Building the Digital World that Young People Deserve

Priorities for the Online Harms Bill

5Rights Foundation
October 2020

5RIGHTS FOUNDATION

Contents

Foreword

Taking action to tackle online harms has been a flagship promise of successive governments since 2017, but progress has been tempered by interruptions to the parliamentary timetable, the decision to leave the European Union, frequent changes of leadership in the DCMS, and most recently COVID-19.

COVID-19 has impacted in unimaginable ways on all aspects of life, including the digital lives of young people. Remote schooling, lock-down, social distancing rules, and the cancellation of activities outside the home have all conspired to increase the time young people spend online five-fold.[1] With this increase in time spent online has come increased harms,[2] a greater understanding of the risks young people face, and the urgent need to mitigate them.[3]

The need to address online harms faced by children was already critical. Now it is absolutely imperative.

The urgency felt by teachers, parents, young people, and civil society must be heeded by government. The upcoming Online Harms Bill is an opportunity to deliver on the promise of successive ministers to make the "UK the safest place in the world to be online."[4] This promise will only be fulfilled if the values that we live by in other parts of public and private life, also apply when our young people are online.

The proposals contained in this document require industry to recognise in practice – what every citizen already understands – that children are not adults. Young people need a digital world that responds to the vulnerabilities associated with their age, one that takes a proactive view of their safety, and one that takes a principled stance on their wellbeing. The measures called for here are proportionate, pragmatic and practical – and importantly, they are focused on age-inappropriate product and feature design not a censorious blacklist of proscribed content. They require a new world in which the sector seeks to identify the risks of their products and services and takes steps to mitigate those risks before they reach children, not after harms have been inflicted – a process that is accountable under the law.

The assumption of this document is all who play a part in the value chain of the digital world have a responsibility to young people; but it is simply a fact that a handful of companies occupy a predominate position in the lives of the young and as such must be a particular focus of regulators. We call on government to take a pragmatic and proportionate approach to the bill, not by excluding technologies and particular harms, but by ensuring all parties and all players identify and mitigate the risks they cause, whether spreading small risks at large scale, allowing poor design to accumulate harm and habituate dangerous practices, or by tolerating catastrophic outcomes for young people as a result of products and services that are not fit for purpose.

1 Covid-19: Lockdown measures and children's screen time, *House of Lords Library*, June 2020.
2 NCA predicts rise in online child sexual abuse during coronavirus pandemic, *Guardian*, April 2020.
3 5Rights Response to the Home Affairs Committee Inquiry into Home Office Preparedness for COVID-19, *5Rights*, May 2020.
4 Making Britain the safest place in the world to be online, *UK Government*, October 2017.

The Online Harms Bill offers a singular opportunity to prioritise young people's needs and rights in the design of the digital world, it is something that successive UK governments have promised. We owe it to our young people to act decisively, swiftly and comprehensively to build the digital world they deserve.

Baroness Beeban Kidron OBE
Chair 5Rights Foundation

About 5Rights Foundation

5Rights Foundation develops new policy, creates innovative projects and challenges received narratives to ensure governments, regulators, the tech sector and society understand, recognise and prioritise children's needs and rights in the digital world. In all of our work, we maintain and advocate that a child or a young person is anyone under the age of 18, in line with the UN Convention on the Rights of the Child.

Our work is pragmatic and implementable, allowing us to work with governments, intergovernmental institutions, professional associations, academics, and young people across the globe to build the digital world that young people deserve.

Executive Summary

- One billion young people are online, making up one in three internet users worldwide. Yet the digital world still largely fails to acknowledge their presence, preferring instead to treat 'all users equally'.
- The digital world is not optional for young people. It is essential, therefore that digital services and products treat young people according to their age and takes account of the needs of young people, from the design stage.
- In any area where the digital world affects young people, foreseeable risks must be identified and mitigated.
- The Online Harms Bill must capture any and *all* online services that create risks or facilitate or cause harms to children, including private messaging.
- Regulated services must be required to conduct regular Child Impact Assessments to reveal known harms, unintended consequences and emerging risks. These must feed into a continual drive to improve knowledge of the risks children face.
- Regulated services must also ensure their services meet minimum standards laid down by the regulator, provide age-appropriate default settings and account for the impacts of their algorithms.
- Regulated services must publish their terms in ways their users can understand and be accountable for upholding their own community guidelines, terms and conditions, and privacy notices.
- The harms within the scope of the Bill must include known harms, both legal and illegal. But the regulation must be flexible enough to take account of new and emerging harms as well.
- The Regulator must act in the best interests of children at all times and have regard to the UK's obligations under the United Nations Convention on the Rights of the Child.
- Many service providers will need to establish the age of their users in order to give young people the specific protection to which they are entitled. Those that don't will have to provide services appropriate for all users, including the very youngest.
- The Regulator must enforce the duty of care and be provided with sufficient resources to overcome the current asymmetry of arms.
- The Regulator must have an escalating ladder of enforcement powers, including significant corporate fines, and the power to require closure or change of individual features or services which pose a risk to children. Its powers must be on the face of the Bill and guidance and codes of practice must have a statutory footing.
- Government should develop open systems through which independent researchers can access privately held data to understand what works in keeping children safer online.
- The Online Harms Bill is an opportunity for the UK to show global leadership and secure its reputation for keeping children safe, both now and into the future.

Introduction

One billion young people are online, making up one in three internet users worldwide.[5] Yet the digital world still largely fails to acknowledge their presence, preferring instead to treat 'all users equally.' That is, to treat all users as adults.

Even where digital services and products do make design decisions to cater for young people's specific needs, these accommodations overwhelmingly come after the fact, bolted-on to pre-existing services rather than as a result of prior assessment of the possible impacts of products and services on the wellbeing and safety of young people. Often these changes will emerge as a hurried response to a discovered flaw, only coming to light once a young person has been harmed.

The digital world is not optional for young people, it is their gateway to education, information, entertainment, health services, and mediates their relationships and experiences. It is essential therefore that digital services and products treat young people according to their age. The digital technology they use is built by a series of conscious decisions that optimise its operation. Those decisions should take account of the needs of young people, from the design stage. But far too often they do not.

The technology sector has a responsibility to young people and must be held to account for the influence it exerts over their welfare and development. Independent regulators must develop minimum standards that reflect the needs and rights of young people, and these must be linked to mechanisms which ensure consistent implementation of those standards.

The introduction of strong, proportionate and thoughtful legislation would allow an independent regulator to hold companies to account, not for the behaviour of their users, but for the ways in which their own services and products facilitate or amplify risk, and for the efforts they make to identify and mitigate that risk. No environment is risk free, and no childhood is without upset or challenge, but since the digital world mediates all aspects of young people's lives, it must be a place that they can inhabit as children, not as underage adults.

The UK has an enviable reputation on the world stage for the steps it is taking to ensure a safe and secure online world and there is an international groundswell in the EU, the US and other jurisdictions towards increased regulatory intervention. The Online Harms Bill is an opportunity to show global leadership and secure that reputation both now and into the future.

It will not be adequate for the Online Harms Bill to offer a narrow interpretation of what is harmful or create a regulator without sufficient resource or authority. *Ultimately, the success of the Online Harms Bill will be judged not on what is excluded or included but on its impact on the lived experience of young people.*

5Rights Foundation speaks specifically on behalf of and is informed by the views of young people. Therefore, our comments reflect, and are restricted to, how the Online Harms Bill should protect users under the age of 18. However, we recognise that many of our approaches are relevant to other user groups and we welcome any efforts that

5 One in Three: Internet Governance and Children's Rights, *UNICEF*, 2016.

government makes to make the digital world more equitable for all user groups, particularly the vulnerable.

Opportunity vs Cost

The anxiety that innovation and the UK economy will suffer if it introduces regulation to make a safer and more equitable digital world is often cited as a barrier to regulation. This assessment is a natural reaction but does not adequately take account of the ever-increasing costs of failing to act.

