Response to 'Online harms' consultation

01 July 2019

Contents

Consultation response by RPC (Reynolds Porter Chamberlain LLP)

This consultation response is submitted on behalf of RPC's Tech Group[1]. RPC is an international law firm with offices in London, Bristol, Hong Kong and Singapore. RPC acts for a range of clients from start-ups to multi-nationals in the technology, media, and retail sectors and for their insurers. RPC's Tech Group provides specialist advice on regulation, content liability, commercial contracts, outsourcing, data protection, cyber, intellectual property, e-commerce, and investigations and disputes. Views expressed in this submission do not necessarily represent those of our clients.

Nicola Cain
Partner, RPC
01 July 2019

[1] With special thanks to Marlon Cohen, Oliver Murphy, Victoria Noto, Anna Greco, Rachael Ellis, and Charlie Gould at RPC for their assistance in conducting research and preparing this response.

Executive Summary

RPC's Tech Group welcomes the opportunity to engage with Government in relation to the proposals to regulate online harms as set out in the Online Harms White Paper[2], and endorses the Government's stated aim of ensuring that the UK develops and maintains a *"vibrant technology sector"* and its commitment to a *"free, open and secure internet"* and to the protection of *"freedom of expression online"*.

We agree that the law does and should extend to the internet and that internet users and platforms are obliged to comply with the law and to act responsibly and with due regard for one another, as would be expected in any public space. Indeed, the largest online service providers already take a multitude of steps to protect users and third parties from harm on their services, from imposing standards on users to proactively identifying certain types of content and removing other content upon notification. However, that does not justify treating the internet as a space which requires more stringent regulation than any other, or for speech or conduct which would not be unlawful in a public place to be treated as if it were unlawful when it takes place online, or for the operator of the space to be held liable for the enforcement and censorship of that lawful speech or conduct which is nevertheless deemed undesirable in that space by the government at any particular time.

We are concerned that at a time when the Government is seeking to promote the importance of free speech and the free exchange of ideas around the world, including those which question authority, its efforts to mitigate online harms will inadvertently provide a veil of respectability to undemocratic regimes which would seek to stifle and censor lawful content.

While we address the specific questions posed by the consultation at the conclusion of this response, we first address a number of issues raised by the proposals in the White Paper which we believe ought to be given further consideration before any policy response is finalised and implemented. In summary:

- We would support the commissioning of further analysis and research as to the prevalence and scale of the risk posed by online harms, in order that any policy response may be targeted and proportionate;
- We would similarly support the introduction of mechanisms for monitoring the impact of any policy initiative to tackle online harms to enable further analysis and research including as to whether any detrimental impact outweighs their benefit;
- The Government should refrain from imposing expectations or obligations on online service providers in the absence of clear evidence demonstrating that this is necessary and proportionate;
- While we note the Government's concern that the current approach to regulation of the online space is fragmented, there are numerous proposals currently in the process of being implemented or consulted upon which are relevant to tackling the issue of online harms and in the absence of central co-ordination there is a real risk of developing conflicting policies, over-burdening industry and losing the opportunity to understand the impact of any changes;
- The proposed breadth of application of the proposals should be narrowed to focus on providers which are dedicated to the sharing of user-generated content, to the exclusion of media websites, retail websites (including online marketplaces), and search engines amongst others;
- We would invite the Government to consider whether smaller, not for profit or low risk online service providers should fall within the scope of the proposals;
- We are extremely concerned by the inclusion of not only unlawful content within the scope of the proposals but also content which is lawful but considered undesirable, particularly in circumstances where judgements as to

[2] Online Harms White Paper, CP 57, April 2019: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/793360/Online_Harms_White_Paper.pdf

what is appropriate will be devolved to a regulator or to private entities in relation to what is essentially a moral or ethical issue on which there may not be any consensus within society. We consider that such an approach would be detrimental to the UK's stated commitment to promoting free speech around the world and contrary to the Article 10 right to freedom of expression and information. If Government wishes to regulate online speech, it should seek to legislate to do so, making the relevant speech unlawful, rather than seeking to impose an obligation on private actors to identify, adjudicate upon and remove lawful material. We note that the Law Commission intends to consult on amendments to communications law offences in 2020;

