digital Economy Act delegated responsibility for this area to the British Board of Film Classification (BBFC) in line with their other responsibilities to regulate in this area. However, the BBFC's certification scheme for providers of age-verification technologies was voluntary, which would have resulted in non-secure providers using this new compelled system to harvest individuals' most sensitive personal data.

115. Unless these problems are addressed, the system will suffer from the same flaws and will create inherent privacy risks for adults online.

PART 7 - OFCOM'S POWERS AND DUTIES IN RELATION TO REGULATED SERVICES

Clause 83 - Duties in relation to strategic priorities

116. Clause 83 sets out further executive powers at the disposal of the Secretary of State; the ability to issue a statement of strategic priorities which Ofcom must "have regard to". Coupled with a catalogue of similar executive powers issued throughout the legislation, this provision means that the Secretary of State will have an excessive level of influence over the newly appointed regulator, Ofcom, and as such, the limitations of expression online.

Clauses 91-109 - Powers to require information

117. Ofcom's powers to require information, as set out in clauses 91-109 constitute a mechanism by which the regulator can investigate and issue penalties against companies for any non-compliance with the new regulatory regime. This function is performed by issuing an intermediary with an "information notice" as set out in clause 92.

118. Clause 91 requires regulated services to name a designated "senior manager" upon request. Such an individual is then bound by reporting obligations to the regulator and takes on a degree of personal liability for the conduct of the organisation in dispatching its relevant duties.

119. Clause 99 states senior managers would commit an offence if they fail to comply with an information notice. Penalties include a custodial sentence of up to 2 years or a fine.

120. During the Bill's passage through the House of Commons, a number of Conservative MPs sought to extend this provision further, so that tech executives would have direct liability for any so-called harmful material on their sites under the threat of custodial prison sentences.[38]

38 Tech bosses could face jail after Tory MPs revolt on bill

121. At a risk of serious punishment from individual criminal liability, platforms will endeavour to unscrupulously remove content on their sites. Coupled with broad definitions and a low threshold of acceptable expression, these measures would guarantee widespread censorship online.

122. These measures create a devastating example internationally and will embolden authoritarian actors around the world to impose criminal liability on companies' senior management. These powers could be read to justify the imprisonment of social media executives overseas, a practice which is already being undertaken or threatened by some state actors, including the governments of China, India, Russia and Turkey[39].

123. Of particular concern is where the legislation states that a person or senior manager would commit an offence, if in response to being issued an information notice, the person:

> *(a) provides information which is encrypted such that it is not possible for OFCOM to understand it, or produces a document which is encrypted such that it is not possible for OFCOM to understand the information it contains*

124. This is a deeply problematic subclause and will have the effect of dissuading online intermediaries from encrypting services (particularly messaging services). Encryption is used by a variety of services to keep information private and should not be seen as inherently harmful but actually something that is used to keep individuals safe and secure online. This move would likely be welcomed by malign actors around the world.

125. The Government have also made changes to these clauses which would give the regulator the power to issue an information notice to inform whether the regulated service may be instructed to use scanning technology to detect certain kinds of illegal content.

Clause 110 - Notices to deal with terrorism content or CSEA content (or both)

126. The Bill makes repeated references to different types of "technology" that regulated services may use to guarantee compliance with their relevant duties. This is often an endorsement of algorithmic content moderation tools which surveil users' online activity and make blunt, inaccurate and often biased judgements on the permissibility of online expression.

127. Clause 110 sets out a mechanism for Ofcom to mandate online intermediaries to use technology of this kind. 110 (2) states:

Published, BBC News, 17 January 2023, https://www.bbc.co.uk/news/uk-politics-64298338
39 Dixit, P. Twitter Unblocked Accounts That Criticized India's Government. Now, Its Employees Are Being Threatened With Jail Time Unless It Blocks Them Again, Buzzfeed, 3 February 2021, https://www.buzzfeednews.com/article/pranavdixit/india-threatens-twitter-jail

A notice under subsection (1) that relates to a regulated user-to-user service is a notice requiring the provider of the service to do either or both of the following—

(a) use accredited technology to identify terrorism content communicated publicly by means of the service and to swiftly take down that content;

128. This provision constitutes the state forcing a platform to scan everything users post on that site, using tech companies as privatised online police, in search of "terrorism content". Genuine terrorist material online, which constitutes a security threat to the public, should not be only dealt with by companies in Silicon Valley but the police and other security bodies.

