Stakeholder Input Report:

What are the views of the tech industry and civil society groups regarding the upcoming Online Harms Bill?

A report for Parliamentarians – the first stage in a series of actions by PICTFOR, with input from the tech community, responding to, and discussing Online Harms

Contents

This is not an official publication of the House of Commons or the House of Lords. It has not been approved by either House or its committees. All-Party Parliamentary Groups are informal groups of Members of both Houses with a common interest in particular issues. The views expressed in this report are those of the group.

Foreword

About PICTFOR

The Parliamentary Internet, Communications and Technology Forum (PICTFOR) is the largest and most active All-Party Parliamentary Group (APPG). We bring together engaged Parliamentarians with key industry stakeholders to discuss issues on tech, communications and internet policy, as well as many more issues facing the sector.

Facilitating frequent, high-level discussions between policymakers and thought leaders across the sector, PICTFOR creates unique opportunities to exchange ideas and enhance Britain's competitiveness as a digital economy, whilst championing those members of the tech sector who have a social purpose.

Never before has the combination of these two values been more important for the digital economy in the UK. The escalation of the COVID-19 pandemic has highlighted just how vital the tech, communications and internet sectors have become to our nation's physical, mental and economic health.

PICTFOR and the Online Harms Agenda

The Parliamentary Internet Communications and Technology Forum (PICTFOR) is proud to facilitate regular and diverse discussions on policy relating to the internet, communications and technology. As an All-Party Parliamentary Group, our annual events programme features recurring events on cyber security, diversity in the sector, tech skills, connectivity, 5G, driving growth, attracting investment into the UK industry, health tech and many other issues.

Our Parliamentarian, industry and stakeholder members play a vital role in curating the themes of these events, regularly feeding into the events programme to reflect the concerns and priorities of thought leaders and policy makers. Since the announcement of the Online Harms White Paper in April 2019, PICTFOR members have expressed their desire to have a forum in which they can debate and scrutinise any proposed legislation.

Our programme has been impacted by recent developments: our Online Harms discussion has been scheduled twice before and was postponed due to the prorogation of Parliament, and then again due to the COVID-19 pandemic. We, our parliamentary Vice-Chairs, and of course our industry members, are all glad to be able to hold this discussion now, at a pivotal time for the tech industry.

The COVID-19 pandemic and the associated social restrictions have changed the way we engage online as well as offline. Many people now work from home, we socialise virtually, see GPs remotely, and a significant amount of education now takes place online. As our internet dependence has increased, so has our potential to be exposed to online harm. Now, more than ever, we must develop legislation that protects our freedoms while protecting us from harm in the digital world. The ensure the policy developed to tackle this issue is effective, realistic and fair, it must draw on insights from across the sector, as well as from civil society groups, academics and Parliamentarians.

PICTFOR discussions offer an invaluable insight into the workings and concerns of the most influential and important tech organisations operating throughout the UK. We welcome the level of engagement from the sector on this topic and would like to thank all those who have contributed to this report. We look forward to seeing industry and stakeholder groups engage further with government and Parliament as we legislate toward a safer environment for all.

Darren Jones MP – PICTFOR Co-chair

Baroness Neville-Rolfe DBE, CMG –PICTFOR Co-chair

The Rt Hon Lord McNally – PICTFOR Treasurer

Executive Summary

For this report, PICTFOR received 13 submissions from organisations across the tech sector and civil society, more than half of these came from PICTFOR industry members. These submissions reflect the diversity of our members and the wider tech community, with contributors including tech accessibility charities, trade bodies and some of the world's largest corporations.

PICTFOR members and the tech sector as a whole, continue to be invested in Online Harms legislation; PICTFOR has been encouraged by the level of engagement from the sector, the willingness to collaborate with Parliament, and the detailed and insightful contributions around this topic.

This report is the first of a number of actions PICTFOR will take to facilitate the sector's response and engagement on the issue of Online Harms. The next action will involve a virtual roundtable event with leading parliamentary voices, officials, and of course thought leaders from across the sector.

List of Contributors

We would like to express our gratitude to the following organisations that have contributed to this report:

Contributors that are PICTFOR Members

- BCS
- BT
- Good Things Foundation
- Google
- Huawei
- Internet Association
- TikTok

Contributors that are not members of PICTFOR

- Antisemitism Policy Group
- Carnegie
- Internet Watch Foundation (IWF)
- NSPCC
- Open Rights Group
- 5Rights Foundation

BCS – Insights & Contribution

The purpose of the BCS is to promote and advance the education and practice of computing for the benefit of society. We bring together industry, academics, practitioners, and Government to share knowledge, promote new thinking, inform the design of new curricula, shape public policy and inform the public. As the professional membership and accreditation body for IT, we serve around 60,000 members, in the UK and internationally. We also accredit the computing degree courses in universities across the UK and offer a range of widely recognised professional and end-user qualifications.

BCS welcomes the opportunity to submit evidence to the All-Party Parliamentary Group (APPG) Parliamentary Internet, Communications and Technology Forum (PICTFOR) call for evidence regarding Online Harms.

Duty of Care
BCS welcomes and urges the need for effective and reasonable social media regulation, however it is important users and organisations are aware of their respective responsibilities in the use of social media platforms. The term 'duty of care' that is aimed at the technology industry requires significant clarification.

The proposals that the Online Harms White Paper (OHWP) suggest leave room for debate about who is regulated and consequently who withholds the duty of care – the platform or the regulator. To enforce the legalities that lie with a duty of care, there needs to be clarity around whether this equates to the tort of negligence in civil law, or whether the term 'duty of care' is being used with less legal precision, without sufficient consideration to its consequences or meaning. If this does not gain clarity it will lead to regulatory uncertainty and the potential for online harms to continue without anyone being held accountable or for those who are taking their responsibilities seriously taking steps that do not align with unclear expectations.

Definition
What does 'online harms' mean? Online harms can mean numerous things to a number of different people, dependant on the type of online activity, what a victim might be experiencing and what a platform might be permitting in terms of content. For legislation to provide suitable punishment for those causing harm to others online, the law needs to identify specific forms of harm to tackle specific cases which requires significant debate.

For example, the Law Commission is currently doing a good job reviewing image-based abuse to improve the legislation, but they are specifically saying that they will not look at minors. Consequently under 18's are still being criminalised for sending explicit images based upon outdated legislation and the harm they might be subjected to as a result in the non-consensual sharing of an intimate image might be exacerbated by a fear of non-disclosure due to fears of criminalisation. This is a critical issue and requires immediate work to ensure we are using the law in the correct way to safeguard children and prevent them from being subject to a loophole in the law.

Safeguarding and digital skills
To navigate the online world safely, people need Digital literacy combined with critical thinking However 11.9 million people do not possess essential digital skills that are vital to protecting themselves properly online. Over half of the 4.1 million adults who are offline in the UK are from a low-income household[1], and supporting these adults to gain their digital skills is essential for their children's online safety – as they also need to understand what dangers their children face.

[1] https://www.goodthingsfoundation.org/news-and-blogs/blog/do-you-know-trusted-source

Providing children with digital literacy and critical thinking from a young age is required to provide them with the knowledge to live in the online world safely. If a child is from a lower income family they are more likely to be targeted online as are likely to have fewer digital skills and will therefore be more vulnerable to online harms in their many forms.

Safeguarding the whole of society from online harms is the way we are able to make the internet a wholly safer environment and requires buy-in by all stakeholders. That means the whole of society requires basic digital skills focussed on helping them understand how to be safe online.

Online regulation and Brexit

Online platforms have traditionally been exempted from content liability under EU law. Section 15 of the e-Commerce Directive 2000 stipulates that member states cannot impose liability on platforms for the behaviour and speech of their users, under the principle of 'platform not publisher'. The legal separation of platforms from content has formed the bedrock of internet conduct and commerce, but in recent years has faced challenges from civil rights campaigners, interest groups, and now the UK government in the form of the OHWP, published April 2019.

Ultimately, the OHWP seeks to address problems that will only become more salient as technologies become more advanced, and more generations grow up in the internet world. Groups and stakeholders will not cease lobbying for change in this area. The OHWP generally speaks to a noble cause. The growing problems faced by children are invidious and difficult to prevent, given the centrality of social media and personal computer use to vast swathes of society.

The OHWP has some weaknesses and has received criticism. In terms of those relevant to Brexit, one prominent criticism has been to point out that the scope of regulation is unclear and likely to be too broad. The OHWP's scope is named as companies which provide platforms for users to share user-generated content or interact with other users, including but not limited to social media platforms, discussion forums, search engines and messaging services. This ambitious attempt to regulate online speech in a very general sense seems bound to fail.

By specifying the remit of Ofcom for legal content suppression, the government has demonstrated some commitment to free speech and free flow of information, which the e-Commerce Directive has staunchly protected. A focus on systemic change rather than mandating takedown of individual pieces of content has clarified the government's position on online harms.

The OHWP has stipulated that new regulation would be compatible with the e-Commerce Directive. Heather Burns, a tech policy consultant, has stated that the e-Commerce Directive and its replacement is the most important Brexit issue for digital professionals. Navigating this intersection will be crucial for the future of tech innovation in the UK given the vast numbers of platforms that could be subject to new forms of regulation.

