

## CARNEGIE UK TRUST RESPONSE TO UN COMMITTEE ON THE RIGHTS OF THE CHILD (UNCRC) CONSULTATION ON THE DRAFT GENERAL COMMENT ON CHILDREN'S RIGHTS IN RELATION TO THE DIGITAL ENVIRONMENT

November 2020

1. We welcome the opportunity to submit our views to this consultation. This note builds on the consultation response we submitted on the Concept Note for the General Comment and draws out the relevance of our work on a statutory duty of care for online harm reduction to the principles and objectives set out in the draft General Comment.
2. Since early 2018, Professor Lorna Woods and William Perrin have been working, under the aegis of Carnegie UK Trust, on a detailed regulatory proposal for online harm reduction. It draws on well-established legal concepts to set out a statutory duty of care backed by an independent regulator, with measuring, reporting and transparency obligations on the companies. An orientation towards the outcome (harm) makes this approach futureproof and necessarily systemic. We propose that, as in health and safety regulation companies should run their systems in a proportionate, risk-based manner to reduce reasonably foreseeable harm.
3. We have built up a wide, diverse network of organisations who are supportive of, and influenced by, our work (including children's campaigners, such as the NSPCC and 5 Rights Foundation). Our work has also influenced much of the recent thinking among Parliamentary committees, government advisers and civil society groups in the UK and has been referenced by international organisations – such as the UN Secretary-General High-Level Panel on Digital Cooperation<sup>1</sup> - and think tanks. We have also contributed to the deliberations of the High-Level Panel – in person, at their workshop in London in December 2018, as well as via their consultation process and subsequent working groups.<sup>2</sup>
4. The UK government's Online Harms White Paper proposals from April 2019 drew heavily on our work. As we await their final proposals this autumn, we continue to work to ensure that the forthcoming legislation is robust, proportionate and systemic. The UK now has an opportunity to set the benchmark in relation to regulating big platforms to make the online world safer for all users. Action to tackle online harms and protect the rights of children will be most effective if it is international and multilateral. We therefore continue to support the commitment of the UNCRC to lead action in this area and are delighted to submit an update on our work for consideration.
5. As we set out in our submission in 2019, our approach respects the European approach to free speech while also protecting people from harm caused by the operation of commercial social networks. We feel that the Convention strikes an important balance for children between free speech and being protected from harmful speech, a factor often overlooked in policy debate. The draft general comment enhances this well for the digital age.

1 [https://www.un.org/en/content/digital-cooperation-roadmap/assets/pdf/Roadmap\\_for\\_Digital\\_Cooperation\\_EN.pdf](https://www.un.org/en/content/digital-cooperation-roadmap/assets/pdf/Roadmap_for_Digital_Cooperation_EN.pdf)

2 Our submission can be viewed here: [https://d1ssu070pg2v9i.cloudfront.net/pex/carnegie\\_uk\\_trust/2019/03/01101237/UN-High-Level-Panel-Digital-Cooperation.pdf](https://d1ssu070pg2v9i.cloudfront.net/pex/carnegie_uk_trust/2019/03/01101237/UN-High-Level-Panel-Digital-Cooperation.pdf)