The issues that young people face online do not stay online. They impact on our schools, social services, criminal justice system and NHS, as young people experience record levels of self-harm,[6] eating disorders,[7] anxiety,[8] identity theft,[9] bullying,[10] online sexual abuse[11] and exploitation. In contrast to the now-accepted "polluter pays" principle, the financial costs are borne by the whole of society and the emotional costs by young people.

The much smaller costs associated with identifying and mitigating risk in advance — simply a cost of doing business— would be welcomed by many.

General Principles & Definitions

Definition of children and young people
A child is a young person under the age of 18.[12] Particular services and products may have additional legal, statutory or voluntary age restrictions. But the UK's obligations under the Children Act (1989), as a signatory to the UN Convention on the Rights of the Child and as a member of the OECD require public bodies to consider and provide protections to all children.[13] This document also uses the term "young people" to refer to all persons under the age of 18.

Services, Platforms, Devices and Products
It is vital that in any area where the digital world affects young people, foreseeable risks are identified and mitigated. In this document, we use "services" to encompass services, platforms, devices and digital products (physical and virtual) that impact young people.

Duty of Care
The Government has signalled that the Online Harms Bill will introduce a duty of care, and our model implementation for that duty is found in Annex A. However, if such a duty is withdrawn or narrowed to the extent that *any* of the subsidiary duties described below is no longer in scope of the proposed duty of care, then the specific duty or duties should be present in some other part of the Bill. The duties and requirements set out below are both interconnected and cumulative, and without *all* of them the Online Harms Bill will fail to protect young people.

6 Major rise in non-suicidal self-harm in England, study shows, Susan Mayor, *BMJ Open,* 2019.
7 Incidence of anorexia nervosa in young people in the UK and Ireland: a national surveillance study, Petkova H, Simic M, Nicholls D, *et al, BMJ Open,* 2019.
8 Anxiety on rise among the young in social media age, *Guardian*, February 2019.
9 Identity theft isn't just an adult problem. Kids are victims, too, *CNBC,* April 2018.
10 Screen time is up – and so is cyberbullying, *National Geographic,* October 2020.
11 Child sex abuse: Record number of images dealt with, charity says, *BBC*, January 2020.
12 Article 1, UN Convention on the Rights of the Child, *OHCHR*, November 1989.
13 *Ibid,* Article 5.

The Four Cs

The focus of policymakers, journalists, and activists is often on content risks alone, to the exclusion of other prevalent risks. There are four broad categories of risk, and each should be considered by the Government as being of equal importance as it formulates the Online Harms Bill.

- *Content* – A young person is exposed to harmful material (e.g. age-inappropriate content, pornography, extreme and real-life violence, discriminatory or hateful content, disinformation, content that endorses risky or unhealthy behaviours such as anorexia, self-harm, suicide).
- *Contact* – A young person participates in activity with a malign actor, often, but not always, an adult (e.g. sexual exploitation, grooming, harassment, stalking, blackmail, unwanted sexual advances, location sharing).
- *Conduct* – A young person is involved in an exchange, often, but not always, peer-to-peer, as either a perpetrator or victim, sometimes both (e.g. bullying, sexting, revenge porn, trolling, threats and intimidation, peer pressure, loss of control of digital legacy/footprint).
- *Contract* – (also referred to as commercial risks) A young person is exposed to inappropriate commercial contractual relationships or pressures (e.g. compulsive use, gambling, targeted advertising, hidden costs, unfair terms and conditions, loss of control of personal data).

Children's Rights

Young people have existing rights under the United Nations Convention on the Rights of the Child (UNCRC),[14] that are interconnected and interdependent. The UK has undertaken to uphold and observe these rights[15] and in all matters that affect them a child's 'best interests'[16] should be the primary consideration. This means that where children's rights come into conflict with other rights, for example those of adults or corporations, a child's rights must be considered of primary importance. The Committee on the Rights of the Child (CRC) is currently formulating a General Comment on the Digital World[17] that will codify the convention for the digital world (to be adopted in 2021). The Online Harms Bill must offer no lesser standard than that provided by the UNCRC.

Future Proofing

In order to be relevant as the digital world innovates, the Bill must take a principles-based and systemic approach to risk. Whilst it will inevitably reference specific known harms, the Bill and the Regulator must be flexible enough to take account of new and emerging risks. A requirement for regulated services to conduct regular Child Impact Assessments to reveal known harms, unintended consequences and emerging risks must be put at the heart of the Bill. Crucially, the results of these Child Impact Assessments, research conducted by the Regulator and real-life experience from platforms and users must feed into a continual drive to improve our knowledge of the risks children face.

However, the need to future-proof must not be a recipe for regulatory paralysis. While individual services or devices might come or go, the nature of the risks they pose to

14 *Ibid.*
15 *Ibid.* Article 2.
16 See 'Best interests of the child,' *Age Appropriate Design Code*, ICO, January 2020.
17 The Draft General Comment on children's rights in relation to the digital environment can be found here. It is open for consultation until 15 November 2020.

young people and therefore the principles which should govern their use are no longer a mystery. We have learned a lot since the original laissez-faire rules of the internet were formulated in the mid to late 1990s. It is unacceptable to limit our ability to act on the clear and present dangers against the possibility that something completely unforeseeable might happen at some indeterminate point in an infinite future.

Corporate culture

The huge growth of the tech sector in the last twenty years is welcome, as are the conveniences provided by digital technology. But this cannot be at the expense of harming or putting young people at risk. The balance and proportionality that the Government seeks must shift the culture in the sector to take responsibility for the impact of their services and offer the support and safety that we afford young people in other areas of their lives *and* provide for the vulnerabilities associated with childhood.

Education

5Rights strongly supports the call for more digital literacy, but it will never be sufficient to rely on "online safety" education to keep young people safe. It is inappropriate to try to educate young people to use a world which systematically asks them to act beyond their maturity and puts them at risk, and it is dangerous to make them responsible for aspects of design over which they have no control.

Parents

Services and products which encourage greater parental engagement with young people's online lives should be welcomed, but this does not relieve the online service of responsibility. The presence of parents who have the time or the understanding and skills to mitigate risk cannot be assumed. Some children do not have parents, many children do not have parents with sufficient skills, and some services are designed in ways that even the most engaged and knowledgeable parent cannot change. Services must be responsible for aspects of their design which impact on the wellbeing and safety of young people.

Young people's views

Young people are enthusiastic users and early adopters of digital services and products. They want greater regulation and protections, but not in a manner that cuts them out from the digital world. Young people remain articulate and creative critics of the services they use and consistently ask for protections to be designed into the system as a norm.[18] It is important that in the course of creating regulation and in designing online products and services that *meaningful* consultation with young people is undertaken and their voices are reflected in the Bill.

18 5Rights Foundation will shortly be publishing a consultation with 700 children from 26 countries, which captures their views of the digital world.

Services in scope of regulation

It is vital that the scope of the regulation captures any and *all* online services[19] that may create risks or facilitate or cause harms we identify in this paper. The White Paper[20] proposed that the scope be limited to services "that allow, enable or facilitate users to share or discover user-generated content (UGC), or interact with each other online." This is far too narrow, as it would lead to significant vectors of harm being excluded from the regulation. It would be a catastrophic failure if the Online Harms Bill, with the stated objective of preventing online harm, knowingly or not, failed to mitigate risk or respond to harms that routinely affect young people by defining the scope in a way that excludes key areas. While all online services must be in scope of the Bill, it is important to acknowledge the gatekeeping and monopoly power[21] of some services that set the culture and standards of design. These companies may have additional responsibilities and regulators must pay specific attention to their impact on both the user and the ecosystem. Some examples that fall outside of the too-narrow scope suggested in the white paper are below.

- *Online gaming*. 59% of 5-15-year-olds in the UK now play online games (71% of boys and 48% of girls).[22] The risks young people may be exposed to online gaming are not limited to those associated with chat functions[23] or video-sharing features, but extend to inappropriate financial/commercial pressure, bullying or hate speech,[24] and exposure to unsuitable content.

- *Search services.* These form the backbone of most users' day to day digital interaction and provide easy access to content and services which pose high risks to young people such as child sexual abuse material (CSAM),[25] health misinformation[26] and conspiracy theories, as well as consistent failure to identify paid for content.[27]

- *Online Pornography*. A focus on services with user-generated content only would allow professional pornography sites to escape the scope. This is discussed further below.

