LIBERTY

LIBERTY'S BRIEFING ON THE ONLINE SAFETY BILL FOR SECOND READING IN THE HOUSE OF COMMONS

APRIL 2022

ABOUT LIBERTY

Liberty is an independent membership organisation. We challenge injustice, defend freedom and campaign to make sure everyone in the UK is treated fairly. We are campaigners, lawyers and policy experts who work together to protect rights and hold the powerful to account.

Liberty provides policy responses to Government consultations on all issues which have implications for human rights and civil liberties. We also submit evidence to Select Committees, inquiries and other policy fora, and undertake independent, funded research.

Liberty's policy papers are available at libertyhumanrights.org.uk/policy.

CONTACT

SAM GRANT
Head of Policy and Campaigns
samg@libertyhumanrights.org.uk

JUN PANG
Policy and Campaigns Officer
junp@libertyhumanrights.org.uk

CONTENTS

INTRODUCTION

1. Advances in technology have significantly influenced the ways in which we understand our rights and freedoms, and the very societies in which we live. In particular, the internet has become a primary frontier for free expression and the exchange of ideas, and a crucial site of public participation and democratic engagement. Revolutionising the manner in which we communicate, debate, dissent, organise, and form opinions, it remains 'history's greatest tool for global access to information'.[1] Yet the contemporary internet has also facilitated the proliferation of hate and oppressive speech, spread of viral propaganda intended to manipulate or undermine democratic institutions, and ubiquitous collection of data and mass surveillance.

2. The question of internet regulation – whether and how it should be done – has arisen in response to growing public concern over the different ways in which the internet can enable harm to individuals and wider communities. But it is impossible to consider this question without first situating it within wider global and national structures of power and control, in a world where private companies function alongside and, in some cases, compete with one another and with state governments to exercise control over key arenas of public and political participation. Governments around the world have attempted to legislate for internet regulation in different ways; there is no universal standard, and criticisms have arisen in relation to each of these approaches.[2] Private internet companies themselves have developed complex internal constitutions, in the form of terms and conditions, to which they hold individuals who use their services to account.

3. The key question underlying internet regulation is how different aspects of online space are and should be designed, and to what end. To borrow Lawrence Lessig's formulation, "whether a part of cyberspace—or the Internet generally—can be regulated turns on the nature of its code. Its architecture will affect whether behaviour can be controlled… its architecture is its politics."[3] The secondary question is whether this design can and should be achieved by states via national legislation (or international organisations via international law), private companies via changes in terms and conditions, by societal forces via norm shifts, or combinations of the above. These questions do not have easy

[1] Report of the UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression (6 April 2018) UN Doc. A/HRC/38/35 available at https://documents-dds-ny.un.org/doc/UNDOC/GEN/G18/096/72/PDF/G1809672.pdf?OpenElement
[2] Conrad, J., *China Cracks Down On Its Tech Giants. Sound Familiar?,* Wired, 29 July 2021, available at: https://www.wired.com/story/china-cracks-down-tech-giants-sound-familiar/;
[3] Pg. 24, Lessig, L. (2006), *Code 2.0,* Basic Books.

normative answers, not least because it is unclear who can legitimately make and in turn be held to account over these decisions.

4. The Online Safety Bill (OSB) is the UK's attempt to answer both questions. The product of a years-long consultation process, starting with the initial Online Harms White Paper, a first draft of the legislation, and subsequent pre-legislative scrutiny by the Joint Committee on Human Rights (JCHR), the fundamental aim of the OSB is to protect internet users from harm. The OSB imposes duties of care on internet service providers to enable them to achieve this objective. The current communications regulator OFCOM will be given new oversight powers in relation to these duties and be empowered as the UK's regulator for online safety. OFCOM's functions will ultimately be subject to directions and regulations on the part of the Secretary of State.

5. The aim of protecting internet users' safety is a laudable aim, and indeed, states do have an obligation to protect the safety of those within their jurisdiction. However, the concept of 'safety' is broad, and legislation and policies that hinge on concepts of 'safety' tend to take a parochial approach to human rights, concerned fundamentally as they are with mitigating and managing risk. They are also more likely than not to seek to centralise control: to borrow from Lessig, "[I]f some architectures are more regulable than others—if some give governments more control than others—then governments will favour some architectures more than others."[4]

6. We acknowledge that there are endemic and systemic problems with the way key forums of public participation and engagement are designed, that are rooted among other things in the economy of the internet and the interests of different stakeholders including private companies and state governments. This briefing does not express a view as to whether the internet should be subject to regulation; instead, our aim is to highlight some of the most important rights issues that we believe are engaged by the Online Safety Bill (OSB).

7. **Overall, we caution against the use of 'safety' as the prism through which to conceptualise regulation. To the extent that the OSB seeks to change the architecture of aspects of the internet in the UK in order to enhance people's abilities to express themselves freely online,[5] its foundations must be firmly rooted in human rights principles. In particular, we are concerned about the OSB's expansive definitions of 'illegal content' and 'legal but harmful' content as well as its potential to erode end-**

[4] Pg. 24, Lessig, L. (2006), *Code 2.0,* Basic Books.
[5] World-first online safety laws introduced in Parliament, *Department for Digital, Culture, Media and Sport and the Rt Hon Nadine Dorries MP*, 17 March 2022, https://www.gov.uk/government/news/world-first-online-safety-laws-introduced-in-parliament

to-end encryption and undo the right to anonymity, for the effects this will have on individuals', including children's, right to freedom of expression and privacy. We also express our reservations regarding the extent to which the OSB guarantees OFCOM's regulatory independence and potential negative ramifications of the proposed new harmful communications offence.

8. Taking into consideration the wider political context, the OSB demonstrates the importance of defending the Human Rights Act[6] and existing data protection[7] and privacy laws[8] against the Government's ambitions of overhaul and reform, to ensure that individuals continue to be able to rely on these safeguards while exercising their rights on the internet, and to protect the health of our democracy. In particular, the Government's current plans regarding the HRA, including the proposal to alter the way the principle of 'proportionality' is applied, will have a significant impact on – and indeed, may contradict – the proposals within the OSB. Other proposals to reform the HRA, including to give Article 10 (the right to freedom of expression) a higher status than other rights (such as the right to respect for one's private and family life under Article 8), may have unintended consequences on many of the safety duties contained within the OSB, while also scuppering the potential for enhanced legislative protections for internet users' data and privacy, which is crucial to addressing the root causes of online harms. None of these issues have been mentioned in the Government's publications regarding the OSB, even though they are evidently of significant concern. **We urge parliamentarians to defend the Human Rights Act to ensure that there exist robust safeguards for individuals' ability to participate freely in the internet, including their rights to freedom of expression and privacy.**

UNDERSTANDING THE 'DUTY OF CARE' FRAMEWORK

9. To achieve its aim of protecting internet users' safety, the OSB imposes new duties of care on internet services to manage and deal with different categories of content, including illegal content, child sexual exploitation and abuse (CSEA) and terrorism content, and legal but harmful content (which we consider in greater detail below).