129. Schedule 5 sets out the suspected terrorism offences that platforms must use technology to scan for. They include section 12(1A) of the Terrorism Act 2000 which makes it a criminal offence to express an opinion or belief supportive of a proscribed organisation) and 13(1A) of the Terrorism Act 2000, which makes it a criminal offence to publish an image of the uniform of proscribed organisation. These are complicated offences which law enforcement bodies and courts must make careful judgements on, balanced against their obligations set out in human rights law. They are not offences which Silicon Valley companies' algorithmic systems can definitively identify. This measure will result in the mass surveillance of users online and have a collateral impact on free expression.

130. For the purposes of this provision, the Bill retains some ambiguity about the definition of the word "publicly". Whilst clause 203 offers some guidance on this, it is not apparently clear that the measure could not apply to a large group communicating via a private messaging service.

131. In fact, the Bill makes clear that private messaging channels are not exempt from the scope of the legislation and are therefore bound by many of the duties set out in the Bill, including the provisions set out in clause 110. This is a dangerous direction and will result in growing surveillance online, even in spaces intended for users to hold a private conversation.

132. 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.

133. This incompatibility is laid out explicitly where the Bill grants Ofcom the power to compel intermediaries operating private messaging services to surveil their users using scanning technology. Clause 110 (2) states:

> *A notice under subsection (1) that relates to a regulated user-to-user service is a notice requiring the provider of the service-*
>
> *(b) use accredited technology to identify CSEA content, whether communicated publicly or privately by means of the service, and to swiftly take down that content.*

134. Regulated services do have an opportunity to appeal a notice served under this provision.

135. Government changes to this section also reintroduced Ofcom's ability to issue a "warning notice" and widen the provisions in this area to cover search services in a further expansion of the the Bill's state-mandated private-actor fulfilled surveillance powers.

136. It is vital that terrorism and CSEA content are removed from the internet. However, tackling such content does not require entire encrypted channels to be compromised, sacrificing the security, safety and privacy of billions of people. 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.

137. 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.[40] It is not unreasonable to expect that such technology would be escalated in time, put to use in other areas and result in increased surveillance of individuals' private messages.

138. The legal powers in the Bill appear to create mechanisms to mandate this technology at a mass, suspicion-less scale. This is the view of a leading human rights and technology barrister, Matthew Ryder KC, expressed in a legal opinion commissioned by Index on Censorship. In the opinion, Ryder set out his view that measures in the Bill would grant Ofcom a wider remit on mass surveillance powers of UK citizens than bodies such as GCHQ.[41] This would undermine the presumption of innocence and the principle that surveillance should be based on suspicion.

---

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

41Legal opinion by Matthew Ryder KC and Aidan Wills on the human rights implications of client-side scanning, November 2022: https://www.indexoncensorship.org/wp-content/uploads/2022/11/Surveilled-Exposed-Index-on-Censorship-report-Nov-2022.pdf

139. As with other areas of the Bill, one of the 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.

140. 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.

In order to protect the right to privacy, powers to force intermediaries to scan private messaging channels should be removed from the Bill. Accordingly, Peers should support the amendment to clause 110 in the name of Lord Clement-Jones.

Clauses 118-130, confirmation decisions and penalty notices

141. Clauses 118-130 set out the processes at the disposal of Ofcom to police the regulatory regime and issue penalties for non-compliance with the duties set out in previous parts, including operational safety duties.