- *Business-to-business services*. The Government's initial consultation response states "Business-to-business services have very limited opportunities to prevent harm occurring to individuals and as such will be out of scope of regulation." Yet, evidence shows that business-to-business (B2B) messaging and file-transfer services are a key method of distributing[28] child sexual abuse material. To exclude B2B services from the outset would create a significant (and bizarre)

19 As noted above, this includes services, platforms, devices and digital products (physical and virtual) that impact young people.
20 Transparency, trust and accountability (3.22), *Online Harms White Paper,* February 2020.
21 Investigation of Competition in Digital Markets, Majority Staff Report and Recommendations, *Subcommittee on Antitrust, Commercial and Administrative Law of the Committee on the Judiciary,* September 2020.
22 Children and Parents: media use and attitudes report 2019, *Ofcom*, 2020.
23 Roblox: 'I thought he was playing an innocent game,' *BBC News*, May 2019.
24 Free to Play? Hate, Harassment, and Positive Social Experiences in Online Gaming, *Anti-Defamation League,* July 2019.
25 Child sexual abuse material is available "within three clicks on open web", *National Crime Agency*, February 2020.
26 Online Information of Vaccines: Information Quality, Not Only Privacy, Is an Ethical Responsibility of Search Engines, Ghezzi, P., Bannister, P., Casino, G., Catalani, A., Goldman, M., Morley, J., Neunez, M., Prados-Bo, A., Smeesters, P., Taddeo, M., Vanzolini, T. and Floridi, L., *Frontiers in Medicine*, 2020.
27 Ads Tied to Web Searches Criticized as Deceptive, *The Wall Street Journal,* October 2014.
28 The NetClean Report 2016, *Netclean*, 2016.

hole in the regulation, undermining its protection of young people from sexual exploitation and abuse.

- *Online shops.* The White Paper refers to the "sale of illegal goods" and to the "illegal sale of weapons to young people online." Given that online shops are particularly likely to be involved in the occurrence of these (and related) harms, they must be within scope of the regulation. This is further discussed under "sale of age-restricted items" below.

The Online Harms Bill must set its sights on reducing the risk of harm. *Restricting scope by the nature of the service rather than the nature of the risk is antithetical to its purpose*. The duty of care is a broad principle which should apply to all parts of the internet value chain which may pose a risk of harm to young people.

Scope of the duty

- For the purposes of the Act, all companies providing online services to users in the UK must be 'regulated services.'
- Regulated services must be designed and operated in a way which ensures that, as far as reasonably practicable, young people are not put at risk of harm from their operation or use.
- A service's obligations under the duty of care must be appropriate and proportionate to the nature and risks associated with the service it provides. For the vast majority of online services, the regulatory burden should be minimal or immaterial (for example those dealing with sectors which rarely impact young people, such as agriculture or construction), while for services that pose a greater risk, the burden would be greater.

Duties on regulated services

We anticipate that the Bill will provide for several duties placed specifically on regulated services. The primary duty to design and operate a service in a way which ensures that, as far as reasonably practicable, young people are not put at risk of harm from their operation or use must be supported by a number of specific duties.

Duty to conduct Child Impact Assessments

The underlying issue for young people's experience of the digital world is that the majority of the products and services they use are not designed with their specific needs and vulnerabilities in mind. As a result, they are routinely exposed to demands that are beyond their maturity and risks which could and should have been mitigated at the design stage. A duty to conduct a child impact assessment before launching a service or new feature (or making significant changes to an existing service) would ensure that young people's best interests are a primary consideration in the design and delivery of digital services that engage with them or which impact on them.

Impact assessments are a common and established means of identifying the future consequences of a current or proposed action.

- Under the Town and Country Planning Regulations 2017,[29] developers are required to conduct an Environmental Impact Assessment before submitting a planning application to the local authority.
- Under the Data Protection Act[30] and the statutory Age Appropriate Design Code,[31] a Data Protection Impact Assessment (DPIA) must be carried out to assess and mitigate any risks to the rights and freedoms of young people arising from data processing. This must be done before processing any high-risk data and if the DPIA identifies any high risks which cannot be mitigated, then the Information Commissioner must be consulted before any data is processed.

> **Risk accumulates through a series of design choices. While each design choice may seem benign, in combination they accumulate risk to unacceptable levels.**
>
> **An illustration of how risks to children accumulate on common platforms can be found in 5Rights project Risky by Design: https://www.riskyby.design**

Routine Child Impact Assessments would give the provider of a regulated service the necessary information on the likely risks to young people and how to mitigate them. Services that do not engage with young people or whose services do not pose any risk will not have to take any further action.

Providers of regulated services must:

- conduct a Child Impact Assessment in relation to their services (and individual features) in order to identify the risks that their operation or use may pose to young people;

29 The Town and Country Planning (Environmental Impact Assessment) Regulations 2017.
30 Data Protection Act 2018.
31 'Data protection impact assessments.' *Age Appropriate Design Code*, ICO, January 2020.

- take measures to minimise or eradicate any risks identified in the Child Impact Assessment;
- keep a record of the Child Impact Assessment and any action taken as a result;
- regularly review the Child Impact Assessment and the effectiveness of any risk-mitigation measures taken;
- make available to the regulator any data used to inform the Child Impact Assessment and the action taken as a result, and collect and provide data on the effectiveness of the action taken;
- inform the regulator of any emerging concerns, including those that may have industry-wide relevance; and
- have regard to any guidance on Child Impact Assessments produced by the Regulator.

Duty to ensure services meet minimum standards

Regulated services must ensure that their service meets the terms of minimum standards published by the Regulator. These minimum standards should be systemic in nature and be useful to innovators and commercial companies covering such issues as: impact assessments, presentation of published terms, age assurance, moderation, and other such matters that the regulator deems necessary.

Duty to publish and uphold published terms

A service's published terms (normally consisting of at least community rules, privacy notice, cookie policy, and Terms and Conditions and often many more) together set out the agreement between the service and the user. They should allow young people and their parents to understand the nature and practices of an online service and anticipate the risks it might pose.

Currently they are incomprehensible, rarely read and poorly upheld.[32] They are presented at times and in ways that encourage agreement without engagement and take an implausible amount of time to read through.[33] Crucially, young people and parents have no guarantee that the terms set out meet any agreed standard of protection or behaviour, or if what they do offer in the way of 'rules' will be routinely applied.

5Rights research shows that 82% of British parents agreed that internet companies should be held accountable in law for how well they uphold their own community guidelines, terms and conditions, and privacy notices.[34] The UK's recently introduced Age Appropriate Design Code includes a requirement for service providers to uphold their own terms, explaining that young people "should be able to expect the service to operate in the way that you say it will, and for you to do what you say you are going to do."[35]

To be fit for purpose, published terms must meet minimum standards,[36] be understood by the user, and be routinely upheld. The presentation of information suitable for young people of different ages should be as routine, varied and ubiquitous as accessibility information (for instance, by taking a layered approach to providing information, and/or presenting separate, child-friendly information).

32 Most Online 'Terms of Service' Are Incomprehensible to Adults, *VICE News*, February 2019; We Read 150 Privacy Policies. They Were an Incomprehensible Disaster, *New York Times*, June 2019.
33 Social Site Terms Tougher Than Dickens, *BBC News*, July 2018.
34 5Rights YouGov poll: Parents' views on internet and child data protection regulation, 2019.
35 Policies and community standards, *Age Appropriate Design Code*, ICO, January 2020.
36 Human Rights due diligence, *UN Guiding Principles on Business and Human Rights*, 2011.

Providers of regulated services must:

- publish community standards and other published terms that meet the minimum standards[37] set out by the Government and Regulator;
- ensure that published terms are presented in ways that are truthful, easily understood and accessible to young people accessing the service, at the time or times when they are most likely to engage;[38] and
- put in place clear processes to ensure that their service's community standards and other published terms are upheld, including what action will be taken if they are violated.