[6] See: Liberty's response to the Ministry of Justice's consultation on Human Rights Act reform: A Modern Bill of Rights, March 2022, available at: https://www.libertyhumanrights.org.uk/wp-content/uploads/2022/03/Libertys-response-to-the-Ministry-of-Justices-consultation-%E2%80%98Human-Rights-Act-Reform-A-Modern-Bill-of-Rights-March-2022.pdf
[7] Open Rights Group, *Weakening privacy will fuel online harms,* 29 September 2021, available at: https://www.openrightsgroup.org/blog/weakening-privacy-will-fuel-online-harms/
[8] See: DCMS, *Data: A new direction,* available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1022315/Data_Reform_Consultation_Document__Accessible_.pdf. See also Liberty's response to the consultation: https://www.libertyhumanrights.org.uk/wp-content/uploads/2021/11/Libertys-response-to-the-Department-of-Digital-Culture-Media-and-Sports-consultation-into-Data-A-new-direction.pdf

Content is defined as "anything communicated by means of an internet service, whether publicly or privately, including written material or messages, oral communications, photographs, videos, visual images, music and data of any description" (clause 189).

10. The OSB will affect the following internet services: user-to-user services, in which users can upload and share content which may be encountered by another user, e.g. social media sites; search services, e.g. search engines; and providers of regulated services that publish pornographic content. The Bill sets out services that are exempt from regulation (Schedule 1), with the Secretary of State retaining the ability to exempt new categories of services if they consider that the risk of harm to UK users presented by these services are low. **This briefing focuses on the duties imposed on user-to-user services and search engines.**

Duties of care

11. The following duties of care are imposed on all **user-to-user services**, such as social media sites: duties relating to illegal content risk assessment, illegal content, content reporting, complaints procedures, freedom of expression and privacy and record-keeping and review.

12. In addition to these duties, if a user-to-user service is "likely to be accessed by children", there will be duties to conduct children's risk assessments and to protect children's online safety. Clause 33 defines three ways in which a service may be defined as "likely to be accessed by children":

    i. A children's access assessment carried out by the provider of the service concludes that it is possible for children to access the service or a part of it (clause 33(2)); and there is a significant number of children who are users of the service or that part of it or the service is of a kind is likely to attract a significant number of users who are children (clause 31(3)); or

    ii. Where the provider of the service fails to carry out the first children's access assessment required by the OSB (clauses 33(4) and 32(1)); or

    iii. Where, following an investigation into a failure to comply with a duty set out in clause 32, OFCOM determine that a service should be treated as likely to be accessed by children.

13. OFCOM will have the power to designate certain regulated user-to-user services as Category 1 or Category 2B; and regulated search services as Category 2A, based on

threshold conditions set out in regulations made by the Secretary of State (Schedule 10). Category 2A and Category 2B services will be given the duties set out above. **Category 1 services, which will encompass services with "the largest audiences and a range of high risk features,"[9] will be given additional duties relating to 'legal but harmful' content**, including to undertake adults' risk assessments, protect adults' online safety, empower adult users, protect content of democratic importance, protect journalistic content, and protect individuals' rights to freedom of expression and privacy.

14. **Search services** (including search engines) are subject to similar duties of care as user-to-user services. They will be given duties relating to illegal content risk assessments, illegal content, content reporting, complaints procedures, freedom of expression and privacy, and record-keeping and review.

Enforcement

15. The OSB empowers OFCOM to act as the regulator for online safety, subject to the direction of the Secretary of State. OFCOM will be required to produce **codes of practice** to assist service providers in the discharge of their duties (clause 37).

16. If service providers fail to discharge the above duties and other duties set out in the OSB (the full list of enforceable duties is set out in clause 111), they may face enforcement action. The procedure for enforcement is set out in Part 7, Chapter 6, and involves OFCOM giving service providers a series of notices and decisions. Service providers will be given opportunities to make representations, and may be required to take steps to resolve issues and/or pay penalties in relation to issues identified in OFCOM's notices. The maximum level of penalties is set out in paragraph 4, Schedule 12 as the greater of £18 million or 10% of the provider's qualifying worldwide revenue for the person's most recent complete accounting period.

17. If a service provider continuously fails to comply with a requirement imposed by OFCOM or to pay a relevant penalty, OFCOM may apply to the court for **a service restriction order** requiring the provider to terminate the operation of its services in the UK, among other potential requirements (clause 123). If OFCOM considers that a service restriction order is or would be insufficient to prevent significant harm arising to individuals in the UK, then it may make a request to the court to make an **access restriction order**, that may require the service provider to take steps to prevent internet users from being able

[9] Department for Digital, Culture, Media and Sport, *Online Safety Bill: factsheet,* 17 March 2022, available at: https://www.gov.uk/government/publications/online-safety-bill-supporting-documents/online-safety-bill-factsheet

to access the service (clause 125). OFCOM can also make requests for these orders on an interim basis provided certain conditions are met (clause 124 and clause 126).

18. OFCOM will have the power to require that service providers provide it with any information that it requires to exercise its functions (clause 85), including to assess compliance with any of the illegal content and 'legal but harmful' content duties, by way of an information notice. If the person being given an information notice is an entity, then it will need to name a senior manager in its response to the notice (clause 87). Clause 92 sets out the new offences in connection with information notices, including if a person fails to comply with a requirement of an information notice (clause 92(1)); and if, in response to an information notice, a person provides information that is false and at the time of providing it, the person knows or is reckless as to whether it is false (clause 92(3)).The named senior manager will be liable for the relevant information offence if the entity has committed an offence and that manager has failed to take all reasonable steps to prevent that offence being committed (clause 93). The penalties for committing information offences are set out in clause 96, and range from a fine to imprisonment for a term not exceeding two years or a fine (or both), depending on the offence committed.

ILLEGAL CONTENT

19. 'Illegal content' is defined in clause 52(2) as meaning "content that amounts to a relevant offence". These 'relevant offences' correspond to existing criminal offences, set out in various Schedules. Priority illegal content, which falls within illegal content, means terrorism content, CSEA content, and content that amounts to an offence specified in Schedule 7 (clause 52(7)); such content attracts a greater level of scrutiny and regulation. Situations in which user-generated content will "amount to a relevant offence" are set out in clause 52(3).

20. Under the OSB, both user-to-user services and search services will be subject to illegal content duties. Crucially, internet services will have a duty 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" (clause 9(2) and clause 21(2)). **'Priority illegal content' is subject to an even higher level of regulation**: user-to-user services will have a duty to operate a service using "proportionate systems and processes designed to **prevent individuals from encountering priority illegal content**"; "**minimise the length of time** for which any priority illegal content is present"; and "after being made aware of any illegal content, **swiftly take down such content**" (clause 9(3)). Search services will have a duty to operate a system using proportionate systems and

processes designed to **minimise the risk of individuals encountering priority illegal and other illegal content** (clause 21(3)).

21. User-to-user services and search services will also be required to undertake illegal content risk assessments (and take appropriate steps to keep this assessment up to date) (clause 8 and 23), to provide information about how individuals are to be protected from illegal content in the terms of service (clause 9(5) and clause 21(5), to apply these measures consistently (clause 9(6) and clause 21(6), and to provide information about any proactive technology that is used to achieve this aim (clause 9(7) and clause 21(7)).

