5RIGHTS FOUNDATION

5Rights analysis of the Online Safety Bill

March 2022

Five years in the making, the Online Safety Bill is a major piece of legislation that has the power to transform the digital world as we know it by establishing a new regulatory framework to tackle online harms. On 17 March the Bill was brought to Parliament. Here are initial thoughts from 5Rights.

***

Without decisions to fast-track provisions for children, it could be three more years before children receive the safeguards they've been promised from the Online Safety Bill. Despite its length – at 213 pages plus explanatory notes – the revised Online Safety Bill leaves a great deal to be worked out. Some of the most important issues at stake are deferred to secondary legislation or are left to Ofcom to clarify in multiple codes of practice. This would mean years of delay for children to get the safe digital world they were promised.

Services in scope

The government has rejected the Joint Committee's recommendation to bring the scope of the Bill in line with the Age Appropriate Design Code[1], which would have brought all services that create risk for children under the new regime, not only user-to-user and search services. Sites with provider-generated content that create risks for children, including many gore and pro-anorexia sites, will remain out of scope of regulation. Not only this, but companies will have to navigate a complex patchwork of digital regulation, and Ofcom and ICO will face greater challenges enforcing the Online Safety Bill alongside the Age Appropriate Design Code. This wholly undermines the government's commitment to giving children the highest levels of protection and to making the UK the safest place in the world to be online.

We welcome the inclusion of pornography sites in the revised Bill (clause 67) and particularly the requirement to protect children from encountering pornographic content through the use of age assurance technology. The impact of this duty will depend entirely on Ofcom setting binding minimum standards of security, efficacy and privacy for these technologies.

The inclusion of pornography providers shows there is no reason why the government cannot bring into scope all services likely to be accessed by children. As a result of the Age Appropriate Design Code, all information society services (ISS) are already required to conduct an assessment of whether their services are likely to be accessed by children for the protection of children's data. Those services that can demonstrate they are not likely to be accessed by children will have no further action to take, and those that are likely to be accessed will already have to undertake a child risk assessment. Extending the scope of the Online Safety Bill would only bring in services that already have obligations under the Age Appropriate Design Code, simplifying compliance for companies and ensuring regulatory harmony.

[1] Age Appropriate Design Code, Information Commissioner's Office

Age assurance

The requirements for age assurance are strengthened in the revised Bill, and there is welcome recognition that these systems must not be implemented at the cost of user privacy (clause 11(3)). Providers can only say that it is not possible for a child to access their service if they have systems or processes in place, "for example, age verification, **or another means of age assurance"** that achieve the result that children are not normally able to access the service or part of it (clause 33(2)). A service needs to meet the duties in relation to children if it has or is likely to have a "significant" number of child users (significant here meaning a significant proportion of the total number of UK users). This is the same wording as found in the draft Bill, but a subclause has been added to clarify that this calculation should be based on evidence about who actually uses a service, rather than who the intended users of the service are (clause 31(4b)). There is no indication of who provides this evidence – Ofcom, providers themselves or independent researchers.

> **60% of UK children aged 8-11 have a profile on at least one social media service, despite most social media having a minimum age requirement of 13.**
>
> **- Ofcom research, 2022[2]**

We are pleased that the government has indicated in their response to the Joint Committee that Ofcom will be setting out its expectations for how these systems should operate in codes of practice, but there is no evidence of this in the Bill itself. It is notable – and disappointing – that Ofcom is tasked under the Bill with producing guidance for ID verification, but there is no mention at all of similar work that needs to be undertaken for age assurance. An age assurance code of practice must be fast-tracked and backed up by binding standards that establish rules of the road. Otherwise we'll be waiting another 3 years for services to have age assurance in place, and when they finally do, they may not be proportionate, secure, effective or privacy-preserving.

Risk assessments

As per the draft Bill, regulated services likely to be accessed by children have a duty to carry out a children's risk assessment. These clauses have been expanded to require services to consider the risks to children with certain characteristics or belonging to certain groups (clauses 10(6d) and 15(5b)) This is encouraging, but could be strengthened with reference to protected characteristics set out in the Equality Act 2010.

