CENTER FOR DEMOCRACY & TECHNOLOGY

June 29th, 2022

Chairs, Online Safety Bill Committee
House of Commons
London SW1A 0AA

The Center for Democracy & Technology (CDT) respectfully submits the following comments pertaining to the Online Safety Bill (OSB) currently under your consideration. We extend our thanks to the Joint Committee on the Draft Online Safety Bill for the invitation to our Europe Director to testify before the Committee when it visited Brussels in November 2021. At that time, we shared with the Committee our analysis and advocacy positions as they related to the EU Digital Services Act, given the similarity of issues covered by the Online Safety Bill. CDT is a non-profit public interest advocacy organization dedicated to advancing human rights and civil liberties in Internet and technology law and policy. CDT has offices in Washington, DC and Brussels and regularly engages in policy advocacy in the US and Europe concerning freedom of expression online, intermediary liability laws, corporate transparency and accountability, and communications privacy. While CDT appreciates the gravity of the problems that the OSB is meant to address, its filtering mandates will threaten the right to freedom of expression and undermine the availability of end-to-end encrypted technologies.

CDT understands that helping internet users stay safe online is, and should be, a priority for the UK government. Yet, this interest cannot justify disproportionate measures that threaten users' right to privacy and freedom of expression in online communications. As currently drafted, the OSB casts proportionality aside by constructing a robust and intrusive system of mandated mass surveillance and censorship.

1. **OSB filtering obligations threaten freedom of expression:**

Though the bill gestures towards protecting freedom of expression online, its stipulations ultimately encourage censorship by platforms. The OSB creates a new "duty of care" framework which governs how online intermediaries handle illegal content, as well as content that is legal but is deemed harmful. Requiring online intermediaries to police individuals' lawful speech violates the fundamental right to freedom of expression and is inconsistent with the rule of law and principles of due process. Moreover, such obligations provide overwhelming incentives for platforms to engage in over-censorship in order to avoid the risk that they will be found not to be meeting their duty of care. The OSB places undue responsibility on platforms' shoulders to determine what is legal yet unacceptable online. Platforms may take down a wide range of lawful content, viewing it as the safest approach to compliance. If they do not do so, they risk triggering the harsh penalties set forth in the bill. The government should remove from the bill any obligations for online service providers to police lawful speech.

The bill's imposition of general and proactive content monitoring obligations constitutes a serious infringement upon users' rights to free expression. In Section 104, the OSB stipulates that user-to-user services must use accredited technology to identify and remove publicly-communicated terrorism

content, as well as publicly and privately communicated Child Sexual Exploitation and Abuse (CSEA) content. In addition, under Section 117, OFCOM will have the power to order websites to use "proactive technology" to address illegal content, children's online safety, and fraudulent advertising. General monitoring obligations such as this clash with international legal standards protecting free expression. For example, Article 15 of the EU's ECommerce Directive restricts EU Member States from imposing general obligations which force platforms to monitor online speech. Though the UK is no longer bound by the ECommerce Directive, the OSB errs in ignoring such international precedent, and this type of general monitoring requirement will have dire consequences for free expression in the United Kingdom.

Prior research has demonstrated the major technical difficulties that arise in creating robust and proactive systems and processes which will be able to separate content that may be illegal or harmful from other content generated by users. Encouraging companies to rely on proactive algorithmic review overlooks the fact that content monitoring often requires contextual analysis in order to properly determine whether it may be illegal, harmful, or acceptable. In our report "Mixed Messages? The Limits of Automated Social Media Content Analysis," we explain that "Today's tools for automating social media content analysis have limited ability to parse the nuanced meaning of human communication, or to detect the intent or motivation of the speaker. … Without proper safeguards, these tools can facilitate overbroad censorship and biased enforcement of laws and of platforms' terms of service." The bill appears to assume a degree of accuracy in automated tools which current technology cannot achieve, resulting in a disconnect between technical reality and the bill's requirements.

The bill does attempt to safeguard free expression by enacting special protections for journalistic content, news publisher content, and content of democratic importance. Unfortunately, these protections introduce even more uncertainty to a bill already laden with ambiguity. Platforms will be forced to make difficult real-time decisions about whether something qualifies as journalistic content or whether it serves a democratically important function. This pulls providers of user-to-user services in two distinct and diverging directions; on the one hand, they will be worried about failing to proactively address harmful content, while on the other hand, they will be concerned about violating the provisions on free speech. The bill lacks clear guidance on how to balance these obligations, leaving platforms in a perilous position.

2. **CSEA scanning obligations threaten the use of end-to-end encrypted services:**

As the previous section discussed, the OSB's filtering obligations in Section 104 and Section 117 harm peoples' right to freedom of expression online. Section 104(2)(b) presents a unique set of additional threats to human rights, as it pertains to content scanning obligations aimed at private communications. Section 104(2)(b) gives the Office of Communications (OFCOM) the power to order user-to-user services to use accredited technology to scan both public content and private communications in order to identify CSEA content online. The OSB's scanning obligation pertaining to private communications fundamentally undermines the use of end-to-end encrypted services. Companies offering end-to-end encrypted services have no access to the content communicated

between users; only the sender and the intended recipients can view content that is sent over an encrypted platform. Experts have long held that to enable scanning, companies will have to compromise end-to-end encryption by introducing backdoors. This requirement fundamentally undermines the very purpose of encryption, which is giving users a secure way to communicate with the assurance of privacy. In sum, the coexistence of end-to-end encrypted services and mandated content scanning is an impossibility.