22. While criticism of the OSB has primarily focused on its approach to 'legal but harmful' content, **we are concerned by the way that 'illegal content' has been defined and the risks this definition and related duties pose to the rights to freedom of expression.** First, the OSB's definition of "illegal content" refers to an extremely **wide list of relevant offences**, ranging from offences related to terrorism (Schedule 5) and CSEA (Schedule 6), to acts intended or likely to stir up racial hatred under the Public Order Act 1986, assisting suicide under the Suicide Act 1961, causing or inciting prostitution for gain under the Sexual Offences Act 2003, and assisting unlawful immigration under the Immigration Act 1971, among many other offences listed under Schedule 7. Apart from the offences in Schedules 5, 6, and 7, there is also a residual and broad category of offences 'of which the victim or intended victim is an individual (or individuals)', which will fall under the 'illegal content' category.

23. Second, the OSB gives the Secretary of State **the power to amend Schedules 5, 6, and 7, including to add offences to the list.** Notably, the Secretary of State may **add offences** to Schedule 7 (i.e. non terrorism and non-CSEA offences) by way of regulations passed under the draft affirmative procedure (clause 176), subject to certain limited exceptions (clause 176(5)). The SoS's exercise of their discretion is limited only by the following criteria: they must consider it appropriate to add an offence because of the prevalence of content that amounts to that offence on internet services; the risk of harm to individuals in the UK presented by that content; and the severity of the harm (clause 176(4)). The Delegated Powers Memorandum to the Bill justifies this power on the basis that the online safety framework "needs to be able to adapt to new harms", especially "if new criminal offences are to be created elsewhere in legislation that the government also wishes to designate as priority illegal content, so that companies would be required to

take steps to search for, remove and limit people's exposure to content relating to that criminal behaviour."[10]

24. Notwithstanding the additional safeguards provided by the draft affirmative procedure and the stipulation of conditions that must be satisfied prior to a new offence being added, **the SoS's wide power to designate additional categories of content as 'illegal', including 'priority illegal', remains highly concerning**. This is because under this provision, the Government of the day could further expand the list of relevant offences to suppress speech it simply does not like. For example, a future Secretary of State could seek to insert additional public order offences (such as those it is attempting to pass through the Police, Crime, Sentencing and Courts Bill and other anti-protest legislation) into Schedule 7, on the basis that there is a "great prevalence" of content on social media sites advocating people to engage in disruptive protest; this could result in a degree of harm to individuals in the UK (for example, if the disruption causes alarm or distress to individuals); and this harm is deemed to be severe. This would ultimately stifle the ability of individuals to engage in certain forms of political discussion and organising online by removing this content from internet platforms altogether.

25. Raising similar concerns over the level of discretion granted to the Secretary of State to determine 'illegal content' in the draft OSB, the JCHR noted: "[g]iven that illegal content will in most cases already be defined by statute, [the power to designate content relating to an offence as priority illegal content] should be restricted to **exceptional circumstances**", and only after consultation with the JCHR and subject to the affirmative procedure.[11]

26. Our concern over the breadth of these offences is exacerbated by the peculiar way that content will be determined to 'amount to' a relevant offence. First, the draft OSB had previously required the service provider to have 'reasonable grounds' to believe that content amounted to a relevant offence, which is a term recognised in law as requiring some objectivity. The current iteration of the Bill lacks such a requirement. We appreciate that during pre-legislative scrutiny, the JCHR considered evidence that it would be difficult for service providers to apply the objective test required by a 'reasonable grounds' test, however, we are concerned that the removal of reference to an objective assessment might render such judgements **entirely subjective**.

[10] Paragraph 369, Delegated Powers Memorandum: Online Safety Bill, https://publications.parliament.uk/pa/bills/cbill/58-02/0285/e02721600delegatedpowersmemorandumelay.pdf
[11] Paragraph 148, Joint Committee on Human Rights, *Draft Online Safety Bill: Pre-legislative scrutiny report*, available at: https://publications.parliament.uk/pa/jt5802/jtselect/jtonlinesafety/129/129.pdf

27. Clause 52, which sets out situations in which content can 'amount to' a relevant offence, is also **worryingly vague**. For example, in the case of a user-to-user service, one of the situations in which content can amount to a relevant offence is when "the use of the words, images, speech or sounds, *when taken together with other regulated user-generated content present on the service*, amounts to a relevant offence (emphasis added)". While confusingly phrased, we are concerned that this could mean that even if an individual's words do not themselves amount to a relevant offence, their words plus the existence of other content on a platform (that the individual may not have themselves created), could amount to an offence for the purposes of the OSB. In other words, **the definition of 'illegal' content would be wider online than offline**.

28. This is exacerbated by the fact that many of the relevant offences are interpreted very widely and their meanings have been contested in the criminal courts. Gavin Millar QC, in his evidence to the JCHR on the draft OSB, said: "… applying the statutory wording of most modern criminal offences to the facts is a difficult and technical exercise. It is one which police, CPS and courts often get wrong. This is both because of the flexibility of the language that is used and because of detailed nature of the drafting in most of our contemporary criminal offences."[12] The English Collective of Prostitutes (ECP) highlights that women have been prosecuted under the offence of "causing or inciting prostitution for gain" for helping friends build a website or place an advert.[13] Four asylum seekers who helped steer small boats across the Channel recently won their appeal against their convictions for "assisting unlawful immigration" under s.38 of the Immigration Act 1971.

29. **Ultimately, we are concerned that the illegal content duties might lead private companies to pre-emptively take down significant amounts of content that appear to be 'illegal' even if it is not, in order to avoid being penalised. This will not only have negative ramifications for freedom of expression online but may also have knock-on effects for the rights of those who currently suffer the brunt of different forms of criminalisation.** For example, the ECP has warned that increased takedowns of sex workers' online advertisements will increase sex workers' exposure to violence, as they will be increasingly forced to work on the street and have less control over their working conditions.[14] A similar law passed in the US that resulted in tech firms pre-emptively taking down sex work advertisements led 60% of sex workers to accept less safe clients,

[12] Paragraph 141, Joint Committee on Human Rights, *Draft Online Safety Bill: Pre-legislative scrutiny report*, available at: https://publications.parliament.uk/pa/jt5802/jtselect/jtonlinesafety/129/129.pdf
[13] English Collective of Prostitutes, *Briefing: Online Safety Bill – Criminalising sex workers' online adverts will undermine safety*, 3 March 2022, available at: https://prostitutescollective.net/briefing-online-safety-bill/
[14] English Collective of Prostitutes, *Briefing: Online Safety Bill – Criminalising sex workers' online adverts will undermine safety*, 3 March 2022, available at: https://prostitutescollective.net/briefing-online-safety-bill/

according to the organisation Coyote.[15] This is an important example of how freedom of expression is not antithetical to, but a crucial way of, keeping people safe.

