Joint Committee on the Draft Online Safety Bill

GLOBAL PARTNERS DIGITAL

Global Partners Digital submission
September 2021

About Global Partners Digital

Global Partners Digital (GPD) is a social purpose company dedicated to fostering a digital environment underpinned by human rights.

Introduction

We welcome the opportunity to provide a submission to the Committee on the government's Draft Online Safety Bill (Draft Bill). GPD recognises the legitimate desire of the government to tackle illegal and harmful content and behaviour online, and we believe that some of the provisions in the Draft Bill are reasonable and sensible. Based on our analysis, however, we believe that many provisions of the Draft Bill, if taken forward in their current form, would pose significant risks to individuals' right to freedom of expression and privacy online and could be inconsistent with the United Kingdom's international human rights obligations. As such, our evidence is primarily targeted towards the Committee's question, "*Does the proposed legislation represent a threat to freedom of expression, or are the protections for freedom of expression provided in the draft Bill sufficient?*".

In this response, we relay our concerns and make a series of recommendations on how the Draft Bill could be revised and amended to mitigate these risks. We believe these recommendations, if incorporated into the final Bill, will help safeguard freedom of expression and privacy online.

Summary of analysis of the Draft Bill

- **Illegal content**: We agree with the principle that what is illegal offline should be illegal online. Indeed, this is already the case. The Draft Bill goes far beyond this principle however, proposing a radically different mechanism for responding to illegal content online. The provisions essentially delegate to services decisionmaking as to what is illegal under UK law, incentivise the over-removal of content where its legality is not clear, and provide limited safeguards or oversight compared to restrictions on freedom of expression in the offline environment. The approach also undermines the role that the criminal justice system plays in ensuring accountability and deterrence to others. As such, clause 41 should be amended to remove the catch all definition of "illegal content" and include an exhaustive list of the most serious types of illegal content. Services should not be required or expected to make determinations of the legality of content themselves, and clause 9(3)(d) should be removed or amended to require services to remove content "swiftly" only where a determination by a court that that particular piece of content is illegal has taken place.

- **Content that is harmful to adults**: We firmly believe as a matter of principle that content which is legal for adults should not be restricted online. If Parliament considers that there are certain forms of online speech which adults should not be able to post, share or receive, these should be clearly and precisely defined and prohibited through primary legislation, whether criminal or civil, with such prohibitions also applying to equivalent speech offline. Any approach which leads to the removal of certain types of speech online, but not offline, would amount to a two-tier system of speech regulation for which we cannot see any

justification. Clause 11 and all other provisions in the Draft Bill relating to content that is harmful to adults but which is legal should ideally be removed or, alternatively, significantly reworded to make clear that there is no requirement or expectation on services to remove such content.

- **Content that is harmful to children**: We recognise that there are certain forms of content which are generally legal, but which may not be suitable for children and which are regulated in other areas of life. It is also important to recognise that children, too, have a right to freedom of expression and should be able to generate, share and access content which might be upsetting or offensive. This is an extremely sensitive area, complicated by the fact that responses should be tailored for children of different ages and with different levels of maturity. It is also important that in seeking to protect children, adults or not unduly restricted from generating, sharing or accessing content. The approach of providing a single definition in clause 47 should replaced with new clauses which clearly and precisely define each of the different types of content that is harmful to children. Clause 26 should be re-worded to ensure that platforms are not forced with an invidious choice of either using age verification or restricting adults' right to freedom of expression.

- **General monitoring:** It has long been a basic principle of the regulation of online services that they should not be required to undertake general monitoring of the content on those platforms. This principle, given legal status while the United Kingdom was a member of the European Union via Article 15 of the E-Commerce Directive, has been recognised as critical for the protection of the right to freedom of expression. While the United Kingdom has left the European Union, this should not result in a lower level of protection of the right to freedom of expression than the European Union. However, the Draft Bill contains no prohibition on general monitoring and certain provisions can be read as requiring it. For the avoidance of doubt, a new clause should be inserted into the final Bill which prohibits services from undertaking general monitoring of content and from actively seeking facts or circumstances indicating that content is illegal or harmful to adults or children.

- **Private communications**: The right to be able to communicate privately with others, free from interference, is a critical element of both the rights to freedom of expression and the right to privacy. It is fundamentally important that, in seeking address illegal and harmful content online, that our ability to communicate privately is not undermined. The Draft Bill, however, provides only very limited safeguards and protections for private communications. even that limited safeguard appears to have been further watered with the Draft Bill making almost no reference to protections for private communications, in contrast to many other jurisdictions where they are entirely out of scope. Private communications should be explicitly outside the final Bill's scope through amendment to the definition of "content" in clause 137(1) or by expanding the list of excluded services in clause 39(2) and Schedule 1 to include all private communications and storage services, including interpersonal communications services.

- If, however, the scope of the final Bill does include private communication services, it should provide for more meaningful safeguards to ensure that compliance with any duties does not undermine the right to privacy. At a minimum, this should include (i) clarity that compliance any duties does not require platforms to remove (or to refrain from introducing) private communication services or privacy-enhancing technologies such as end-to-end encryption; and (ii) an explicit requirement that OfCom may not require any such measures to be taken in its codes of practice.

- The Draft Bill contains provisions that have never been proposed in any democratic country, the potential for services to be mandated to use certain forms of "technology" to

identify illegal content on both public and private parts of the service. We have significant concerns about their inclusion due to their inconsistency with any prohibition of general monitoring (see above), their levels of inaccuracy, and disproportionate impact on persons from minority groups, and the precedent this sets for other governments in authoritarian countries to demand that companies use technology to scan private and encrypted communications for content illegal under their laws. In authoritarian countries this could mean content that was merely critical of the government, related to LGBT+ individuals or religious minorities. Clauses 63 to 69 should be removed from the final Bill.