- Any extension of obligations to remove content to encapsulate lawful speech must be subjected to the utmost scrutiny, and in practice online service providers must be provided with a margin of discretion in relation to the removal of such content;
- We do not consider that online service providers should be responsible for determining the veracity of content or that material posted from prisons should fall within the scope of the proposals;
- We would invite the Government to consider and consult upon whether the criminalisation of harmful conduct by individuals could provide a proportionate alternate response to the issue of online harm;
- While we reject the proposal for the imposition of a 'duty of care', which has a specific legal meaning and implications, we would welcome an expectation that online service providers act responsibly toward their users and others;
- If a duty of care were to be imposed, it is imperative that this would not compromise the application of the safe harbour provisions in the E-Commerce Directive, which would exclude a pre-moderation approach to harmful content; this is also unlikely to be technologically possible through the use of AI in relation to the broad range of harms identified in any event;
- We would not support the inclusion of an obligation applicable to online service providers globally to support UK law enforcement investigations as part of any duty of care, which has the potential to place online service providers in conflict with local law obligations;
- We welcome the Government's proposals for transparency in relation to the prevalence of harmful content but would encourage the publication of a wider range of information, such as the identity of those seeking the removal of content;
- In relation to transparency in relation to algorithms, we would encourage accountability rather than transparency to be overseen by the Centre for Data Ethics and Innovation;
- We would encourage a system of regulation which aims to centralise resource, improve digital literacy, disseminate best practice and avoid duplication of regulation;
- We would invite the Government to consider and consult upon whether a system of self-regulation, which already works effectively in relation to the advertising and press industries, could provide a proportionate response at no cost to the taxpayer;
- Any regulator which is required to oversee and adjudicate in relation to lawful but undesirable speech must be well-versed in regulating freedom of expression issues;
- We do not support the imposition of civil liability for breaches of the proposed statutory duty; and,
- We do not believe that the proposed enforcement tool of imposing fines on individual directors will be effective and anticipate that it will be a significant disincentive to investment in the UK, risking the UK's industrial strategy of leading the global technological revolution.

Evidence base

In implementing and overseeing the new regulatory framework, it is stated that the new regulator is intended to *"take a risk-based approach, prioritising action to tackle activity or content where there is the greatest evidence or threat of harm"*. In order to do so, it logically follows that there must already be a catalogue of evidence upon which to judge those risks.

It is apparent, however, that it is envisaged that the regulator is expected to develop and enforce codes of practice prior to that evidence base being collated, with proposals throughout the White Paper for the regulator to:

- *"work closely with UK Research and Innovation (UKRI) and other partners to improve the evidence base"*[3];
- *"use evidence of the actual incidence of harms on different services and the safety track record of different companies to prioritise its resources"*[4];
- *"use its powers to conduct thematic reviews, undertake targeted horizon scanning and investigate specific issues to develop its understanding of the risk landscape"*[5];
- *"run a regular programme of user consultation, in-depth research projects"*[6]; and,
- *"work with companies to ensure that academics have access to company data to undertake research, subject to suitable safeguards"*[7].

This is recognised in the White Paper, which acknowledges that it seeks *"to engage with the widest possible audience on our proposals…from industry, civil society, think tanks, campaigners and representatives"*, and requesting *"not just their opinions, but also the supporting facts and reasoning to inform the evidence base for the development of our final proposals"*[8].

This is necessary because in its response[9] to the White Paper's predecessor, the Internet Safety Strategy Green Paper[10], the Government recognised that *"research has shown significant gaps in existing evidence, not least because online harms can change rapidly, and many key trends are too new to have been the subject of longitudinal studies. Our upcoming programme of work on internet safety will include undertaking new research, on which we will be working closely with UK Research and Innovation"*. It is not clear what research the Government has commissioned and reviewed between May 2018 and April 2019 and how this has contributed to the current proposals. We suggest that the Government's review should include bot UK-based evidence and evidence from countries in which similar legislation has been enacted, in terms of its effectiveness and consequences (both positive and negative).

It is a matter of concern, however, that the White Paper not only makes proposals, apparently in the absence of such evidence, but indicates that the Government expects measures to be taken immediately to address the issues raised and not to wait until a regulator is appointed or has drawn up codes of conduct. This not only risks over-zealous measures being taken, in a manner which is not being monitored or overseen, but also presents a latent threat that if inadequate measures are taken this will result in the introduction of more stringent regulation.

[3] White Paper, p. 9
[4] White Paper, p. 54
[5] Ibid.
[6] White Paper, p. 56
[7] Ibid.
[8] White Paper, p. 95
[9] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/708873/Government_Response_to_the_Internet_Safety_Strategy_Green_Paper_-_Final.pdf
[10] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/650949/Internet_Safety_Strategy_green_paper.pdf

We note, for example, that recent research conducted by researchers at the University of Oxford[11], testing hundreds of thousands of datasets related to adolescent well-being and digital technology use, found that *"in one dataset, for example, the negative effect of wearing glasses on adolescent well-being is significantly higher than that of social media use"*.

While it is right to take steps to understand the background to isolated but high profile incidents, it is important to ensure that these do not form the basis for a reactionary response which is neither based on evidence nor the risk of harm and has the potential to cause far greater harm to society.

Regulatory action and enforcement must not be based on a mere perception of harm. There must be concrete evidence regarding each harm's prevalence, cause, nature and impact before attempting to assess risk and to formulate an appropriate strategy for handling such content. A failure to identify these elements could create an environment of over-regulation in which lawful content is removed and freedom of speech is unnecessarily impinged.

We welcome the Government's commitment to ensuring that the regulator takes an evidence-based approach to regulatory activity and suggest that the regulator's first task should be to audit the evidence base and perform a gap analysis before commissioning further research by trusted independent institutions. Accordingly, we would recommend that the Government refrain from raising expectations upon online service providers, further developing policy or having the regulator develop the proposed statutory codes of practice unless and until research has been conducted which confirms that the policy is a necessary and proportionate response to the issue, which balances the risk of harm to individuals against the rights of society as a whole.