30. More broadly, we are concerned that the illegal content duties will effectively empower private companies to make judgments as to what is and isn't 'illegal content'. This will set a worrying precedent, not least in **ceding greater power to private companies, running counter to the OSB's objectives.** As we explain below, we do not believe that the cross-cutting duties about freedom of expression and privacy will be a sufficient safeguard.

31. The wide net cast by the list of relevant offences (and the potential for the Secretary of State to add to these offences), in combination with the vague way that content is defined as 'amounting to illegal content' is likely to create significant difficulties for internet service providers in discharging their duties. Hashing algorithms that have been used to detect CSEA images from an existing database of images,[16] for example, are unlikely to be effective at identifying if certain words and images constitute 'assisting unlawful immigration'. **Service providers will inevitably need to develop more sophisticated AI and automated-decision making systems in order to comply with their duties and moderate content at scale, even as existing systems have come under fire for being biased and giving rise to inaccurate and unfair outcomes.** On the other hand, in order to discharge their duties to have complaints procedures and to have due regard to freedom of expression (among others), service providers may also seek to hire greater numbers of human content moderators, including through outsourcing, in order not to fall afoul of heavy penalties. We believe that as yet, not enough attention has been paid to the human cost of the OSB's regulatory regime, specifically, **the potential entrenchment and expansion of hidden and exploitative working conditions worldwide.**[17]

ADULTS' ONLINE SAFETY: 'LEGAL BUT HARMFUL'

32. A cornerstone of the OSB is the creation of a category of content that is legal but deemed to be harmful to adults ('legal but harmful' content). The OSB defines harm as "physical

[15] Oppenheim, M., *Violence against sex workers will surge under new online safety bill, campaigners warn,* 18 March 2022, available at: https://www.independent.co.uk/news/uk/home-news/sex-workers-online-safety-bill-b2039044.html
[16] Thorn, Introduction to Hashing: A Powerful Tool to Detect Child Sex Abuse Imagery Online, 12 April 2016, available at: https://www.thorn.org/blog/hashing-detect-child-sex-abuse-imagery/
[17] The technology rights organisation Foxglove is currently supporting Daniel Motaung, a former content moderator working at one of Facebook's outsourcing companies in Kenya to challenge the poor working conditions, pay, and lack of mental health support given to those in this role. See: Foxglove, *NEW CASE: Foxglove supports Facebook content moderator sacked for leading workers to form a trade union in Kenya,* 16 February 2022, available at: https://www.foxglove.org.uk/2022/02/16/foxglove-supports-facebook-content-moderator-sacked-kenya/

or psychological harm" (clause 187). When referencing harmful content, the OSB will consider harm arising from or that may arise from the nature of the content itself; the fact of its dissemination; and the manner of its dissemination (for example, content repeatedly sent to an individual by one person or by different people) (clause 187). Examples of legal but harmful content that have been discussed in relation to the OSB include hate speech that does not account to a criminal offence and content depicting self-harm.

33. The Secretary of State will be given the power to define 'legal but harmful' content for children and for adults by way of secondary legislation (clauses 53(2) and 54(3) respectively). User-to-user services that are Category 1 services are given additional duties in relation to legal but harmful content. One such duty is the risk assessment duty, which includes a list of matters that must be considered under clause 12(5). After conducting this risk assessment, Category 1 services will need to set out in their terms of services how they intend to deal with this content, which may include **taking down the content, restricting users' access to the content, limiting the recommendation or promotion of the content, or recommending or promoting the content** (clause 13(4)). Category 1 services will also have a duty to provide features which adult users may use or apply if they wish to increase their control over harmful content.

34. Article 10 of the ECHR, as protected by the Human Rights Act 1998, provides that "everyone has the right to freedom of expression." The right encompasses ideas themselves—including information that might shock, offend, or disturb, journalistic freedom, commercial information, and political and artistic expression. **Freedom of expression is a fundamental right and vital to any healthy democracy**: the ability to openly share our views and to access information is key to informed public debate and essential to public accountability and transparency in government.[18] More widely, Article 10 is important because it reinforces and enables people to access their other rights by facilitating dialogue and advocacy.

35. **Liberty supports a broad interpretation of Article 10, although we also acknowledge that Article 10 is a qualified right, meaning that restrictions on Article 10 may sometimes be necessary and proportionate.** Indeed, while the right to freedom of expression is universal, the harm that can occur through its exercise is not evenly felt: marginalised communities that already suffer from the effects of structural inequality can be further harmed through certain forms of expression. For example, hate speech—that is, forms of expression that target people's protected characteristics—can cause

[18] See: Handyside v. United Kingdom (Application no. 5493/72)

individuals to quit their jobs, leave their homes, avoid public places, curtail their own speech, and otherwise modify their behaviour. The prevalence of hate speech in society can also contribute to a climate of intolerance in which forms of discriminatory behaviour are seen as acceptable, with negative impacts on marginalised communities. A report by Amnesty International analysing tweets that mentioned women MPs in the run up to the 2017 General Election found that the 20 Black, Asian, and Minority Ethnic MPs received 41 per cent of the abuse, despite making up less than 12 per cent of those in the study, prompting widespread discussion over the impact of hate speech on the health of our democracy.[19] Organisations working in the violence against women and girls sector have highlighted the impact of online abuse on women.[20] This gives rise to a need to consider how best to balance Article 10 rights with the rights of marginalised communities and individuals, for example, their right to be free from discrimination.

36. **Liberty will only support State restrictions on freedom of expression if they are lawful, an effective means of achieving a legitimate aim, and proportionate. We believe that the criminal law - tightly constrained - should be the starting point for regulation for potentially harmful online activity.**

37. **We are concerned that the 'legal but harmful' category set out in the OSB is inadequately prescribed by law and risks disproportionately infringing on individuals' right to freedom of expression and privacy.** In particular, we are concerned about the wide definition of online harm as meaning "physical or psychological harm" (clause 187).[21] This is an extremely low threshold, and encompasses innumerable kinds of harm, the extent of which in our view far exceeds the qualifications on Article 10 provided by the ECHR and HRA.

38. This wide definition is exacerbated by the fact that the Secretary of State has **extensive latitude** to define what constitutes legal but harmful content for children and adults respectively (clauses 53(2) and 54(2)), with limited constraints set out in statute (clause 55). For children, the SoS may designate content as harmful to children if they consider "there is a material risk of significant harm to an appreciable number of children presented by content of that description that is regulated user-generated content or search content" and they deem it is appropriate for the relevant duties in the OSB to

[19] Amnesty International, 'Black and Asian women MPs abused more online': https://www.amnesty.org.uk/onlineviolence-women-mps [accessed 30 November 2021]
[20] End Violence Against Women coalition, Online safety committee fail to name violence against women in recommendations for new law, https://www.endviolenceagainstwomen.org.uk/parliament-committee-misogynistic-abuse-online/
[21] This definition is different to the definition of 'harm' for communications offices. Clause 150 provides that the definition of 'harm' in relation to the harmful communications offence is "psychological harm amounting to at least serious distress".

apply. For adults, the SoS may specify content as harmful to adults if they consider that "in relation to regulated user-to-user services, there is a material risk of significant harm to an appreciable number of adults presented by content of that description that is regulated user-generated content." "Appreciable number" is not defined in the OSB.