- **Services' duties to respect freedom of expression and privacy**: We recognise that some effort has been made in the Draft Bill to ensure that services within scope consider the rights to freedom of expression and privacy. However, those duties are weak, and certainly far weaker than the other duties in the Draft Bill, simply requiring services to "have regard" to these rights. Services must be given the confidence and ability to ensure that they are not forced or pressured to take action which they consider would undermine their users' right to freedom of expression and privacy. Clauses 12(1) and 23(1) should be re-worded to strengthen this duty and give it equivalence to those other duties to which they are subject.

- **Regulatory oversight:** Much of the actual impact of the final Bill will be determined by its enforcement by OfCom, with the potential for risks to freedom of expression and privacy to be mitigated through proportionate and human rights-respecting regulatory enforcement and oversight. It is therefore essential both that OfCom has a clear and explicit steer towards such an approach and is given sufficient independence to be able to enforce the legislation without interference from the government. In both respects, the Draft Bill falls short. OfCom's duties with respect to respecting freedom of expression should be strengthened and extend to all of its regulatory duties rather than simply the steps recommended in codes of practices. Those provisions which most seriously undermine OfCom's independence, and risk political interference in its day-to-day work, should removed, including clauses 33, 34(6), 57 and 113.

Framework for analysis

Our analysis of the Draft Bill is based on international human rights law, specifically the International Covenant on Civil and Political Rights (ICCPR), ratified by the United Kingdom in 1976. Article 19 of the ICCPR guarantees the right to freedom of expression, including the right to receive and impart information and ideas of all kinds regardless of frontiers. Article 17 of the ICCPR guarantees the right to privacy and provides that "no one shall be subjected to arbitrary or unlawful interference with his privacy, family, home or correspondence". Restrictions on the right to freedom of expression or privacy guaranteed under international human rights law are only permissible when they can be justified. In order to be justified, restrictions must meet a three-part test, namely that: (1) restrictions are provided by law; (2) restrictions pursue a legitimate aim; and (3) restrictions must be necessary and proportionate, which requires that the restriction be the least restrictive means required to achieve the purported aim. The United Kingdom has similar obligations under the European Convention on Human Rights (respectively Articles 10 and 8).

It is important to remember that the United Kingdom's obligation to ensure that these rights are not unjustifiably restricted exists both in relation to restrictions which stem from the actions of the state itself as well as those caused by third parties, such as private companies. As such, it makes no difference from the perspective of the individual affected whether any restrictions are imposed and enforced directly by the state (e.g. through creating criminal offences which are enforced by the police and the courts) or through third parties, particularly when the third party is acting in order to comply with legal obligations.

With respect to the actions of private companies specifically, the United Nations Guiding Principles on Business and Human Rights (UNGPs) makes clear that a state's international human rights obligations include establishing a legal and policy framework which enables and supports businesses to respect human rights. Principle 3 notes that this general obligation includes ensuring "that (...) laws and policies governing the creation and ongoing operation of business enterprises, such as corporate law, do not constrain but enable business respect for human rights".

Given the impact that online platforms have upon the enjoyment and exercise of the rights to freedom of expression and privacy, the government has a clear obligation to ensure that these rights are respected by these platforms. This includes ensuring that legislation and other measures do not constrain online platforms' ability to respect the right to freedom of expression or privacy themselves, nor should they directly or indirectly constitute a restriction on the enjoyment and exercise of those rights by those that use such platforms.

Our analysis of the regulatory measures proposed in the general scheme are based on these frameworks. Given the limited existing interpretation and case-law of these frameworks as they apply to measures comparable to those proposed in the general scheme, we also make reference, as appropriate, to Recommendation CM/Rec(2018)2 of the Council of Europe's Committee of Ministers to member States on the roles and responsibilities of internet intermediaries (Recommendation CM/Rec(2018)2),[1] and relevant commentary from the UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression (the UN Special Rapporteur). These guidelines and commentaries provide detail on the obligations of states with respect to the protection and promotion of human rights in the digital environment, with a particular focus on any legal frameworks that apply to internet intermediaries.

Though not a framework for the purpose of our analysis, we note that United Kingdom has, through its membership of the Freedom Online Coalition, signed up to a number of commitments which are relevant to the subject. These includes commitments made in the "Recommendations for Freedom Online, Adopted in Tallinn, Estonia on April 28, 2014 by Ministers of the Freedom Online Coalition":

> "We, the members of the Freedom Online Coalition
>
> 4. Dedicate ourselves, in conducting our own activities, to respect our human rights obligations, as well as the principles of the rule of law, legitimate purpose, non-arbitrariness, effective oversight, and transparency, and call upon others to do the same,
>
> (...)
>
> 6. Call upon governments worldwide to promote transparency and independent, effective domestic oversight related to electronic surveillance, use of content take-down notices, limitations or restrictions on online content or user access and other similar measures, while committing ourselves to do the same".[2]

[1] Council of Europe, Recommendation CM/Rec(2018)2 of the Committee of Ministers to member States on the roles and responsibilities of internet intermediaries, 7 March 2018.
[2] Recommendations for Freedom Online, Adopted in Tallinn, Estonia on April 28, 2014 by Ministers of the Freedom Online Coalition, available at: https://www.freedomonlinecoalition.com/wp-content/uploads/2014/04/FOC-recommendations-consensus.pdf.

More recent commitments were made in the Freedom Online Coalition's "Joint Statement on Internet Censorship":

> "In 2017, the world witnessed state-sponsored Internet censorship in various forms: states have manipulated and suppressed online expression protected by international law, have subjected users to arbitrary or unlawful surveillance, have used liability laws to force ICT companies to self-censor expression protected by international law, have disrupted networks to deny users access to information, and have employed elaborate technical measures to maintain their online censorship capabilities. Further unlawful efforts included state censorship in private messaging apps and systematic bans of news websites and social media. Likewise certain states have introduced or implemented laws which permit executive authorities to limit content, on the Internet broadly and without appropriate procedural safeguards. Individuals who may face multiple and intersecting forms of discrimination, including women and girls, often faced disproportionate levels of censorship and punishment.
>
> (...)
>
> The FOC firmly believes in the value of free and informed political debate, offline and online, and its positive effects on long term political stability. The Coalition calls on governments, the private sector, international organizations, civil society, and Internet stakeholders to work together toward a shared approach - firmly grounded in respect for international human rights law - that aims to evaluate, respond to, and if necessary, remedy state-sponsored efforts to restrict, moderate, or manipulate online content, and that calls for greater transparency of private Internet companies' mediation, automation, and remedial policies."[3]

Full human rights analysis of the Draft Bill

Illegal content (clauses 9, 21 and 41)

We agree with the principle that what is illegal offline should be illegal online (subject to the caveat that that which is illegal offline is consistent with international human rights law and standards). Indeed, that is already the case. A person who commits a criminal offence through their activity online (whether generating or sharing particular content or behaving in a particular way) should be treated in the same way as a person who commits a criminal offence through their offline activity. That it can be more difficult for the police to investigate and collect sufficient evidence for prosecution in relation to offences committed online is no reason to depart from this longstanding principle.

