Response to the UK Government's Public Consultation on the Online Harms White Paper

INTRODUCTION

On behalf of the Center for Data Innovation (datainnovation.org), we are pleased to submit comments in response to the open consultation on the "Online Harms White Paper," published in April 2019 by the UK government, which calls for a new regulatory framework for online safety.

The Center for Data Innovation is the leading think tank studying the intersection of data, technology, and public policy. With staff in Washington, D.C., and Brussels, the Center formulates and promotes pragmatic public policies designed to maximize the benefits of data-driven innovation in the public and private sectors. It educates policymakers and the public about the opportunities and challenges associated with data, as well as technology trends such as predictive analytics, open data, cloud computing, and the Internet of Things. The Center is a non-profit, non-partisan research institute affiliated with the Information Technology and Innovation Foundation.

GENERAL COMMENTS

Our response addresses a number of selected questions, through which we argue that despite the paper's intent to promote free speech, and to make the UK both the world's safest space to be online and the best environment for digital businesses to grow and innovate, if implemented, this proposal will restrict legitimate online content without due process, hurt digital businesses, limit access to information and, by creating state regulation of speech, damage freedom of expression for millions of users.

In particular, there are three main problems with the proposal. First, the definition of "online harms" is too broad, and by including content and activity such as trolling or disinformation, it would outlaw content that is generally protected speech in Western democracies.

Second, the scope covers Internet companies of all types, from social media platforms, file hosting sites, public discussion forums, to messaging services and search engines. As a result, the proposed framework will impose liability on companies for content they may only be hosting, caching, or transmitting. This proposal's vague definition and overly broad scope will expose companies to greater uncertainty in the online environment. In addition, the paper proposes to impose severe sanctions in case of non-compliance, including personal and criminal liability on senior management. As a result, companies will significantly err on the side of caution—with self-censorship and slower innovation as consequences.

Third, the paper proposes to create a regulatory body—an online sheriff—in charge of regulating content and activity on the Internet. Affording a single authority the power to unilaterally define a code of conduct for online speech and enforce these rules is a dangerous threat to access to information and free speech because of a lack of check and balance in a decision-making process that will have a far-reaching impact. In addition, although this regulator would be taking a "proportionate approach," the paper fails to describe what this would mean in practice. Finally, the paper does not clarify if this regulator would have a strong commitment to safeguarding freedom of expression.

Policy proposals aiming at regulating the Internet often reflect policymakers' crusade to "fix" it, as though the Internet were the cause of all harms, and lack consideration for the realities of the online ecosystem. The recent series of controversial attempts by EU policymakers and EU member states to implement rules for the Internet, such as the EU copyright directive, the EU online terrorist content proposal, the German law on hate speech, and the French law on disinformation, should have inspired others to take a more cautious approach. Unfortunately, the "Online Harms White Paper" published in April 2019 by the UK government, has not.

CONTRIBUTION

1. Activities in Scope of Regulation

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?

First, we recommend clarifying the definition of "online harms" which, in the current proposal, is too broad and vague. The paper indeed describes a broad range of harms in vague terms, and explicitly proposes to censor activities, materials, and behaviors that are "not necessarily illegal." In particular, the paper notes that "inaccurate information (…) can be harmful." By this definition, speech by many politicians would need to be censored online, as they often convey inaccurate information in the service of advancing their policy goals. It suggests regulating harms with a less clear definition, such as intimidation, disinformation, the advocacy of self-harm, and trolling the same way as harms with a clear definition such as child sexual exploitation, terrorist activity, modern slavery, and cyberstalking. Yet some of these activities with a less clear legal definition may be better addressed with a different legislative response, such as counter speech. Moreover, harm is a subjective concept. A study published in May 2019 by Ofcom and British data protection authority ICO reports that 61 percent of respondents had a "potentially harmful experience online" in the last 12 months, but the survey was based on a broad definition which fails to distinguish between "mildly annoying" and "seriously harmful" experiences.

The UK government should fix this lack of clarity in the definition of harms, as laws that use catch-all terms to define a concept that is subjective and amorphous could go as far as making legitimate content potentially illegal. It will make it difficult for companies to know which type of

content the future regulator could censor, would subject companies to "look-and-feel" type assessment, and would create an incentive to restrict and remove content including material that is perfectly lawful but could be considered "harmful."

Second, the scope of the proposal is overly broad. It covers Internet companies of all types, including social media platforms and messaging services, retailers that allow user reviews, as well as file sharing websites, cloud storage services, non-profit organizations, and public discussion forums. This means that the proposed regulator could ban user-generated content such as any shared electronic document, but also the "comments" section on news websites and the "letters to the editor" from their readers. These types of regulations would disadvantage UK online services and media companies and make it less likely for other businesses to expand to the UK.

Instead, UK policymakers should consider mutually supportive approaches such as co-regulation and co-governance between government and Internet companies, to ensure that none of the parties involved bears the burden of being both judge and jury. It is through voluntary measures and collaboration with the European Commission that a number of technology companies have been making rapid and significant progress to combat online harms in the last two years, for instance through the Code of Conduct on countering illegal hate speech online, and the Code of Practice on disinformation. And in June 2019, Facebook handed over to French judges the identification data of users suspected of hate speech on its platforms, in the context of a collaboration with the French government.

In addition, as harmful online content is a cross-border issue, fragmented policies across member states will introduce more complexity and uncertainty as they may conflict with other countries' laws, and will be economically damaging for companies. Better solutions lie in coordination with other nations and regions—especially the EU after the UK leaves the union.