When conducting risk assessments, service providers will need to consider how the design and operation of the service reduces or increases risk, including the business model and governance, but now they must also factor in "use of proactive technology" and "measures to promote media literacy and safe use of the service" (clauses 10(6h) and 25(5e) This cannot be used to give companies a get out of jail free card. Risk assessments must assess risk. They should not be used as a checklist of existing safety measures.

While the requirement for services to conduct risk assessments is central to the Bill, the government has stated it will not require Ofcom to mandate the way in which these

[2] Children and Parents: Media use and attitudes report, OFCOM, March 2022

assessments must be conducted or establish criteria against which services can assess risk. This will fail to give companies the certainty they need and put additional burdens on Ofcom to establish if a risk assessment is adequate.

The "size and capacity" of the service provider are still determining factors when assessing whether a measure is proportionate to meet the safety duties protecting children, in addition to levels of risk and potential harm (clause 11(9b)). This could create a significant loophole for smaller companies to exploit. Small does not mean safe. A service with a smaller number of users can still cause significant harm and a company with a small turnover or workforce can still reach a vast number of users. Focusing on size creates the opportunity for companies to restructure to avoid regulation rather than to prevent harm. As in many other sectors, designing safe products and services is simply the price of doing business for any trading company. Smaller providers need support to comply, not permission to harm.

Illegal content and new criminal offences

The requirements relating to illegal content have been strengthened so all regulated user-to-user services now have a duty to "prevent" individuals encountering illegal content, not just to minimise its presence. Illegal content is defined as content that amounts to a "relevant offence" (clause 52(2)), which include terrorism offences, offences related to child sexual exploitation and abuse (CSEA), and priority offences as listed in Schedule 7. These are assisted suicide, threats to kill, public order offences, harassment, stalking and fear of provocation or violence, drugs and psychoactive substances, firearms and other weapons, assisting illegal immigration, sexual exploitation, sexual images (including revenge porn), proceeds of crime, fraud, financial services and inchoate offences.

Three new communications-based criminal offences based on the Law Commission's recommendations[3] are set out in Part 10. These are a harmful communication offence to capture communications intended to cause harm without a reasonable excuse, a false communications offence to criminalise the sending of messages with the intention to cause non-trivial psychological or physical harm, and a threatening communications offence where a message conveys a threat of death or serious harm. A new 'cyberflashing' offence (clause 156) will also be included under the Sexual Offences Act 2003 as part of the new online safety regime. While the criminalising of these activities is welcome, the government have not committed to any corresponding investment in resources for the police, law enforcement or judiciary, calling into question whether such cases will ever reach the courts or result in convictions.

Definition of harm

In response to the Joint Committee's concern that the Bill focuses only on content and not on other types of harmful activity, the government stated that "all online activity is facilitated by content and, therefore, by imposing duties on services to address illegal and harmful content, the Bill will cover both activity and content."[4] Content is defined in the Bill as "anything communicated by means of an internet service, whether publicly or

[3] Modernising Communications Offences, A final report, Law Commission, July 2021

[4] Government response to the Joint Committee report on the Draft Online Safety Bill, March 2022, p21

privately, including written material or messages, oral communications, photographs, videos, visual images, music and data of any description" (clause 189). This definition does not reflect the full range of online activity that can create risks beyond user-to-user interaction.

Harm is defined as "physical or psychological harm" that is caused by the nature of the content, the fact of its dissemination and/or the manner of its dissemination, for example, content repeatedly sent to an individual by one person or by different people (clause 187). Harm may arise in circumstances where a user acts in a way that causes or increases the likelihood of harm to themselves or to others. This includes where individuals act in a way that "is related to that other individual's characteristics or membership of a group." This is encouraging and recognises that certain groups may be more at risk of harm or experience harm differently to others.