At the same time, and consistent with this principle, it is also reasonable to expect that online platform who have knowledge that certain content on their platforms is illegal under UK law remove it expeditiously.[4]

---

[3] The Freedom Online Coalition, Joint Statement on Internet Censorship, available at: https://freedomonlinecoalition.com/wp-content/uploads/2018/05/FOC-Joint-Statement-on-Internet-Censorship-0518.pdf.

[4] Recognising that the UK comprises a number of different legal jurisdictions, the submission uses the term "UK law" to refer to all of the legal frameworks in the UK, whether UK-wide or the English and Welsh, Scottish or Northern Irish legal frameworks.

The Draft Bill, however, goes far beyond this principle, proposing a radically different mechanism for responding to illegal content online. The provisions essentially delegate to services decisionmaking as to what is illegal under UK law, incentivise the over-removal of content where its legality is not clear, and provide limited safeguards compared to restrictions on freedom of expression in the offline environment.

An overly broad approach to illegal content

The Draft Bill does not anywhere set out an exhaustive list of the types of illegal content within its scope. While it does do so with respect to criminal offences relating to terrorism or child sexual exploitation and abuse (CSEA), the definition of illegal content in clause 41(4) includes any "offence (…) of which the victim or intended victim is an individual (or individuals)". That the government itself has been unable to produce a list of which criminal offences fall within this definition demonstrates the breadth of this task: the victim of the vast majority of criminal offences is an individual or individuals.

Even considering just terrorist and CSEA offences in England and Wales, Schedules 2 and 3 of the Draft Bill list over 30 different criminal offences, not including inchoate offences such as attempting or conspiring to commit an offence; encouraging or assisting the commission of an offence; and aiding, abetting, counselling or procuring the commission of an offence. Including Scotland and Northern Ireland, there are around 100 offences. It is likely that the total number of offences that would be covered by the broad definition in clause 41(4) would among to several hundred, if not thousands.

Many of these criminal offences are extremely broad, and include offences relating to communications which are "grossly offensive", "indecent"; "obscene" or "false" (see section 127 of the Communications Act 2003 and section 1 of the Malicious Communications Act 1988). The Law Commission has recognised that the scope of these particular criminal offences over-criminalises and "permits prosecutions that, absent the careful prosecutorial guidance we have seen, would be so great in number as to swamp the criminal justice system, and may nonetheless constitute an unjustifiable interference in freedom of expression".[5] In addition to prosecutorial guidance, the police and the Crown Prosecution Service have a duty under section 6 of the Human Rights Act 1998 not to act in a way incompatible with the European Convention on Human Rights, which includes the right to freedom of expression. This means that many instances of speech which are technically illegal are neither investigated nor prosecuted as to do so would constitute an interference with the right to freedom of expression. The Draft Bill, however, contains no equivalent to any such prosecutorial guidance, or to the duty under section 6 of the Human Rights Act 1998, meaning that content which is illegal under these broad offences would be removed, even though its offline equivalent would not be prosecuted.

In short, it is wholly unreasonable to expect online platforms, including small platforms, to take action in respect to hundreds, possibly thousands, of different criminal offences, many of which are extremely broad in scope, when they are undertaking illegal content risk assessments (clause 7(1), mitigating and managing the risk of these offences being committed (clause 9(2)), specifying in terms and conditions what action they take in relation to different types of illegal content (clause 9(4)) and making determinations as to whether individual pieces of content are illegal (clause 9(3)(d)). If nothing else, there is a real risk that in trying to take action in relation to hundreds, if not thousands, of broadly-worded criminal offences, online platforms will fail to prioritise the actions that are needed to address the most severe forms of illegal content that cause the most harm.

[5] Law Commission, "Modernising Communications Offences: A final report", HC 547, 2021, Para 1.6.

The approach also contrasts with that taken in other jurisdictions. Germany's NetzDG sets out an exhaustive list of criminal offences within scope, as does the European Union's proposal for a Digital Services Act. Canada's recently published proposals for tackling illegal content online lists just five priority types of content which are illegal under Canadian law.

The final Bill should, instead of covering all criminal offences, focus on the most serious forms of illegal content; these should be listed explicitly in the Bill. This can be done by removing the catch-all provision in clause 41(4)(d) and limiting the definition of illegal content to the existing forms of terrorist and CSEA content listed in Schedules 2 and 3 to the Draft Bill alongside a proportionate number of further criminal offences which should all be listed explicitly. This list should not, however, include the communications-related criminal offences in section 127 of the Communications Act 2003 and section 1 of the Malicious Communications Act 1988 until these have been amended in line with the recommendations of the Law Commission.

Recommendation 1: Clause 41(4)(d) of the Draft Bill should be deleted.

Recommendation 2: Clause 41(4)(c) of the Draft Bill should be replaced with an exhaustive list of further criminal offences (beyond those listed in clauses 41(4)(a) and (b)). This list should comprise a reasonable and proportionate number of criminal offences focusing on the most severe forms of illegal content that cause the most harm. It should not include the communications-related criminal offences in section 127 of the Communications Act 2003 and section 1 of the Malicious Communications Act 1988 until these have been amended in line with the recommendations of the Law Commission.