BT - Insights & Contribution

Summary

Technology is changing the world in all sorts of wonderful ways, but for many of us that sense of change can also be scary. We live in a time of proliferating data, sensors and intelligent machines. The ways we live, work, travel and stay healthy are going to continue to change long after the current pandemic is over. At BT we're excited and ambitious about what it could all mean. But we also know the pace of change can be jarring - disorientating and hard to understand.

That's understandable given the clear risks and harms caused by illegal and damaging online content, in particular for children, who can be vulnerable to abuse and exploitation. People worry about what they might experience online, from harassment to disinformation and fraud. For some, it's enough to put them off going online at all.

There is a noisy debate about how far Government should intervene in digital markets, whether through regulation or taxation. Many companies and commentators claim to know what's best for people or what the public want, but too often these views aren't evidence based.

Ofcom has published some very thorough analysis of the prevalence of different sorts of online harm and offence, and the extent of people's concern. But there has been very little published research that tests what ordinary citizens think about how such things should be regulated and why.

BT has 30 million customers right across the UK with our three brands – BT, EE and Plusnet. We want to understand what they think there is to hope or fear from technology: we want to listen to them, and to give them a voice in these debates.
That's why we asked Demos to conduct in-depth research into public attitudes to the different ways that the worst excesses of internet behaviour could be tackled, which was the focus of the Government Online Harms White Paper last year.

The research found that there is real consensus over the need for online service and social media companies to better protect those they impact negatively, especially children. The headline results are striking and suggest that the Government has a very clear mandate to be bold with its forthcoming legislation. In a representative sample of more than 2,000 people:

- More than half (53%) have themselves experienced online harm
- Over 80% think the grooming, bullying and sexual exploitation of children are big problems for society
- A majority think that responsibility for fixing that is shared – between different sorts of companies, politicians, police, regulators and individuals themselves
- For the most serious crimes, the great majority – again over 80% - place most responsibility with the social platforms and with Government and regulators
- 77% favour 'age verification' measures to keep kids away from certain sites
- Over 75% support the blocking of entire websites if companies fail to take the right steps to prevent online harm

None of this means the task of framing legislation is easy, because there are some complicated trade-offs involved. For example, if both social media companies and regulators are to be held responsible for preventing harm, where does the platform's responsibility stop and when does the regulator take over? Is it right to include the contents of private messaging services in the scope of any future regulation: at what point should an individual user's right to privacy give way to a need to protect the wider safety of children?

Within BT, from time to time we've had our own internal debates about how to approach these questions. In fact, that was one of the reasons we commissioned this research in the first place – because we could all agree that the public's voice should be the loudest and the most important.

The research shows that many people are willing to make some sacrifices to their individual online liberties for the sake of the wider community's protection and security. For example:

- 64% want to stop online user anonymity because they think it allows more harmful behavior
- 65% say people should not be free to express themselves online if what they say causes harm or serious distress to someone else
- 58% would be happy to see barriers to harmful content put up, even if that means accidentally censoring some non-harmful content

However, it's important to say that these responses also seem to be less clear-cut than others. When tested in focus groups, people raised concerns about who should be able to decide what is harmful, and how. Perhaps counter-intuitively, some of those who have themselves been direct victims of online harm are less likely to see it as a big problem for society, or to propose the strongest action. But there are also some who decide to leave online spaces entirely, given their experiences there.

The most divisive question Demos asked was about whether it should be possible for government agencies to access the contents of private messages between two people, in order to identify and prevent the most serious crimes, such as child abuse and terrorism. Respondents were split roughly 50/50.

When people were given more time to talk this through in groups, they were quick to focus on issues such as the boundaries that could be put around access, and the checks and balances that could ensure the authorities targeted any extreme intervention of this sort at the most extreme examples of criminal behaviour. In other words, they got straight to the sort of debate you might expect to hear in Parliament.
All of which leads me to make two very broad suggestions to DCMS on the back of this research.

First, press on and publish a draft Bill as soon as possible – there is a strong public demand for this legislation; the principles are very clear; and if appointed as the regulator Ofcom will do a great job.

At the same time: invite more open scrutiny and deliberation on the really tricky questions, about how to regulate in a way that provides both freedom and protection to all online citizens in the UK. These are fundamentally democratic questions, and many of the right answers are likely to come from further research and consultation with citizens themselves.

Demos's summary of their findings is here:

Their summary of focus groups:

Full Polling results:

Good Things Foundation – Insights & Contribution

About Good Things Foundation

Good Things Foundation is the UK's leading digital and social inclusion charity. We bring together a network of hyperlocal community partners across the country (the Online Centres Network), working together to reach those who need support. Through our online learning platform, Learn My Way, we've supported more than three million people to gain digital skills since 2010. Our network partners are all independent of Good Things - they are small community centres, local charities supporting people with disabilities or unemployment, homeless shelters, job clubs, libraries, Age UK centres, et al. We call it a 'big club with a shared vision' - a vision of a world where everyone can benefit from digital. Our focus is on adults - especially those who are more likely to be older, isolated, or face wider social, health and economic challenges.

Setting a vision

We share the Government's vision of making Britain the safest place in the world to be online; and belief in the importance of a vibrant tech sector to drive economic prosperity. As the Online Harms White Paper set out: 'Innovation and safety online are not mutually exclusive'. We see building trust in the digital economy and new technologies as essential for the UK; but remain deeply concerned that not enough is being done to protect, empower and support people – particularly those with no or limited digital access, skills or confidence. As our Digital Nation 2020 infographic reminds us: 9 million UK adults struggle to use the internet independently. There is a clear need for levelling up: only 18% of people in the North East use the internet fully, compared to 49% of people in the South East. Concerns about online safety and security (especially fraud and scams, data privacy and security) remain high - and can prevent people from going online or using the internet fully (e.g. for online banking, or online government or health services); concerns about data practices are also growing.

What is needed

(1) Investment in essential digital skills for older and vulnerable citizens. In our Digital Blueprint, we call for three things:

- A Great Digital Catch Up to kickstart our economic recovery and build back better through ensuring that every community has a place to get digital skills support;
- A Data Poverty Lab to find innovative sustainable solutions to data poverty;
- A Digital Strategy for Everyone - recognising that digital is now essential across so many aspects of our daily lives.

Without these, we question whether the UK has the necessary conditions - social, cultural, educational and economic - for citizens to play their part in protecting themselves and each other against online harms - whether criminal harms, or those which may be deemed legal but harmful.

2) Clarity about consumer or economic harms. A strong case has already been made by leading stakeholders (e.g. UK Finance (2019); Money and Mental Health Policy Institute (2020); Carnegie UK Trust (2020)) for these harms to be brought within the scope of Online Harms legislation. This has taken on increased urgency, given growth in online fraud/scams and identity theft; the financial and health impacts to people's lives as well as the economy; and the reality that the design choices which enable these to spread are the same as around other harms in scope. Irrespective of whether scope is broadened to include consumer harms, it is clear that more effective protections for consumers are required.

3) Continued evolution of approaches to digital skills and confidence - so we empower people to build their digital resilience, understanding and active digital citizenship. The current government is right to prioritise digital skills for and within the workforce, including basic digital skills. But evidence from a range of sources (Centre for Date Ethics and Innovation (2020) Demos (2019), Doteveryone (2020), Good Things (2018), Kennedy et al (2020), Lloyds Bank (2020), Ofcom/ICO (2020), Yates et al 2020) indicate the need to evolve our understanding and practice:

- **Public fear and mistrust of tech** companies, the internet and how organisations use our data; this risks undermining an economic recovery powered by data and improved innovation in using AI for societal benefit. Experience of everyday online harms can mean people step back from using technologies, and lose trust in data practices, as well as experiencing financial, emotional and health impacts. Too much 'thin data' reduces data quality and may compound algorithmic bias.

- **Additional barriers to being safe online** for new or 'limited' internet users - who only access the internet on small screens; cannot afford home broadband; rely on less secure free WiFi; have low digital, data and media literacy. These factors increase vulnerability to online harms (legal and criminal). Evidence shows a correspondence between these factors, age, low educational attainment and living on a low income.

- **A reality gap between how safe we believe we are** and what we do to protect ourselves. This is unsurprising and alarming. It highlights a need for simpler, consistent messages; and for a media literacy strategy which is future-proofed - blending media literacy with digital literacy and data literacy - so people can learn how to navigate the internet safely. As the Centre for Data Ethics and Innovation has flagged, digital exclusion and lack of digital health literacy are risks for the future success of innovation and expansion of online and mobile banking and digital health services, such as NHS apps or GP online services.

- **An opportunity to empower and support people** - especially those facing the increased risks noted above - so they can find support (in their local community and online) about protecting themselves from harm; and to become active digital citizens - supporting friends, family and others to be safe online. Campaigns like Friends Against Scams, and the networked power of community organisations in the Good Things network of 'online centres', are strong starting points for creating this cultural as well as technological change.

- **A potential levy to support public education.** The government has recognised gaps in support for adults - for themselves, and in their role as parents and carers - but it is unclear how this will be bridged. Online safety is a shared responsibility between government, companies and citizens. We want to see greater priority afforded to the critical task of empowering and supporting citizens to protect themselves and each other. Some stakeholders have proposed a levy on tech companies to pay for more public education and awareness around media, data and digital literacy. We would be interested in the views of other PICTFOR members on this, and on the central importance of digital inclusion to the UK's next Digital Strategy, and in making the UK the safest place in the world to be online.

