Internet Association Initial Response To The Online Harms White Paper

1. About Internet Association

Internet Association (IA) is the only trade association that exclusively represents leading global internet companies on matters of public policy. IA has over 40 members, in areas ranging from search (e.g. Google), to hosting platforms, social media (e.g. Facebook and Twitter), digital services and beyond.[1] 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.

2. Introduction

As the Government's Online Harms White Paper argues, the internet is now an integral part of everyday life and often a powerful force for good.[2]

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,[3] and multiple estimates have found that internet services create significant consumer surplus for ordinary individuals.[4] 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.[5]

Like the internet, the introduction of the printing press, newspapers, radio, television and computer games all created significant concerns about their wider impact on society. In each case, it took time for society to best understand which problems were most pressing, and how best to solve them in a way that didn't undermine either the wider strengths of the medium or the importance of free speech in a liberal democracy.

The fundamental strength of the internet is its openness, and the unprecedented ability it gives to everyone to have a voice. At the same time, the sheer scale of the internet means that the same regulatory tools and models that we used in other forms of media are fundamentally ill-suited to online services. Every day, around 570,000 hours of video are uploaded to YouTube, 350 million photos uploaded to Facebook,[6] 500 million new Tweets added[7] and 2 billion WhatsApp messages sent.[8] This

[1] IA membership list here: https://uk.internetassociation.org/our-members/

[2] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/793360/Online_Harms_White_Paper.pdf

[3] https://www.ofcom.org.uk/__data/assets/pdf_file/0022/117256/CMR-2018-narrative-report.pdf

[4] https://www.pnas.org/content/116/15/7250 , http://www.publicfirst.co.uk/wp-content/uploads/2018/10/GoogleImpact2018.pdf

[5] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/507081/2904936_Bean_Review_Web_Accessible.pdf

[6] https://www.brandwatch.com/blog/facebook-statistics/

[7] https://www.dsayce.com/social-media/tweets-day/

[8] https://www.zdnet.com/article/whatsapp-now-one-billion-people-send-55-billion-messages-per-day/

sheer scale shows how platforms have unlocked human creativity, but also makes it impossible to take the same approach to content moderation as taken with previous technologies like TV or radio. No editor or collection of reviewers can moderate every piece of content without significantly restricting who has a voice on the internet.

2.1 Internet Companies Take Significant Action To Address Harms

Solving these problems is not easy, but as an industry we are committed to reducing harms as much as is currently possible – and what is possible is increasing every month. Nobody wants the internet to be a place where anyone feels unsafe, or users are misled. Internet companies take meaningful steps to protect their users from harm on their services. Initiatives include:

- Investing significant resources in both human content moderation and, partnering with third sector organisations and think tanks, developing machine-learning technology to detect and remove harmful material more quickly.
- Working with groups like the Counter Terrorism Internet Referral Unit, and forming the Global Internet Forum to Counter Terrorism (GIFCT) to curtail the spread of terrorism and violent extremism online.
- Partnering with a number of organisations across the globe, including the Internet Watch Foundation, to work together to remove harmful CSAM from the internet.
- Forming internal online safety councils and designating employee teams to improve online safety and promote a productive and welcoming environment online.
- Creating clear pathways for people to report inappropriate or harmful content, so that it can be addressed under companies' terms and conditions.
- Investing in fact-checking services and using AI and other technology to tackle false information.
- Investing in educating users about how online services operate and how to make the best use of them. Efforts to educate people on what is appropriate on online platforms helps guide behaviour and can help minimise the need for moderation.

This is not to say that a simplistic focus on the amount of money invested in trust and safety initiatives or the number of content moderators employed by a company is the best or only means of assessing an organisation's commitment to safety. Indeed, different models of content moderation are used in industry to good effect – for example some services use a more community-based system of moderation – and as the White Paper recognises there should be different expectations on companies depending on their particular circumstances. Regulation should be proportionate in terms of scale of harms prevalent on a service, and also in terms of the economic development stage and size of the platform.

Nevertheless, this is a hard problem, and no system of moderation, whether algorithmic or human, centrally-managed or community-based, will be perfect. Some of the discussion around the release of the White Paper has appeared to suggest that harm only persists on the internet because internet companies do not care or are not trying hard enough to tackle it. We fundamentally reject this characterisation – as set out above, internet companies take significant steps to address harms on their services and know there is more to be done. However, we do not believe it is reasonable to claim that there is an easy way for internet companies to instantly eliminate all online harms.