[11] 'Social media's enduring effect on adolescent life satisfaction': https://www.pnas.org/content/116/21/10226

Regulatory environment

The White Paper raises the concern that there currently exists *"a fragmented regulatory environment which is insufficient to meet the full breadth of the challenges we face"*[12] online. Examples given of such supposedly fragmented regulation include: GDPR and the Data Protection Act enforced by the ICO; the Electoral Commission's oversight of the activity of political parties, and other campaigners, including activity on social media; forthcoming age verification requirements for online pornography; the Equality and Human Rights Commission's oversight of the Equality Act 2010 and Freedom of Expression; Ofcom's existing oversight of video-on-demand services; the revised EU Audiovisual Media Services Directive, which will introduce new high-level requirements for video sharing platforms such as YouTube; the Gambling Commission's licensing and regulation of online gambling; and, the Competition and Markets Authority's (CMA) enforcement of consumer protection law online. The E-Commerce Directive also regulates liability imposed on online intermediaries.

A number of these regulatory interventions continue to be in development or have not yet been implemented. There is a risk of numerous regulatory initiatives being developed in isolation and without proper regard for how they will interact. For example, a few days after the launch of the White Paper, the Information Commissioner launched her consultation on *"16 standards that online services must meet to protect children's privacy"*[13], focusing on age appropriate design for online services, whether or not they are specifically directed at children. This includes a proposed obligation on online services which are likely to be accessed by children (regardless of the nature of the content of the website, or to whom it is targeted, and therefore applying equally to a sports site or website providing educational resources for children) to comply with the code unless robust age-verification mechanisms are in place. This encourages the age gating of content on the internet and the collection of additional personal data concerning individuals, reducing individual privacy of the majority by way of effectively surveilling the use of the internet and placing large databases of private information in the hands of a small number of companies providing age verification services. The practical difficulties in such an approach are evidenced by the numerous delays to date to the requirement that online pornography be age gated as required by the Digital Economy Act 2017.

The online environment is not solely subject to regulation, but is equally subject to criminal law in the same way as conduct which occurs offline as well as to online specific offences. These include offences under the Malicious Communication Act 1988, the Communications Act 2003, Protection From Harassment Act 1997, Protection of Children Act 1978, Public Order Act 1986, Racial and Religious Hatred Act 2006, and the Criminal Justice and Courts Act 2015, as well as substantive offences and offences at common law.

In seeking to implement numerous wide-reaching proposals over a short space of time supported by a limited evidence base creates the risk of conflicting and burdensome requirements to be implemented within a short space of time, the impact of which cannot be measured or understood. Any consideration of the regime governing online conduct must take into account both existing criminal law and regulatory schemes and proposals must be developed in tandem and implemented over a time period which not only allows for online service providers to prepare properly for their introduction but also allows their impact to be analysed.

Notably, it is not suggested that the proposed regulation of online harms will replace current or future regulation (and nor do we consider that realistic). Instead, the proposals will add to the regulatory frameworks and regulations, which will become even more complex when the UK seeks to impose domestic regulation on international services.

[12] White Paper, p55
[13] https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2019/04/ico-launches-consultation-on-code-of-practice-to-help-protect-children-online/

Entities falling within the scope of the proposals

The White Paper states that it is intended that entities falling within the scope of the proposals will be *"companies that allow users to share or discover user-generated content or interact with each other online",* applicable to both public and private channels albeit with a separate regime for private channels; this is regardless of the size of the company or the nature of its activities.

This would appear to encapsulate social networking sites, messaging services, search engines, content sharing sites, dating sites, messaging boards and blogs, media organisations providing comment functionality, digital lockers, online marketplaces and retailers offering review functionality.

We welcome the indication given by the Culture Secretary, following an approach by the Society of Editors, that *"the press and media as a whole are not the target of any new regulation in this area"* and will *"not affect press freedom"* or *"duplicate efforts"* in relation to user comments on media sites. We would invite the Government to explicitly exclude such sites from the scope of the proposals.

The potential application of the proposals to search engines, which do not host the originating content and have no direct relationship with the poster and are therefore a step removed from hosts of online content, presents particularly acute difficulties given the volume of material that would potentially be captured and the dangers in essentially removing catalogue cards from the digital equivalent of the Dewey Decimal system. We would suggest that these online service providers should also be excluded from the remit of the proposals.

While it is proposed that the regulatory framework would provide *"clarity for the wide range of businesses of all sizes that are in scope of the new regulatory framework but whose services present much lower risks of harm, helping them to understand and fulfil their obligations in a proportionate manner"*, we are concerned at the red-tape this will introduce in circumstances where there is a lack of evidence that such services propose any real risk. For example, we are also concerned at the potential for retailers to fall within the scope of the proposals. User reviews on retail websites do not appear to be considered to present a real risk of online harm, and we would suggest should therefore be clearly excluded from the scope of the proposals. We also anticipate that such a broad application of the regulatory regime would also result in issues of enforceability which could undermine confidence in the regime.