Google – Insights & Contribution

Keeping users safe

We take the safety of our users very seriously, and we are committed to ensuring the small proportion of inappropriate content that appears on our platforms is addressed as quickly as possible.

Google is supportive of content regulation and the Government's goal of keeping users safe online and we want to work with you on a new model of content governance that is accountable, transparent and sustainable. That said, **we haven't waited for regulation to address problematic content online**. We feel a great responsibility to our users when they place their trust in us to deliver them trustworthy, helpful information that meets their needs. Core to our mission is a focus on the relevance and quality of the information we present to users. In different ways across our different platforms, we strive to connect people with 'high-quality information'; the most useful, trustworthy, and helpful content at the moment a person needs it. At the same time, we work to prevent user and societal harm and limit the reach of 'low-quality information'; content that strays furthest from those qualities.

- Our mission requires us to strike a careful balance between the free flow of information and social responsibility. The product, policy, and enforcement decisions we make are guided by an inclination toward openness and accessibility, respecting user choice, and building products and services for everyone. We strongly agree that all these efforts need to be made in a transparent and accountable manner, and this is why we have launched a whitepaper detailing information quality and content moderation efforts, as well as our regular transparency reporting efforts.
- Over the years, we have made significant investment in technology and human resources, and we have engaged with policymakers in the UK and around the world on the appropriate oversight for content sharing platforms, such as social media and video sharing sites.

For each product, we have a specific set of rules and guidelines that are suitable for the type of platform, how it is used, and the risk of harm associated with it. Our teams tackle a huge spectrum of online abuse, from scams, to abhorrent content, including child sexual abuse material (CSAM) online. Understanding the different parameters of the products we serve is vital to our work and policy development. Given that breadth, our team is diverse, comprising product specialists, engineers, lawyers, data scientists, ex-law enforcement officials and others. They work hand-in-hand around the world and with a global network of safety and subject matter experts.

Our approaches to tackling harmful content is built around four complementary levers:

- **Remove:** We set reasonable and responsible rules for each of our products and services and take action against content and behavior that violates them. We take tens of millions of actions every day against content that does not abide by the 'rules of the road' for one or more of our products. In Q2 2020, we removed 11.4 million videos globally for violating our community guidelines - over 10.8million of which were first flagged by machines rather than humans. More than half of the videos removed after being flagged by machines had never received a single view. We also removed over 2.1 billion comments that breached our guidelines, of which around 99% were flagged by our automated systems This is a fraction of the billions of comments posted on YouTube each quarter.

- **Raise** we elevate high-quality content and authoritative sources where it matters most. To determine whether a piece of content is useful, we must first try to understand a user's intent. Our systems then look for signals that can help determine the expertise, authoritativeness and trustworthiness of relevant web pages on that topic, so that we can prioritise the most appropriate sources available. We are constantly improving these ranking systems. In 2019 we ran over 464,065 experiments with trained external Search Raters and live tests, resulting in more than 3,620 improvements to Search.
- **Reduce** We work to reduce the spread of potentially harmful information wherever we feature our recommended content. In 2019, we launched over 30 different changes to our recommendations systems on YouTube in order to reduce recommendations of borderline content and harmful misinformation. The changes cause a 70% decrease in watch time from non-subscribed recommendations in the United States.
- **Reward** We set a high standard of quality and reliability for publishers and content creators who would like to monetise or advertise their content. We have no desire to derive revenue for ourselves, or for any other business, from harmful content or behaviour. For example, over the course of 2019:
  a. We removed more than 2.7 billion bad ads from our systems.
  b. We took action against almost 1 million bad advertiser accounts.
  c. We terminated over 1.2 million publisher accounts for violations of our policies.
  d. We removed ads from over 21 million pages that are part of our publisher network for violations of our policies.

Our goal is to achieve both accuracy and scale in our work. That's why we have people and technology working together - and we invest heavily in both. We now have **over 10,000 people across Google working on content moderation and removal** on our platforms. This includes reviewers who work around the world, 24/7, speak many different languages and are highly skilled.

To complement our internal efforts, we work with many talented experts and organisations across the technology industry, government, and civil society to ensure that we are doing everything we can to set the right policies, establish industry best practices, and get ahead of emerging challenges. We do this in part by relying on a community of partners to help us identify. content that violates our policies, seeking the advice of subject-matter experts as we craft and update policies, and working with industry partners to share best practices and cutting-edge technology – for instance, within the Global Internet Forum to Counter Terrorism (GIFCT). While this remains an important challenge, we are optimistic about the progress we have made on our own services and working together with other companies and governments. It's important that additional frameworks governing online speech be carefully balanced, clear, and fit for purpose. We look forward to continuing to work with governments to that end.

Online Harms regulation

Effective oversight of content moderation practices can also play a complementary role. Throughout the internet's history, industries, policymakers, and civil society have worked on codes of practice to guide appropriate behavior by online services.

The scrutiny of lawmakers and others often improves our products and the policies that govern them. It's sometimes claimed that the internet is an unregulated "wild west," but that's just not the case. Many laws and regulations have contributed to the internet's vitality: competition and consumer protection laws, advertising regulations, and copyright, to name just a few. Existing legal frameworks reflect trade-offs that help everyone reap the benefits of modern technologies, minimise social costs, and respect fundamental rights. As technology evolves, we need to stay attuned to how best to improve those rules.

As content sharing services like social media and video sharing sites have become more important to public discourse, oversight methods will continue to evolve as well, so as to better review platforms' efforts in light of best practices. We think new forms of oversight can work well when they focus on a specific, clearly defined problem and build on a principles-based approach that touches on:

- **Clarity:** Content-sharing platforms are working to develop and enforce responsible content policies the establish baseline expectations for users and articulate a clear basis for removal of content as well as for suspension or closure of accounts. But it's also important for governments to draw clear lines between legal and illegal speech, based on evidence of harm and consistent with norms of democratic accountability and international human rights. Without clear definitions, there is a risk of arbitrary or opaque enforcement that limits access to legitimate information.

- **Suitability:** It's important for oversight frameworks to recognise the different purposes and functions of different services. Rules that make sense for social networks, video-sharing platforms, and other services primarily designed to help people share content with a broad audience may not be appropriate for search engines, enterprise services, file storage, communication tools, or other online services, where users have fundamentally different expectations and applications. Different types of content may likewise call for different approaches. Transparency - Meaningful transparency promotes accountability. We launched our first Transparency Report more than eight years ago, and we continue to extend our transparency efforts over time. Done thoughtfully, transparency can promote best practices, facilitate research, and encourage innovation, without enabling abuse of processes.

- **Flexibility:** We and other tech companies have pushed the boundaries of computer science in identifying and removing problematic content at scale. These technical advances require flexible legal frameworks, not static or one-size-fits-all mandates. Likewise, legal approaches should recognise the varying needs and capabilities of start-ups and smaller companies.

- **Overall quality:** The scope and complexity of modern platforms requires a data-driven approach that focuses on overall results rather than anecdotes. While we will never eliminate all problematic content, we should recognise progress in making that content less prominent. Reviews under the European Union's codes on hate speech and disinformation offer a useful example of assessing overall progress against a complex set of goals.

- **Cooperation:** International coordination should strive to align on broad principles and practices. While there is broad international consensus on issues like child sexual abuse imagery, in other areas individual countries will make their own choices about the limits of permissible speech, and one country should not be able to impose its content restrictions on another.

With the right regulations in place, the UK can ensure that it benefits from the advantages promised by digital innovation, including jobs and growth across the UK, while creating a world-leading regime to ensure users are safe online.

Huawei – Insights and Contribution

Implications for privacy and online safety:

- For families, connectivity brings and will continue to bring benefits for homes and everyday lives. However, as a result, homes will become ever more porous, as connected devices continually collect data. Families seem willing to give this data, yet this means their homes have become places where they no longer feel entirely private. Present for most families, this low level of trust and privacy conflicts, but seemingly does not outweigh their desire to buy technology and to evolve their homes. Tech corporates need to do more voluntarily to ensure that their connectivity, devices and online spaces are safe for families and offer more transparency about what their data is being used for. With the growing use of voice technology, we should not readily accept that the erosion of privacy and the commercialisation of personal data is the price families pay for convenience and an easier life. Parents have concerns about Voice Assistants recording what they say but seem much less aware of the data being continually collected and permanently stored through growing technologies like VR-headsets. With VR, intimate personal information has already started to be given away, and while driving innovation for the consumer, there needs to be greater clarity on who else is benefitting so families can crucially make more informed choices about what products they adopt.

- The risks inherent in connected technologies to home safety, security and privacy are not perceived by the family to be their responsibility. Many children and teenagers access information and have social contact privately, without parental intervention. With more new types of connected tech emerging, there will be more reliance on stakeholders, legislators and regulators to ensure that at a macro-level everyone and particularly children will be safe and secure. The response to the Online Harms White Paper, the Age Appropriate Design Code and the increased role of the regulator intends to increase the safety of digital spaces. However, it is yet to be seen how compliance to regulation will be effectively enforced, and it will always lag behind technological innovation and user appropriation.