Duty to provide age-appropriate default settings

Many digital products and services offer the lowest privacy settings by default, which puts young people at unnecessary risk, for example by automatically making their profile public,[39] or enabling adults they do not know to send them private messages.[40]

Online users overwhelmingly stick with the default privacy and safety settings.[41] Low-privacy default settings present an additional risk when using livestreaming and video sharing services that make young people's profiles available to a large number of people they don't know. On some platforms, this potential audience for a young person's content can be measured in many tens of millions of users.[42] Children as young as 10 routinely use popular livestreaming sites and report being contacted by unknown adults via the chat function. Often they are alone in their bedrooms,[43] allowing groomers to gain an insight into a child's hobbies and interests, who then use this information to establish rapport and build a relationship.[44] The NSPCC reported that 6% of children who livestream had received requests to change or remove their clothes.[45] The current combination of low privacy by default and ineffective age-gates and multiple risky design features that are integral to many of the platforms popular among young people.[46]

Young people have called on companies to make accounts private by default and describe confusion and frustration when trying to navigate privacy settings, particularly when trying to keep up with app updates.[47] Automatically setting defaults to high-privacy and safety offers a systemic way of reducing unnecessary risk and support to young people. When young people try to lower their privacy, warnings should pop up to explain likely risks.

Settings that support privacy and safely, for example those that restrict users' access to sensitive content, or give users additional, pop-up safety information should *always be on by default* for young people. Mandatory risk assessments and minimum

37 That is, the standards set out by the legislation or Regulator in relation to the harms in scope of the legislation.
38 The Institute of Electric and Electronics Engineers (IEEE) will create a standard for Age Appropriate Published terms, having recognised the gap between service offerings and children's unique needs and rights to be able to understand and engage with published terms. 5Rights' work with the standard pushes for digital service providers to better measure their own published terms to determine whether or not they are age appropriate.
39 For example, Instagram sets users profile as public by default when registering.
40 Although it appears in the form of a "message request" when users receive a private message from someone they do not follow, strangers are still allowed to send private messages to children by default.
41 Previous research on users' Microsoft Word settings found that less than 5% of surveyed users had changed any settings at all.
42 For example, as of August 2020, TikTok had over 100 million active monthly users in Europe.
43 Children stream on Twitch where potential predators find them. *WIRED,* 30 July 2020.
44 Understanding grooming discourse in computer-mediated environments, *Nuria Lorenzo-Dus, Cristina Izura, Rocío Pérez-Tattam*, Discourse, Context & Media, Volume 12, 2016, pp 40-50, June 2016.
45 Livestreaming and Video-Chatting, *NSPCC,* Snapshot 2.
46 Explored further in Risky by Design, *5Rights Foundation,* July 2020.
47 Children's data and privacy online: Growing up in a digital age. Research findings, *Stoilova, M., Livingstone, S. and Nandagiri,* R, London: London School, 2019.

standards of privacy for under 18s would transform the unnecessary risks faced by young people.

Regulated services must:

- ensure that young people are given the highest level of privacy and safety, by default;
- ensure high privacy settings are easy to maintain;
- ensure that young people are not incentivised to reduce their privacy or safety on the service in order to access unrelated features (e.g. through bundled consents, or nudges that encourage lower privacy);
- offer warnings and advice that promote privacy when young people attempt to lower settings;
- have regard to any guidance on default settings produced by the Regulator; and
- detail and explain the decisions they have taken in relation to the default settings provided to young people, including evidence on the impact of different default settings, in their Child Impact Assessments.

Some features are never appropriate for young people, for example, enabling direct messaging from stranger adults[48] or publicly broadcasting to 'everyone'. In these cases, it is not enough to disable them by default, instead, they should not be made available to young people at all.

Duty to account for algorithms

It is not possible to determine the true nature and impact of a digital service without understanding its algorithms (and associated automated system technologies such as Machine Learning), and how they amplify, encourage, moderate and/or discourage particular behaviours and outcomes.

The White Paper proposed that companies be required "to demonstrate how algorithms select content for children, and to provide the means for testing the operation of these algorithms." Assessing the impact of algorithms must not be limited to systems that "select content for children," but be reframed to cover 'algorithms that may impact on young people.' Recommendation algorithms are embedded within most platforms young people use. 70% of views on YouTube[49] are a direct result of its recommendation algorithm, and 80% of content hours watched on Netflix[50] come from its recommendation algorithm. These algorithms are characterised as 'personalisation' for the user based on what they (and "similar" users) have watched, shared or interacted with previously, however it is increasingly clear that recommender algorithms amplify more extreme content (such as disinformation) and are responsible for spreading harmful content. This can have devastating impacts on young people, leading to videos containing suicide imagery being *actively recommended* or 'personalised' and delivered to children on platforms such as TikTok (and many others).[51]

Young people spend much of their time online on sites where algorithms pose risks. The use of algorithms is not exclusive to recommending content, but can also be used for example, to recommend children as 'friends' or 'followers' to adults (exposing young people to contact risks) or to recommend videos of partially clothed

48 For example, TikTok banned under-16s with registered accounts from private messaging in April 2020.
49 YouTube's AI is the puppet master over most of what you watch, *CNET*, January 2018.
50 How Netflix uses Big Data to Drive Success, *Inside Big Data,* January 2018.
51 How Social Media Companies are Fighting to Remove Graphic Content after TikTok's Viral Suicide Video, *Independent,* September 2020.

prepubescent children to adults who have viewed similar content previously.[52] In the case of the latter the child is not an 'active participant' but still needs protection from a system that can recommend them or their content to a perpetrator.

Algorithms are integral to the experience of the user but the basis on which they 'optimise' the user experience is opaque to anyone outside the company. Without algorithmic oversight it is increasingly impossible to ascertain the nature, presence or responsibility for harms experienced by young people.

Regulated services must:

- assist the regulator in understanding the purpose and policies of the algorithm by identifying and assessing the data used to train the algorithm (and how it was collected), analysing the source code and/or statistical model in use, assessing the impact of the algorithm, and conducting its own tests on how the algorithm operates in practice and over time;[53]
- assist the regulator by making available those who design and supervise algorithms and automated systems in order to question the potential for risk, and or the potential for risk mitigation with regard to children's safety and other published guidance;
- have regard to any guidance on default settings produced by the Regulator;
- detail in their Child Impact Assessments the risks their algorithms pose to young people and what risk-mitigating action they have taken; and
- take immediate action when algorithms put children at risk.

Duty to establish age of users

Companies may choose to ensure that their service is suitable for all users by meeting all the duties required, in which case establishing the age (or age range) of their users may not be necessary. However, many service providers will need to establish the age of their users in order to give young people the specific protection to which they are entitled.

A 'risk-based' approach to age assurance – similar to that introduced by the Age Appropriate Design Code – is required to strike a balance between ensuring that young people can be protected from risk and ensuring that services are not unduly burdened if they pose little or no risk to young people. The level of certainty that a service must obtain about the age of their users will depend on a range of factors, including: whether the service engages with young people, in what numbers, and the age-range of those young people; the nature of the service including both the content it hosts and the range of activities it enables (including the presence of specific features that are known to be risky for young people); as well as the steps it has taken to mitigate risks to its users, including young people.[54]

Where the service serves content or activities which are clearly illegal to provide to children, then the most robust certainty about the age of their users will be necessary.

Services must use age assurance that is privacy-enhancing and easy for young people to engage with. In all circumstances, the information gathered to assess the age of a child must not be used, stored or shared for any other purpose.

52 On YouTube's digital playground, an open gate for pedophiles, *New York Times,* June 2019.
53 Algorithm Inspection and Regulatory Access, *Institute for Strategic Dialogue,* April 2020.
54 5Rights will be publishing a paper on methods of Age Assurance in late autumn 2020.

If a regulated service has actual or constructive knowledge[55] of the age of a child user and fails to provide them with the requisite level of protection, it must be deemed to have failed in its duty of care.

Regulated services must:

- establish the age of their users with a level of certainty appropriate to the risks to young people inherent in their service;
- have regard to any guidance on age assurance produced by the Regulator; and
- account for the efforts they have made to establish the age of their users, and the effectiveness of those efforts, in their Child impact Assessments.