## Our view on the draft General Comment

6. We have continued to develop our thinking on the application of the duty of care to online harm reduction. Of particular relevance to the UNCRC's right-based work is the comprehensive paper by Professor Woods which sets out how our proposal is compatible with fundamental rights, in particular the right to freedom of expression.<sup>3</sup> We welcome the General Comment and the guidance it provides to states with regard to their obligations under the Convention in the light of the opportunities, risks and challenges for children's rights in the digital environment. The set of rights in the draft Comment cover the right areas to ensure that children can fully realise the benefits of the digital world, as well as being protected from its potential harms. The guidance to States under each key group of rights is clear and the particular considerations that are specific to protecting the rights of children are well explained.
7. As we set out in our previous submission, our proposals have focused on developing a regulatory framework that takes account of the right of all users, but particularly vulnerable groups such as children, to be protected from a range of harms as well as to protect their health and wellbeing. It fits neatly within a rights-based framework, particularly in delivering the "protection rights" envisaged in the general comment, such as those set out in section VII (violence against children) and in section VI, para 55 re the need to protect children from harmful material. We have very much supported the work led by Baroness Kidron here in the UK to develop and implement the Age Appropriate Design Code, which we see echoes of in para 56 with regard to the guidance that "age/content-based systems designed to protect children from age-inappropriate content should be consistent with the principle of data minimisation". Para 57, with its reference to the need for States to "ensure that digital providers comply with relevant guidelines, standards and codes, enforce their own community content rules and provide sufficient moderation to meet their published terms" is important in protecting children and vulnerable users from harm. But, as the ongoing debate in the UK about the handling of "legal but harmful" activity under the proposed Online Harms legislation demonstrates, leaving such judgements on the appropriate policies and terms and conditions solely to the platforms will not be effective without regulatory accountability and oversight of the enforcement of those policies.
8. While we acknowledge in our work that education and digital literacy is a vital supporting component to protect children's rights online more broadly, as we set out in our previous submission, our focus has been on the need for urgent legislative measures to ensure that social media and other online providers take action at a system level ("by design") to ensure that their services are designed in such a way to protect their users so far as possible from reasonably foreseeable harm, so allowing children to access the opportunities afforded by digital services concomitant with their individual competence and developmental stage. Our approach we hope will encourage of a range of services appropriate to different age-groups/stages of development. As a corollary, the policy debate we hope moves from being principally about parental control and responsibility to oversee the child's internet use. While parental oversight remains important, the child's right to develop his or her own preferences and interests distinct from those of the parents should also be protected. Children, and other vulnerable groups, should not be put in the position of having to defend themselves.

<sup>3</sup> <https://www.carnegieuktrust.org.uk/publications/doc-fundamental-freedoms/>

## In summary: a statutory duty of care for social media

9. We would like to take this opportunity to remind the UNCRC of the substance of our proposal, and its alignment with the aims of the Draft General Comment in its focus on systems, design and rights. Our proposal sets out how a statutory duty of care would require most companies that provide social media or online messaging services to protect people from reasonably foreseeable harms that might arise from use of those services. This approach is risk-based and outcomes-orientated. A regulator would ensure that companies delivered on their statutory duty of care and it would have sufficient powers to drive compliance.
10. Social media service providers should each be seen as responsible for a public space they have created, much as property owners or operators are in the physical world. Everything that happens on a social media service is a result of corporate decisions: about the terms of service, the software deployed and the resources put into enforcing the terms of service.
11. In the physical world, it is accepted in many jurisdictions that those owning or operating buildings and spaces are responsible for them; in the UK, Parliament has long imposed statutory duties of care upon property owners or occupiers in respect of people using their places, as well as on employers in respect of their employees. Variants of statutory duties of care also exist in other sectors where harm can occur to users or the public. A statutory duty of care is simple, broadly based and largely future proof. It identifies the objective – harm reduction – and leaves the detail of the means to those best placed to come up with solutions in context: the companies who are subject to the duty of care. A statutory duty of care returns the cost of harms to those responsible for them, an application of the micro-economically efficient ‘polluter pays’ principle. The E-Commerce Directive permits duties of care introduced by Member States.
12. Harms should be identified by the legislature. These should include harms to children arising criminal activity directed towards them (e.g. grooming); accessing age-inappropriate content as well as emotional harm arising eg from cyber-bullying (especially where there is a racial, religious, gendered element). This latter point is important in trying to ensure equal access to the benefits of online services.
13. The regime would cover reasonably foreseeable harm that occurs to people who are users of a service and reasonably foreseeable harm to people who are not users of a service. This is particularly important in relation to children’s rights, where risks to rights such as the protection of privacy and identity as well as to health and wellbeing (for example through cyber-bullying) might take place on platforms where the individual that is being harmed is not themselves a user. We have recently been involved in discussions with Parliamentarians in the UK on the rise of “capping” – where child abusers and groomers take screenshots during explicit video calls and live streams with minors, and then circulate them widely amongst abuser networks, or use them to blackmail or coerce children. Platforms have responsibility both where the content originated and also on the platforms or services where groups of sex offenders congregate.