- As legislation will not offer blanket protection – digital literacy and education about appropriate online behaviours and relationships remain essential for all. To truly be the safest place in the world to be online – which is the stated ambition of DCMS – all users, children, teenagers, adults and seniors must have the education, training and skills to enable them to take responsibility for their online lives. This will require a public education campaign – the need of which has been ably demonstrated by the lockdown. Educating families about digital wellbeing has never been more critical, both now and well into the future of the technological family.

Internet Association – Insights and Contribution

(The below is the introduction of Internet Associations 2019 DCMS submission)

IA represents over 40 of the world's leading internet companies and is the only trade association that exclusively represents leading global internet companies on matters of public policy. IA's mission is to foster innovation, promote economic growth, and empower people through the free and open internet – in November 2018 IA established a London office to constructively engage in the internet public policy debate in the UK.

We are firm believers in the benefits that technology brings to everyday life and the economy, and for the potential that internet innovation has to transform society for the better. IA economic analysis shows that the internet sector contributes £45 billion to the UK economy each year, and is responsible for nearly 80,000 businesses and around 400,000 jobs. Recent IA polling found that 82 percent of British people believe that the internet had "made their lives easier and more enjoyable."

As the OHWP argues, the internet is now an integral part of everyday life and often a powerful force for good. Thanks to the internet, we now have unprecedented access to information, entertainment, communication and a vast range of new goods and services – which has created a more informed, connected and productive society. According to Ofcom data, the average person now spends 24 hours a week online, and multiple estimates have found that internet services create significant consumer surplus for ordinary people. Many of these services are provided to consumers free of charge, with the recent Bean Independent Review of UK Economic Statistics estimating that including the value created by free internet services in GDP would boost growth by 0.35 – 0.66 percentage points a year.

IA believes that the internet sector needs a balanced policy and regulatory environment to continue, and grow, its contribution to the UK economy, consumers and society in the future. The internet will drive 21st century prosperity, but there is a risk to this potential if policies and regulations are introduced which will damage the ability of the internet sector to: 1) drive UK economic growth; 2) provide services that people value highly; and 3) make a positive contribution to society.

IA has previously proposed a number of regulatory policy principles which we believe can help deliver this balanced environment, and IA and our members will continue to work constructively with policymakers and regulators on these important issues as the White Paper process continues.

TikTok – Insights & Contribution

TikTok strongly supports the government's plans to legislate on online harms. It is our view that the social media and content sharing industry need to operate within a clear framework established by Parliament. Major strides have been taken through self-regulatory initiatives, but the Online Harms Bill will go further by laying the foundations for regulatory system of cooperation between companies, users, parents, government and the new regulator. This will ultimately have the effect of enhancing safety and security online, which is TikTok's top priority.

Promoting a safe and positive experience is our priority

TikTok is a space designed to inspire creativity and bring joy. Our top priority is to promote a safe and positive experience so that everyone can be free to express their creativity.

Our Community Guidelines reflect our values and explain the kind of behaviour we expect from our community. We enforce these rules using a combination of technologies and thousands of safety experts around the world.

We develop pioneering safety policies and features, such as restricting direct messaging and live-streaming to over 16s only, and restricting use of TikTok’s gifting system for users under 18. We also help parents to manage their teen’s TikTok experience through Family Pairing. We actively promote these features to increase our users' awareness and use of them. Over the past year, we have engaged with our parents of teenage users as well as users themselves through a series of focus groups so that we hear and respond to their concerns first hand. As part of these discussions we explored parental attitudes towards issues including age assurance, preferred methods for receiving safety information and parental attitudes toward screen time.

We collaborate with others in the industry to make the digital world safer and kinder ~~safer~~ for everyone. We work with organisations such as Internet Matters, Mind, the Internet Watch Foundation, BEAT, The Holocaust Educational Trust, Stonewall and the WeProtect Global Alliance, and sign up to initiatives like the Voluntary Principles to Counter Online Child Sexual Exploitation and Abuse.

As a newer platform, we are uniquely positioned to set the bar for safety, and we will never stop working to make TikTok an ever safer platform for our community so they can feel free to express themselves.

Our approach to online harms

TikTok's approach to the online harms framework has four pillars:

1. Assume responsibility

The current system has a flaw that needs to be fixed. In the simplest terms, if a platform looks for harmful and illegal content, it could become liable; if it does not look, it cannot become liable. This acts as a disincentive against platforms that want to invest in and develop techniques to detect, review and remove videos that contain illegal content. That is why TikTok supports the development of 'good Samaritan clause' which will give legal clarity to allow and encourage platforms to proactively remove illegal content. We would like the online harms framework to be aligned with this approach.

2. Be transparent and open

TikTok believes that all companies should disclose their algorithms, moderation policies, and data flows to regulators. Rather than waiting for regulation to come, TikTok has taken the first step by launching a Transparency and Accountability Center for moderation and data practices. Experts can observe our moderation policies in real-time, as well as examine the actual code that drives our algorithms. We believe that the codes of conduct under the new online harms framework should include provisions on transparency of moderation practices, and access to and scrutiny of algorithms.

3. Collaborate across platforms

TikTok believes that "harmful but legal" content should not be dealt with in the same way as illegal content. This is because harmful content is more frequently open to debate and perception is shaped by context. Platforms do, however, need to assume responsibility for preventing and removing "harmful but legal" content.

This means putting in place mechanisms to collaborate between companies to spot harmful trends early and to coordinate interventions. That is why on 21st September TikTok wrote to other technology companies proposing the establishment of a global coalition to protect against harmful content. This would take the form of a Memorandum of Understanding (MOU) that would encourage companies to warn one another of violent, graphic content that has appeared on their own platforms, along the lines of the cooperation mechanisms that already exist for child sexual abuse material.

4. Regulate in the most impactful way

Each social media and content sharing platform is different and is enjoyed differently by users. Regulation needs to reflect this - a single, uniform solution will not be effective for each and every platform. That is why TikTok supports a principle based approach in the Online Harms Bill where common principles and outcomes to protect users are established in law and then the new regulator works closely with different platforms to ensure they have the best systems in placed to protect their users online. The penalty system needs to reflect this too. Companies should face enforcement action for systemic failures, rather than for individual, isolated cases.

Antisemitism Policy Trust – Insights & Contribution

The Online Harms Bill: Top Lines

One of the most important aspects of the forthcoming Online Harms Bill is the proposed introduction of a regulator and a statutory Duty of Care that addresses reasonably foreseeable harms.

Regulator:

While we are supportive of the formation of the online harms regulator and are satisfied that Ofcom can fulfil the role as proposed, we are concerned that this new regulator should work alongside, and in partnership with, other existing regulators in the online space. For example, we have worked closely with the British Board of Film Classification (BBFC) and respect its unparalleled understanding of harmful content. Sara Khan, the Lead Commissioner for the independent Commission for Countering Extremism has already suggested that her commission could help the new regulator in defining and understanding extremist content. That would be a welcome step.

Duty of Care:

We are concerned that **fulfilling the proposed statutory Duty of Care should mean companies and bound to follow a number of Codes of Practice** (or having systems equal to them) and that it would be **insufficient to only require companies to have minimum Terms and Conditions** (the Government's White Paper made clear that this is not sufficient). Bitchute (think YouTube but filled with extreme content) is but one example of a so-called 'alternative' platform with poor Terms that should be borne in mind as part of this policy development.

The White Paper and interim response suggest that the Home Secretary will have oversight of the Codes relating to Child Sexual Exploitation and Abuse, and on Terrorism. These are given special status in the proposals. However, given there are already advanced measures in place to address these harms, **giving additional prominence to the Hate Crimes/Harms Code of Practice would be welcome**. These codes are due to set out systems-level requirements for companies, in a proportionate way, and our hope is that not just illegal, but also legal but harmful content will be captured.

Penalties:

Contravention of the Duty of Care will result in penalties. Fines are all well and good but in some cases, social media companies are worth Billions of dollars. To that end, **in extremis, individual liability for senior management of companies in scope should be introduced,** as exists in the financial sector, and similar to that in Health and Safety legislation. Other penalties like a public adverse behaviour warning would also be welcome.

1) Online harms - Offline harms:
The online world can no longer be considered in a silo. Extremism, Terrorism and Policing require online and offline responses, too often these are poorly co-ordinated, if considered at all. Any policy developed by government should be designed for contiguity between online and offline elements.

2) Holocaust Denial (a legal harm):

Holocaust denial is a good example of a legal harm. As a society, we are in danger of outsourcing our values to individuals and companies in California and elsewhere. Harmful content online must be considered in respect of the impact it has. These new technologies cannot be considered through existing matrixes of decision making and regulation.

3) Coronavirus and antisemitism:

The adaptability of online hate actors, and their responsiveness to particular situations can be ahead of technological abilities to cope. With facial manipulation software and other new inventions on the horizon, responsiveness, and address for reasonably foreseeable harms will be critical.

4) Misogyny and antisemitism:
Intersectional harms are rising online, as this briefing proves. The intersection of identities and ability of online actors to impact in intersectional ways means we require a new approach to understanding hate online. The Law Commission's current review of hate crime law will give us an opportunity to raise these issues but thought should be given to the way in which the law might require change to capture harms like this.

5) Big Data/hidden harms report:
This report explores search and post trends on Google and the far-right website Stormfront. It evidences that small changes by a technology company can lead to significant address for harm. The responsibility for such companies to have a duty of care to their users is clear.

6) Online Harms white paper
Though now overtaken by recent developments, this was the Trust's initial response to the Online Harms White paper which might be of some interest.