Duty to train staff

The Online Harms Bill must ensure that 'not knowing' or 'failing to consider' young people's rights and welfare in the design and distribution of products and services, is no longer acceptable.

New norms will be established by the Online Harms Bill and companies should be required to provide training to ensure they are understood by staff. There is a long-established precedent in the Health and Safety at Work Act 1974[56] (on which the online duty of care is partly modelled), which requires employers to provide free of charge "information, instruction, training and supervision as is necessary to ensure, so far as is reasonably practicable, the health and safety at work of his employees." This principle should be used to realise young people's rights and provide for their safety across the sector.

Regulated services must:

- provide effective training to all staff in the design and governance chain (including developers, engineers, UX designers, product managers, and others) on young people's rights, vulnerabilities at different stages of development and the range of risks and harms they may experience online as a result;
- ensure that such training is not restricted to known harms but engenders a broader understanding of how young people use technology and how technology impacts on their rights and wellbeing;
- have regard to any guidance on training produced by the Regulator; and
- account for the efforts they have made to train their staff, in their Child impact Assessments.

Duty to co-operate with the Regulator

It is essential that regulated services co-operate with the Regulator, in order to ensure that it can carry out its functions effectively. This duty of co-operation is well established in other regulated sectors, for example in Section 63 of the Data Protection Act 2018.

In order to combat emerging harms, there must also be a duty on regulated services to report any new or unanticipated harm arising on its platform. This duty again has precedents, for example where data processors are required to report any serious personal data breaches to the ICO[57] or where providers of financial services are

55 Constructive knowledge is defined as the "knowledge that one using reasonable care or diligence should have, and therefore that is attributed by law". It is a stronger standard, and more in keeping with a 'duty of care', than 'actual knowledge', which only includes information of which a person is demonstrably and consciously aware.
56 Health and Safety Work Act 1974.
57 Data Protection Act 2018, S108.

required to report any concerns about customer using their service to launder money to the money laundering concerns to the Financial Conduct Authority.[58]

Specifically, there must be an identified individual or individuals within the regulated service, responsible for complying with the duties placed upon it. This is based upon the analogous duty in Section 69 of the Data Protection Act, which requires a Data Controller to appoint a Data Protection Officer with "expert knowledge of data protection law and practice."[59]

Regulated services must:

- co-operate with the Regulator in the performance of the Regulator's tasks;
- respond with alacrity and transparency to the Regulator's information-gathering and investigatory powers;
- provide to the Regulator information on significant harms which arise through the operation or use of the platform, whether previously identified through a Child Impact Assessment or not;
- comply with any statutory guidance and codes of practice issued by the Regulator; and
- identify to the regulator a senior manager responsible for complying with the duty of care, who has been nominated as understanding the risks and harm young people experience online.

58 Money Laundering, Terrorist Financing and Transfer of Funds (Information on the Payer) Regulations 2017.
59 Data Protection Act 2018.

Types of harm

A risk is a harm that is yet to happen. A harm is a risk that has been realised. It is crucial that the Bill sets its sights on preventing harm. The Bill and the following regulations will inevitably identify specific harms from which young people must be protected. But to deliver the digital world that young people deserve, the Bill must take a systemic approach to anticipating and mitigating risks and negative outcomes for young people. It should be noted that while these harms focus on individuals, they often have wider impacts on whole communities and classes of young people, especially when amplified by platforms which are designed to encourage sharing of shocking and outrageous content.

Illegal content and activity affecting young people
Providers must be expected to prevent illegal harms which cause significant added harm to young people from occurring on, or being facilitated or amplified by, their services and to take swift action if it does.

These illegal harms must be clearly in scope and the regulator must be given all the necessary powers to ensure compliance with the law. Where there is ambiguity over existing law - for example, self-harm material that acts as a gateway to or explicitly encourages suicide, or loot boxes that have all the features of gambling but are aimed at young people – these must be explicitly cited in the Bill to give clarity and purpose to the application of the law in the online context.

For young people, 'on and offline' is not a concept that they understand. Their lives move seamlessly between the two and laws and cultural expectations must apply in both.

- *Creation or sharing of child sexual abuse material.* In 2019 the National Center for Missing and Exploited Children (NCMEC) received 16.9 million reports of child sexual abuse or exploitation, containing 69 million photos and videos.[60] This includes sexual imagery generated by young people themselves. Self-generated imagery now accounts for a third of the webpages the Internet Watch Foundation have identified as hosting child sexual abuse material. More than three quarters of these images are of children aged 11-13.[61] The Bill must ensure that detection and take down rates are increased and must prevent services implementing technology that blocks these efforts.

- *Commission, arrangement, or facilitation of offences against young people under the Sexual Offences Act 2003* e.g. causing a child to watch a sexual act, causing a child to engage in sexual activity, engaging in sexual activity in the presence of a child, sexual communication with a child, and trafficking a child for the purposes of sexual exploitation.[62] Service designs that facilitate these harms should be in scope of the Bill.

- *Cyberstalking and online harassment under the Protection of Freedoms Act 2012.* Young people are especially vulnerable since they may not be able to recognise cyberstalking and harassment in the context of 'dating' for example, but often do not get police support. Nearly 48% of teenage survey respondents

60 Exploited children statistics, *NCMEC*, 2020.
61 The dark side of the selfie, *IWF*, January 2020.
62 Sexual Offences Act 2003.

from the Survey on Teen Relationships and Intimate Violence reported that they had experienced stalking behaviours, including a partner going through their online account.[63] There should be greater resources provided for awareness and enforcement.

- *Encouraging or assisting suicide under the Suicide Act 1961*. This includes content which is pro-anorexia, or encourages self-harm.[64] A 2019 study found that "those who viewed self-harm on Instagram during their lifetime tended to show more self-harm and suicidality-related outcomes," and that of those who had been exposed to self-harm content on Instagram, just 20% intentionally searched for this content.[65] Specifically online services that recommend such content to young people, rather than host it, should clearly be defined as having 'encouraged' suicidal behaviour in young people.

- *Content which glorifies terrorism under the Terrorism Act 2006.* Young people[66] particularly those known to be vulnerable, are a known target of terrorist radicalisers and recruiters. While this material is clearly illegal, great efforts should be made not to criminalise young people and resources made available to support them.

- *Hate crime.* Race-based hate crime against young people increased by 22% in the three years prior to 2019, although hate crime is believed to be significantly underreported. Half of 12-15s say they have seen something hateful about a particular group of people in the last year –up from a third in 2016– and less than half of young people who see hateful content online report it.[67] Recommendation systems that normalise and spread hate crime must be identified and defined as illegal under the Bill.

- *Selling of age-restricted items (e.g. alcohol, knives, prescription drugs, tobacco).* Tests by National Trading Standards officers in 2019 found that online retailers failed to prevent the sale of weapons to young people on over 40% of occasions.[68] Amazon has also been found to suggest knives to young people buying school rucksacks as part of their 'frequently bought together' feature.[69] Recommendation systems that normalise and encourage the carrying of knives must be identified and defined as illegal under the Bill.

- *Gambling (including loot boxes and other 'gambling-like' features and betting on esports tournaments).* The Gambling Act 2005 makes it an offence to 'invite, cause, or permit a child or young person to gamble.' However, paid-for 'random reward' features like loot boxes do not currently meet the definition of gambling set out in the Act. 91% of young people said there are loot boxes available in the games they play, with 40% having paid to open one. Of young people in the UK

63 Almost half of US teens who date experience stalking and harassment, *The Conversation,* August 2020.
64 Suicide Act 1961.
65 Effects of exposure to self-harm on social media: Evidence from a two-wave panel study among young adults, Arendt, Scherr, Romer, *New Media and Society,* May 2019.
66 School closures putting vulnerable pupils at risk of radicalisation by extremists online, warn police, *The Telegraph*, June 2020.
67 Children and Parents: media use and attitudes report 2019, *Ofcom*, 2020.
68 NTS National Statistics for Test Purchasing Knives to Under 18's, 2019.
69 Amazon's 'frequently bought together' feature suggests 14-year-old buys knife with his school rucksack, *The Telegraph,* September 2019.

who have played an online game, 76% say that online games 'try to make you spend as much money as possible'. [70] A Bristol University study shows that over a quarter of those engaging with esports betting tweets are children under the age of 16, suggesting esports gambling may be as attractive to children as the computer games themselves.[71] The Bill must define 'gambling-like' features routinely deployed online as being subject to existing law.