Forcing platforms to determine the legality of content

Even with a more limited number of criminal offences, a further problem with the Draft Bill is that it forces online services to make determinations as to whether particular pieces of content on their platforms are illegal. This is particularly the case as a result of clause 9(3)(d) which provides that all services must operate the service with systems and processes designed to "swiftly take down" illegal content whether they are alerted by a person of its presence or become aware of it in any other way. This means that when a service becomes aware of a piece of content that may be illegal under UK law, they will have to take it down swiftly if it is illegal, and the onus is on the platform to make that determination.

Determining whether a particular piece of speech is illegal or not is not simple. While there are certain criminal offences where it *is* relatively straightforward (such as with respect to CSEA material), for many criminal offences it is not.

To take just one example, section 1 of the Terrorist Act 2006. This offence is committed where a person publishes a statement and, at that time, intends members of the public to be directly or indirectly encouraged or otherwise induced by the statement to commit, prepare or instigate acts of terrorism or Convention offences.[6] There is also a requirement that the person who publishes the statement is reckless as to whether members of the publicly "will be directly or indirectly encouraged or otherwise induced by the statement to commit, prepare or instigate such acts or offences".

Further, section 1 provides that "statements that are likely to be understood by members of the public as indirectly encouraging the commission or preparation of acts of terrorism or

[6] A "Convention offence" is either one of the offences listed in Schedule 1 to that Act (and there are dozens listed) or "an equivalent offence under the law of a country or territory outside the United Kingdom".

Convention offences include every statement which (a) glorifies the commission or preparation (whether in the past, in the future or generally) of such acts or offences; and (b) is a statement from which those members of the public could reasonably be expected to infer that what is being glorified is being glorified as conduct that should be emulated by them in existing circumstances". Section 1 goes on to say that "the questions how a statement is likely to be understood and what members of the public could reasonably be expected to infer from it must be determined having regard both (a) to the contents of the statement as a whole; and (b) to the circumstances and manner of its publication".

This is just one criminal offence and, as noted above, the Draft Bill currently covers hundreds, of not thousands. The complexity of this offence, and the various contextual factors that need to be considered in order to determine whether a particular statement is illegal or not, is not a task that an online platform can or should be expected to undertake. Traditionally, police officers would be given weeks to collect the necessary evidence, and court proceedings would last hours or days, with expert judges making the decision. However, online platforms are expected to make these sorts of determinations with no legal expertise "swiftly". It is inevitable that, as a result, they will make rushed and incorrect decisions frequently. Given the fact that regulatory action is far more likely to be taken if platforms fail to remove enough content than if they take down too much, it is a recipe for the censorship of legal speech.

A further problem with this approach is that it undermines the critical role that the criminal justice system plays in ensuring accountability for the commission of criminal offences and the deterrent role that it plays. To take a recent example, Scott McCluskey was convicted earlier this year of posting racially abusive posts on Facebook relating to black England footballers. He was sentenced to a 14-week suspended sentence and required to wear an electronic tag for 40 weeks. If McCluskey's post had simply been deleted by Facebook, it is quite possible that he would not have been investigated, prosecuted and sentenced. There would have been no accountability for his criminal behaviour, and none of the deterrence to him or others that comes with a criminal prosecution.

While it is reasonable to expect online platforms to remove content which has been identified as illegal by an authoritative body, such as a court, it is wholly inappropriate to expect them to make those determinations themselves. The approach also undermines the role that the criminal justice system plays in ensuring accountability and deterrence to others. Provisions which require this should therefore be removed from the final Bill.

Recommendation 3: Clause 9(3)(d) should be removed from the Bill and, if necessary, replaced with a provision that requires the service to remove content "swiftly" where an authoritative determination by a court that that particular piece of content is illegal has taken place and the service has been informed as such.

Content that is harmful to adults (clauses 11 and 46)

We firmly believe as a matter of principle that content which is legal for adults should not be restricted online. If Parliament considers that there are certain forms of online speech which adults should not be able to post, share or receive, these should be clearly and precisely defined and prohibited through primary legislation, whether criminal or civil, with such prohibitions also applying to equivalent speech offline. Any approach which leads to the removal of certain types of speech online, but not offline, would amount to a two-tier system of speech regulation for which we cannot see any justification.

On that basis, we are concerned that there are provisions in the Draft Bill relating to content which is harmful – but entirely legal – to adults in clause 11. We can do no better than the repeat

the way this concern was outlined by the House of Lords Communications and Digital Committee which, in its recent report, "Free for all? Freedom of expression in the digital age", which stated that:

> "We do not support the Government's proposed duties on platforms in clause 11 of the draft Online Safety Bill relating to content which is legal but may be harmful to adults. We are not convinced that they are workable or could be implemented without unjustifiable and unprecedented interference in freedom of expression. If a type of content is seriously harmful, it should be defined and criminalised through primary legislation. It would be more effective—and more consistent with the value which has historically been attached to freedom of expression in the UK—to address content which is legal but some may find distressing through strong regulation of the design of platforms, digital citizenship education, and competition regulation."[7]

Recommendation 4: Clause 11 and all other provisions in the Draft Bill relating to content that is harmful to adults but which is legal should be removed.

If provisions relating to legal content are retained, the risks to freedom of expression must be mitigated through significant rewording of those provisions. At the outset, we recognise that the obligations in clause 11 would only apply to Category 1 services, however, it is not clear at this stage which services will be designated as Category 1 meaning that the obligations could still apply to a large number of services. In any event, from the user's perspective, what content one is able to generate, share or access should not matter on whether the service is large or small.

More importantly, we understand that the government's intention is that Category 1 services will not be required by the Draft Bill to *remove* content that is harmful to adults for the purpose of clause 11. However the current wording of clause 11(2) requires such services within scope to set out in their terms of service whether content that is harmful to adults (priority or otherwise) "is to be dealt with by the service" and regulatory enforcement action may be taken if they do not deal with it according to those terms of service. We believe that this wording is ambiguous and suggests that the services will have to take some kind of action in relation to content that is harmful to adults in order to be considered as having "dealt" with it. This ambiguity could be addressed by clarifying the wording of the clause to make clear that it does not require removal.