7) 'Alternative' sites
Sites like Bitchute, Telegram, 4chan and 8Chan are responsible for significant harms online. It is important that any policy to address digital harms not be concentrated on mainstream sites like Facebook and Twitter alone.

Carnegie – Insights & Contribution

1. This submission summarises the main headlines from our work on the development of a statutory duty of care for online harms reduction and the areas which we feel should be a priority for focus ahead of the Government's publication of its final proposals. We have provided links to supporting material for reference and look forward to joining the event on the 3rd November.

2. Our work to develop a statutory duty of care for online harm reduction began in 2018, was developed through a series of blog posts, roundtables and submissions to Parliamentary and other consultations and set out in full in our full reference paper of April 2019. Our proposal is for social media companies to design and run safer systems – not for government to regulate individual pieces of content. Companies should take reasonable steps to prevent reasonably foreseeable harms that occur in the operation of their services, enforced by a regulator. 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. The celebrated 'move fast and break things' method has had its day. Broadcast regulation demonstrates that a skilled regulator can work to assess harm in context, regulate it and balance this with maintaining free speech. Proportionality in regulation allows for innovation and market entry by SMEs.

3. We have published a draft Online Harm Reduction Bill to demonstrate that such a regime could be legislated for in quite a simple way. We dovetail online harms into an existing, proven regulatory regime (the Communications Act 2003). We are also supporting Lord McNally on his Online Harm Reduction Regulator (Report) Bill – a short paving Bill that would give powers to Ofcom to prepare for the introduction of an online harms regime – and await a date for its Second Reading in the Lords.

4. The Covid19 pandemic and subsequent lockdown have seen increased prevalence of online harms, whether through an increase in child abuse and grooming online, large volumes of scams and fraud, rising levels of online abuse targeted at minority groups and the spread of conspiracy theories and disinformation, which has led to real world harms (such as the destruction of 5G masts) and an increase in anti-vaccine sentiment. We have attached a briefing note which sets this out further.

5. Had a systemic statutory duty of care already been in place, as per our proposals, it would have required platforms to be accountable for the design of their platforms and the actions taken to limit the spread and reach of these harms. We set out below, with some background material, our suggestions for areas which need urgent Parliamentary and political focus to ensure the legislation, when introduced, is as effective as it can be.

6. "**Legal but Harmful":** the Government has still not made clear how it will address significant harms arising as a result of activity that is not illegal (like terrorism or child sexual abuse content, which is definitively in scope). It has repeatedly referred to the need for platforms to "enforce their own terms and conditions" and framed the decision as something that potentially impacts freedom of speech. Professor Woods, in a comprehensive paper published in December 2019, has set out how a systemic duty of care that bites at the level of platform design rather than that of individual pieces of content, is compatible with fundamental rights, including freedom of expression.
7. A further issue that arises in relation to having a weak obligation on platforms to enforce their own policies is that changes to those policies and Ts&Cs only tends to occur after the harm has occurred

and at such a scale that civil society and political pressure is brought on the platforms to act. We have seen this recently in relation to Twitter's actions to change elements of their design to reduce the risk of a repeat of the spread of viral disinformation that occurred during the 2016 Presidential election and the slew of piecemeal policy changes introduced by Facebook to address the societal harms caused by the spread of, amongst other things, QAnon-co-ordinated conspiracy theories, holocaust denial and anti-vaxx content. The Secretary of State's recent evidence to the DCMS Select Committee did not appear to go any further than suggest that platforms would submit their Ts&Cs to the regulator for approval and then be judged on them. This is not enough. Judging the platforms by their own policies does not equate to accountability for harm. Without regulation, they can just change their policies back. It is also not clear by what standards Ofcom will be judging the Ts&Cs – there's a risk of a race to the bottom if base standards are not clear – or that companies then retreat from standards set. A systemic duty of care would require the platforms to be accountable for the reasonably foreseeable risk of harm before it is caused to users and society, not in response to a regulator's judgement on the companies' performance once the harm has occurred.

8. **Misinformation and Disinformation:** the Government's recent response to the DCMS Select Committee inquiry into Misinformation in the Covid 19 infodemic suggests that, despite the impact on public health of the spread of false health claims during the pandemic and the clear risks from anti-vaxx campaigns to the effective roll-out of a vaccine, this will not be in scope of their Online Harms legislation. Under a systemic duty of care, we see no reason why public health harms should not be addressed, as Will Perrin sets out in this blog post from April 2020.

9. We have also published a longer piece on how Online Harms regulation that focuses on a platform's system design would address the "infodemic". In relation to harms to democracy, we have also joined forces with civil society organisations to make the case for this also to be included within a systemic duty of care.

10. **Consumer harms and scams:** despite the scale of financial losses and emotional distress incurred by users of social media who fall victim to scams and fraud (Action Fraud estimates 85% of the losses of £2.3bn in the year to June 2020 were "cyber-enabled"), the Government is currently refusing to include this in scope of the Online Harms legislation. Its view is that the new legislation should not duplicate or conflict with existing work by government, regulators and other bodies; for example, law enforcement being led by the Home Office and economic crime being dealt with by HM Treasury. Evidence given to the Home Affairs Committee in the summer from the Centre for Economic Crime and Action Fraud suggests otherwise: existing law enforcement and regulatory oversight are powerless to deal with the scale of fraud and scams and representatives of both organisations called for Online Harm legislation to be extended in this regard. We have set out detailed thinking on how this might be delivered, with a system of interlocking regulation that enables existing agencies and regulators to work together to tackle consumer and economic harms, with Ofcom empowered to act on evidence passed to them by competent authorities without the risk of duplication of responsibilities.

Internet Watch Foundation – Insights & Contribution

The IWF supports the ambitions laid out within the Online Harms White Paper and commends the Government for leading the way in publishing proposals for an online regulatory framework. We believe that the principles-based approach adopted by the White Paper is essential in ensuring that any legislation remains relevant due to the fast-paced nature of innovation and change in the digital sector. However, The White Paper still has some challenges to work through in providing clarity to the new regulator, industry and others so that the proposed new regulatory framework will not have unintended consequences. The internet is, by nature, extremely difficult to regulate; its global nature, low costs to access, and vast amounts of content evade traditional models of policy making. Regulating the online space raises serious questions ranging from privacy and security to human rights and freedom of expression. As such, it is crucial that any regulator remains truly independent from Government, and any political interference.

Building on Best Practice
The UK has some strong and effective mechanisms in place for protecting its users from online harms, and it is critical that successful models are not swept aside in the new regulatory environment. In 2018, only 0.04% of known child sexual abuse material was found to be hosted within the UK, and we have some of the fastest take down times in the world. The UK Government must build on existing legislation and regulation, utilising best practice and existing technical knowledge to deliver an effective approach that does not negatively impact the UK's vibrant digital economy. The IWF is concerned with the impact that a regulator levy would have on the funding of current voluntary initiatives. The regulator must be careful not to prevent industry from funding impactful organisations. If industry cannot fund current voluntary initiatives, such as the IWF and our partners at the UK Safer Internet Centre (UKSIC), this could result in the removal of fewer child sexual abuse images, decreased levels of awareness amongst users, and increased pressure on law enforcement agencies. The IWF has successfully operated within a very specific field for 23 years, and we could consider extending our remit to include grooming and live streaming, in consultation with our Members.

Technical Expertise
As a partnership organisation, the IWF believes that collaboration is central to sustainable and impactful legislation. We urge the Government and the new Regulator to engage with industry and other relevant experts, such as the IWF, in a framework and codes of practice that are technically viable. We would encourage the Government to build a shared forum in which industry members can disclose issues on their platforms and build solutions. The IWF provides a safe, secure and trusted forum for companies of all sizes and from different sectors to discuss issues with child sexual abuse material on their platforms, and we believe this is a crucial function which must be continued and further supported in the future. This would also be a strong arena for the regulator to engage with start-ups and provide high quality advice and support for building platforms that are safe by design. We would encourage any company to engage with the IWF as early as possible and believe that the Government should mandate through the Code of Practice industry taking relevant IWF services.

International Cooperation
Further thought needs to be given to how the Regulator will interact with companies based outside of the jurisdiction of the UK. The new regulatory environment must work in tandem with existing legislation at a transnational level, such as the E-Commerce Directive and Child Sexual Abuse Directive. Furthermore, several of our Members have stated that it would be difficult to create technical solutions for just the UK regulatory environment, as their operation is based globally. The IWF believes the UK government should look to establish an international standard for categorising CSEA, possibly raised at the five-eyes ministerial

forum and in discussion with the European Union. We are also concerned that the new regulatory framework may hamper the companies that deploy our services internationally and potentially our portal programme which provides a vitally needed place to report in the most under-developed countries in the world. Education and Prevention The IWF is calling for a national prevent campaign to be launched aimed at 18 to 24 year old men, who our research has found to be the most likely group to stumble across child sexual abuse material on the open web and least likely to report it. We firmly believe that more should be done to prevent people from ever seeing indecent content in the first place. In an increasingly digital world, users need to be empowered to keep themselves and their children safe online. We believe that there needs to be a better approach to sex and relationships education in schools and a more open dialogue with girls aged 11-13. Self-generated imagery now makes up one third of all the child sexual abuse content which we remove from the internet: 82% of which features 11-13 age range, and 99% of which is girls. However, given the complexities of regulating the online environment, we believe that the Regulator would not have the capacity to focus appropriate time and energy into education and awareness. Rather, the IWF recommends that this should be kept in the remit of charities with a good understanding of engagement in schools and professionals, such as our partners at the UK Safer Internet Centre (UKSIC). Such organisations should be provided with support and funding from Government to deliver this crucial service.