Legal content and activity that is harmful to young people

The following behaviours and activities are legal for adults but are known to pose risks to young people and are in violation of their existing rights. As such, services must be expected to be proactive in preventing young people's exposure to them.

When considering these legal harms, the emphasis on harm frequently settles on the extreme. This framing does not adequately understand young people's experience. In reality few young people suffer acute harm but these so-called lesser harms are cumulative, impact differently on different young people and offer gateways to serious harm.[72] The social and financial cost of these harms has yet to be adequately calculated, but it is clear that the burden on education, health, mental health, police, local (council) support services, as well as the individual, family and community costs are rising exponentially.

Many of these harms could be usefully excluded by introducing child risk assessments and as a result of adhering to minimum standards by online products and services that engage with young people.

Including these legal but harmful risks with the Online Harms Bill and requiring mandatory Child Impact Assessments will significantly reduce the pathways to extreme harm and reduce the need for expensive provision of support across all areas of young people services.

- *Promotion of eating disorders (if not covered above).* As of 2019, content encouraging eating disorders is the top cause of concern online for 10-16-year-olds. 29% of young people in that age-range said they have viewed content online that encouraged eating disorders, ranging from 22% of 12-year-olds to 44% of 15-year-olds.[73] The harms associated with eating disorders are life-threatening for young people, life-altering for their families, and the associated costs increasingly unsustainable for schools and health services. There are clear links between pro-eating disorder content and pathways to pro-suicide content.[74] The Bill must ensure that services no longer actively recommend such content.

- *Misinformation/disinformation, including public health harms.* Young people are more likely than adults to come across misinformation online[75] and they are more vulnerable to its impacts. Half of young people have found it hard to know what is true or false about Covid-19.[76] The Royal Society of Public Health found that 40% of parents in the UK with children under five have been exposed to

70 The Rip Off Games: how the new business model of online gaming exploits children, *ParentZone*, August 2019.
71 'Biddable Youth': Twitter sports and esports gambling adverts: action required to protect children, *University of Bristol*, August 2019.
72 Risky by Design, *5Rights Foundation,* July 2020.
73 Anorexia overtakes cyberbullying as top source of online concern among ten to 16-year-olds, *The Telegraph*, 2019.
74 Pro-Self-Harm and the Visibility of Youth-Generated Problematic Content, *Boyd, Danah, J. Ryan and Alex Leavitt,* January 2011.
75 Covid-19 news and information: summary of views about misinformation, *Ofcom*, July 2020.
76 Ofcom, May 2020.

anti-vax messages on social media.[77] Misinformation is often conceived as scandalous rather than harmful, but in the cases of young people failing to social distance[78] or parents failing to vaccinate children[79] it results in extreme illness and avoidable deaths.[80] Similarly, religious and racial misinformation amplified by online services[81] can have serious impacts on children's physical and mental health, is a cause of conflict in schools and communities and is used as recruiting tool by extremists, leading to recent concerns from enforcement officers about the recruiting of young people by, among others, the far right.[82] The Bill must ensure that services no longer actively recommend such content.

- *Consumer harms, including fraud and scams.* Young people are at high risk of online fraud, as they are less likely to be able to distinguish determine the bona fides of online discounts and offers. They are also disproportionately targeted to become 'money mules', where their bank accounts are used to conduct illegal activity.[83] Young people also suffer from other consumer harms, such as ticket touting (already illegal, but rarely enforced online[84]), or from failure to identify paid for content from influencers.[85] Existing regulatory frameworks[86] for consumer harm could form a useful starting point for minimum standards encapsulated in published terms that are then upheld.

- *Vilification of women, particularly public figures.* Women/girls in general face more abuse in the digital world. New research from Plan International in their largest ever global survey on online violence revealed that one in five girls (19%) have left or significantly reduced use of a social media platform after being harassed, while another one in ten (12%) have changed the way they express themselves.[87] Women in public life are more likely to be subjected to violent or sexualised online abuse,[88] leading to concerns about a dissuasive effect on girls wanting to pursue careers that may bring them into the public eye. For example, the 2018 Girls' Attitudes Survey by Girlguiding UK found that 34% of girls are put off politics by the way female politicians are represented in the media.[89] This is especially likely to impact girls from poorer or minority backgrounds.[90] This is a particularly insidious harm for girls, since it makes the digital world appear hostile on grounds of gender and race, pushing back years of effort to tackle discrimination more widely. Existing equality and human rights laws could form a useful starting point for minimum standards encapsulated in published terms that are then upheld.

77 Fear of side effects number one reason for choosing not to vaccinate, *Royal Society for Public Health*, January 2019.
78 The causes and consequences of COVID-19 misperceptions: understanding the role of news and social media, *Harvard Kennedy School Misinformation Review*, June 2020.
79 Measles cases at highest for 20 years in Europe, as anti-vaccine movement grows, *The Guardian*, 2018.
80 At least 60 people have developed complete blindness after drinking methanol as a believed cure of coronavirus. From COVID-19-Related Infodemic and Its Impact on Public Health: A Global Social Media Analysis, *The American Society of Tropical Medicine and Hygiene,* August 2020.
81 Full Fact, The Full Fact Report 2020: Fighting the causes and consequences of bad information, p95, April 2020.
82 Far right recruiting children on YouTube, The Times, October 2020
83 It could be you: stemming the tide of financial fraud in the UK, *Policy Network/Natwest*, November 2017.
84 FAQs and the Facts About Ticket Touting, *Fanfair Alliance.*
85 What Is Influencer Marketing and How Does It Target Children?, *Marijke De Veirman, Liselot Hudders, and Michelle R. Nelson*, December 2019.
86 FCA Mission: Approach to Customers, *Financial Conduct Authority.*
87 Abuse and harassment driving girls off Facebook, Instagram And Twitter, Plan International, October 2020
88 Intimidation in public life, *Independent Committee on Standards in Public Life*, December 2017.
89 Girls Attitudes Survey 2018, *Girlguiding UK*, 2018.
90 Sexism aimed at public figures is holding back poor and black young women, *Independent,* March 2019.

- *Online bullying* is an issue frequently raised by young people. It is closely associated with offline bullying and has similar long term, often very serious, impacts on physical and mental health. The Regulator must require robust and effective anti-bullying measures including minimum standards of moderation, speed of redress and spread of content, building on the outputs of the Royal Foundation Taskforce on the Prevention of Cyberbullying.[91] These standards should be encapsulated by the regulator as minimum standards for published terms that are then mandated to be upheld.

- *Online pornography* currently has no effective, or systemic controls[92] to prevent young people from accessing it routinely, or more urgently, it showing up in their world unwanted[93] and unsought. Youth and adolescent exposure to pornography has been found to lead to a diminished understanding or respect for sexual consent and a permissive and/or supportive attitudes towards sexual violence and rough sex.[94] One in five 11-12-year-olds wish to copy acts (including dangerous ones) they'd seen in pornography, rising to two in five 13-16-year-olds.[95] Over half of 11-16-year-olds in the UK have been exposed to online pornography, and the majority of young people's first-time seeing pornography, in many cases as young as seven or eight-years-old, is accidental.[96]

  The most recent legislative attempt to keep pornography from young people was in the Digital Economy Act 2017. This Act required providers to ensure that pornography was not routinely made available to young people, but its scope was too limited. In particular, it did not cover content on social media sites, now one of the primary gateways for young people to access explicit material.[97] Implementation of that Act was cancelled on the grounds that the protections would be encompassed within the Online Harms Bill.