End-to-end encrypted services are used across the world for communication by members of marginalized and vulnerable populations, many of whom are at risk of government censorship and persecution. This is why David Kaye, the then-U.N. Special Rapporteur on freedom of expression, urged governments not to mandate surveillance "backdoors" on encryption services. In a 2015 report, Kaye argued that "States should promote strong encryption and anonymity. National laws should recognize that people are free to protect the privacy of their digital communications by using encryption technology and tools that allow anonymity online … States should avoid all measures that weaken the security that people may enjoy online, such as backdoors, weak encryption standards and key escrows."

Eliminating the functionality of end-to-end encrypted services by mandating scanning of private messages sets a dangerous precedent that may serve as impetus for repressive regimes around the world to crack down on encryption. End-to-end encrypted services are crucial for protecting people such as journalists and political dissenters who may face persecution for their news articles and ideas when repressive governments surveil communications. As such, this regulation has ramifications for the human rights of people across the world who rely on private and secure communications for their own safety. As a global leader, the United Kingdom should set precedent that safeguards human rights both within and outside its borders.

End-to-end encrypted services can actually help keep children safe online by securing the content they send and receive, shielding children's communications from bad actors who may otherwise attempt to hack and exploit their online activity. As a UNICEF report explains, children's "digital devices and communications contain personal information that could compromise both their privacy and safety if it fell into the wrong hands … Children's digital communications constitute a record of calls, texts, web searches and images, which is private and potentially sensitive information that could be used for threats or blackmail. The application of robust encryption means that this information can be more secure." Thus, while the OSB purports to make the internet safer for children, its effect on end-to-end encrypted services puts children online at risk by making their private communications vulnerable to interception.

Indeed, a human rights impact assessment conducted by Business for Social Responsibility (BSR) in 2019 at Meta's request suggested that Messenger Kids implement end-to-end encryption for the protection of users who are children because their content is at heightened risk for abuse. The report also found that encryption is fundamental to human rights and reaffirmed that client-side scanning is fundamentally incompatible with E2EE messaging services. In opening the door to content scanning of

CENTER FOR DEMOCRACY & TECHNOLOGY

private communications, the OSB ultimately threatens the protections afforded to children by end-to-end encryption.

Though the bill promises the use of "accredited technology," the details about what technology would qualify are absent. Last year, Apple abandoned its plans to introduce scanning capabilities to its products after a coalition of more than 90 U.S. and international organizations dedicated to civil rights, digital rights, and human rights warned that "algorithms designed to detect sexually explicit material are notoriously unreliable. They are prone to mistakenly flag art, health information, educational resources, advocacy messages, and other imagery." Though the bill's accreditation process may be meant to inspire confidence in scanning technology, accreditation cannot fix the demonstrated shortcomings of such tools.

Fortunately, other effective options for addressing CSEA materials exist aside from scanning the content of users' communications, as proposed in clause 104(2)(b). In our report ,"Outside Looking In: Approaches to Content Moderation in End-to-End Encrypted Systems," we detail numerous methods of content moderation that can be implemented while preserving encryption and user privacy. In that report, we concluded that "technical approaches for user-reporting and meta-data analysis are the most likely to preserve privacy and security guarantees for end-users. Both provide effective tools that can detect significant amounts of different types of problematic content on E2EE services, including abusive and harassing messages, spam, mis- and disinformation, and CSAM." Research by Riana Pfefferkorn of the Stanford Internet Observatory confirms that content-dependent techniques such as content scanning are not a "silver bullet," and highlights the utility and widespread use of other approaches to identifying CSEA materials such as user reporting. The aforementioned alternatives constitute more proportional policy solutions to the issue of CSEA content than the intrusive scanning mandates proposed in the bill, and we urge your consideration of these alternatives.

If, instead, Parliament determines that the scanning mandates should remain in the bill, we urge that it be amended to specify additional factors for OFCOM to consider before issuing technology notices to platforms. In Section 105, the bill urges OFCOM to consider what is "necessary and proportionate" before giving notice, and lists some factors which are relevant to the decision. Unfortunately, the listed factors leave out critical considerations, which the bill should be amended to include. These considerations include the security of data, costs and practicality of implementation, and potential for abuse of the technology both locally and internationally. This will afford OFCOM needed room for discretion in applying scanning mandates that greatly infringe upon peoples' privacy in their online communications.

Thank you very much for considering our views.

[For further information, please contact Greg Nojeim, Director of the CDT Security & Surveillance Project at gnojeim@cdt.org or Jessie Miller, Security & Surveillance Intern at jmiller@cdt.org.]