142. Most notably in this section, clause 124 sets out one of the most chilling measures in the entire Bill when it comes to damaging individuals' privacy online. The provision gives Ofcom the power to mandate the use of "proactive technology" to identify and remove any kind of content the platform believes could be illegal or content which is deemed to be harmful to children.

143. Once again, this would result in the unprecedented surveillance of all activity of potential millions of users, on a suspicionless basis, on the platform in question. This is an entirely disproportionate response to the problem at hand.

144. Furthermore, these kinds of proactive technologies which use AI to scan and detect expression or images often have high rates of inaccuracy and incorporate a range of systemic biases, making them inappropriate tools for identifying illegal or harmful content in contexts where their decisions directly impact individuals' freedom of expression.

145. Accompanying clauses 118-130, Schedule 13 sets out that a failure on the part of a platform to fulfil its relevant duties of care could result in a fine of up to £18m or 10 per cent of annual global turnover, depending on which is higher.

146. It is unprecedented for the Government to seek to punish technology companies for essentially failing to act as effective law enforcement auxiliaries and even for failing to censor or demote lawful content. Given the financial and reputational costs that could be incurred if these proposals go ahead, there will be a chilling effect that will motivate companies to monitor, demote and censor expression over-zealously.

Clauses 131-135 – Business disruption measures

147. In terms of penalties, the Bill goes even further and clauses 131-135 give Ofcom license to seek service restriction orders (e.g. forced removal from the app store) or Access Restriction Orders (ISP blocking), either of which must be approved in court. The proposal for search engine, intermediary and ISP blocking is severe and is a fundamental threat to free expression.

148. Clause 131 gives Ofcom the power to apply to a court for a service restriction order. This can be sought where a company fails to comply with its obligations under the relevant duties or where there is a failure that is coupled with a "risk of harm" to individuals using that service.

149. Such measures would target ancillary services which support the platform in question and could include hosting providers or ad servers. Clause 132 allows Ofcom to apply for such an order on an interim basis.

150. Clause 133 gives Ofcom the power to seek, from a court, permission to impose an access restriction order, where a service restriction order "was not sufficient to prevent significant harm arising to individuals in the United Kingdom" or if issuing a service restriction order is not deemed sufficient to prevent "harm".

151. This involves the full blocking of a service so that it may not be accessed by users in the UK. Clause 134 gives Ofcom the power to seek such a measure on an interim basis.

152. 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 Chinese-style 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 is at odds with liberal democratic values.

153. Concerns about service restriction orders and access restriction orders were also raised by Article 19 in its response to the draft Bill. Addressing what the group described as "disproportionate sanctions", it stated:

> *"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."*[42]

154. The wide range of punishments set out in this section, are excessively severe and are designed to pressure intermediaries to implement their operational safety duties in an overbearing manner. In the event that the measures set out in clauses 131-135 should ever be drawn upon, they would be a direct violation of the right to freedom of expression. Blocking access to a major intermediary in the UK would prevent many citizens from freely expressing themselves and would inhibit the free flow of information in this country. Such measures are more commonly associated with authoritarian regimes and have no place in a liberal democracy.

Clause 139 - Advisory committee on disinformation and misinformation

155. Clause 139 of the legislation states that:

> *(1) OFCOM must, in accordance with the following provisions of this section, exercise their powers under paragraph 14 of the Schedule to the Office of Communications Act 2002 (committees of OFCOM) to establish and maintain a committee to provide the advice specified in this section.*

156. The committee must include executives from the tech sector and "persons with expertise in the prevention and handling of disinformation and misinformation online". The committee does not have to include members who have expertise in protecting human rights or freedom of expression. The committee must publish a report 18 months after being establish and publish "periodic reports" from here on in.

157. While the Government have now removed provisions regarding legal but harmful speech from the Bill, they have repeatedly made clear their intention to use this new regulatory system to clamp down on misinformation

42 UK: 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/

and disinformation online. Given the way in which these terms can be politicised, this is deeply concerning from a freedom of expression perspective.