39. We fear that one of the potential implications of the 'legal but harmful' category is that it may be further expanded **to encompass forms of speech that the Government of the day simply does not like.** The Delegated Powers memorandum to the OSB provides that the purpose of enabling the Secretary of State to define 'legal but harmful' content is to "allow the government to respond rapidly to the changing nature of online services and the risks to children and adults online, ensuring that new, currently unforeseen harms can be dealt with as quickly as they emerge."[22] It states that regulations to designate these categories of content will be subject to the draft affirmative resolution procedure, although in some cases the Secretary of State can in urgent cases use the made affirmative resolution procedure (clause 179(3) to (7)). **While we welcome the acknowledgement that enhanced scrutiny is necessary for such wide-ranging powers to define what constitutes 'legal but harmful' content, we do not feel it is sufficient as a safeguard for the risks to freedom of expression posed by this category.**

40. The ultimate upshot of these wide definitions is that service providers may opt to take down or otherwise restrict access to vast swathes of content that appear to be legal but harmful, to avoid being penalised. Takedown is not the only measure that service providers can take under the Bill, but the fact that it is presented as one of the primary measures means that there is a risk that **service providers will opt for it as a matter of course** – indeed, there are well-documented examples of large social media platforms taking down content for allegedly violating their terms of service, even if this is not actually the case.[23] More widely, there is ongoing debate about whether certain kinds of content should be taken down: for example, Instagram's proposed policy to "not allow any graphic images of self-harm", which resulted from the tragic death of Molly Russell who died by suicide and whose Instagram account contained distressing material about depression and suicide, has prompted concern from people who say they post pictures of their self-harm scars as a way to combat stigma and speak openly about mental health issues.[24]

[22] Paragraph 91, https://publications.parliament.uk/pa/bills/cbill/58-02/0285/e02721600delegatedpowersmemorandumelay.pdf

[23] Plymouth Hoe is a historic part of Plymouth. In 2021, Facebook identified 'Plymouth Hoe' as an offensive term, and took down numerous people's posts which included reference to 'Plymouth Hoe', on the basis that they breached bullying rules. It was subsequently forced to apologise. See: BBC News, *Facebook apologises for Plymouth Hoe 'error',* 27 January 2021, available at: https://www.bbc.co.uk/news/uk-england-devon-55827981

[24] Bramwell, K., *Instagram: 'I don't want people to be ashamed of their scars',* BBC News, 30 May 2019, available at: bbc.co.uk/news/health-48431858

We are concerned that restrictions on free speech may fall disproportionately on marginalised communities whose expression may already be targeted and policed online, for whom being able to engage in and participate in online spaces is particularly important.

41. Measures such as takedown stand to have a corrosive effect on access to information, if it results in news content being taken down from a user-to-user or search service. The Society of Editors has highlighted that because of the time-sensitive nature of current events, by the time that a complaint is made about news content being taken down, it may be too late.[25]

Safeguards

42. The OSB itself recognises that there is a risk that its duty of care approach will erode internet users' rights to freedom of expression and privacy. For this reason, it includes in clause 19 cross-cutting duties about freedom of expression and privacy that apply to the various duties on service providers regarding illegal content, children's online safety, adults' online safety, content reporting, and complaints procedures.

43. It is important to note that the Government is simultaneously consulting on its overhaul of the Human Rights Act. **Any erosion of rights protections in the HRA will inevitably have a knock-on effect on the degree to which the limited protections within clause 19 are able to effectively safeguard internet users' rights, and must be resisted.**

Freedom of expression

44. Clause 19(2) provides that when deciding on and implementing safety measures and policies, internet service providers will have "a duty to have regard to **the importance of protecting users' right to freedom of expression** within the law". Nonetheless, we believe the 'due regard' safeguards for freedom of expression are unlikely to be a sufficiently robust safeguard to protect individuals' rights. This is because having 'due regard' to a right sets a low threshold, that is likely to be easily discharged by service providers with little material impact on how they deal with content.

45. The OSB contains further carveouts for Category 1 services, providing that such services must have due regard to **"content of democratic importance" and "journalistic content"**, ostensibly in acknowledgement of the fact that journalism and political

[25] Paragraph 295, Joint Committee on Human Rights, *Draft Online Safety Bill: Pre-legislative scrutiny report*, available at: https://publications.parliament.uk/pa/jt5802/jtselect/jtonlinesafety/129/129.pdf

expression receive a higher level of protection than everyday speech under the ECHR.[26] **Nonetheless, we believe these safeguards are lacking.** Regarding the former, it is unclear how "content of democratic importance" will be defined – will it depend on who is generating the content (i.e. politicians), or the content of that speech? If it is the former, the organisation Legal to Say, Legal to Type summarised concerns aptly when it said that this could create a two-tier system with "free speech for journalists and politicians, and censorship for ordinary citizens."[27] If it is the latter, the decision as to whether content would substantively qualify as being "of democratic importance" is likely to be highly subjective and contested.

46. In respect of the "journalistic content" exemption, the National Union of Journalists has criticised the OSB for potentially giving service providers additional powers to define what constitutes journalism, in ways that might have a detrimental impact on media plurality online.[28] Other organisations have flagged the definition as potentially being overly narrow. For example, the Professional Publishers Association has flagged the definition as potentially excluding consumer magazines and business media if they do not focus on current affairs (clause 16(8)).[29] These sections appear unchanged from the draft OSB, meaning that the concerns expressed throughout the pre-legislative scrutiny process regarding the potential inefficacy of these safeguards remain unaddressed.

Privacy

47. The right to respect for one's private and family life is protected by Article 8 of the ECHR and the HRA, and it encompasses protections for one's personal data and correspondence. Article 8 undergirds the UK's data protection framework, which is an important tool for people to hold public and private bodies to account over potential breaches of their rights. Overall, it is a vital safeguard for individuals' dignity, protecting individuals' lives against interference from the state but also private bodies.

48. Clause 19(3) of the OSB provides: "When deciding on, and implementing, safety measures and policies, a duty to have regard to the importance of protecting users from **a breach of any statutory provision or rule of law concerning privacy** that is relevant to the use or operation of a user-to-user service (including, but not limited to, any such provision or

[26] R v BBC, ex p ProLife Alliance (2003) UKHL 23
[27] Paragraph 297, Joint Committee on Human Rights, *Draft Online Safety Bill: Pre-legislative scrutiny report*, available at: https://publications.parliament.uk/pa/jt5802/jtselect/jtonlinesafety/129/129.pdf
[28] National Union of Journalists, Written evidence submitted by the National Union of Journalists (Online Safety Bill Pre-Legislative Scrutiny), September 2021, available at: https://committees.parliament.uk/writtenevidence/39385/html/
[29] Paragraph 287, Joint Committee on Human Rights, *Draft Online Safety Bill: Pre-legislative scrutiny report*, available at: https://publications.parliament.uk/pa/jt5802/jtselect/jtonlinesafety/129/129.pdf

rule concerning the processing of personal data)." As above, we do not believe that a 'due regard' duty is sufficient to safeguard against intrusions into individual users' privacy data rights, given the low threshold needed to discharge this duty.