It is not clear at present how it is envisaged that the framework would apply to smaller or low risk online service providers, or how principles or reasonableness and proportionality will be applied, and this will create uncertainty for business. A regime which has limited applicability only to larger for-profit sites would be worthy of consideration.

In reality, any obligation imposed on smaller online service providers without Government taking a central role in assisting such companies to achieve compliance, for example by making online tools available, providing template documentation etc is likely to prove futile. We note that in other jurisdictions, such as in Germany, the equivalent legislation[14] is only applicable to for profit social media platforms with a minimum of 2 million users and that media and messaging companies are explicitly exempt. In Australia, the Office of the Children's eSafety Commissioner manages a centralised cyber bullying complaint scheme and abuse image portal[15]. Any such centralisation and pooling of resource would also reduce barriers to entry which could otherwise have an anti-competitive impact on tech start-ups and is to be encouraged.

[14] The Network Enforcement Act (or "NetzDG") was passed in June 2017 and came into effect on 1 January 2018
[15] The Enhanced Online Safety Act came into effect in July 2017

Content and activities falling within the scope of the proposals

It is proposed that not only unlawful content but also content that is deemed to be "harmful" including content which "*threatens our way of life in the UK, either by undermining national security, or by reducing trust and undermining our shared rights, responsibilities and opportunities to foster integration*" ought to be the subject of regulation and, as a consequence, censorship.

Examples provided of harmful content which is not unlawful but which the White Paper suggests should nevertheless be the subject of regulation includes cyberbullying and trolling, extremist content and activity, coercive behaviour, intimidation, disinformation, violent content, advocacy of self-harm and the promotion of Female Genital Mutilation.

Taking the example of violent content, it is not clear what type of violent content it is considered ought to be removed. Is it intended that this would include the content of a trailer for an action film that might be circulated online, or UGC footage of an amateur boxing match, for example? This typifies the difficulties inherent in seeking to regulate lawful content which is nevertheless considered undesirable. Nor is it clear how concepts such as cyberbullying and trolling would be defined and applied and when it is anticipated that a joke or teasing will cross into unacceptable conduct. The lack of certainty inherent in seeking to regulate these undefined concepts is contrary to the rule of law. For example, would robust and lawful criticism and ridicule of a politician, which would be considered to be high-value protected speech, be considered content which that individual is entitled to have removed?

It is also not clear whether it is intended that what is harmful in such contexts is to be judged objectively or subjectively and whether the harmful nature of content must be discernable on the face of the content itself. Content that would obviously be deemed harmful if it were to appear on an educational website targeted at children may not be harmful if it is made available in the different context of a gated website targeted at consenting adults. Nor does the White Paper specify whether it is anticipated that online service providers will be entitled or expected to engage with the original posters of content in order to notify them of the fact of receipt of a complaint and ascertain their position as to the complaint.

It is said to be deliberate that no definition is provided of what constitutes harmful content, or what the characteristics of harmful content are in order to ensure that the principles can be applied as deemed necessary: *"this list is by design, neither exhaustive nor fixed. A static list could prevent swift regulatory action to address new forms of online harm, new technologies, content and new online activities"*. Such an approach undermines the UK's vigorously defended tradition of free speech, reflecting values of tolerance and pluralism. This right is protected in law by Article 10 of the European Convention on Human Rights which extends not merely to information and ideas that are favourably received or are regarded as inoffensive but at its most potent protects those that offend, shock or disturb the State or any sector of the population[16].

While intended to provide flexibility, the proposal to permit the regulator to develop its own rules around online speech without the scrutiny of Parliament or within specific confines of legislation would constitute an inappropriate delegation of power to a government regulator which would be less readily scrutinised.

It would place private entities in the invidious position of being expected to adjudicate on the propriety of lawful content, faced with the prospect of enforcement action if they fail to take steps to remove material, and will therefore encourage a conservative approach which will result in an even broader range of material being the subject of restriction. This will inevitably have a chilling effect on freedom of speech. It is not apparent from the White Paper whether it is envisaged that a margin of discretion would be afforded to online service providers, i.e.

[16] Handyside v. the United Kingdom judgment of 7 December 1976, § 49

whether they would only be held to have failed to comply with their obligations if they could not reasonably have believed that the relevant content was in accordance with the relevant code.

At a time when tyrannical and undemocratic regimes would seek to censor free speech, in all public spaces including online, and the Government has re-emphasised its commitment to securing media freedom around the world, any attempt to restrict and regulate lawful speech is liable to have disastrous consequences, setting a precedent which would permit such regimes to implement censorship while hiding behind a veil of respectability by reference to the UK's approach. We are concerned that this will undermine the Government's endeavour to achieve *"global approaches for online safety that support our democratic values, and promote a free, open and secure internet"*.

We refer, for example, to Vietnam's Law on Cybersecurity which was passed in 2018 and came into force at the beginning of this year. In the name of preventing and dealing with infringements of cybersecurity, this prohibits the publication online of embarrassing information, information which distorts or defames the people's administrative authorities, and information which is insulting to famous people or leaders, inter alia[17], and permits the paralysation or restriction of cyberspace where deemed warranted in the interests of national security or particularly serious harm to social order and safety.