158. It should generally not be the place of a private company to assess and then instruct their users as to the "reliability" of the information and news sources they access. This is a highly subjective task best fulfilled by internet users themselves, who can optionally conduct wider research or access fact-checking websites online. This is much easier online than it is in a library and offline public spaces. The critical faculties of members of the public are not the responsibility of tech companies. Nor are tech companies best placed to judge the "reliability" of information.

Part 8 - APPEALS AND SUPER-COMPLAINTS

Clause 150 - Power to make super-complaints

159. Clause 150 creates a super-complaints system within the regulatory framework. It is set out as follows:

(1) An eligible entity may make a complaint to OFCOM that any feature of one or more regulated services, or any conduct of one or more providers of such services, or any combination of such features and such conduct is, appears to be, or presents a material risk of—

(a) causing significant harm to users of the services or members of the public, or a particular group of such users or members of the public;

(b) significantly adversely affecting the right to freedom of expression within the law of users of the services or members of the public, or of a particular group of such users or members of the public; or

(c) otherwise having a significant adverse impact on users of the services or members of the public, or on a particular group of such users or members of the public.

160. As with many of the provisions within the Bill, this is a well-intended inclusion that has a number of fundamental flaws. The fact that only "eligible entities", who will meet criteria set out by the Secretary of State, may make super complaints is a major limitation. With a degree of executive discretion, the Secretary of State could refine such a group of potential complainants to those of their own choosing. Moreover, that this function is not open to all members of the public means that certain groups and individuals will have a greater degree of influence over the permissibility of speech than others.

PART 9 – SECRETARY OF STATE'S FUNCTIONS IN RELATION TO REGULATED SERVICES

Clause 157 - Secretary of State's guidance

161. Part 9 holistically groups the powers of the Secretary of State, many of which have been mentioned in previous sections of this briefing. These include the powers to issue a statement of strategic priorities and other powers to direct Ofcom.

162. However, clause 157 introduces a new power awarded to the Secretary of State which gives the office holder of the day a greater degree of power to influence the regulatory system, by giving "guidance" to Ofcom.

163. Clause 157 (1) states:

> *The Secretary of State may issue guidance to OFCOM about—*
>
> *(a) OFCOM's exercise of their functions under this Act,*
>
> *(b) OFCOM's exercise of their powers under section 1(3) of the Communications Act (functions and general powers of OFCOM) to carry out research in connection with online safety matters or to arrange for others to carry out research in connection with such matters, and*
>
> *(c) OFCOM's exercise of their functions under section 11 of the Communications Act (media literacy) in relation to regulated services.*

164. The joint committee of both Houses of Parliament, tasked with undertaking pre-legislative scrutiny of the Bill recommended the removal of this clause in full. In their report, the Committee stated:

> *"The powers for the Secretary of State to ... give guidance to Ofcom give too much power to interfere in Ofcom's independence and should be removed."*[43]

In revising the draft Bill, the Government have failed to adhere to the Committee's recommendation and as such, Parliament must remove this provision.

PART 10 – COMMUNICATIONS OFFENCES

165. Part 10 constitutes substantial changes to the UK's communications offences. Current offences as set out under the Malicious Communications Act 1988 and Communications Act 2003 have often received criticism for criminalising expression which is "grossly offensive". However, the proposed

43 Joint Committee on the Draft Online Safety Bill, Report of Session 2021–22 , 10 December 2021, https://committees.parliament.uk/publications/8206/documents/84092/default/

revisions, based on recommendations by the Law Commission, have received widespread criticism for the threats that they also pose to freedom of expression. In particular, English PEN described the thresholds for criminality set out in new offences, such as the harm-based offence, as "broad and ambiguous" and as such, likely to do damage to the right to freedom of speech.

166. In a Government announcement last year, the Secretary of State confirmed that previous clause 151 which introduced a new communications offence, criminalising "seriously distressing" speech would be removed from the Bill.[44] Given the breadth of speech that this new offence would have criminalised, this is welcome. However the Government have opted to retain the proposed "false communications offence" which itself poses threats to free expression in the UK.