2. The Regulator And Its Approach

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

The UK government should clarify a number of statements in the document. For example, the paper suggests that the proposed regulator will take a "proportionate" approach, taking into account the size and capacity of firms, yet it is unclear what this would mean in practice. For instance, will the regulator avoid imposing certain limits on platforms such as 4Chan, 8Chan, or Gab, because of their small size? The government should also clarify the definition of "online harms" because vagueness will hand more power and discretion to the proposed regulator.

The UK government should not equally apply the proposed regulation to all types of Internet companies. By imposing a "duty of care"—a concept based on the precautionary principle making companies "responsible for their users' safety and tackling harms caused by content or activity on their services"—the proposal would impose liability on companies for content they may only be hosting, caching, or transmitting, and introduce more uncertainty. For example, compliance may be difficult for providers which provide end-to-end encryption to their users.

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

If the UK is to introduce a new regulatory framework for online content, it should apply to all businesses in the relevant categories, regardless of size. If the UK is so concerned with the harmful effects of certain kinds of speech, then it makes no sense to exempt smaller businesses and startups from their duties and from penalties.

Moreover, rather than introducing individual liability for senior executives at companies and introducing legal uncertainty, UK policymakers should aim to support a digital environment where businesses focus more on protecting users from online harms rather than on protecting themselves from unpredictable sanctions. Policymakers should incentivize businesses to contribute to a healthier digital environment by letting them deploy more freely their technological solutions to these harms.

The new regulator would be a single authority in charge of defining a code of conduct for each type of harm identified. Companies will likely have to err on the side of caution, and overly restrict user speech, since the proposed regulator may impose unpredictable and severe sanctions. For example, new sites may ban comment forums dealing with sensitive topics such gender roles, in anticipation that commenters may post messages that may distress some of their users. To address the risk of censorship that will likely arise, the proposal would need to delineate what this "duty of care" would mean for the various Internet companies, the measures they would need to take, the framework to fairly assess compliance, and the type of enforcement required.

In addition, the proposed duty of care may not be compatible with the eCommerce Directive, in which Article 15 already reserves a right for member states to require "duties of care, which can reasonably be expected of [online intermediaries]….in order to detect and prevent certain types of illegal activities."

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

Addressed in our response to Question 9.

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

The government should not levy fees and charges on businesses to recoup the setup and operating costs of a new regulator. One risk is that it would impose fees on only certain types of businesses, and not all the businesses and industries a regulator would be in charge of regulating, which would be discriminatory. Moreover, a regulator that depends on industry penalties to be able to conduct its role will not be independent.

3. Enforcement

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?

The proposal should clarify, in case a company breaches its duty of care, the type of online harms that would lead to the enforcement measures subject to this consultation, such as incriminating senior managers, forcing online services to remove third-party content from search results, and requiring UK ISPs to block sites. Not all these measures should apply to every potential harm, especially for sites that are not dedicated to illegal activity, and the paper should propose various levels of enforcement depending on the type of harm.

Imposing personal and criminal liability on senior executives constitutes a high financial and reputational risk which could incentivize companies to proactively remove content. Depending on the type and severity of harms, such sanctions may not be necessary and civil penalties should suffice. As the paper fails to precisely delineate the scope of the future regulator and to define online harms thoroughly, the unpredictability and severity of these sanctions are a significant threat to businesses, but also to innovation in the UK.

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?

Requiring companies based outside the UK and EEA, but who have a legal presence in the country, to appoint a nominated representative in the UK or EEA would impose significant compliance costs on many businesses. In particular, some foreign startups may simply find it easier to avoid the UK market. Moreover, other countries may impose similar “reciprocal” rules on UK businesses doing business abroad, which could quickly drive up costs for UK businesses.

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?

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

Companies should have a right to appeal against the regulator's decisions and (in response to Question 14a) should be able to use this mechanism without restriction. But this right will be trumped by the lack of clarity around the definition of "harms"—it would be difficult for companies to demonstrate something may not be causing harm, given "harm" is not properly defined.

4. Technology As Part of The Solution

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

AI can be a powerful tool to automatically detect and respond to intentionally false or harmful content. Internet companies have accelerated their efforts, for instance by running hundreds of thousands of experiments on their algorithms, creating thousands of changes with each adjustment being tested against human reviewers. Despite the formidable opportunity it provides as a solution, AI is not a "silver bullet." Algorithms can identify patterns in how content is spreading, flag the sudden surge of a type of content and link it back to its source, or detect certain types of harmful content with high levels of accuracy. AI can also be used to identify "deepfakes" and other forms of disinformation. But it is not error-free, it is not able to understand context, and it does not understand human traits of humor such as sarcasm. Government-funded research can accelerate progress in these areas, such as by producing new models and training data. In addition, governments can partner on this type of research with other countries committed to upholding similar values.

5. Empowering Users

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

The government should and can do more to help people manage online safety, but it should do so primarily by investing in education and awareness-raising campaigns. Indeed, it is important to recall that companies can use technologies such as automated filters to moderate content, but that these may not be effective in addressing all situations, all types of content or speech equally. As a result, there needs to be a cautious and proportionate, rather than excessive, use of these

technologies in content moderation. Heavy-handed regulation that leads companies to take proactive measures so as to avoid sanctions will be counterproductive and distract them from the purpose of protecting their users and their freedoms. Instead, policymakers should support the efforts of Internet companies to collaborate and thwart harmful activity across online services.