Content that is harmful to children is divided into three categories: "primary priority content", "priority content", to be defined in secondary legislation by the Secretary of State, and "non designated content" which presents a "material risk of significant harm to an appreciable number of children in the United Kingdom" (clause 52(4c)). This replaces the previous definition of harm as that which has a "significant adverse physical or psychological impact on a child of ordinary sensibilities". The change in language shifts the focus from impact onto reach, and in so doing appears to introduce a higher threshold for harm. The Bill and explanatory notes do not include a definition of "appreciable number" which will instead be determined by the Secretary of State. To capture risks to children beyond content (as discussed previously) this section should be rewritten entirely and harms to children defined according to the 4 Cs of online risk[5], and listed in schedule 7 alongside the named priority offences.

The government maintains in its response to the Joint Committee that setting out priority harms to children in secondary legislation is the best approach, to allow for adequate research and consultation with child safety groups. Harms to children are well evidenced, researched and documented. To claim more time is needed to allow for research and consultation in this area is to ignore the reams of research and evidence from academics, civil society organisations, youth advisory groups and Ofcom itself, in the three years since the government published the Online Harms white paper. Setting out harms on the face of the Bill will not only give certainty to regulated companies about the risks they will need to mitigate, but reassurance to those concerned that freedom of expression will be under threat if regulations are made by the Secretary of State that will not undergo the same level of parliamentary scrutiny.

Safety by Design

Despite its insistence that this is a "systems and processes" Bill, the government has brought forward a revised Bill that does not seem to address many of the functionalities and features that create risks to children, such as dark patterns or nudges, focusing mainly on harmful content. The definition of functionality (clause 186) is identical to the draft Bill – focused only on those features which facilitate interactions between users – and so do not cover risks created by features of a service which, for instance,

[5] Livingstone, S., & Stoilova, M. (2021). The 4Cs: Classifying Online Risk to Children. (CO:RE Short Report Series on Key Topics). Hamburg: Leibniz-Institut für Medienforschung

encourage extended engagement, lower privacy settings or place users under commercial pressure, such as:

- Notifications on by default
- Targeted advertising
- Lootboxes
- Endless scroll
- Autoplay
- Time-limited content
- Pay-to-play

Teenagers describe getting into "a TikTok trance" as a result of features such as infinite scroll: "once you start, you can't stop."[6] **- 5Rights research**

91% of children say there are loot boxes available in the games they play, 40% have paid to open one.[7]

It is however encouraging that the definition of functionality in relation to search services has been updated to include predictive search (autocomplete). The risks of this particular function are explored in our latest Risky by Design case study on recommendation systems.

Codes of Practice

In addition to codes for CSEA and terrorism, Ofcom will need to produce codes of practice to cover the 'relevant duties' for services, including for child online safety. These codes will set out 'recommended measures' that companies can take. If a service decides not to follow these steps, Ofcom must assess whether the alternative measures taken are sufficient to meet the duties and protect privacy and free speech. This is essentially a 'comply or explain' model, as seen in financial reporting, and will create much more work for Ofcom when assessing compliance.

While we are pleased that the Bill instructs Ofcom to produce codes of practice for CSEA and child online safety, but the codes will only be effective if they are mandated, and carry more weight than 'recommended measures' which companies can choose to implement. In response to the Joint Committee, the government said "there is no obvious precedent of codes being binding… any binding rules would have to be in legislation rather than codes." This is simply untrue. Statutory codes exist under the Children and Families Act 2014 to provide for children and young people with special educational needs.[8] Similarly, the government is proposing a statutory code under the new pro-competition regime to promote open choices, fair trading and trust and transparency.[9] The government need only look at the lack of compliance with the Video sharing platform regime enforced by Ofcom for evidence that non-statutory guidance does not deliver the aims of the legislation.

[6] Pathways: How digital design puts children at risk, 5Rights Foundation, July 2021, p52
[7] The Rip-Off Games: How the new business model of online gaming exploits children, ParentZone, 2019
[8] Children and Families Act 2014
[9] Government Proposals for a new pro-competition regime for digital markets, July 2021

Binding codes do not need to be inflexible or overly prescriptive. They would simply set out the agreed expectations for compliance from companies while remaining technology neutral and futureproof. Without them, tech companies will be left to decide for themselves what is best for their users and Ofcom will not have the authority they need to enforce the regime.