Scope

The IWF supports the Government’s attempts to ensure that the Online Harms White Paper is a comprehensive document. However, we believe that it is not feasible to expect any regulator to have the capacity to address all 29 harms outlined in the White Paper. Rather, we recommend that the Government works with existing technical experts working in this field, such as the IWF, to address such a range of multi-facetted harms. Additionally, we are concerned that the 29 harms in scope will demand significant resources, disproportionately affecting small companies and start-ups. We are concerned that this regulatory environment has been primarily designed for social media companies, and then later extended to encompass a much broader range of companies that make up the internet ecosystem. We believe the Regulator must clarify both the harms in scope and its expectations, clearly setting out the responsibilities of industry. Further clarification is needed regarding the duty of care on industry, private communications, and how regulation will interact with existing legislation. Ultimately, our wish is that the best thing is done for children who have been sexually abused, and then been further violated by having the images and videos of their abuse shared online. They are the priority and focus of our mission and protecting them should remain central in any new regulatory environment.

NSPCC – Insights & Contribution

How to win the Wild West Web
Six tests for delivering the Online Harms Bill – a summary

In the coming weeks, the Government will decide on legislation that could finally protect children from online abuse. If it acts with urgency and ambition, it can secure an Online Harms Bill that delivers tough but proportionate regulation, and that sets a global standard.

But if the measures fall short, children will continue to face avoidable harm. One in five UK internet users[2] will face online abuse that continues to increase in both scale and complexity. The cost of industry inaction will continue to be felt by children, families and society.[3]

After a year in which they have faced unprecedented online risks, fuelled by the public health emergency but driven by the long-term failure of self-regulation, it couldn't be clearer that children deserve better than the status quo.

Earlier this year, the Prime Minister told participants at his Hidden Harms Summit he was determined to take tough but necessary action to hold social media companies to account. He heard the words of a mother whose 12-year-old daughter Freya was subject to online abuse: "Our children should be safe in their bedrooms, but they're not. They should be safe from messages from strangers if their accounts are on private, but they're not."

The NSPCC has led the campaign for a social media regulator – with companies subject to a legally enforceable Duty of Care that requires them to identify reasonably foreseeable risks, and address them through systemic changes to how their services are designed and run.

This report reaffirms the case for action – but it is clear the Government will only deliver on its ambition to make Britain the safest place in the world online[4] if it is bold and ambitious in its plans. If regulation is poorly designed, or the regulator isn't given the powers it needs, children will continue to face otherwise preventable harm.

Last year, in conjunction with Herbert Smith Freehills, the NSPCC published clear proposals for a regulatory model.[5] We now set out a series of tests that the Online Harms Bill must meet if it is to deliver for children, and against which the Government's commitments should be judged.

If it meets each of these tests, the result will be a highly effective regulatory regime, and a Duty of Care that gives children long overdue online protections.

Six tests for the Online Harms Bill

1. **An expansive, principles-based Duty of Care**

Statutory regulation must be tough but proportionate, and it should deliver the strongest possible protections from abuse for children. This means the Duty of Care must be realised through a principles-based approach which is broad, future proof and that applies expansively.

[2] Data from the Information Commissioner's Office
[3] The Center for Humane Technology maintains a Ledger of Harms that lists the 'negative impacts of social media that do not show up on the balance sheet of companies, but on the balance sheets of society.'
[4] Department for Digital, Culture, Media and Sport (2017) Internet Safety Strategy Green Paper. London: DCMS
[5] NSPCC (2019) Taming the Wild West Web. London: NSPCC

In the event that harm occurs, a platform would breach its Duty of Care if it failed to demonstrate sufficiently rigorous processes to identify or mitigate reasonably foreseeable harm, or if children had been put at material risk as a result of systemic failures that could reasonably have been addressed.

The Government must resist calls for a more prescriptive, and by implication less ambitious, approach. It is precisely because the Duty of Care requires platforms to assess the risks on their own sites, not just to follow a tick box set of remedies, that regulatory requirements will be hardwired into platform decision making – and that significant cultural change will be achieved.

2. Tackling online child abuse

The regulator must demonstrate an ambitious and determined focus on tackling online child abuse.

Ofcom will rightly be judged on how effectively it disrupts both online grooming and the production and distribution of child abuse images. It must prove capable of responding to constantly evolving abuse and highly agile threats.

Despite the understanding that tackling child abuse requires an emphasis on only illegal material, there are significant problems with abusive images that may not meet the criminal threshold, but which have significant potential to cause harm, signpost to illegal material, or re-victimise the children involved. More proactive processes to respond to such images will be required, which should include consistent takedown processes.

Platforms should have a duty to collaborate on child abuse risks, and should be subject to enhanced regulatory measures for high-risk design features that increase the risk of technology-facilitated abuse, including livestreaming, private messaging and end-to-end encryption.

3. Tackling legal but harmful content

In its interim response to the white paper, the Government set out differentiated expectations for illegal content, and that which is legal but causes harm.[6] This effectively requires platforms to only adopt clear policies on legal but harmful material, and enforce them effectively.

The regulator must adopt a child-centred and harm-based approach to legal but harmful content. Its regulatory approach must take decisions that are appropriately balanced against freedom of expression, but that respond to the very significant potential for harm that comes from platform mechanisms that promote or algorithmically suggest harmful content, including suicide, self-harm content and preparatory child abuse images. These clearly require an effective regulatory response, and the Government has positive obligations to protect children online.

Child users should receive protection that is proportionate to the likely harm caused. The Government must therefore ensure that any differentiated Duty of Care does not result in companies facing a perverse incentive to adopt weaker community standards, because in turn it will result in less onerous regulatory requirements.

In accordance with a risk -based approach, the regulator should signal its intention to apply enhanced regulatory scrutiny on content that is likely to be harmful to children.

[6] Department for Digital, Culture, Media and Sport (2019) Online Harms White Paper. London: DCMS

4. Transparency and investigation powers

Comprehensive transparency powers are crucial to the regulator's success. Unless Ofcom has robust investigatory and information disclosure powers, there will be a clear information asymmetry - and this could mean it is forced to take decisions on low quality evidence, or is less inclined to propose more ambitious measures.[7]

It is not enough to rely on industry transparency reporting. Such arrangements will only be beneficial if they provide significant and interrogable information, compared to existing approaches that are widely dismissed as a form of 'transparency theatre'.[8]

Platforms should face new information disclosure duties, including a requirement to proactively disclose to the regulator any information it could reasonably expect to be informed about, and to 'red flag' cases where failings could put children at risk. To embed a safety-by-design approach, sites should be required to undertake a risk assessment if they plan to introduce new services or amend their existing ones.

5. Criminal and financial sanctions

If the regulator is to effectively hold platforms to account, it requires comprehensive enforcement powers that ensure companies comply with the Duty of Care. Both platforms and senior managers must be liable to financial and criminal sanctions.

The powers available to the regulator must clearly correspond to the size and scope of the companies it regulates. We support GDPR equivalent fines, but for the largest companies the deterrence value of such fines is at best unclear. The Government must therefore commit to both corporate and senior management liability.

The Bill must introduce a Senior Managers scheme that imposes personal liability on directors whose actions consistently and significantly put children at risk. For the most serious of failings, the threat of personal prosecution should apply.

Industry groups have fiercely opposed personal liability, but the case for criminal sanctions in providing incentives to take action is compelling.[9]

6. User advocacy arrangements

As part of the regulatory settlement, it is essential there are effective arrangements in place for civil society to represent children's interests in regulatory debates. It will be necessary for civil society to support the regulator in understanding often complex child abuse risks; provide high-quality evidence of a sufficient regulatory threshold; and to demonstrate areas of concern or non-compliance.

[7] This is likely to be particularly apparent in respect of ex ante measures. Beverton-Palmer, M et al (2020) Online harms: bring in the auditors. London: Tony Blair Institute for Global Change

[8] Douek, E. (2020) The rise of content cartels: Urging transparency and accountability in industry-wide content removal decisions. New York City: Knight First Amendment Institute, Columbia University

[9] This is perhaps best expressed by Twitter's CEO Jack Dorsey. Asked why his payments company Square, of which he is also the CEO, seems to operate more smoothly he said: 'We had to get every single thing right. There's a lot of regulation around payments. If you do something wrong, you go to jail.' Comments made in a January 2019 interview with Rolling Stone https://www.rollingstone.com/culture/culture-features/twitter-ceo-jack-dorsey-rolling-stone-interview-782298/

Perhaps most crucially, the regulator is unlikely to deliver the strongest possible outcomes for children unless there is a strong civil society counterbalance to well-resourced industry interventions.

This is particularly important given that some companies might seek to frustrate or delay the regulator's work, and the heavily limited potential for children to exercise the redress options that the Online Harms White Paper proposes for adult users.