14. Central to the duty of care is the idea of risk. If a service provider targets or is used by a vulnerable group of users (e.g. children), the risk of harm is greater and service provider should have more safeguard mechanisms in place than a service which is, for example, aimed at adults and has community rules agreed by the users themselves (not imposed as part of Terms of Service by the provider) to allow robust or even aggressive communications. In that instance, however, the platform operator should ensure that it has rigorous safeguards in place to ensure that its service is not used by children. We argue throughout our work that content takedown or indeed age verification is not the only tool in the box – indeed, on its own it is inadequate for the scale of the problem. Designing in appropriate tools can give children greater freedom to explore and experiment online. This approach might maximise the potential for a rights balance that is not a zero sum game. In that regard, we would refer to the 2019 report on hate speech<sup>4</sup> from the UN Special Rapporteur on Freedom of Expression where he set out the different ways, as well as content deletion, that information flows can be affected. For example:

*“restricting its virality, labelling its origin, suspending the relevant user, suspending the organization sponsoring the content, developing ratings to highlight a person’s use of prohibited content, temporarily restricting content while a team is reviewing, demonetizing, minimizing its amplification, interfering with bots and coordinated online mob behaviour, adopting geolocated restrictions, and even promote counter-messaging”.*

15. Regulation should be proportionate, mindful of the size of the company concerned and the risk its activities present. Small or low-risk companies should not face an undue burden from the proposed regulation. However, some groups are sufficiently vulnerable (e.g. children) that any business aiming a service at them should take an appropriate level of care, no matter what its size or newness to market. Again, taking a “by design” approach means that baking in harm reduction to the design of services from the outset reduces uncertainty and minimises costs later in a company’s growth.

## How would regulation work?

16. We recommended an independent regulator (we cite Ofcom in the UK, and the Government has indicated that it is “minded” to appoint them as the online harms regulator when they confirm their proposals). It is important however that the regulator is independent and is evidence-based in its decision-making. The regulator should be given substantial freedom in its approach so as to remain relevant and flexible over time. We suggest the regulator employ a harm reduction method similar to that used for reducing pollution: agree tests for harm, run the tests, the company responsible for harm invests to reduce the tested level, test again to see if investment has worked and repeat if necessary. If the level of harm does not fall or if a company does not co-operate then the regulator will have sanctions.
17. In a model process, the regulator would work with civil society, users, victims and the companies to determine the tests and discuss both companies harm reduction plans and their outcomes. The regulator would have the power to request information from regulated companies as well as having its own research function. The regulator is there to tackle systemic issues in companies and, in this proposal, individuals would not have a right of action to the regulator or the courts under the statutory duty of care.

<sup>4</sup> <https://www.ohchr.org/EN/Issues/FreedomOpinion/Pages/ReportOnlineHateSpeech.aspx>

18. The regulator needs effective powers to make companies change behaviour. We propose large fines set as a proportion of turnover, along the lines of the GDPR regime. We have also made suggestions of powers that bite on Directors personally, such as fines given the way charismatic founders continue to be involved in running social media companies.
19. The industry develops fast and urgent action is needed. This creates a tension with a traditional deliberative process of forming legislation and then regulation. We are urging the UK government to find a route to act quickly and bring a duty of care to bear on the companies as fast as possible. In the meantime, we welcome the fact that the UN's work on the draft General Comment is proceeding with requisite focus and urgency to address the potential harms that arise to children's rights in the digital sphere and are happy to support their proposals.

Professor Lorna Woods  
William Perrin  
Maeve Walsh

November 2020

(Contact: [maeve.walsh@carnegieUK.org](mailto:maeve.walsh@carnegieUK.org))