We consider that if Government wishes to regulate online speech, it should seek to legislate to do so, making the relevant speech unlawful rather than seeking to impose an obligation on private actors to identify, adjudicate upon and remove lawful material. We note that in other jurisdictions, such as Germany, legislation in this field is restricted to unlawful content and we would endorse a similar approach. We acknowledge that the Law Commission has now been asked to undertake phase 2 of the Abusive and Offensive Online Communications project to make recommendations for reform of communications offences and whether the criminal law can better address online harassment by groups of people[18]. We understand that a consultation paper on wider issues is intended to be published in 2020 and we consider that any requirement to impose obligations in relation to lawful content should await the outcome of that consultation and any ensuing legislation.

In relation to the proposed regulation of disinformation, the White Paper presents contradictory information, stating that while the focus shall be *"on protecting users from harm, not judging what is true or not"*, which would clearly be an inappropriate obligation to impose on private companies who may not have the knowledge or expertise to determine the veracity of content, it is also acknowledged that *"there will be difficult judgement calls associated with this"*, which suggests that online service providers may be expected to seek to analyse content. We do not consider that any obligation on online service providers to determine the veracity of content is appropriate.

A further example of *"harmful"* content identified in the White Paper is material uploaded from prison. It is clearly not the nature of the content itself, which could be entirely anodyne, but the location from which it is published which purports to constitute a harm. While we appreciate that victims of crime and their families would be outraged to see such content, we do not consider that it is appropriate to seek to impose an obligation to remove evidence from the public domain of the failure to effectively enforce prison rules and that it would be more appropriate for the Government to seek to eradicate connected devices from prisons.

We agree that dual regulation should be avoided and that harms arising as a result of breaches of data protection legislation ought to be excluded from the scope of these proposals.

[17] Article 16, Law on Cybersecurity
[18] https://www.lawcom.gov.uk/law-commission-to-undertake-phase-2-of-the-abusive-and-offensive-online-communications-project/

Proposed duty of care

It is not only proposed that certain types of content ought to be the subject of removal, but that the mechanism for enforcing this should be for online service providers to be subjected to a statutory obligation to take *"reasonable steps to keep users safe and prevent other persons coming to harm as a direct consequence of activity on their services"*, which is described as a *"duty of care"*. It is stated that *"All companies in scope of the regulatory framework will need to be able to show that they are fulfilling their duty of care"*.

We find the concept of the application of a duty of care to the online environment conceptually difficult. The concept of a duty of care in this context was proposed by William Perrin and Professor Lorna Woods[19], who seek to equate the online sphere and the relationship between online service providers and their users with that of the occupier of a building toward visitors (and, to a lesser degree, trespassers) or to that between employers and employees in the context of safety at work. These create an obligation to provide a reasonably safe environment over which the entity has control. With respect, we do not accept that these are analogous situations.

Unlike what is proposed for the online environment, the provision of a reasonably safe environment for employees and visitors would not generally require steps to be taken to prevent malicious actors from causing harm to each other. Nor would liability be imposed for a concept as nebulous as 'harm' which does not necessarily reach the threshold of causing a recognised psychiatric injury, whether or not that standard is applied objectively or subjectively (which is not clear from the White Paper). Furthermore, the duty only applies to an easily identifiable category of individuals, i.e. employees and visitors, whereas the proposed online duty of care would apply not only to users but also to third parties. For these reasons we consider that the application of the recognised legal concept of a duty of care to the online environment is unworkable and inappropriate.

Even if such a duty could be formulated in a manner which could provide online service providers with sufficient certainty, and notwithstanding the commitment in the White Paper that *"The new regulatory framework will increase the responsibility of online services in a way that is compatible with the EU's e-Commerce Directive, which limits their liability for illegal content until they have knowledge of its existence, and have failed to remove it from their services in good time"*, it is difficult to envisage how the practical steps necessary to comply with such a duty would not effectively deprive online service providers of the so-called safe harbour defences. This would place the most responsible online service providers in the invidious position of risking losing their statutory defences and being exposed to civil liability or alternatively to face sanction by the regulator. Again, we note that in Germany the obligations imposed on online service providers do not include proactive steps to be taken, thus preserving the ability to rely on the E-Commerce Directive.

For these reasons, while we would welcome an expectation that online service providers should owe a responsibility to users, we would not endorse the imposition of a statutory duty of care.

We contrast the approach proposed by the Government with that adopted in New Zealand which, rather than seeking to impose liability on online service providers (other than where they are on notice of harmful online content), instead criminalises harmful conduct by individuals; a person will commit a criminal offence if (a) the person posts a digital communication with the intention that it cause harm to a victim; and (b) posting the communication would cause harm to an ordinary reasonable person in the position of the victim; and (c) posting the communication causes harm to the victim. We consider that this presents a credible alternative approach, and one which is itself worthy of consultation.