In order to create a 'level playing field' for child users, and secure the regulator's focus on child abuse risks, the Government should commit to statutory user advocacy arrangements for children, funded by the industry levy. This mirrors established user advocacy arrangements in many other regulated sectors, reflects the urgency of the child abuse threat, and responds to the inherent vulnerability of children as users of internet services.[10]

The urgency of Coronavirus

The importance of the Online Harms Bill has never been clearer than during the pandemic. The magnitude of child abuse risks vividly underlines why tech firms must finally be held accountable for the harm caused by their sites.

Lockdown created a perfect storm for online abuse. We don't yet know the true scale of online abuse during the pandemic, but we do know young people spent longer on platforms with fewer moderators.[11] We also know that offenders viewed Covid-19 as an opportunity to target often vulnerable and lonely children.[12]

No one could foresee the circumstances of the pandemic, but when the perfect storm rolled in, tech firms hadn't fixed the roof. The failure to design basic child protection into their services, and invest sufficiently in technology that could disrupt abuse, meant that social networks could be exploited ruthlessly.

But the pandemic also showed this could have been different. The same platforms that have dragged their heels on child abuse for many years responded impressively to the disinformation threat, rolling out design features in days that we were previously told might not be possible in months.[13] Platforms can act quickly and comprehensively when required.

[10] For example, Recital 38 of the General Data Protection Regulation states that 'children merit specific protection as they may be less aware of the risks, consequences and safeguards concerned and their rights in relation to [online services].'

[11] Digital, Culture, Media and Sport Select Committee (2020). Second Report of Session- Misinformation in the COVID-19 Infodemic. London House of Commons

[12] Europol (2019) Exploiting Isolation: Offenders and victims of child sexual abuse during the Covid-19 pandemic. The Hague: Europol

[13] Platforms rushed out new design features to frustrate the spread of misinformation, for example WhatsApp introduced new limits on the forwarding of user messages

Open Rights Group – Insights & Contribution

Briefing on the Online Harms Bill

for PICTFOR

Open Rights Group are grateful for the opportunity to provide our input for PICTFOR. Open Rights Group are a digital rights campaigning organisation. We seek to help build a society where rights to privacy and freedom of speech online are respected, protected and fulfilled. We have over 20,000 engaged supporters across the United Kingdom. We operate an evidence-based policy, guided by respect for fundamental human rights.

ORG has actively engaged with government on the Online Harms framework. Our previous submissions include:

1. Response to Consultation Paper, 2019
2. Internet Regulation, Part I and Part II, 2018-2019
3. Blocked: Collateral Damage in the War against Online Harms, 2019
4. DNS Security: Getting It Right, 2019

In addition to the issues we raised in those reports, we call PICTFOR's attention to the following concerns which have become clearer since the publication of the interim response to the Online Harms framework in February. We look forward to continuing to engage with government and PICTFOR on these issues.

Data and democracy

A shared point of concern for both ORG's work on Online Harms, as well as our work on data and democracy (adtech, privacy, and electoral interference), is the proposed extension of the Online Harms "duty of care…to preventing generic harm to our democracy". As with all aspects of Online Harms, it is unclear what exactly this means, as "harm" is extremely hard to determine when looking at content and speech rather than physical acts.

The invocation of democracy extends the remit of proposed Online Harms legislation to include the democratic character of our society and what is acceptable democratic discourse, or even what is true or untrue. Ofcom and Internet companies should not be drawing the line around 'truth' in the guise of reducing risks to individuals.

Ofcom's remit would reach into areas including relations with hostile states and active cyberwarfare, while outsourcing not a small amount of those frontline battles to everyday site administrators and managers under the "duty of care".

The potential for both abuse and mundane errors created by these regulatory and statutory obligations is, by our account, more likely to be a hindrance than a help to UK democracy. Democracy relies on trust that individuals can freely express their views without state or corporate interference. Placing governance of democratic norms into the hands of a state regulator and large corporations invites the belief that democracy is being 'managed', and errors will compound such a view. Government should not task Ofcom and private companies with regulating online democratic debate, nor should it compel private information society services to face a legal obligation to manage the limits of our democracy.

Privacy / Encryption

We strenuously object to proposals calling for private and personal messaging to fall within the scope of the Bill. Private communications must not be subject to government surveillance based on the assumption that those communicating privately are doing so to traffic in CSAM.

Some proposals include a call for the regulator (Ofcom) to create some sort of "licensing" system for the use of encryption, where they would have the power and authority to order a company to stop using encryption if that company could not prove it had met the highly subjective standards of the "duty of care".

It is clear that consumers are worried by the ability of companies to read private messages, to use them for commercial purposes, and also for such messages to leak. Given the vast and sensitive nature of private messaging, it makes absolute sense for companies to follow consumer demand and provide them with safe, encrypted messaging systems.

CSAM clearly needs to be tacked, and perpetrators brought to book. However, sacrificing general safety measures is not the way to achieve this. We are clear that *any intentional undermining* of encryption, even for legitimate purposes, weakens *everyone's* security online, and it must not be controlled, be made permissible, or revoked by a regulator.

Management Liability

We remain concerned about the provisions in the White Paper, as amplified by several media outlets, which call for criminal liability to be imposed on the managers and directors of companies which fail to achieve the standards of the "duty of care". While these provisions clearly target two or three specific high-profile individuals based in the United States, they will miss those targets badly and hit everyone else instead:

1. Management liability will create "collateral censorship", where site administrators feel they have no choice but to take down what may be *perfectly legal and harmless* content out of fear of personal arrest;
2. It will create 'free speech martyrs' whose arrests and criminal trials will, on principle, draw far more sympathy to their causes, and the infringing content which triggered their arrests, than government intended; and
3. It will render the UK a "no-go" zone for the professional experts whose talents are needed to tackle online harms *on and within the platforms where they most frequently occur,* and will also dissuade anyone, in their right mind, from starting an online business of any sort.
4. It sets a very poor global example. Arresting people for speech-related offences, and regulating domestic speech through foreign global companies, are tactics employed by governments with little respect for human rights. The UK should not give them cover by using similar excessive measures.

Online harms will not be tackled by turning the issue into a "winnable" series of *ad hominem* witch-hunts, arrests, and trials. Nor will they be tackled by criminalising everyday site administrators, managers, and moderators for the ways that members of the public misuse their tools and services, and to inevitably arresting and prosecuting some of those individuals because the law says someone must be prosecuted.

5Rights Foundation – Insights & Contribution

Building the Digital World that Young People Deserve

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.

Introduction

5Rights welcomes the Online Harms Bill to reflect the HMG commitment to making the UK the safest place in the world to go online. The digital world is not optional for young people. It is their access to information, play, connectivity, and education. What was urgent before the Covid-19 pandemic is now critical, with remote-schooling, lockdown, and social distancing measures all conspiring to increase young people's time spent online by five-fold[14].

All actors in the digital world's value chain have a responsibility to young people to identify the risks that the digital world poses and to act decisively and swiftly to mitigate them. Ultimately, the government's Bill will be judged not on what is included or excluded, but on the difference it makes to the lived experiences of young people. We owe it to our young people to build the digital world they deserve.

Priorities for the Online Harms Bill

The Online Harms Bill must:

- **Require services to identify risks and mitigate them swiftly and robustly at the design stage before services and products are distributed, via the use of Child Impact Assessments.** *Impact assessments are a common and established means of identifying the future consequences of a current or proposed action, for example, the requirement of a Data Protection Impact Assessment (DPIA) via the Data Protection Act and the statutory Age Appropriate Design Code.*
- **Introduce minimum standards for online services covering issues such as impact assessments, presentation of published terms, age assurance, take down mechanisms, moderation etc.**
    - *Other such measures that the regulator deems necessary must be included, and these standards should be systemic in nature.*
- **Mandate that those services that engage with young people provide an age appropriate service by default.**
- **Require digital services to provide clear, concise, and age-appropriate terms of services and other published rules that meet minimum regulatory standards.** *These published terms must be subject to oversight and enforcement, to ensure that they uphold children's rights in accordance to the UN Convention on the Rights of the Child.*
- **Give information powers to the regulator for algorithmic oversight, including identifying and assessing the data used to train the algorithm (and how it is 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.**

[14] Covid-19: Lockdown measures and children's screen time, *House of Lords Library*, June 2020.

  - *Algorithmic oversight is essential to ascertain the nature, presence, or responsibility for harms experienced by young people in the digital environment. These automated system technologies are embedded within most digital platforms that young people use.*
- **Require digital services to assess the age of users or provide a service suitable for mixed audiences that include young people.**
  - *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 that services are not unduly burdened if they pose little or no risk to young people.*
- **Create a public interest access to service's private data for research purposes**.
  - *This should also include access to algorithms and internal procedures but 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.*
- **Give sufficient enforcement powers, independence and resource to the regulator to make it an effective guardian of the sector.**

Types of Harms

There has been much debate about the nature of harm, dividing it into the categories of *illegal* and *legal but harmful*. This framing does not adequately reflect the reality of young people's lives, and the unique risks they face in the digital world. If the Online Harms Bill continues to allow routine appearances of pornography in remote learning platforms, teenagers to be targeted by cosmetic surgery[15] and diet content[16], child sexual abuse offenders and groomers to be recommended videos of pre-pubescent children[17], or services to recommend stranger adults as "friends"[18] or "followers" to young people, then the Online Harms Bill will have failed to fulfil its promise to make the UK the safest place to go online.