Recommendation 5: If clause 11 is retained in the final Bill, clause 11(2) should be reworded as:

"(2) A duty to specify in the terms of service what action, if any, the service takes in relation to content that is harmful to adults."

Further, if provisions relating to content that is harmful to adults are retained in the final Bill, it is essential that these are as clear and narrowly defined as possible. The definition of "harmful" in clause 46(3) falls far short and is near impossible to properly understand. Each of the elements of the definition – "reasonable grounds to believe", "material risk", "having, or indirectly having", "a significant adverse physical or psychological impact" and "an adult of ordinary sensibilities" – requires a degree of careful consideration in order to make assessments; combined, it is difficult to see how any service can confidently make determinations as to what determine whether content is "harmful". This difficulty is

[7] House of Lords Communications and Digital Committee, "Free for all? Freedom of expression in the digital age", 2021, Para 182.

exacerbated by clauses 46(4) to (7) which provide further gloss on certain elements in the definition in clause 46(3). There is a real risk that this confusing definition could lead to the removal of content which may be shocking, offensive or disturbing, simply on the basis that for some users it could amount to causing an adverse psychological impact.

That trying to provide a single definition of content that is "harmful to adults" is a fruitless task is reflected by the fact that no other democratic government proposing to regulate online platforms has sought to do so. Many – such as Germany, Austria and Canada – have not sought to regulate legal content at all. Those which have done so have at least provided clear and precise definitions of the types of content where action needs to be taken, for example the Irish Online Safety and Media Regulation Bill which defines "harmful content" as include "material which is likely to encourage or promote eating disorders" and "material which is likely to encourage or promote self-harm or suicide or provides instructions on how to do so" (with an the content was part of philosophical, medical or political discourse).

In any event, if provisions relating to content which is harmful to adults are retained in the final Bill, it is critical that (i) the definition of such content is clear and precise, and (ii) the duties on services are limited to ensuring that there is clarity in their terms of service

Recommendation 6: If clause 11 is retained in the final Bill, the approach of providing a single definition in clause 46 should replaced with new clauses which clearly and precisely define each of the different types of content that is harmful to adults.

Content that is harmful to children (clauses 10, 22 and 45)

In contrast to content which is legal for adults, we recognise that there are certain forms of content which are generally legal, but which may not be suitable for children and which are regulated and restricted in other areas of life, such as pornography and violent content. We also accept that it is reasonable for the government to seek to ensure that children are not exposed to content online which they would not generally be able to access or see for example, in magazines, on television or films.

That being said, it is also important to recognise that children, too, have a right to freedom of expression and should be able to generate, share and access content which might be upsetting or offensive. This is an extremely sensitive area, complicated by the fact that responses should be tailored for children of different ages and with different levels of maturity. It is also important that in seeking to protect children, adults or not unduly restricted from generating, sharing or accessing content.

On that basis, we are concerned that the provisions in the Draft Bill relating to content that is harmful to children risk both disproportionately restricting children's right to freedom of expression and creating an online environment in which either all content is child-appropriate, thus undermining adults' right to freedom of expression, or adults are forced to verify their age – providing personal and sensitive information – before they are able to access the service.

Disproportionate impacts on children's right to freedom of expression

The definition of content that is harmful to children is near-identical to the definition of content that is harmful to adults, save that the clauses use the word "children" instead of "adults". As we set out above, this definition is almost impossible to understand and potentially extremely broad in scope. There is a vast amount of content that could be considered as potentially having "a significant adverse (…) psychological impact on a child of ordinary sensibilities" but which should not be censored for all children, particularly older children. Videos of police brutality and

violence at a time when many young people are concerned about racial injustice for example, or discussions, images and videos relating to mental health, self-harm, gender reassignment surgery or domestic violence may all be upsetting or even distressing, but still an important part of a child's understanding of these issues and their personal development, especially when looking for information online to help understand their own experiences and development.

It is critically important that in seeking to protect children from content which is objectively inappropriate and harmful, that children are not prevented from accessing information which is important to them, even if it has some psychological impact by causing distress. The attempt to define all forms of content harmful to children through a single definition should be replaced with an exhaustive list of specific harms which are clearly defined.

Recommendation 7: The approach of providing a single definition in clause 47 should replaced with new clauses which clearly and precisely define each of the different types of content that is harmful to children.

The risks to adults' right to freedom of expression and privacy

Unlike the physical world where people can be segregated on the basis of age with relative ease, the online environment does not easily allow this. Most online services and websites are equally accessible to persons of all ages with the service and the service or website does not verify a person's age before they are able to use the service. The general exception to this is where the service is one which is only appropriate for those above a certain age to use, and so may require verification of some sort that the user is above that age (for example many pornography sites and social media platforms), although the form of verification may simply be self-identification.

In seeking to address the concern that it is too easy for children to access services which may contain content which is inappropriate, the Draft Bill creates a situation whereby *all* websites and services will be presumed to contain such content. This is due to the way that clause 26 approaches assessments to determine whether a particular service is "likely to be accessed by children", thus triggering the duties in clause 10.

Under clause 26(1), all services must carry out an assessment to determine whether it is possible for children to access the service or any part of it, and, if they are, whether the "child user condition" is met in relation to the service or any part of it. The reality is, however, that unless the service uses some form of age verification, then children will always be able to access the service. Indeed, clause 26(3) provides that a service can only conclude that it is not possible for children to access it, or a part of it, "if there are systems or processes in place that achieve the result that children are not normally able to access the service or that part of it". This means age verification.

Clause 26(4) goes on to say that the "child user condition" is met if "there are a significant number of children who are users of the service or part of it" and "the service, or that part of it, is of a kind likely to attract a significant number of users who are children". If this condition is met, then the service is one "likely to be accessed by children". But it is not possible for a service to know the age of the people who are users of the service unless they record people's age, and this too means some form of age verification.