Although identified in the context of the regulator's powers to require transparency reporting from online service providers, we note that the Government identifies the use of the *"Proactive use of technological tools, where appropriate, to identify, flag, block or remove illegal or harmful content"*. While this may be appropriate and feasible in relation to certain types of extremist content, for example using the tools developed as part of the ASI Data

[19] https://www.carnegieuktrust.org.uk/blog/reducing-harm-social-media-duty-care/

Science project, or those to tackle child exploitation material, the application of such automated upload filters is less relevant to the broader concepts of harmful content and would risk producing significant false positives which would need to be reviewed and released by human moderators. Given the uncertain scope of what will be considered to be harmful content, it is not clear how automatic filters could be deployed effectively. It is difficult to see how this would be a proportionate approach and would operate as an impermissible restraint on free speech. Even with post-moderation activities there is a risk of malicious reporting being used to silence legitimate speech.

We note the proposal that the duty of care will include an obligation to *"Support law enforcement investigations to bring criminals who break the law online to justice"*. While we are aware that many online service providers, regardless of whether or not they are based in the jurisdiction, endeavour to co-operate with law enforcement investigations and prosecution where they are permitted to do so by the local law under which they operate, we are concerned that a requirement to support law enforcement investigations as part of the duty of care would serve to circumvent jurisdictional issues and would place such providers at risk of being required to break local laws or face sanctions. This would clearly be unacceptable.

Transparency

The White Paper proposes that transparency and reporting obligations would be imposed on online service providers requiring that they outline the prevalence of harmful content on their platforms and what countermeasures they have put in place. The largest online service providers already take steps to demonstrate transparency in their removal of content, as the White Paper acknowledges.

Given the vague concept of what constitutes harmful content, we anticipate that this could only be fulfilled by reference to the number of complaints received, and the number of items of content removed. We would also welcome transparency reporting on the identities of those seeking the removal of content. For example, in the context of its reporting on the removal of content pursuant to the right to be forgotten[20], Google identifies the categories of content requested for de-listing, and the category of individual requesting de-listing, i.e. whether they are a minor, private individual or a government official or politician. We consider that the publication of such information will be important in enabling the analysis and review of the impact of any legislation in this area and would therefore suggest that this best practice approach form the basis for the Government's proposals.

It is not clear what the Government envisages when it proposes that online service providers should report on *"emerging"* harms as well as *"known"* harms, and appears to suggest that these private companies should themselves be identifying further material outside of categories already identified of lawful content which ought to be prohibited from the online space. This would further devolve responsibility for setting the parameters of acceptable speech online and would be contrary to Article 10 of the European Convention on Human Rights.

While we acknowledge the benefit that could be derived from requiring the publication of transparency reporting in relation to the enforcement of terms and conditions, given the variety of approaches that may be taken by online service providers, we would not expect that such statistics would provide any useful basis for comparison as between services.

It is also proposed that the regulator could require the publication of information on the *"impact of algorithms in selecting content for users"*. Taking the Google search engine, for example, Google already publishes information about how its search works[21]. Ranking systems are proprietary, complex and are comprised of multiple algorithms and often use artificial intelligence to continually improve as well as human interventions, so these are not entirely static processes. Factors or signals that will be relevant to an individual's search results may include (dependent on their privacy settings) a user's location, the specific search terms entered, the relevance of pages to those terms, previous account activity, and whether the search appears to relate to an ongoing event, such as a sport score or airline arrivals information. In any given case however, these may be weighted differently depending on the combination of factors being taken into account. Meaningful transparency in relation to these issues is therefore difficult to achieve and runs the risk of being abused by malicious actors who would seek to manipulate the system.

While we would welcome any steps to improve public understanding of algorithms and artificial intelligence as part of wider digital literacy initiatives, it is important to recognise that transparency does not necessarily lead to accountability. We would therefore advocate an approach developed in conjunction with the Centre for Data Ethics and Innovation to promote best practice in the development and testing of algorithms to encourage effective accountability.

[20] https://transparencyreport.google.com/?hl=en_GB
[21] https://www.google.com/intl/en_uk/search/howsearchworks/algorithms/

Enforcement and redress

It is proposed that an independent regulator will be established to enforce the statutory duty of care, and develop codes of practice, inter alia.

It is perhaps surprising that the government has turned immediately to propose the imposition of a new independent regulator, rather than considering whether self-regulation may be appropriate in this context. We note, for example, the success of the Advertising Standards Association, or the Independent Press Standards Organisation, which are free of charge to the taxpayer and yet maintain the ability to impose fines and other sanctions (either directly or by reference to a government regulator) in the event of non-compliance. We consider that the viability of such an approach ought to be given due consideration and consulted upon before an expensive exercise of creating a new statutory regulator is embarked upon. In this scenario, the role of Government would be to work with the self-regulator to improve education and awareness and to work with industry to harness and exploit technological developments.

The largest online service providers, who would be best placed to comply with the proposed regime, already take steps to facilitate transparency and complaints, for example, and impose terms and conditions and standards of conduct upon their users for the benefit of all. These measures should be encouraged to enable best practice to be harnessed and shared throughout the industry.