Alongside company efforts, there is also a role for government to provide industry and the public

guidance on matters relating to public discourse, based on our laws and culture. There is a growing view that moral and ethical judgements around content should not be simply delegated to private companies, and that in some circumstances the government should be more clear on the standards required. The expectations for behaviour online should be the same as for behaviour offline; and public institutions have a key role to play in establishing those norms, for example by following through on police investigations of criminal online harms, or providing online citizenship education through PSHE lessons in schools.

We need to keep working together to make the internet safer. In order to further reduce online harms, we will need private and public sectors to work together – and to make use of a variety of tools, including technology, public education, the law and better regulation.

2.2 Internet Association Online Harms Regulatory Principles

As part of our engagement with government, IA and its member companies have said that there is a role for targeted, proportionate online safety regulation. However, in order to be effective and not endanger the wider benefits created by the internet, we believe that this regulation should:

- Be targeted at specific harms, using a risk based approach;
- Provide flexibility to adapt to changing technologies, different services and evolving societal expectations;
- Maintain the intermediary liability protections that enable the internet to deliver significant benefits for consumers, society and the economy;
- Be technically possible to implement in practice, and also take into account that resources available for this type of activity vary between companies;
- Provide clarity and certainty for consumers, citizens and internet companies;
- Recognise the distinction between public and private communications.

3. Problems With The Online Harms White Paper

At present however, we are worried that the current proposals in the Online Harms White Paper are not sufficiently targeted or proportionate to the harms they are designed to minimise. As other commentators have argued, there is a real risk these proposals hurt the British tech sector, worsen the quality of internet services for ordinary consumers, undermine privacy, and produce a chilling effect on freedom of speech.

At the same time, we believe that we should collectively focus on addressing harms holistically – tackling societal problems both at their root in the real world, as well as in their online manifestation. We believe that a broader policy approach, looking both offline and online, will lead to better outcomes for society.

Further, we are concerned that while the White Paper states that the proposals are “compatible with the EU’s e-Commerce Directive”, in aggregate, the framework undermines the intermediary liability protections that have enabled the internet to deliver benefits to the UK. One particular worry is that in practice, the only way to meet many of the demands of the new obligations will be through the introduction of mandatory filtering.

The internet has flourished in part because platforms permit users to post and share information without

fear that those platforms will be held liable for third-party content. Dilution of intermediary liability protections would encourage internet companies to engage in over-censorship for fear of being held liable for content, with a consequential impact on freedom of speech. Intermediary liability protections also play a critical role in driving economic growth, by enabling new companies to invest and launch new services in the UK and enabling existing companies to innovate, scale and grow their businesses.

In particular, we worry that:

1. "Duty of Care" has a specific legal meaning that does not align with the obligations proposed in the White Paper, creating legal uncertainty, and would be unmanageable;
2. The scope of the services covered by regulation needs to be defined differently, and more closely related to the harms to be addressed;
3. The category of "harms with a less clear definition" raises significant questions and concerns about clarity and democratic process;
4. The proposed code of practice obligations raise potentially dangerous unintended consequences for freedom of expression;
5. The proposed measures will damage the UK digital sector, especially start-ups, micro-businesses and small- and medium-sized enterprises (SMEs), and slow innovation.

3.1 "Duty of Care" has a specific legal meaning that does not align with the obligations proposed in the White Paper, creating legal uncertainty, and would be unmanageable.

Following the suggestions of Woods and Perrin (2018,2019),[9] the White Paper proposes a new statutory duty of care on internet companies "to take reasonable steps to keep their users safe and tackle illegal and harmful activity on their services." This duty will be overseen by an independent regulator, which in turn will create new codes of practice outlining the "systems, procedures, technologies and investment, including in staffing, training and support of human moderators, that companies need to adopt". When these codes don't exist, companies will be expected to use their own judgement to "take action proportionate to the severity and scale of the harm in question."