The combination of these provisions is that all services will be presumed to be likely to be accessed by children unless they use age verification to determine the age of users and, if they so decide, to prevent children from accessing the service. This places online services in a situation where they either make the entirety of their service child-friendly (and thus restrict any form of content which is potentially harmful to children, even if perfectly legal and

appropriate for adults) or introduce age verification and use this to determine the age of users and restrict children.

Neither of these is appropriate: for platforms that choose not to use age verification, the service will have to restrict adults' freedom of expression to avoid non-compliance with the duties under clause 10, or it will have to collect personal and sensitive information (such as credit card details or copies of identification documents) from all its users at a time when there is general public concern that online services already collect and store too much personal information.

The choice for users is equally unappetising: a user will either have to give their personal and sensitive information away every time they want to use an online service, massively increasing their exposure to data breaches and leaks, or they will only be able to use services which are safe for children, rendering them unable to generate, share or access content which is perfectly lawful and legitimate. For adults who rely upon using services anonymously for legitimate reasons, such a whistleblowers, journalists and human rights defenders, the risks involved in providing identification will be significant, potentially rendering them unable to use the service.

The final Bill should not force such an invidious choice on online services and users. Instead, it should take a more proportionate approach which raises the threshold before a service is required to comply with the duties under clause 10. This can be achieved through amending clause 26(4) and the requirements for the child user condition to be met.

Recommendation 8: Clause 26(4) should be re-worded as follows:

(4) The "child user condition" is met in relation to a service, or a part of a service, if—

> (a) the service, or that part of it, is designed for children, or
> (b) the regulated service has reasonable grounds to believe that the service, or that part of it, attracts a significant number of users who are children.

General monitoring (clauses 9 and 10)

It has long been a basic principle of the regulation of online services that they should not be required to undertake general monitoring of the content on those platforms. Within the EU, this principle was given legal status via Article 15 of the E-Commerce Directive which prohibits member states from imposing "a general obligation on providers, when providing the services covered by [the Directive], to monitor the information which they transmit or store, nor a general obligation actively to seek facts or circumstances indicating illegal activity." In his recent opinion in the case of *Republic of Poland v European Parliament and Council of the European Union*, Advocate General Saugmandsgaard Øe highlighted how this principle is critical for the protection of the right to freedom of expression:

> "102. It is true that preventive measures for monitoring information are generally regarded as particularly serious interferences with freedom of expression on account of the excesses they may entail. Those preventive measures are, in principle, disapproved of in a democratic society, on the ground that, by restricting certain information even before its dissemination, they prevent any public debate on the content, thus depriving freedom of expression of its very function as a vehicle for pluralism. For those reasons, as the applicant points out, many Member States prohibit the general prior control of information in their respective constitutions.
>
> 103. Those considerations are fully relevant with regard to the Internet. As the applicant submits, the Internet is of particular importance to the freedom to receive and impart

> information and ideas. That is the case, more specifically, in respect of large social networks and platforms, which, by enabling anyone to upload the content they wish and the public to access it, are 'unprecedented' tools for exercising that freedom. In that respect, those platforms play a role in a form of 'democratisation' of the production of information and, although managed by private operators, they have in fact become essential infrastructures for online expression. In the current state of forms of communication, the right to freedom of expression therefore entails, in particular, the freedom to access those platforms and express oneself on them, in principle, without interference by public authority.
>
> 104. If those authorities were to impose, directly or indirectly, on intermediary service providers which control those infrastructures for expression the obligation preventively to monitor, in general, the content of users of their services in search of any kind of illegal, or even simply undesirable information, that freedom of communication would be called into question as such. In my view, the 'essence' of the right to freedom of expression, as provided for in Article 11 of the [EU Charter of Fundamental Rights], would be affected.
>
> 105. In that context, Article 15 of Directive 2000/31 is, in my view, of fundamental importance. By providing that intermediary providers cannot be made subject to a 'general obligation … to monitor the information which they transmit or store', that provision prevents online information from being subject to general preventive monitoring, delegated to those intermediaries. In so doing, it ensures that the Internet remains a free and open domain.
>
> 106. For that reason, I am inclined to regard the prohibition laid down in Article 15 of Directive 2000/31 as a general principle of law governing the Internet, in that it gives practical effect, in the digital environment, to the fundamental freedom of communication."[8]

While the United Kingdom has left the European Union, this should not result in a lower level of protection of the right to freedom of expression. Indeed, in its proposed Digital Services Act, the European Commission does not touch Article 15 of the E-Commerce Directive and, indeed, includes a new provision which provides that "No general obligation to monitor the information which providers of intermediary services transmit or store, nor actively to seek facts or circumstances indicating illegal activity shall be imposed on those providers" (Article 8).

However, the Draft Bill contains no prohibition on general monitoring. It is quite possible to read the requirements in clauses 9 and 10 to monitor all content generally in order to identify content which is (or might be) illegal (or harmful to children). Indeed, clauses 9(3)(a) to (c) and 10(3) border on requiring general monitoring, by requiring services to operate those services in a way which minimises the presence of content, minimises the length of time for which it is present, minimises its dissemination, and protects children from encountering it. To ensure that there is no dilution of the right to freedom of expression following the UK's departure from the European Union, it is essential that a prohibition on general monitoring be included in the final Bill.

Recommendation 9: A new clause should be inserted into the final Bill which provides that "Nothings in this Act requires a service to generally monitor content, nor actively to seek facts or circumstances indicating that content is illegal or harmful to adults or children."

[8] Republic of Poland v European Parliament and Council of the European Union, Case C-401/19, Opinion Advocate-General Saugmandsgaard Øe, delivered on 15 July 2021.