Clause 160 - False communications offence

167. Clause 160 seeks to update existing communications offences regarding the issuing of false communications. The offence is set out as follows:

(1) A person commits an offence if—

(a) the person sends a message (see section 154),

(b) the message conveys information that the person knows to be false,

(c) at the time of sending it, the person intended the message, or the information in it, to cause non-trivial psychological or physical harm to a likely audience, and

(d) the person has no reasonable excuse for sending the message.

168. A "likely audience" is defined as follows:

if, at the time the message is sent, it is reasonably foreseeable that the individual—

(a) would encounter the message, or

(b) in the online context, would encounter a subsequent message forwarding or sharing the content of the message.

169. Where the likely audience constitutes a number of individuals, it is not necessary for the person to have intended to cause harm to any one of them

44 New protections for children and free speech added to internet laws, DCMS, 28 November 2022, https://www.gov.uk/government/news/new-protections-for-children-and-free-speech-added-to-internet-laws

in particular. The maximum sentence could be up to 51 weeks or a fine (or both).

170. We do not believe this proposed new offence is suitable in its current form and could have a detrimental impact on freedom of expression. We are concerned that the harm threshold is intolerably low. In our view "non-trivial emotional harm" is too broad in scope.

171. Additionally, the offence does not stipulate based on whether a communication is sent in private or in public which could have the effect of criminalising expression where an individual online perceives the communication to have caused them (non-clinical) psychological harm.

172. English PEN has identified that many of the areas of concern with regard to specific and targeted false communications can be deal with by other areas of law including:

- *The torts of defamation and malicious falsehood.*
- *The tort established in Wilkinson v Downton [1897] 2 QB 57*
- *The Protection from Harassment Act 1997 c.40*
- *Common law rules developed for the Offences Against the Person Act 1861*[45]

173. A wider application of the newly created criminal offence to tackle "disinformation" would be a particularly dangerous step and would have a disastrous impact on freedom of expression.

174. Clauses 162 and 164 create new offences regarding threatening communications and cyber-flashing which Big Brother Watch does not object to. The Government have also stated their intention to add in other new criminal offences intended to prevent encouraging self-harm and sharing "downblousing" or deepfakes without consent – however, the text of these amendments has not been published.[46]

PART 11 - SUPPLEMENTARY AND GENERAL

Clause 180 – Extra-territorial application

45 HARMFUL ONLINE COMMUNICATIONS: ENGLISH PEN, CONSULTATION RESPONSE, https://www.englishpen.org/wp-content/uploads/2021/01/Online_Communications_Consultation_Response_English_PEN.pdf

46 New protections for children and free speech added to internet laws, DCMS, 28 November 2022, https://www.gov.uk/government/news/new-protections-for-children-and-free-speech-added-to-internet-laws

175. Clause 180 describes the extra-territorial application of the legislation and its application of the regulatory system to services which are available in the UK but located out of the country.

176. This speaks to one of the core issues with this regulatory system and attempting to introduce strict unilateral content-moderation rules that apply in a certain jurisdiction despite the open and interconnected nature of the internet. This provision could directly inhibit the free flow of information to the UK, which is inherent to the right to free expression. Such a free flow of information could be threatened if a service deems the regulatory framework too complicated or risky to operate to users in the UK or, if the legislation directly conflicts with domestic legislation from the jurisdiction within which it is based.

CONCLUSION

177. The Online Safety Bill poses a greater threat to freedom of speech in the UK than any other law in recent memory. In this briefing, we have set out some of the key threats that this legislation poses to fundamental rights in the UK.

178. It is vital that parliamentarians continue to consider the impact on the rights 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 and it is vital that the legislation continues to be altered in order to mitigate the most damaging elements.

179. In order to protect the rights to freedom of expression and privacy, as a minimum, **peers should support the amendment in the name of Lord Moylan engaging clause 65 and the amendment in the name of Lord Clement-Jones, engaging clause 110.**