As is clear in its design, and made explicit in the third party reports calling for its introduction, much of the inspiration for the duty comes from the Health & Safety Act 1974 and Occupiers' Liability Act 1957. It is understandable why policymakers have looked to these examples as models for online harm, but nevertheless, we also need to be clear about the limitations of these comparisons and in particular the ways in which online harms are not like the physical risks faced in the workplace:

- Risks to physical health and safety are clearly defined, while online harms are much more ambiguous. It is easy to ascertain when an employee has suffered a physical accident, and there are now well established methodologies for quantifying their seriousness in monetary terms. By contrast, many of the harms targeted by the white paper, such as cyberbullying and trolling, extremist content or disinformation, have a much less clear definition or boundary with other types of speech.
- Without clear definitions, it is hard for companies to perform their own cost-benefit analysis and risk assessment. A key element of the current health and safety regime is that while companies

9 https://d1ssu070pg2v9i.cloudfront.net/pex/carnegie_uk_trust/2019/04/08091652/Online-harm-reduction-a-statutory-duty-of-care-and-regulator.pdf

are encouraged to follow standard industry codes of practice, they are also allowed to, in effect, perform their own cost-benefit analysis of what risks are worth reducing. This is much harder for online harms, leaving companies only able to follow agreed guidelines by the regulator – or to act as conservatively as possible. This is a particular risk for new or smaller companies, and could potentially act as a significant chilling effect on innovation.

- There is no perfect technological solution that can completely eliminate the risk of all online harms. In health and safety law our real worry is negligent employers – while many of the solutions and practices needed to ensure safety are relatively straight forward. When a practice is especially dangerous and risk is impossible to fully eliminate, governments tend to introduce more specific regulations specifically for that sector. For many online harms, by contrast, it is much less clear what the appropriate response in return is required.
- For online harms, we need to consider not only the role of companies, but also the role and responsibilities of individuals using the online service. In health and safety we are concerned mainly (with some exceptions) with the relation between the employer and employee, but for online harms we need to look at the responsibilities of both the online services *and* the users that upload the harmful content itself – and the direct interactions between users themselves.

Expert lawyers – for example Graham Smith (2018)[10] – have also raised detailed concerns about adopting a duty of care concept from health and safety / occupiers' liability and applying it to online harms. For example, Smith argues that many harms that may be encountered online are of a different nature from the safety-related dangers in respect of which occupier-related duties of care are imposed in a physical public space; and that, unlike dangers commonly encountered in a physical place, the kind of online harms that it is suggested should be within the ambit of a duty of care typically arise out of how users behave to each other rather than from interaction between a visitor and the occupier itself.

There are good reasons why we do not introduce duties of care whenever we have a policy goal. We do not introduce a duty of care for newspapers to ensure that their readers are perfectly informed, or a duty of care around environmental sustainability. While these are important societal goals, the way we have decided to tackle them with is with specific regulations setting out very clearly the role of companies. There are often hard trade-offs between different societal goals such as privacy, safety and innovation – and these need to be actively considered during the White Paper process.

3.2 The scope of the services covered by regulation needs to be defined differently, and more closely related to the harms to be addressed.

While we speak of *the* internet, in reality this covers a vast array of different services, run by companies and organisations of all sizes. The White Paper makes a deliberate choice to be as comprehensive as possible – and explicitly included in scope are companies and organisations "of all sizes" that "allow users to share or discover user-generated content or interact with each other online". The paper argues that this comprehensive approach is necessary to ensure that the new framework does not just shift harmful speech onto less regulated platforms.

In practice, however, there are significant differences between a social media platform, a search engine, a review website, an online marketplace, a fan forum, a cloud service, or a private messaging service. Moreover, there needs to be closer consideration of where harms are actually occurring online and then regulatory focus should be targeted at where the harm exists, rather than across a wider range of

[10] https://inforrm.org/2018/10/23/take-care-with-that-social-media-duty-of-care-graham-smith/

services.

On private vs public communications, seeking to create a backdoor to read private messages would severely undermine individual privacy and severely undermine civil liberties. Just as the Royal Mail is not responsible for the contents of every letter, or telephone operators the contents of every call, we do not think it is proportionate to seek to regulate private communications in the same way as public communications. In addition, many of the most popular messaging services are end-to-end encrypted, making it technically impossible for platform owners to monitor the content of messages.

While the White Paper explicitly recognises that private communications should be treated differently, it gives little guidance on what this will mean practically – or how 'private' is to be defined. There should not be a fundamental change to how private communications can be monitored without a serious public debate on the implications.

3.3 The category of "harms with a less clear definition" raises significant questions and concerns about clarity and democratic process.

Rather than provide a definitive list of harms it seeks to tackle, the White Paper instead only definitively sets out what harms are not in scope: harms to organisations, breaches of privacy or security, and harms suffered from using the dark web rather than open internet. The paper argues that providing a static list would "prevent swift regulatory action to address new forms of online harm".