In any event, we consider that any proposed regulator must be well-versed in regulating freedom of expression issues, which will be imperative given the range of lawful content which could fall within the scope of the duty.

It is proposed that the regulator's powers will not only include the issuing of fines against entities but will also extend to imposing fines on individual directors and the power to take measures to disrupt and ultimately block the activities of infringing online service providers.

While we would not anticipate that it would arise in relation to the vast majority of online service providers, it is not clear how the imposition of fines on individual directors would work in practice in circumstances where the individual resides and conducts business outside the jurisdiction; the risk of such fines would make it advisable for entities and their directors to operate outside the jurisdiction which would jeopardise the UK's stated aim of building a *"vibrant technology sector"*.

We also note that it is suggested that *“if the regulator has found a breach of the statutory duty of care, that decision and the evidence that has led to it will be available to the individual to use in any private action”*. In some cases, this would contradict the Supreme Court's finding in the case of Rhodes v OPO [2015] UKSC 32 that *"it is difficult to envisage any circumstances in which speech which is not deceptive, threatening or possibly abusive, could give rise to liability in tort for wilful infringement of another’s right to personal safety"*.

Even in relation to speech which is deceptive, threatening or abusive, there is no suggestion in the White Paper that the breach of the duty must have led to actual harm surpassing a minimum threshold in order to create an entitlement to compensation. The potential for such a right to be misused is significant; given the range and volume of content created every day, to hold online service providers liable not merely for fines and other sanctions but potentially also to civil liability and the costs associated with that would impose a significant burden on online service providers and on the courts. Any civil liability which is to be imposed must therefore be reserved for instances of significant actual harm.

Responses to consultation questions

Question 1: This government has committed to annual transparency reporting. Beyond the measures set out in this White Paper, should the government do more to build a culture of transparency, trust and accountability across industry and, if so, what?

We welcome the proposals to encourage transparency reporting, which many of the largest online service providers already voluntarily and proactively publish. We do not consider, however, that transparency reporting should apply only to the prevalence of 'harm' and that it should be expanded to include information regarding the identities of those seeking to have content removed.

Question 2: Should designated bodies be able to bring 'super complaints' to the regulator in specific and clearly evidenced circumstances?

We consider that the restrictions to be imposed on the bodies entitled to bring such complaints should include a requirement that organisations not seek to profit from such complaints, can demonstrate not only a legitimate interest in the regulation of online service providers, but also an expertise in this area and an entitlement to claim to represent affected individuals. We anticipate that the basis on which organisations are permitted to bring representative claims for alleged breaches of the GDPR, as set out at Article 80(1) GDPR and s187 Data Protection Act 2018, could provide a starting point.

Question 2a: If your answer to question 2 is 'yes', in what circumstances should this happen?

We do not consider that it would be appropriate to permit 'super complaints' to be brought by third parties in connection with individual instances of alleged breaches of the obligations imposed on online service providers, and that any complaint should be based on allegations of systemic failure to comply with the relevant code of practice.

Question 3: What, if any, other measures should the government consider for users who wish to raise concerns about specific pieces of harmful content or activity, and/or breaches of the duty of care?

As set out above, we would not support the imposition of a duty of care or the extension of obligations beyond unlawful content. However, in relation to such content, and given that it is currently intended that the proposals will apply to sites of all sizes, it may be desirable for a central reporting function to be created in relation to allegedly harmful content which could then be identified on sites with a view to minimising the duplication of effort amongst providers and reducing barriers to entry.

Question 4: What role should Parliament play in scrutinising the work of the regulator, including the development of codes of practice?

Given the breadth of discretion proposed to be granted to the regulator in developing codes of practice, particularly in relation to harmful content, we consider that it is imperative that Parliament scrutinise any proposal to require the removal of lawful content which is defined as harmful.

Question 5: Are proposals for the online platforms and services in scope of the regulatory framework a suitable basis for an effective and proportionate approach?

As set out above, we consider that the scope of the proposals is overly broad and is likely to capture and therefore pose unnecessary red-tape on a number of online service providers who do not pose any material risk of permitting online harm on their services. We would exclude media organisations, online retailers, search engines and digital lockers from the scope of these proposals.

Question 6: In developing a definition for private communications, what criteria should be considered?

We consider that private communications channels should include channels which not only permit one to one but one to many communications, such as messaging services. We consider that the data protection and wider privacy

implications of any other approach would be impractical. We note that many such channels are already the subject of end-to-end encryption which would render proactive monitoring impossible in any event.

Question 7: Which channels or forums that can be considered private should be in scope of the regulatory framework?

Question 7a: What specific requirements might be appropriate to apply to private channels and forums in order to tackle online harms?

We consider that proposals relating to the enforcement of terms and conditions in relation to content reported to be harmful could potentially be imposed on private communications channels.

Question 8: What further steps could be taken to ensure the regulator will act in a targeted and proportionate manner?

We consider that improving the evidence base on relation to online harms is imperative to achieve targeted and proportionate regulation.

Question 9: What, if any, advice or support could the regulator provide to businesses, particularly start-ups and SMEs, comply with the regulatory framework?