Yet all of these harms and practices happen across the UK, and all fall into the 'legal but harmful' category. Many of the risks of the digital world are not at the hands of bad actors but are the cumulative[19] and unintentional outcomes of digital service design. **The Online Harms Bill, with a duty of care at its heart, must ensure that the design of services that engage with young people have accounted for their age and vulnerabilities in advance and by default.**

Closing Remarks

The Online Harms Bill should not seek to identify and regulate individual pieces of content or conduct that might offend or be harmful to young people. What is imperative is that the Bill acknowledges and tackles the norm of commercial pushes that deliberately push behaviours and content that are not in children's best interests, including but not limited to persuasive design and aggressive datafication[20], through which millions of young people are recommended health misinformation or are able to stream live from their bedrooms to millions of strangers. The Online Harms Bill will never be robust and future-proof if it does not carve out the importance of regulating the automated systems employed by digital services and platforms as well as providing moderation, support and enforcement of illegal activity.

The young people that we work with are early adopters and enthusiastic users of digital services. They, and we, want to be part of a digital future that accepts their age and meets them with the generosity and care they deserve.

[15] 'Irresponsible' lip filler advert banned for encouraging young girls to undergo cosmetic procedure, *The Independent*, February 2019.

[16] Why is TikTok advertising dangerous content to teenage girls? *Rolling Stone*, July 2020.

[17] On YouTube's Digital Playground, an Open Gate for Pedophiles, *New York Times*, June 2019.

[18] How Facebook Made Those Eerie "People You May Know" Suggestions, *Slate,* December 2018.

[19] A more detailed illustration of how risks to young people accumulate on common platforms can be found in 5Rights project Risky by Design.

[20] "It's None of Their Business!" Children's Understanding of Privacy in the Platform Society, *Professor Sonia Livingstone OBE*, 2020, and Addressing the Needs of Children in the Digital Environment, *Elettra Ronchi, Andras Molnar, and Lisa Robinson*, 2020. Both available as part of *5Rights* publication, Freedom Security Privacy: The Future of Childhood in the Digital World.

A full outline of 5Rights priorities for the Online Harms Bill can be found here.

Which – Insights & Contribution

Introduction

As the importance of the digital world continues to grow in our everyday lives, it is vital that consumers can continue to make the most of the many opportunities this transformation presents. Fundamental to this will be ensuring that online platforms are as simple, fair and safe as possible for the people using them.

The services provided by online platforms that host user-generated, third party content are now an essential part of daily life for most people. Consumers benefit significantly from the convenience and connectedness they offer, and the considerable choice on their sites, that allows people to shop, socialise and work with greater ease than ever. This reliance on digital services has increased significantly as a result of the coronavirus pandemic, underlining the importance of online platforms and, in some cases, consolidating sites' significant market power.

It is vital therefore that the responsibilities of online platforms for protecting their users reflect the sites' unique position within the digital markets in which they operate. This is particularly the case as the rapid change in the way people now interact with products and services online has been accompanied by a rise in several different harms that people face when navigating the digital world.

Which? has identified two key areas where harm to *consumers* is being perpetrated through illegal content and activity on online platforms: risks to people's safety, and misleading information that results in worse outcomes for consumers. This harm can manifest in a variety of ways, including new online-specific threats, such as fake reviews, and the re-emergence of traditional consumer harms in a digital setting, such as online scams or the sale of unsafe products through online marketplaces. The lack of competition in some digital markets also exacerbates poor outcomes for consumers.

Key gaps in protecting consumers against harm online caused by illegal content and activity

The success of digital markets depends on regulatory frameworks keeping pace with technological developments and changing consumer behaviour. Currently, there are significant gaps that risk leaving growing numbers of people exposed to harm online. In some instances, such as online scams, the framework is not fit for the digital age, while in others, like the sale of unsafe products, online marketplaces' intermediary status exempts them from any meaningful responsibility.

Under the UK's current regulatory frameworks, the legal responsibilities of online platforms are extremely limited and do not adequately reflect the reality of the situation consumers face. Platforms are generally defined as intermediaries, meaning all that is required of them is to "expeditiously" take down illegal third party content on their sites when they are made aware of it - although no timeframe is given for this. Platforms do not currently therefore have the right incentives to protect their users.

In addition, Which? research strongly suggests there is a gap between consumers' expectations of the responsibilities of online platforms to protect their users and sites' actual legal responsibilities. Many people believe that platforms are legally obliged to do more to protect their users than is required of them by the current regulatory frameworks. This gap can result in consumers having misplaced trust and overconfidence that the sites are protecting them, and could lead to users taking fewer steps to protect themselves online. To address this, the responsibilities of platforms must be brought more in line with consumers' expectations.

To tackle harms comprehensively, online platforms must be part of the solution. To support with this, the legal responsibilities of online platforms should be clarified in relation to illegal content and activity that takes place on their sites involving misinformation and safety. Platforms should be required to introduce proactive measures that prevent dissemination of harmful content and reactive measures to address harmful content once it appears. This combination of measures has the potential to deliver improved outcomes if online platforms are required to continuously demonstrate that the systems and processes they implement result in a reduction in harm. There should also be greater transparency obligations so that consumers know who they are interacting with on the sites, whether content they see is paid-for or organic and if advertisers have been verified.

To ensure that changes to platform responsibility help to deliver a reduction in harm and continued benefit to consumers from trusted digital markets, regulators must be able to effectively enforce these changes. Regulators must be empowered to hold platforms and bad actors on them to account if they fail to keep their users safe, and be given the necessary resources, powers and tools to do so, while working closely with other regulators to ensure a coherent approach that is clear to both platforms and consumers.

Which?'s recommendations for addressing the gaps in online consumer protections

The sale of unsafe products through online marketplaces

There are currently significant gaps in consumer protection when purchases are made from online marketplaces. To address these, existing product safety legislation requires updating to better reflect the role that platforms are best placed to play to keep online shoppers safe. To ensure people are better protected when shopping on online marketplaces, Which? is calling for the following:

- Online marketplaces to have greater legal responsibility for ensuring that consumer products listed by sellers on their sites are safe. This should take the form of a due diligence defence.
- The actions that are required by online marketplaces when unsafe products are identified to be clarified, giving online marketplaces 24 hours to remove unsafe products and responsibility for overseeing recalls.
- Greater transparency obligations so that consumers are clear who they are buying from.
- Stronger regulatory oversight, including an independent product safety regulator and enforcement officers equipped with the appropriate powers and resources to police online marketplaces and the supply networks that underpin them.

Online scams

Online platforms currently have limited responsibility and no legal obligation to protect users against fraudulent and scam content. Although the e-Commerce Directive requires sites to remove illegal content 'expeditiously' when they are made aware of it, many people can still be exposed to fraudulent content and fall victim to scams before a platform has acted to remove the content.

Unless online platforms are required to actively tackle this problem, the harm from scams will continue to grow with devastating consequences for people and, in particular, vulnerable consumers. Online platforms should be given greater legal responsibility for preventing scam content from appearing on their sites, including fake and fraudulent adverts, as well as more responsibility for taking quick action to remove harmful content when it is reported. Platforms should also be required to introduce clearer labelling of paid-for ads and whether advertisers have been verified.

Which? believes the government has a perfect opportunity to deliver these measures to tackle online scams in the upcoming online harms bill, and if not ministers must set out their proposals for further legislative action to effectively protect consumers from online scams.

Fake reviews

The Competition and Markets Authority is investigating how fake reviews are being used to manipulate online shoppers on major websites, and we expect the regulator to take strong action against sites that are failing to protect users on their platforms. However there are gaps in the current regulatory framework that may serve to discourage platforms from taking the necessary steps to crackdown on misinformation on their sites and take strong action against abuses of their policies.

Appendix - links to relevant research and wider reading

Connecting the world to fraudsters? - Which? research into consumer attitudes and behaviour towards scam content on social media sites (October 2020)

The real impact of fake reviews - a behavioural experiment on how fake reviews influence consumer choices (May 2020)

Online marketplaces and product safety - Which? policy paper and consumer insight on online marketplaces and product safety (November 2019)

66% of products tested from online marketplaces failed safety tests - Which? data on consumer organisation safety testing of products purchased on online marketplaces (February 2020)

Which? investigation into scams on social media (April 2020)

Which? investigation into scam adverts on social media platforms and search engines (July 2020)

About Which?

Which? is the UK's consumer champion. As an organisation we're not for profit - a powerful force for good, here to make life simpler, fairer and safer for everyone. We're the independent consumer voice that provides impartial advice, investigates, holds businesses to account and works with policymakers to make change happen. We fund our work mainly through member subscriptions. We're not influenced by third parties – we never take advertising and we buy all the products that we test.

Conclusion

A full video recording of PICTFOR's Online Harms roundtable will made available on the group's website along with minutes of the meeting.

PICTFOR have been proud to facilitate the discussion between the tech sector and Westminster and will endeavour to further this discussion in the coming weeks and months through virtual events and further parliamentary activity.

If you would like to find out more about PICTFOR, this report or would like to attend one of our events, please visit our website or email the secretariat on admin@pictfor.org.uk

This is not an official publication of the House of Commons or the House of Lords. It has not been approved by either House or its committees. All-Party Parliamentary Groups are informal groups of Members of both Houses with a common interest in particular issues. The views expressed in this report are those of the group.