  The Online Harms Bill must now ensure that:

  - pornography is not made routinely available to young people, whatever the nature of the online service or platform it is hosted on;
  - where a regulated service prohibits sharing of pornographic material in its published terms it has a duty to implement appropriate measures to ensure that this is the case; and
  - the regulator has the power to require the implementation of robust age-verification measures where a service has failed to do so and enforcement powers to secure compliance, irrespective of the country of origin of the publisher.

When mitigating harms that are legal for adults, regulated services must be clear about the nature of the service, its age appropriateness and the actions they take to mitigate behaviours that are not permitted. Where a service chooses to engage with

91 Action Plan for The Royal Foundation's Taskforce on the Prevention of Cyberbullying, *Royal Foundation*, 2017.
92 The powers granted to OFCOM under the Audio-Visual Media Services Directive are effective only on UK-based sites, which does not reflect the trans-national phenomenon of pornography use by young people.
93 On first viewing of pornography, young people reported feeling shock and confusion, where younger children were most likely to report feeling "disturbed" by what they had seen. From I wasn't sure it was normal to watch it. London: NSPCC , Martellozzo, E., Monaghan, A., Adler, J.R., Davidson, J., Leyva, R. and Horvath, M.A.H, 2016.
94 The effects of pornography on children and young people, Antonia Quadara, Alissar El-Murr and Joe Latham, *Australian Institute of Family Studies*, 2017.
95 Impact of online pornography on children, *Children's Commissioner for England,* June 2016.
96 *BBFC*, 2019.
97 Research by the British Board of Film Classification (BBFC) showed 46% of 16- and 17-year olds who wanted to view porn used social media platforms, compared with 44% who instead chose dedicated porn websites.

young people, they should either be offered an age appropriate service, and young people and/or their parents should be fully aware of the terms of their engagement. Under no circumstances should it be possible for companies to routinely host suicide sites, or recommend pornography, or fail to flag misinformation – without being clear at the outset that is their corporate decision to host this content. Even where they do host it, they should not be able to recommend it to young people. It must also be part of an informed consent process that is transparent to the child and/or the parents and which meets minimum standards of age appropriate presentation. Where the activity or the content is clearly not in the best interests of young people, companies should take action to restrict their service to only those of an appropriate age.

Duties and powers of the Regulator

The Bill must give the Regulator, independent of government and industry, sufficient resources and autonomy to successfully carry out the following powers and duties. The asymmetry of the resources and power of the tech sector vs traditional regulators is palpable, and the enforcement record of regulatory bodies is worryingly out of step with the evidence of wrongdoing. The Bill must put the regulator's powers and duties on the face of the Bill and give it the power to make statutory guidance. Where companies fail to comply, the regulator must have a sufficient ladder of escalating powers to ensure that it is able to enforce its decisions.

It should be noted that regulation provides a floor below which companies must not fall, but that companies should be free to provide a quality of service that exceeds minimum standards and creates commercial differentials in providing best practice or innovative protections for young people. The success of regulation, and therefore the Bill itself, must be measured by the ability of young people to engage with in the digital world safely and in line with their rights and development capacity.

Overarching duties

- Duty to enforce the requirements on companies set out in legislation and statutory guidance, and to investigate regulated services for breaches (covered further below).
- Duty to support and encourage compliance by industry through the provision of statutory guidance, prepared in consultation with Government, civil society, the public (including young people) and industry.
- Duty to identify and respond to new and emerging risks and to provide guidance as they emerge.
- Duty to prepare an annual report detailing trends in risks and harms to users, alongside the enforcement action taken.
- Duty to draw on the expertise other regulators and to co-operate with them (including the Information Commissioner, Competition and Markets Authority, and the Advertising Standards Authority) when exercising its powers.
- Duty to consult with interested and expert parties in fulfilling its regulatory functions, including NGOs, academics, schools, parents, and young people themselves.
- Duty to facilitate and promote independent research.
- Duty to maintain a register of identified senior managers responsible for compliance with the Act.

Duties to young people

- Duty to act in the 'best interests of children'[98] at all times, and to have regard to the UK's obligations under the United Nations Convention on the Rights of the Child.
- Duty to reflect government and official guidance on providing services to and working with and supporting young people.
- Duty to provide guidance on the preparation by online service providers of Child Impact Assessments.
- Duty to provide guidance on minimum standards for online services that engage with young people, including effective governance, design principles, fairness

98 'Best Interests of the Child', *Age Appropriate Design Code*, ICO, January 2020.

and presentation of terms, child-appropriate moderation and redress (including speed of resolution), age assurance, child impact assessments, and accessibility.
- Duty to continually improve the safety of young people online through carrying out and publishing research and best practice documents, including on issues such as age assurance standards and the presentation of age-appropriate information.
- Duty to encourage innovation in design of services which will enable and encourage active digital participation by young people.
- When undertaking enforcement, duty to act proportionately to the risks that services pose to young people and to recognise reasonable efforts taken to avoid and mitigate risk.

Information-gathering powers
Power to audit, gather evidence and require disclosure of information from providers detailing:

- adherence to their services' own published terms and community standards;
- internal risk assessments, including Child Impact Assessments, and risk mitigation measures;
- internal process to ensure the safety and wellbeing of young people who may be affected by the service;
- the impact and operation of algorithms, including access to the designers and operators of algorithmic and automated systems, the purpose and policies of the systems, the underlying data, and the source codes or statistical models in use (including the power to conduct tests on the operation of algorithms in order to assess their impact); and
- the operation of reporting/complaints mechanisms and data related to the quantity, subject, and outcome of user reports/complaints.

Enforcement powers
The regulator must have a duty to support and encourage compliance. This can be done through advice, guidance and warnings. But where a service does not comply, the regulator will have the power to order redress measures and penalties, including:

- mandatory or enhanced child impact assessments;
- mandatory changes to individual features or service design;
- temporary disabling/closure of individual features, processes or an entire service;
- significant corporate fines (up to 4% global turnover);
- fines on individual responsible directors, where they have not ensured that the service complies with the regulatory regime;
- publication of fair and accurate descriptions of the services they provide; and
- publishing of steps being taken to safeguard children's interests and rights.

These penalties must be imposed proportionately to:

- the risk to young people;
- harms experienced by young people;
- the scale and spread of risk throughout communities of young people;
- the efforts made by the service to avoid or mitigate the risk; and
- the co-operation shown by the service.

The court will have the power to enforce measures and penalties made by the Regulator. Where there are egregious breaches, or where a regulated service has continued to defy the Regulator, stronger sanctions many be needed. Given that these require the careful balancing of rights and duties, it is appropriate that these are reserved to the court.

The court will have the power to:

- order the permanent closure of individual features, processes or an entire service;
- order the withdrawal of payment and banking services from a regulated service;
- order ISPs to block access to the service from the UK (similar to the court's powers under section 97A of the Copyright Design and Patents Act 1988); and
- remove a person from the register of senior managers and issue personal fines.

Miscellaneous provisions

Private communications

These duties must apply to providers of private communications services, in addition to their existing duties under data protection legislation to promote the privacy of their users. There is no treaty which provides for privacy to be an absolute right and the European Convention on Human Rights specifically allows for it to be limited, in order to "prevent…crime or…[protect the] rights and freedoms of others."[99]

The White Paper states "Reflecting the importance of privacy, any requirements to scan or monitor content for tightly defined categories of illegal content will not apply to private channels." The Government's concern not to *unduly* encroach on the privacy of individuals engaging in private communication is understandable, but it cannot be at the expense of protecting young people, in particular from sexual exploitation and abuse, nor should it be framed as a necessary by-product of innovation or growth. Moreover, there is no evidence that it is a requirement of detection and protection requires wholesale invasion of adult privacy. If private communications remain out of scope, the Government would, in effect, be giving the green light to further roll-out of encrypted services, which enable child sexual abuse material to be shared widely, even on mainstream platforms.[100]

The Government should require companies to develop techniques and technologies that preserve user privacy and allow for young people to be protected (whether or not such techniques and technologies would or could be required by a regulator). The statement in the White Paper is unnecessarily restrictive, therefore, and may even have a chilling effect on the development these technologies.