49. **Crucially, we are concerned that the duties created by the OSB risk entrenching the very business models that amplify harm in online environments, by further incentivising service providers to engage in extensive amounts of surveillance and profiling of users.** Without seeking to state the obvious, the purpose of a private internet company is to make profit. In turn, the profitability of a social media company hinges on keeping people engaged in a platform for prolonged periods of time; the longer a user is engaged, the more they are likely to be exposed to and influenced by advertisements, and the more revenue is made for the company. To keep people engaged, social media platforms profile their users and tailor the content they see according to what they have previously engaged with, their personal data (including their geolocation, political beliefs, and sexual preferences), and other behavioural insights gained through surveillance across their internet usage, through techniques such as micro-advertising and through 'rewards' such as more tailored content, likes, and/or engagement.[30][31] The JCHR highlighted in its report that children in particular can be vulnerable to be being targeted with harmful content as algorithms can serve them with progressively more extreme content to keep them engaged.[32] This system design creates the conditions for algorithmic amplification, which is when certain content is selected and ranked higher, and made more visible than others, regardless of the substance of that content.[33] The JCHR found in its pre-legislative scrutiny report that social media platforms' key performance indicators focused on engagement "are maximised regardless of the nature of that engagement or quality of the content that is being engaged with… By making design choices that maximise engagement, service providers therefore exacerbate the presence, spread, and effect of harms.[34]

50. **The OSB risks entrenching and expanding harmful (and potentially unlawful) practices of surveillance and in turn interfering with individuals' Article 8 rights.** This is because

[30] Bhargava, V.R. and Velasquez, M., *Ethics of the Attention Economy: The Problem of Social Media Addiction*, 6 October 2020, available at: https://www.cambridge.org/core/journals/business-ethics-quarterly/article/ethics-of-the-attention-economy-the-problem-of-social-media-addiction/1CC67609A12E9A912BB8A291FDFFE799
[31] Killock, J., *Internet policy is broken,* Open Rights Group, 10 March 2022, available at: https://www.openrightsgroup.org/blog/internet-policy-is-broken/
[32] Paragraph 40, Joint Committee on Human Rights, *Draft Online Safety Bill: Pre-legislative scrutiny report*, available at: https://publications.parliament.uk/pa/jt5802/jtselect/jtonlinesafety/129/129.pdf
[33] Huszár, F. et al., *Algorithmic amplification of politics on Twitter,* PNAS, 21 December 2021, available at: https://www.pnas.org/doi/10.1073/pnas.2025334119.
[34] Paragraph 37, Joint Committee on Human Rights, *Draft Online Safety Bill: Pre-legislative scrutiny report*, available at: https://publications.parliament.uk/pa/jt5802/jtselect/jtonlinesafety/129/129.pdf

it mandates (and effectively gives State backing) to service providers' profiling of users, in the name of protecting them from harm. For example, within the provisions concerning the adults' risk assessment duty, service providers will be required to take into account "the level of risk of harm to adults presented by priority content that is harmful to adults *which particularly affects individuals with a certain characteristic or members of a certain group"* (clause 12(5)(d); emphasis added); a similar measure is required in the context of the children's risk assessment duty (clause 10(6)(d)). We are concerned that such techniques of profiling may disproportionately target people with protected characteristics.

51. In her evidence to the JCHR, the previous Information Commissioner highlighted her concerns over the ways that platforms infer data to direct people to their platforms, and questioned whether these processes are compliant with data protection law.[35] In this context, an add-on safeguard to have 'due regard' to individuals' right to privacy will do little to protect individuals from harm.

52. The OSB seeks to align companies' terms of service according to standards set by the Secretary of State, and to give OFCOM oversight over companies' discharge of their duties. But in failing to address the ways that social media platforms garner their profits – including the personal data harvesting that enables profiling – and how this structures online environments, and given that the largest social media companies who have a monopoly on online environments have little incentive to change their business models, the Government risks viewing the OSB as a panacea for online harms, without actually tackling their root causes.

ERODING ENCRYPTION AND UNDOING ANONYMITY

53. The OSB gives OFCOM the ability to issue service providers with a notice to deal with terrorism content or CSEA content (clause 103). Such a notice will require providers to use "accredited technology" to identify and swiftly take down terrorism and CSEA content respectively. "Accredited technology" is technology which OFCOM or another person appointed by OFCOM has designated as "meeting minimum standards of accuracy in the detection of terrorism content or CSEA content" (clause 105(9)), standards which must be approved and published by the Secretary of State (clause

[35] Paragraph 41, Joint Committee on Human Rights, *Draft Online Safety Bill: Pre-legislative scrutiny report*, available at: https://publications.parliament.uk/pa/jt5802/jtselect/jtonlinesafety/129/129.pdf; see also: Open Rights Group, *Our Adtech challenge: What we won, what we lost and what we do next,* 11 December 2021, available at: https://www.openrightsgroup.org/blog/our-adtech-challenge-what-we-won-what-we-lost-and-what-we-do-next/

105(10)). Service providers may use accredited technology alone or such technology in addition to human moderators in order to fulfil these requirements (clause 103(5)).

54. CSEA is a serious human rights issue with life-changing impacts for children, families, and wider communities. The National Society for the Prevention of Cruelty for Children (NSPCC), among others, has highlighted the increasing scale and complexity of online CSEA, particularly in light of the coronavirus pandemic, and the urgency of reforms to internet regulation to mitigate against the harms of this phenomena.[36] We acknowledge the Government's well-intentioned proposals within the OSB to minimise the prevalence of CSEA content on the internet, which are informed by the recommendations of the JCHR, and we do not underestimate the difficulty of eliminating CSEA content online nor the hard work of civil society groups and campaigners to ensure that tackling CSEA is a part of the Government's agenda.

55. Similarly, we acknowledge that terrorism poses a threat to human rights, the rule of law and democratic values. Nonetheless, we have consistently warned of the extremely broad scope of existing terrorism offences and the harmful effects of the Prevent duty in chilling freedom of expression and violating marginalised communities' rights,[37] and have argued for proportionate and rights-respecting responses to this complex issue.

56. We are concerned about the implications of OFCOM's ability to issue notices requiring service providers to use "accredited technologies" to detect CSEA and terrorism content for data, privacy and free expression rights. First, the requirement for internet service providers to immediately take down terrorism and CSEA content risks **creating a chilling effect for freedom of expression**, as service providers may pre-emptively take down content even if it does not fall within those categories to avoid breaching their duties and incurring heavy penalties. Most internet service providers rely on automated content detection software which is prone to mistakenly flag art, health information, educational resources, advocacy messages, and other imagery, which **exacerbates the risk of inaccurate takedowns at scale**.[38] Not only will this result in people's content being removed, this can actually have counterproductive effects, stopping particular expression from being challenged.