Given the wide scope of activities undertaken on the internet, this effectively gives the proposed regulator enormous discretion to unanimously outlaw or discourage particular activities or types of speech. This is particularly the case in relation to the listed "harms with a less clear definition". We believe it is important that key decisions about internet regulation and freedom of speech are made by Parliament after thorough and reasoned debate.

The paper does provide an initial indicative list of harms that are definitely in scope. It is not clear why all these different problems are being grouped together: terrorist content and activity is very different from online bullying, which is different in turn from excessive screen time. Many harms are already illegal – such as hate crime or selling illegal weapons on the open internet – and despite rhetoric of a "wild west", internet companies are in reality already subject to a multitude of regulatory frameworks, designed to reduce the risk of loss of privacy, discrimination, monopoly, underage access to pornography, copyright infringement or misleading electoral campaigning.

The White Paper does try to take a differentiated approach by envisaging different codes and requirements for different harms. The recognition of the need for different solutions depending on the harm is welcome, however at present a key concern is that the codes – particularly for "harms with a less clear definition" – would be hard to operationalise for a regulator, are too prescriptive, are not sufficiently targeted at where the harms exist, and do not take into account the different functionality of online services.

3.4 The proposed code of practice obligations raise potentially dangerous unintended consequences for freedom of expression.

In its response to the 2012 Leveson Inquiry, the Government rejected the idea of statutory press regulation – arguing in effect that while there were real concerns about the harms created by the press,

that introducing a statutory authority would compromise Britain's democratic traditions. In its last manifesto, the Conservatives confirmed their support of a "free and independent press", and pledged to repeal Section 40 of the Crime and Courts Act, which many have argued will have a disproportionate chilling effect on free speech.

As many civil liberties groups have argued, however, by explicitly including disinformation within scope and referring to the harms created by inaccurate information – "regardless of intent" – the new measures risk introducing state regulation of the press by the backdoor.

This is particularly concerning given that under current plans the regulator is being given independence to decide the list of harms. While the initial list of harms may focus on incitement to terrorism or violent extremism – which is already illegal – it is not hard to see the creation of a slippery slope leading to the shut down of legal speech that people find objectionable. For example, we are already seeing pressure to include books that argue for ideas that many dislike to be in scope of any new duty of care.[11]

Wherever governments want to restrict speech, they should use democratic processes to clearly outline the case. The current proposal risks allowing a regulator to limit access to legitimate information in an opaque and arbitrary manner, in effect banning speech without openly declaring it unlawful or providing clear definitions of what is and is not permissible.

By its very nature, political speech is often controversial – and encourages passionate disagreement. Part of freedom of speech is that we allow people to be wrong, and even to campaign for ideas that may have negative ideas for society. One person's misinformation is another person's dissenting opinion. It is not the role of publishers, platform owners or politicians to adjudicate these arguments – but the market of ideas and democratic debate. To give a relevant parallel, we do not think it would be appropriate to introduce a duty of care for book publishers to avoid harms from reading their works.

Websites that allow consumer reviews provide an interesting example. People's opinions can be very divergent and, in the case of a bad experience, quite negative about a service or a product. The subjects (e.g. the store owner, product creator, etc) of those reviews may claim that content is 'harmful' and, therefore, should be removed. The relevance of reviews for consumers and businesses alike would be at stake if negative reviews that could be potentially perceived as 'harmful' are more systematically not published due to the high legal risk and possible fines. One-size-fits-all regulation to tackle 'hate speech' and/or 'harmful' content online would not only be difficult to apply to the diversity of content and platforms concerned, but would also be at risk of abuse.

The White Paper rightly argues that the regulator should "not be responsible for policing truth and accuracy online". In practice, however, it will be hard for this safeguard to have teeth while maintaining the current focus on online "harms with a less clear definition". While platform providers can perform a helpful moderation role, encouraging their audience to consume a variety of viewpoints, there is a significant difference between this and attempting to ban or censor speech.

11 https://hansard.parliament.uk/Commons/2019-04-30/debates/434AFD6E-7484-47DA-8C1F-4A945896642B/SocialMediaAndHealth?highlight=amazon#contribution-D1CA42D1-0E6B-4FC2-931D-ED8227B5B291

3.5 The proposed measures will damage the UK digital sector, especially start-ups, micro-businesses and small- and medium-sized enterprises (SMEs), and slow innovation.