The initial response to the consultation notes that "overall respondents opposed the inclusion of private communication services in scope of regulation." This opposition was in part the result of a specific campaign by privacy activists and technology companies and does not reflect the views of parents[101] or teachers. For the avoidance of doubt, it should be the stated position of government that companies must not disable or row-back on current methods of detection, *that do not challenge user privacy*, until such time as there are privacy preserving alternatives that are equivalent or better than existing detection methods. Many experts do not accept that privacy and protection is a binary[102] but whatever the framing of the discussion, it undermines the entire purpose of the Online Harms Bill to exempt technology that is known to allow and amplify epidemic proportions of child online sexual abuse.

Additionally, there are a range of requirements that could apply to private channels that do not provoke questions of surveillance. For example, private messaging services should introduce safeguards around adult/child interaction, including blocking unknown adult users from making initial contact with child users; introducing restrictions on sharing of young people contact details; and/ or limits on young people being added to group chats with adults they don't know. Effective reporting procedures and robust age-verification would also make private communications services safer. It

99 Convention for the Protection of Human Rights and Fundamental Freedoms, Article 8.
100 Briefing: end-to-end encryption and child sexual abuse material, *5Rights*, 2019.
101 Research by 5Rights in June 2019 revealed that 78% of parents think that regulation on the use of children's data should apply to *all* online services likely to be accessed by children, rather than just online services that targeted at children specifically.
102 For example photoDNA can be modified and implemented at the point a message/image is sent, contrary to the current approach where it is implemented 'server-side.'

is incumbent on all parties to ensure that privacy concerns are not 'privileged' over young people's right to be protected from violence and or harmful content, but rather seek a detailed set of minimum requirements to protect young people.

A regulatory requirement to prove that new systems are better than existing, should be introduced as a core principle of the Bill to support the innovation and development of new systems of detection among the big players.

Creating independent research access to data

A fundamental issue in the debate around risk and harm is the lack of access to data for the research community. To have to evidence specific harms when the information sits with the service provider hampers progress in finding systemic regulatory solutions to repeated problems. This creates an unacceptable imbalance between civil society and business.

Where data is not already in the public domain, the Bill should create a 'public interest' requirement for digital services to provide access to data, algorithms and internal procedures to independent external researchers. This will require the creation of an independent and accountable central point of access, which is flexible enough to allow researchers to be able to answer basic and applied questions about regulated services, their use and impacts. It must meet the highest standards of data protection and create a 'clearing house' model for ethical data linking, between online life and existing data assets.

This will require some careful work to ensure that the independence of researchers and the IP of companies must be protected in the clearing house process. The Government should now commission research to look carefully at how an independent public research power would best work for all parties as it does in other sectors.

Education

As part of our work, 5Rights Foundation runs deliberative workshops with young people, to develop understanding of the digital environment and how it can better meet their needs. In the workshops young people repeatedly point out that their current digital literacy provision is not adequate, relevant, or effective enough to equip them for the digital age. Young people's concerns are not limited just to their safety online. They have a broad set of needs such as understanding "how technologies… control how they make people behave,"[103] wanting to be "taught how to use the internet to its full capacities,"[104] and learning "what personal information is used by other people."[105] Digital literacy provision in school is too often narrowly focussed on harms, and has little focus on *data* literacy or the impact of growing up in a highly commercialised environment.

In the statutory guidance for Relationships Education, Relationships and Sex Education (RSE) and Health Education, digital literacy makes up just 1 core module out of 8 for both primary and secondary school under the umbrella of 'Physical health and mental wellbeing' and 1 core module out of 5 in the vein of 'Relationships education.'[106] This does not constitute meaningful provision and is out of kilter with the impact of technology on young people's lives.

103 From The 5Rights, by Young People, *5Rights Foundation*.
104 *Ibid.*
105 *Ibid.*
106 Relationships Education, Relationships and Sex Education (RSE) and Health Education, Department of Education, 2020.

Third-party platforms and the tech sector are increasingly stepping in to provide such programs for young people. These programs tend to responsibilise[107] young people and gloss-over risks associated with these same sectors. This means that young people are held responsible for their own well-being online, while being expected to navigate the norm of *commercial* risks in the digital. These points are well-evidenced in the report of the House of Lords Select Committee on Communications on Growing up with the Internet[108] as well as the recent Select Committee on Democracy and Digital Technologies report on Digital Technology and the Resurrection of Trust.[109]

The Online Harms Bill must create a pathway for meaningful data and digital literacy, including teacher training, resources, space in the timetable and where needed extra-curricular and emergency support for young people who experience harm. The Bill must also set out framework for what constitutes the teaching of data and digital literacy including setting out minimum standards and the broad themes that *any provider* private or public, must cover if they are to supply services in educational settings.

Collective Redress

In order to assist the Regulator in overseeing and enforcing compliance with the duty of care, a super-complaints regime should be established, allowing certain expert not-for-profit bodies to take representative action on behalf of users in the UK. Given the complexity of the digital world and the maturity of young people, it is not acceptable to require them to make individual complaints to seek redress for harm. They cannot be said to have meaningful access to justice when platforms routinely ignore complaints from users, even about content and conduct which violates their published terms and it would not be acceptable to expect that young people should go to court to enforce the duties outlined.

There are several models of collective redress available, but the Government should ensure that bodies with experience representing the interests of young people and with specific expertise on young people's rights online can be designated as super-complainants. If a young person cannot reasonably and without cost access the protections provided by the Bill, then their access to justice has been denied.

Limits of Liability

Nothing in this paper is in contravention of the intermediary liability regime[110] that some parts of the tech sector enjoy. By committing to identifying the risks of services design, mitigating those risks and a robust audit and enforcement regime the sector will fulfil its duty of care to young people in a way that protects their freedoms and rights. However, it is worth noting that in many jurisdictions including the US there is increasing disquiet about the nature and impact of technology sector on individuals, civic life, and society more broadly. It is our view that if the sector persistently fails to uphold these minimum requirements, that the government should not only look to fines, and in severe cases impose disruption of service, it should also reassess the liability protections the sector currently enjoys.

---

107 The Rights of the Child in the Digital Environment: From Empowerment to De-Responsibilisation, Professor Dr Eva Lievens, Freedom Security Privacy, The Future of Childhood in the Digital World, 2020.
108 Growing up with the internet, *Select Committee on Communications*, March 2017.
109 Digital Technology and the Resurrection of Trust, *Select Committee on Democracy and Digital Technologies*, June 2020.
110 In the EU and UK: Council Directive 2000/31/EC on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market ('Directive on electronic commerce'), Official Journal L 178. In the US Section 230 of the Communications Decency Act, 47 U.S.C. § 230

Annex A: Model implementation of Duty of Care for young people

Properly fulfilled, the duty of care means anticipating and mitigating risk, in advance and on an ongoing basis, during the design and development of a product or service. The following is an example of what good implementation of the duty of care might look like, in relation to addressing online grooming specifically.

The online service in question is a (hypothetical) social media platform that allows users of all ages (above 13) to interact with each other publicly and privately, including by sharing photos, videos, and livestreams.

The company undertakes a child impact assessment and identifies that users of the platform under the age of 18 are at risk of being contacted and groomed by adult users. In order to mitigate this risk, the company takes the following steps.

- Gives young people and young people default settings that provide them with the highest level of privacy and safety. In particular, young people's profiles and content should not be made visible by default to users they don't know.
- Limit who can comment on a child's live stream or uploaded videos to friends or followers, not 'everyone'.
- Restricts private messaging for young people to ensure that they cannot be contacted privately by people they don't know, and that adult users cannot initiate private contact with young people in any case.
- Gives all young people and young people the option of disabling both comments and private messaging.
- Implements moderation tools to detect grooming patterns and language in interactions with young people.
- Establishes the age of users, going beyond mere self-declaration of age, so that child-specific protections can meaningfully be applied.
- Signposts easy to use and robust reporting tools, encourages reporting and offers child-friendly, swift and decisive responses.
- Regularly assesses the effectiveness of these steps, how they can be improved and the merits of introducing additional steps.

For further information, please contact:
Tony Stower, Director of External Engagement
tony@5rightsfoundation.com