[36] NSPCC, Briefing on the draft Online Safety Bill, September 2021, available at: https://www.nspcc.org.uk/globalassets/documents/online-safety/parliamentary-briefing---draft-online-safety-bill---sept-2021.pdf.
[37] The Home Office, "Prevent Strategy", (June 2011), p.1. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/97976/prevent-strategy-review.pdff
[38] 90+ digital rights organisations ask Apple to drop image surveillance plans, Digital Rights Watch, 26 August 2021, available at: https://digitalrightswatch.org.au/2021/08/26/90-digital-rights-organisations-ask-apple-to-drop-image-surveillance-plans/

57. Second, and crucially, we are concerned that **expanding the use of "accredited technologies" may erode end-to-end encryption (E2EE).** As it stands, certain messaging services such as WhatsApp and Signal are end-to-end encrypted, which means that third parties (such as private companies and state governments) cannot access users' direct messages to one another. All around the world, E2EE has enabled everyone from political dissidents to marginalised communities (such as LGBTQ+ people) to be able to communicate and express themselves without fear.

58. Some groups have expressed concerns over the ways that E2EE has enabled certain forms of abuse, including illegal activities such as fraud. Child protection groups in particular have argued that E2EE must be accompanied by child protections safeguards.[39] In August 2021, Apple announced that it would be installing surveillance software that will conduct on-device scanning in Messages and Photos, with the aim of protecting children and reducing the spread of child sexual abuse material. In response, an international coalition of more than 90 civil society organisations, including Liberty, wrote to Apple, expressing our concern over the ways that eroding E2EE itself may infringe on the rights of children, in particular their privacy and free expression rights.[40]

59. Clause 103 of the OSB stipulates that service providers may be required to detect CSEA content "whether communicated publicly or *privately* by means of the service (emphasis added)". We are concerned about the impact that this will have on people's ability to express themselves and to communicate with others, and the extent to which this might give rise to self-censorship online, with detrimental impacts on individuals' rights. As argued by Open Rights Group, once an "accredited technology" is deployed on a platform, it need not be limited to checking for terrorism or CSEA images; it may easily be expanded to encompass other forms of content. **We are concerned about the slippery slope that may arise, whereby the Government may seek to introduce new categories of content falling under clause 103, under the auspices of which E2EE and users' privacy and freedom of expression may be further eroded.**

60. **Liberty believes that E2EE is a vital way for online users to be safe and for their privacy and free expression rights to be protected.** Attempts to limit or restrict these rights must be necessary and proportionate to a legitimate aim. **Due to the wide scope**

[39] NSPCC, Private messaging and the rollout of end-to-end encryption: The implications for child protection, available at: https://www.nspcc.org.uk/globalassets/documents/news/nspcc-discussion-paper-private-messaging-and-the-roll-out-on-end-to-end-encryption.pdf

[40] Franklin, S.B. and Nojeim, G., *International Coalition Calls on Apple to Abandon Plan to Build Surveillance Capabilities into iPhones, iPads, and other Products,* Center for Democracy and Technology, available at: https://cdt.org/insights/international-coalition-calls-on-apple-to-abandon-plan-to-build-surveillance-capabilities-into-iphones-ipads-and-other-products/

of this duty, and the potential for abuse, we believe that technology notices may result in disproportionate interferences with people's right to privacy and free expression.

61. Related to encryption is the right to anonymity. Anonymity is a vital tool for people to be able to exercise their right to freedom of expression and to privacy; it has enabled survivors of domestic abuse to seek support, whistleblowers to expose exploitative labour practices, and LGBTQ+ people to connect with others in the community. At the same time, anonymity can provide people with cover to engage in abusive and intimidatory speech and behaviour.

62. Clause 57 requires that Category 1 services (the largest user-to-user services, e.g. social media platforms) provide all adult users of the service the option to verify their identity. It states that the verification process may be of any kind, and in particular, it need not require documentation to be provided, and that the service must include clear and accessible provisions explaining how the verification process works.

63. Although user verification is merely an option, and not mandated by the OSB, we are concerned that its inclusion will pave the way for greater use of this technology in the future. Depending on how social media platforms decide to treat verified versus non-verified users, there may emerge a 'two-tier' online environment where these two different groups have vastly different experiences, such that verification becomes the only way to access the full suite of capabilities that a platform has to offer. While no one has the right to use a social media platform, we also believe that people should not have to choose between keeping themselves safe and being able to express themselves in key forums of public conversation.

CHILDREN'S RIGHTS

64. Clauses 11 and 26 set out the safety duties which apply to regulated user-to-user services and search services likely to be accessed by children respectively (the definition of 'likely to be accessed by children' is set out in clause 33). Regulated user-to-user services will have a duty to prevent children from encountering primary priority content or other harmful content, including through the use of age verification or another means of age assurance, with 'age assurance' meaning measures designed to estimate or verify the age or age-range of users of a service. Search services will have a similar duty to operate a service using proportionate systems and processes designed to minimise the risk of children encountering primary priority content or other harmful content (clause 26(3)).

65. **Children have rights to freedom of association, participation, information, and privacy, as well as to protection from harm.** While the OSB presents age verification software as a way of protecting children, it may have the effect of locking certain children from marginalised backgrounds out of participating in online environments, including those for whom the internet may be a source of community, friendship, and knowledge: Internet Matters highlighted in its evidence to the JCHR that some vulnerable children, including care experienced children and young people, may lack acceptable forms of official ID required for service providers' age verification systems, and some young people may be unable to comply with age assurance mechanisms for physical or cognitive reasons.[41]

66. We are concerned that the normalisation of age assurance software will pave the way for the further loss of anonymity online. Indeed, once age verification is rolled out for certain parts of a social media platform, **it is easily rolled out for the rest of the platform,** especially when the duty on Category 1 services to create a user verification option is taken into account.

67. We also agree with Defenddigitalme's assessment that the Bill risks embedding "the commercialisation of childhood" from an early age, and rendering children even more vulnerable to surveillance and targeting by private companies. The normalisation of age and user verification risks requiring everyone to "give up their right to privacy and hand over personal details to commercial companies to create a credential for identifying themselves to websites as not-a-child or a child of a defined age group."[42] Again, this risks further entrenching the root causes of online harms.

OFCOM'S REGULATORY INDEPENDENCE

68. We are concerned about the extent of the Secretary of State's (SoS) control over OFCOM, which we believe may undermine OFCOM's role as an independent regulator.

Strategic priorities

69. In normal circumstances, the SoS will have the ability to set OFCOM's strategic priorities (clause 109(2)), with certain duties to consult with OFCOM and other relevant stakeholders. But the SoS also retains the ability to make directions to OFCOM if it has

[41] Paragraph 230, Joint Committee on Human Rights, *Draft Online Safety Bill: Pre-legislative scrutiny report*, available at: https://publications.parliament.uk/pa/jt5802/jtselect/jtonlinesafety/129/129.pdf
[42] Persson, J., *The online safety bill will do little to protect children,* Schools Week, 28 March 2022, available at: https://schoolsweek.co.uk/the-online-safety-bill-will-do-little-to-protect-children/

reasonable grounds for believing that circumstances exist that present a threat to the health or safety of the public or to national security (clause 146(1), in which case OFCOM must give priority for a specified period to specified objectives designed to address the SoS's identified threats. One example may be the recent Covid-19 pandemic. **We believe that the ability of the SoS to set OFCOM's strategic priorities may undermine the latter's role and expertise as the independent online regulator.**

Codes of practice

70. Under the OSB, OFCOM must submit its draft codes of practice to the Secretary of State (clause 39(1)). If the SOS does not intend to make directions to OFCOM to amend the code, the SOS must lay the draft code before Parliament; the negative procedure will apply prior to the code coming into force (clause 39(4)). The SoS retains the power to make directions requiring OFCOM to amend its codes of practice if they believe that modifications are required for reasons of public policy or in the case of a terrorism or CSEA code of practice, for reasons relating to national security or public safety (clause 40). If the SoS requests OFCOM to amend its codes of practice for reasons of public policy, then the affirmative procedure will apply, and both houses of Parliament must approve the code before it comes into force (clause 41(2)); otherwise, the negative procedure will apply to the amended codes (clause 41(3)).