We consider that the publication of template documents, such as terms and conditions, policies for appropriate conduct, processes and policies for the consideration of complaints would be of value to SMEs.

Question 10: Should an online harms regulator be: (i) a new public body, or (ii) an existing public body?

We believe that consideration should be given to the prospect of self-regulation by the industry rather than immediately moving to a system of public regulation. In the event that the government intends proceed with a system of government regulation, we consider that more important than whether the body should be new or existing, is its expertise to regulate freedom of expression issues.

Question 10a: If your answer to question 10 is (ii), which body or bodies should it be?

Question 11: A new or existing regulator is intended to be cost neutral: on what basis should any funding contributions from industry be determined?

As set out above, we would encourage the government to consider whether a system of self-regulation, at no cost to the taxpayer, would be a more proportionate solution to the challenges raised. Any system of funding ought to exempt the smallest providers, while providing for a scale based on size and nature of website, such that low risk sites were not disproportionately impacted.

Question 12: Should the regulator be empowered to i) disrupt business activities, or ii) undertake ISP blocking, or iii) implement a regime for senior management liability? What, if any, further powers should be available to the regulator?

We have concerns that such a regime would be ineffective or would not, in reality, be used. We would not support the use of such powers other than in the case of the most egregious systemic failings.

Question 13: Should the regulator have the power to require a company based outside the UK and EEA to appoint a nominated representative in the UK or EEA in certain circumstances?

It is not clear in what circumstances it is envisaged this would be appropriate or what the purpose of such a representative would be.

Question 14: In addition to judicial review should there be a statutory mechanism for companies to appeal against a decision of the regulator, as exists in relation to Ofcom under sections 192-196 of the Communications Act 2003?

In the event that a government regulator is appointed, with the wide range of enforcement powers currently envisaged, we consider that a right of appeal would be desirable.

Question 14a: If your answer to question 14 is 'yes', in what circumstances should companies be able to use this statutory mechanism?

We consider that, in the same way as an appeal against a decision of the Information Commissioner under s163 Data Protection Act 2018, such an appeal should allow the review of any determination of fact and that an appeal should be allowed where the decision is either not in accordance with the law or where the appellate court considers that the regulator ought to have exercised their discretion differently.

Question 14b: If your answer to question 14 is 'yes', should the appeal be decided on the basis of the principles that would be applied on an application for judicial review or on the merits of the case?

We consider that the appeal should be decided on the basis of a full review of the merits of the case.

Question 15: What are the greatest opportunities and barriers for (i) innovation and (ii) adoption of safety technologies by UK organisations, and what role should government play in addressing these?

We consider that the lack of a strong evidence base currently acts as a barrier to the innovation and adoption of safety technologies. While targeted solutions could potentially be developed To the extent that obligations are imposed on online services providers of all sizes, we would welcome the government seeking to centralise resources and encourage the development of open source technologies.

Question 16: What, if any, are the most significant areas in which organisations need practical guidance to build products that are safe by design?

Question 17: Should the government be doing more to help people manage their own and their children's online safety and, if so, what?

We consider that it would be desirable for the government to improve digital literacy.

Question 18: What, if any, role should the regulator have in relation to education and awareness activity?

We consider that the regulator should act as the spearhead for education and awareness activity and that the promotion of best practice conduct and the provision of assistance to achieve that should be a focus.

RPC's Tech Group

Jeremy Drew
Partner
+44 20 3060 6125
jeremy.drew@rpc.co.uk

David Cran
Partner
+44 20 3060 6149
david.cran@rpc.co.uk

Keith Mathieson
Partner
+44 20 3060 6486
keith.mathieson@rpc.co.uk

Nicola Cain
Partner
+44 20 3060 6171
nicola.cain@rpc.co.uk

Mark Crichard
Partner
+44 20 3060 6446
mark.crichard@rpc.co.uk

Paul Joseph
Partner
+44 20 3060 6590
paul.joseph@rpc.co.uk

Paul Joukador
Partner
+44 20 3060 6239
paul.joukador@rpc.co.uk

Ciara Cullen
Partner
+44 20 3060 6244
ciara.cullen@rpc.co.uk

Richard Breavington
Partner
+44 20 3060 6341
richard.breavington@rpc.co.uk

Oliver Bray
Partner
+44 20 3060 6277
oliver.bray@rpc.co.uk

Charles Buckworth
Partner
+44 20 3060 6641
charles.buckworth@rpc.co.uk

Ben Mark
Partner
+44 20 3060 6281
ben.mark@rpc.co.uk

Jon Bartley
Partner
+44 20 3060 6394
jon.bartley@rpc.co.uk

Jonathan Crompton
Partner
+852 2216 7173
jonathan.crompton@rpc.com.hk

Tower Bridge House
St Katharine's Way
London E1W 1AA
T +44 20 3060 6000

Temple Circus
Temple Way
Bristol BS1 6LW
T +44 20 3060 6000

38/F One Taikoo Place
979 King's Road
Quarry Bay, Hong Kong
T +852 2216 7000

12 Marina Boulevard
38/F MBFC Tower 3
Singapore 018982
T +65 6422 3000