Complaints systems

All regulated service providers must operate a complaints procedure that is "easy to access, easy to use (including by children) and transparent" (clause 18(2c)). Complaints can be raised concerning harmful content, including illegal content and content harmful to children, a service not complying with a duty in the Bill, content being erroneously taken down or user suspension. Providers will need to state in their terms of service that users can bring a claim for breach of contract if their content is wrongly taken down or restricted (clause 19(4)).

The final route for redress is for an "eligible entity" to issue a 'super-complaint' to Ofcom that identifies more serious or systemic issues that are common to multiple features and companies (clause 140). Those eligible to issue a super-complaint will be defined by the Secretary of State. The government rejected the Joint Committee's recommendation to create an ombudsman, so under the new regime there will be no mechanism that will allow individuals to raise a complaint against a company without going through the courts. Individuals, and especially children, must be given ways to exercise their rights. It cannot be the preserve of the few who can afford to do so through the courts. The government's response to the Joint Committee offers a glimmer of hope – the Secretary of State will be able to reconsider whether independent resolution mechanisms are appropriate at the statutory review, suggesting they are receptive on this issue.

Criminal liability

Despite its claims that the regime will hold tech company bosses accountable and end the era of tech company impunity, the government have only introduced criminal liability for company directors who have failed to comply with information requests. Big tech companies have deep pockets and armies of lawyers. Only the threat of criminal sanctions for failures to protect their users will create the radical cultural shift so desperately needed across the sector, starting from the top.

Access to data for bereaved parents

The government has ignored the pleas of bereaved parents and the Joint Committee to grant parents access to data in cases where a child has died, simply stating that this is "outside the scope of the Bill."[10] This is a truly callous response to the plight of families looking for and routinely denied answers to the circumstances surrounding a loved one's death.

Molly Russell was 14 years old when she took her own life. Her father Ian has spoken about how the content she saw online in the months leading up to her death escalated and encouraged her depression. Frankie Thomas was 15 when she took her life after

[10] "As it stands, disclosure of data relating to a deceased person falls outside the scope of this Bill.", Government response to the Joint Committee report on the Draft Online Safety Bill, March 2022

months of viewing self-harm and suicide material online. Her family struggled to get answers from the providers of services she was using, and they are still denied access to what Frankie was seeing at the time.[11]

> **"What if someone malicious had been in contact with her, or any other child? That person (or those people) is securely hidden, knowing that they will never be exposed, since Instagram will not allow access to a dead child's account by the child's parents."**
>
> **- Judy Thomas, mother of Frankie Thomas**

The Bill must be amended to create a fair process for bereaved parents. Law enforcement and coroners must be granted access to children's data in a timely way to prevent further harm and to offer closure to bereaved parents.

***

The government's approach to the Bill has made it complex and content-focused, when what is required are standards of product safety that can be enforced against agreed criteria. It has also ignored existing research and evidence, and deferred many of the key aspects of the new regime to Ofcom to build out following what we can expect will be costly and lengthy periods of consultation. The sector is innovative and creative, and whilst it has shown reluctance to prioritise children's safety, the changes made in compliance with the UK Age Appropriate Design Code show that when required, they really can design with safety in mind. The Bill should take a safety by design approach and it should act with the speed that is necessary. Each day that they fail to do so, children are coming to very real harm.

We can expect to see some of these issues debated on the floor of the House on 19 April when the Bill receives its second reading. The remaining timetable for the passage of the Bill is unclear, but it's unlikely it will reach the Lords before the autumn or receive Royal Assent before Christmas. What can be guaranteed is some considerable lobbying efforts from tech companies and fierce debates on the battle ground of freedom of speech. The government must now hold its nerve in the face of tech lobbying and deliver the protections it has promised to children.

5Rights will be publishing its calls for changes to the Bill shortly.

For more information, please contact izzy@5rightsfoundation.com

[11] These tech giants led Frankie to kill herself. So why won't they talk to me?, The Times, 27 March 2022