71. We welcome the OSB's provision for heightened scrutiny of the SoS's justifications when they relate to 'public policy'. **Nonetheless, we are concerned that this is a wide category, and risks giving the SoS latitude to undermine the independent judgment of OFCOM. Furthermore, we believe that Parliament should have greater oversight over all of OFCOM's codes of practice, given that wide carve-outs such as for national security reasons also risk being easily abused.**

COMMUNICATIONS OFFENCES

72. The OSB repeals the offences in the Malicious Communications Act 1988 and sections 127(1) and 127(2)(a)-(b) of the Communications Act 2003 and introduces new offences in their place. The three new replacement offences are the harmful communications offence, the false information offence, and the threatening communications offence, which are largely based on the Law Commission's recommendations in its 2021 *Modernising Communications Offences* report. These offences are not limited to the internet or online environments, and can cover a range of communications including

letters. In addition, the OSB creates a new offence targeting cyberflashing. **We focus our attention on the harmful communications offence.**

73. Many organisations, academics, lawyers, and even the police, have been critics of the existing harmful communications offence for its unjustifiably broad scope of criminalisation. For example, a man who tweeted that Captain Sir Tom Moore (who walked around his garden to raise money for the NHS during the coronavirus pandemic) should "burn auld fella buuuuurn" the day after Moore died, was found guilty of sending a "grossly offensive" tweet.[43] The Law Commission's report echoes these concerns, arguing that the existing offence, in its reference to "grossly offensive" or "indecent" content, lacks universally accepted definitions, which in turn means that the offence is susceptible to subjective and inconsistent interpretation. The Law Commission also highlights the important fact that "grossly offensive" and/or "indecent" content is not inherently wrongful, meaning that it should not necessarily be criminalised.

74. Clause 150 sets out the new harmful communications offence, which provides that a person commits an offence if a person sends a message and at the time of sending the message there was a "real and substantial risk that it would cause harm to a likely audience" and "the person intended to cause harm to a likely audience"; and the person had no reasonable excuse for sending the message (clause 150(1)). The definition of 'harm' in the context of this offence is "psychological harm amounting to at least serious distress" (clause 150(4)). An individual is a "likely audience" if, at the time the message is sent, it is reasonably foreseeable that they would encounter the message or, in the online context, would encounter a subsequent message forwarding or sharing the content (clause 150(2)).[44] In deciding whether a person has a reasonable excuse for sending a message, one of the factors the court must consider is whether the message is, or is intended to be, a contribution to a matter of public interest, although this is not determinative (clause 150(5)). A person who commits an offence is liable on summary conviction to imprisonment for a term not exceeding the maximum summary term for either-way offences or a fine (or both); and on conviction on indictment to imprisonment for a term not exceeding two years or a fine (or both) (clause 150(8)). Certain news organisations are exempt from committing this offence.

75. According to the Explanatory Notes, the new harmful communications offence, which is based largely on the Law Commission's recommendations, is intended to criminalise

[43] Barker, D., *Man found guilty of 'grossly offensive' Captain Tom tweet,* STV News, 31 January 2022, available at: https://news.stv.tv/west-central/man-found-guilty-of-grossly-offensive-captain-tom-tweet

[44] In a case where several or many individuals are a likely audience, it is not necessary for the person sending the message to intend to cause harm to any one of them in particular or to all of them (clause 150(3)).

communications based on the potential for harm in a particular context (rather than focusing on whether the communication is 'grossly offensive') and on the culpability of the sender as intending a particular result.[45] The Law Commission contends that the key strength of its proposed, context-specific approach is that "it does not require "universal" standards of offensiveness, indecency or obscenity: it does not carry the risk of juries or magistrates importing their own subjective standards by requiring them to define these terms, and instead requires them only to assess the facts and consider whether, on the evidence, harm was likely."[46]

76. We welcome the long overdue reform of the harmful communications offence. However, we are concerned about the focus on 'likely harm' rather than 'actual harm', given that this would appear to require a hypothetical judgment as to the purported harmfulness of content. Furthermore, we share internet regulation lawyer and expert Graham Smith's concern regarding the potentially wide scope of the offence, given the inclusion of the concept of the "likely audience". This offence appears straightforward in the context of a private communication sent to one person or a particular group of people – the Law Commission provides the example of a tweet directed at a disability charity by means of the '@' function, where the likely audience would be the charity and its followers. Nonetheless, the picture becomes less clear when applied to public posts to a general audience. For example, if someone publishes a public tweet, how will the courts determine if there was a real and substantial risk that it would cause harm to a likely audience and the individual intended to cause such harm, given that their intended audience was the entire Twitter community? There is also the risk that people might attempt to insert themselves into a person's "likely audience" by responding to a tweet with the aim of provoking a response. We are concerned that the above concerns could result in a chilling effect on freedom of expression, including by deterring people from expressing themselves.[47]

CONCLUSION

77. The Government has marketed the OSB as a world-first in online regulation. Nonetheless, the Bill raises numerous serious issues, with the potential to infringe on individuals' rights to freedom of expression and privacy. At the same time, the Government looks set to

[45] Paragraph 628, Online Safety Bill: Explanatory Notes, available at: https://publications.parliament.uk/pa/bills/cbill/58-02/0285/210285en.pdf
[46] Paragraph 2.6, Law Commission, Modernising Communications Offences: A final report, 20 July 2021, available at: https://s3-eu-west-2.amazonaws.com/lawcom-prod-storage-11jsxou24uy7q/uploads/2021/07/Modernising-Communications-Offences-2021-Law-Com-No-399.pdf
[47] Smith, G., *License to chill,* Cyberleagle, 22 November 2021, available at: https://www.cyberleagle.com/2021/11/licence-to-chill.html

overhaul the Human Rights Act as well as the UK's data protection legislation. Not only must protections for human rights be strengthened within the OSB, the wider rights landscape must be preserved so as to provide as an essential safeguard for those seeking to communicate and express themselves both offline and online. **We urge parliamentarians to scrutinise the OSB robustly for its impacts on human rights and to commit to defending the Human Rights Act as a vital safeguard for individuals' freedoms both off and online.**

JUN PANG
Policy and Campaigns Officer
junp@libertyhumanrights.org.uk