Private communications

The right to be able to communicate privately with others, free from interference, is a critical elements of both the rights to freedom of expression and the right to privacy. It is why, in the offline, environment, the Royal Mail does not read our letters, while telephone and mobile communication providers do not listen to our calls, and why our personal conversations are not surveilled by police officers. It is fundamentally important that, in seeking address illegal and harmful content online, that our ability to communicate privately is not undermined. The UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression has detailed how encryption and other privacy-enhancing technologies "provide the privacy and security necessary for the exercise of the right to freedom of opinion and expression in the digital age",[9] and details how they are essential for "journalists, civil society organizations, members of ethnic or religious groups, those persecuted because of their sexual orientation or gender identity, activists, scholars, artists and others to exercise the rights to freedom of opinion and expression".[10]

Recognising this, we had been pleased to see the government in its Online Harms White Paper set out that a different approach would be taken with respect to private communications, one which would not involve any form of monitoring. We were then disappointed to see in the government's full response that this commitment was watered down simply to requiring OfCom to issue codes of practice outlining the systems and processes that companies which set out "what measures are likely to be appropriate in the context of private communications". However, even that limited safeguard appears to have been further watered with the Draft Bill making almost no reference to protections for private communications and, indeed, containing provisions which would potentially mandate scanning of private communications for certain forms of illegal content.

The inclusion of private communications

The Draft Bill makes clear that it applies equally to content which is public or private via its definition of "content" in clause 137(1). The only services providing private communications which are out of scope are emails, and SMS and MMs messages, leaving the potential for the monitoring of many communications which should be considered private, and in particular one-on-one and group chats on interpersonal messaging services, whether provided as a standalone service (such as WhatsApp, Signal and Telegram) or as a feature of a broader service (such as Facebook Messenger). While there are duties on services relating to the right to privacy, these are very weak (see below) and the provisions relevant to OfCom are limited simply to a requirement that, in its codes of practice, OfCom incorporate safeguards for the protection of "the importance of protecting users from unwarranted infringements of privacy".

These safeguards are minimal, and stand in stark contrast to the approach taken in other jurisdictions. For example, in Canada, the government has announced that it will explicitly exclude "private communications" from the scope its new regulatory framework. In Ireland, the Online Safety and Media Regulation Bill contains explicit exceptions "interpersonal communications services" and "private online storage service" when it comes to what can be included in codes of practice. The European Union's proposed Digital Services Act would only apply to content which is disseminated "to the public" and would not apply to content shared within closed groups or interpersonal communication services such as emails or private messaging services.

[9] UN Human Rights Council, Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, David Kaye, UN Doc. A/HRC/29/32, 22 May 2015, Para 56.
[10] Ibid., Para 1.

By failing to provide any meaningful form of protection in relation to private communications, the Draft Bill stands in stark contrast to other legislative proposals being put forward in democratic countries.

Recommendation 10: Private communications should be explicitly outside the final Bill's scope. This could be done by replacing the words "whether publicly or privately" with "publicly" in clause 137(1) of the Draft Bill, or by expanding the list of excluded services in clause 39(2) and Schedule 1 to include all private communications and storage services, including interpersonal communications services".

If, however, the scope of the final Bill does include private communication services, it should provide for more meaningful safeguards to ensure that compliance with any duties does not undermine the right to privacy. At a minimum, this should include (i) clarity that compliance any duties does not require platforms to remove (or to refrain from introducing) private communication services or privacy-enhancing technologies such as end-to-end encryption; and (ii) an explicit requirement that OfCom may not require any such measures to be taken in its codes of practice.

Recommendation 11: If, private communication services are within scope of the final Bill, it should provide for meaningful safeguards to ensure that the right to private communications is not undermined. At a minimum, this should include:

(i) the relevant clauses relating to duties on services should include an explicit provision that compliance those duties do not require those services to remove (or to refrain from introducing) private communication services or privacy-enhancing technologies such as end-to-end encryption.

(ii) the relevant clauses relating to codes of practice developed by OfCom should include an explicit provision that codes of practice must not require or encourage services to remove (or to refrain from introducing) private communication services or privacy-enhancing technologies such as end-to-end encryption.

Technology warning notices

As noted above, the Draft Bill contains provisions that have never been proposed in any democratic country, the potential for services to be mandated to use certain forms of "technology" to identify illegal content on both public and private parts of the service. We recognise that the circumstances under which technology warning notice could be issued is limited and subjected to a number of safeguards, however we have significant concerns about their inclusion in the Draft Bill as a matter of principle.

First, as we discuss above, the prohibition on general monitoring of content is fundamental to ensuring the right to freedom of expression online. It is simply not possible to mandate a service to scan all content on its platform in order to determine whether particular pieces are illegal in a way that does not amount to a requirement to undertake general monitoring. Instead of mandating certain technologies, the government should continue to encourage the participation of online services in existing forums, processes and mechanisms, which share hashes of pieces of content identified by experts as amount to terrorist material (such as the Global Internet Forum to Counter Terrorism) and child sexual exploitation and abuse material (such as the Internet Watch Foundation). This is an area where co-regulation between online services and other stakeholders is proving effective and should be supported rather than over-ridden with unknown forms of technology.

Second, there is simply no technology that exists at the present which make determinations as to whether individual pieces of content are terrorist-related or of child sexual exploitation and abuse material with a sufficiently high degree of accuracy, particularly with regards to the former. Given that there are billions of pieces of content on the largest platforms, even an extremely low error rate of 0.1% would mean millions of pieces of content flagged incorrectly. We also know that image matching error rates are particularly high when it comes to images of persons from minority groups, meaning that they would be disproportionately affected by any content removals.

Third, in relation to the scanning of private content for child sexual exploitation and abuse material, no such technology exists which does not create an unacceptable level of risk to individuals' human rights. The recent case of Apple's proposed technology to identify such material on encrypted images stored on iCloud Photos, which the company immediately withdrew after concerns relating to privacy and other human rights were raised, demonstrates that we are still not yet at a stage where such technology is even close to being acceptable. A key concern for human rights groups is that even if the technology does exist, once one government mandates its use for a particular type of illegal content, this gives licence to governments around the world to demand that companies use that technology to scan private and encrypted communications for content illegal under their laws. In authoritarian countries this could mean content that was merely critical of the government, related to LGBT+ individuals or religious minorities.

Unless and until such technology is developed which does not pose unacceptable risks to the rights to freedom of expression and privacy, provisions which would enable OfCom to mandate its use are premature and should be removed from the final Bill.