The White Paper is keen to emphasise that "innovation and safety online are not mutually exclusive" and proposes to give the new regulator a legal duty "to pay due regard to innovation", pointing to a similar obligation placed on the ICO under the Data Protection 2018. Given its newness, it is hard to test the effectiveness of this particular duty – but more generally, even when the costs are justified, most regulations have some cost burden, as the Government recognises with initiatives like its business impact target or one in-two out rules.

The hardest hit by new regulations are smaller companies and start-ups, who find it harder to absorb the fixed costs of new administrative systems. While the White Paper argues that we can learn from the example of other areas of regulation such as GDPR or Health and Safety rules to reduce the burden on SMEs, it is misleading to argue that these did not create significant costs for businesses. (The Federation of Small Businesses website, for example, reports that "Ensuring your business is compliant with all the relevant health and safety regulations is a time consuming and costly process".[12]) By creating additional costs, new regulations increase barriers to entry, making it harder for disruptive new entrants. Even once launched, companies face a continuing cost to monitor regulatory developments and update their compliance based on any new requirements.

The White Paper argues that the regulator will take a proportionate approach, taking account of the size of companies and the reach of their platforms. This is a good idea – but given that there is no fixed list of harms, and the details of every company are likely to be subtly different, it is going to be difficult to completely remove ambiguity over what is required to be compliant, particularly for companies developing new and innovative products. Given the new proposed liability for significant fines, many companies are going to err on the side of an extremely risk averse approach, which will stifle innovation and chill free expression.

One particular worry is that in practice, the only way to meet many of the demands of the new obligations will be through the introduction of mandatory filtering. Beyond the potential impact on free expression and the risk of taking down perfectly legal content, this kind of content filtering can be extremely expensive to develop and iterate – and often needs to be complemented by significant amounts of human moderation.

Structurally, a duty of care risks embedding a 'precautionary principle' within regulation of the internet – as was explicitly enunciated in the work of Woods and Perrin (2019) calling for a duty of care. While the precautionary principle can be appropriate in areas of high risk or where it is hard to reverse damage, as in environmental sustainability, the precautionary principle is not an appropriate benchmark for all regulation, and taken literally, as leading legal scholars have argued, is likely meaningless – both inaction and action create risks.[13] As many commentators have argued, one of the reasons the internet has delivered significant beneficial new services in the last few decades is an approach of 'permissionless innovation', allowing small teams to try out new ideas rather than have to seek permission from large bureaucracies or a government regulator first.[14]

[12] https://www.fsb.org.uk/benefits/advice/health-safety-advice
[13] https://papers.ssrn.com/sol3/papers.cfm?abstract_id=307098
[14] https://www.mercatus.org/system/files/Permissionless.Innovation.web_.pdf

As well as slowing future innovation, these rules could lead to a worsening of the quality of internet services consumers already receive. As we have already seen to some extent with GDPR, many foreign platforms and publishers will often decide it is easier to block access in the UK rather than incur extra costs or liability. Even for the larger platforms, the loss of effective intermediary liability protection could make it cost prohibitive to keep offering the same kinds of service.

The best way to avoid this chilling effect on innovation is to ensure that any remedies are based around a specific list of harms, with a clear evidence base of harm and a quantified regulatory impact assessment of best practice in reducing risk. This will require the Government to make some hard decisions about trade-offs between different values upfront – but it is better to do this now, than leave potential ongoing ambiguity and the possibility of a slippery slope to continued restrictions.

It will also make it easier for companies to focus resources on innovating on the most important harms. It is in everybody's interests to create a safer internet – and the best way to do that is if we together to keep innovating on new, scalable and cost effective ways of identifying hostile actors, providing users with a balanced range of content, and protecting individual privacy.

We look forward to seeing the government's economic impact assessment of the White Paper proposals and in particular the impact on start-ups and small- and medium-sized enterprises.

4. Conclusion

IA welcomes the opportunity to provide initial comments on the Online Harms White Paper. IA and its member companies are committed to reducing online harms and believe that there is a role for targeted, proportionate online safety regulation. However, IA has a number of concerns with the White Paper proposals, in particular around the proposed Duty of Care, the scope of the services covered, the category of "harms with a less clear definition", impacts on freedom of expression, and impacts on the UK internet economy.

IA supports balanced, proportionate regulation that achieves the joint objectives of protecting people from harm online and ensuring that the internet can continue to deliver benefits to the economy and society. IA has proposed a number of regulatory policy principles which it believes can help deliver this outcome, and IA and its members will continue to work with policymakers and regulators on these important issues as the White Paper process continues.

Internet Association
29 May 2019