Recommendation 12: Clauses 63 to 69 should be removed from the final Bill.

Services' duties to respect freedom of expression and privacy (clauses 12 and 23)

We recognise that some effort has been made in the Draft Bill to ensure that services within scope consider the rights to freedom of expression and privacy. However, those duties are weak, and certainly far weaker than the other duties in the Draft Bill. All clauses 12(1) and 23(1) require from services is "to have regard to the importance of protecting users' right to freedom of expression within the law, and protecting users from unwarranted infringements of privacy, when deciding on, and implementing, safety policies and procedures". A duty to "have regard" to the importance of these two rights will easily be overridden by the far more prescriptive duties set out elsewhere in the Draft Bill relating to the removal of content. As we discuss below, we are unconvinced that the duties on OfCom in relation to freedom of expression and privacy are sufficient either.

Services must be given the confidence and ability to ensure that they are not forced or pressured to take action which they consider would undermine their users' right to freedom of expression and privacy. The duty to "have regard" to these rights will be of little value when challenging demands that they may face from OfCom. This can be remedied by re-wording clauses 12 and 23 to strengthen the duty and give it equivalence to those other duties to which it is subject.

Recommendation 13: Clauses 12(1) and 23(1) should be re-worded as follows:

"(1) A duty to ensure that, when deciding on, and implementing, safety policies and procedures, they do not unjustifiably interfere with the rights set out in subsection (2) (whether held by users or other interested persons), but (where appropriate) to incorporate meaningful and effective safeguards for the protection of those rights.

(2) The rights are -

(a) The right to freedom of expression guaranteed under Article 10 of the European Convention on Human Rights, and

(b) The right to respect for private and family life, home and correspondence guaranteed under Article 8 of the European Convention on Human Rights.

Regulatory oversight

Much of the actual impact of the final Bill will be determined by its enforcement by OfCom, with the potential for risks to freedom of expression and privacy to be mitigated through proportionate and human rights-respecting regulatory enforcement and oversight. It is therefore essential both that OfCom has a clear and explicit steer towards such an approach and is given sufficient independence to be able to enforce the legislation without interference from the government. In both respects, the Draft Bill falls short.

Insufficient consideration given to OfCom's duty to protect human rights

As a public authority, OfCom has a duty under section 6 of the Human Rights Act 1998 not to act in a way which is incompatible with a right protected by the European Convention on Human Rights, including the rights to freedom of expression and privacy. However, as set out throughout this submission, many of the risks to freedom of expression come from the actions of online services in order to comply with their regulatory duties. We believe that OfCom's duty not to act in a way which is incompatible with the rights to freedom of expression and privacy means that it should not act in a way which leads to the services within scope unjustifiably interference with these rights. At present, the Draft Bill contains few provisions which would ensure that this is the case. The most significant provision comes in clauses 31(5) and (6) which provide that, with regards to the codes of practice developed by OfCom under clause 29, any steps recommended therein "must be designed in the light of the principles mentioned in subsection (6) and (where appropriate) incorporate safeguards for the protection of the matters mentioned in the principles". Those principles are "the importance of protecting the right of users and (in the case of search services) interested persons to freedom of expression within the law, and the importance of protecting users from unwarranted infringements of privacy".

This language is relatively weak, however, and risk being wholly undermined by the caveats "within the law" (with respect to the right to freedom of expression) and "unwarranted" (with respect to the right to privacy). Since the Online Safety Bill will itself be "the law", then the protections risk being meaningless since and all interferences with freedom of expression will be required or permitted by "the law" and will also be "warranted". Furthermore, the safeguards only apply when it comes the steps recommended by OfCom in any codes of practice it develops, rather than the full range of regulatory and enforcement duties and powers that Ofcom will have. Clauses 31(5) and (6) should be re-worded, and complemented by a broader general duty on OfCom, with respect to these two human rights.

Recommendation 14: Clauses 31(5) and (6) should be re-worded as follows:

"(5) Steps described in a code of practice under section 29 which are recommended for the purposes of compliance with any of the relevant duties must not require or encourage services to take action which would unjustifiably interfere with the rights set out in subsection (6), but must (where appropriate) incorporate safeguards for the protection of those rights.

(6) The rights are -

(a) The right to freedom of expression guaranteed under Article 10 of the European Convention on Human Rights, and

(b) The right to respect for private and family life, home and correspondence guaranteed under Article 8 of the European Convention on Human Rights.

Recommendation 15: A clause should be inserted into the final Bill as follows:

"OfCom's duty as a public authority under section 6(1) of the Human Rights Act 1998 not to act in a way which is incompatible with a Convention right shall, for the purposes of this Bill, be considered also as a duty not to require regulated services to act in a way which would unjustifiably interfere with the Convention right of any user or other interested person".

A lack of independence

While OfCom will be an independent regulator, many provisions of the Draft Bill give powers to the Secretary of the State to take action which would undermine OfCom's independence in its day-to-day work. The most egregious of these are: the power to direct OfCom to modify codes of practice "to ensure that the code of practice reflects government policy" (clause 33(1)(a)); the power to direct OfCom to review codes of practice (clause 34(6)); to publish a statement "that sets out strategic priorities of Her Majesty's Government in the United Kingdom relating to online safety matters" and to which OfCom must have regard in carrying out its duties (clauses 109 and 57); and to give guidance to OfCom about the exercise of its functions under the Act and to which OfCom must have regard (clause 113).

Giving the government the power to direct OfCom as to how to exercise its functions under the legislation, requiring it to act in line with the government's strategic priorities, to modify OfCom's own codes of practice, and directing OfCom as to when it needs to review them amount to a justifiable degree of political interference in the day-to-day operations of what should be an independent regulator. While it is right that the overall legislative framework is determined by Parliament, including OfCom's general duties (clause 56) and the specific duties set out elsewhere in the Bill, the clauses above go to far and risk undermining OfCom's ability to operate as an independent regulator as opposed to an extension of the government.

Recommendation 16: Clauses 33, 34(6), 57 and 113 should be removed from the final Bill.