Google Australia Pty Ltd
Level 5, 48 Pirrama Road
Pyrmont, NSW 2009
Australia

google.com

Wednesday 19 February 2020

Director
Online Safety Research and Reform Section
Department of Communications and the Arts
GPO Box 2154
Canberra ACT 2601

BY EMAIL: onlinesafety@communications.gov.au

Google and YouTube welcome the opportunity to contribute to the consultation on an Online Safety Act. The Government's proposals have implications for online safety and freedom of expression, innovation, and the distribution of responsibilities for addressing content challenges. These are all topics that we take very seriously.

We believe the Internet has had, on balance, an immensely positive impact on society. Our mission is to organise the world's information and make it universally accessible and useful. We build tools that are a force for creativity, learning, access to information, and much more. They have enabled economic growth, boosted skills and opportunity, and fostered a thriving society.

We recognise, however, that the Internet is also at times exploited by a minority of bad actors. We take the safety of our users very seriously, and we are committed to ensuring that illegal and harmful content that appears on our platforms is dealt with as quickly as possible.

Google is supportive of regulation, where it is carefully crafted and appropriately tailored. And indeed we haven't waited for regulation to address problematic content online. We have made significant investment in technology and human resources, and we have engaged with policymakers in Australia and around the world on the appropriate oversight for content sharing platforms, such as social media and video sharing sites.

We have provided detailed comments and suggestions on several of the proposals made in the discussion paper below, so as to evolve towards a truly effective framework to foster online safety for all Australians. In particular:

- Government should acknowledge that there is a shared responsibility to foster online safety between industry, government, parents / carers, NGOs and civil society.

- The focus of Basic Online Safety Expectations (BOSE) should be on practical best efforts and overall processes, while avoiding being overly prescriptive.

- Any preemptive and preventative action recommended under the BOSE should be coupled with a ‘Good Samaritan’ framework that incentivises companies to take these proactive measures without risking the loss of liability for good-faith missteps in that process.

- Transparency reporting requirements should be flexible, and, if there are to be any sanctions attached to them, they should focus on systemic failures.

- Any expansion to the scope of services subject to both the cyber bullying and cyber abuse schemes should be limited to content sharing services, like social media and video sharing services, which have the principal purpose of helping people to store and share content with the public or other broad audiences, over which the platform provider does not have editorial responsibility. Services that are closed and private by default are arguably already regulated by existing criminal laws.

- If the cyber abuse scheme were to be extended to adults, it is crucial that the definition of relevant content be tied to the Criminal Code.

- Regarding removal turnaround times, we strongly suggest that a more workable standard would be one that instructed online platforms to remove content “with all due speed,” “without undue delay,” or “expeditiously” and without a fixed 24 hour turnaround. We also call attention to the numerous comments made by the eSafety Commissioner that businesses typically do respond expeditiously to requests to remove content.

- The proposed accreditation scheme for safety tools does not provide clear utility. It would entail considerable resources to set up and administer, and would be very slow.

- On the subject of blocking terrorist and extreme violent material online, appropriate legislative instruments already exist to address these issues efficiently, and, to the extent any new instruments are introduced, it is essential that they be narrowly tailored to address only those 'worst of the worst' platforms and services that willfully and systematically fail to respond to valid legal removal requests regarding specific items of identified content.

- For ancillary services, any additional powers should specifically focus on notice-and-takedown of specific illegal material.

- In the context of governance, any increase in the powers and responsibilities of the Office of the eSafety Commissioner should be accompanied by a formal framework of multi-stakeholder oversight into the policy direction and decisions being made by the Office.

Our full submission is set out in the following pages. We look forward to continuing to engage with you during this process.

Yours sincerely,

[redacted]

Samantha Yorke
Government Affairs and Public Policy

Principles for Regulation

We take the safety of our users very seriously, and we are committed to ensuring the small proportion of illegal and harmful content that appears on our platforms is addressed as quickly as possible. We have not waited for legislation to act in tackling illegal or harmful content, and we are committed to doing our part.

Our strategy for tackling illegal and harmful content is tailored to each of our platforms. Across our products, our teams tackle a broad spectrum of online abuse, from scams, like the email allegedly from a 'relative' stranded abroad needing a bank transfer to get home safely, to abhorrent content, including child sexual abuse material (CSAM) online. Understanding the different parameters of the products we deliver is vital to our work and policy development. Given that breadth, our team is diverse, comprising product specialists, engineers, lawyers, data scientists, ex-law enforcement officials and others. They work hand-in-hand around the world and with a global network of safety and subject matter experts. We now have over 10,000 people across Google working on content moderation and removal across our platforms. This includes reviewers who work around the world, 24/7, speak many different languages and are highly skilled.

For each product, we have a specific set of rules and guidelines that are suitable for the type of platform, how it is used, and the risk of harm associated with it. These approaches range from clear community guidelines, with mechanisms to report content that violates them, to increasingly effective artificial intelligence (AI) and machine learning that can facilitate removal of some types of harmful content before a single human user has been able to access it.

We have shared our ideas for approaching oversight of content-sharing platforms in a number of fora.[1] To summarise, we believe that effective regulation should provide legal clarity for platforms; focus on systemic approaches to the relevant issues; and rely on transparency and best practice.

At the same time, online safety is a shared responsibility across society, in which technology companies, governments, child protection advocates, civil society, parents and users all have a role to play. Regulation should take a holistic approach, looking at the roles of all actors both online and offline. For example, tackling terrorist content requires a wide set of actors to develop effective domestic and foreign policies, new security and military tools to more effectively deal with terrorism, and programs to promote an environment of opportunity and inclusion that helps prevent radicalisation.

Objects of the new Act

In turn, we recommend that, in a new Act, the Government should acknowledge these shared responsibilities between industry including online service providers, government, advocates and civil society. Policymakers must consider the full toolkit of approaches to address online safety, beyond simply regulating platforms. For instance, this should include law enforcement efforts, which should focus their efforts directly against users who violate the law. Where users are uploading and sharing illegal content, such as terrorist content or child abuse imagery, platforms should take action to remove that material when they become aware of it. However, this is a mitigation and not preventative; continued law enforcement action is necessary to stop these users from offending and to prevent them from being able to create and share this content online in the first place. We note that the existing powers held by the Office of the eSafety Commissioner to issue end user notices have never been used. Furthermore, a longer term goal should also include behavioural change amongst Internet users; we must work together to identify why individuals engage in anti-social behaviours and develop programs that seek to change these patterns and modes of expression.

[1] https://www.blog.google/outreach-initiatives/public-policy/oversight-frameworks-content-sharing-platforms/

Basic online safety expectations

There is merit in establishing a clear set of societal expectations, which could be couched in the form of Basic Online Safety Expectations (BOSE).

The focus of BOSE should be on practical best efforts and overall processes, while avoiding being overly prescriptive, which would otherwise limit the ability of online platforms to come up with a wide and innovative array of effective approaches that may differ based on the unique properties of each platform.

The overall approach of a new Online Safety Act and corresponding BOSE should endeavour to be pragmatic, reasonable and proportionate. This is particularly important when it comes to proactive harm and abuse monitoring and prevention, as well as transparency. Special regard should be paid to ensuring the appropriate balance with other legal obligations and rights such as freedom of expression.

Preemptive and preventative action should be coupled with a 'Good Samaritan' framework that applies to online platforms. The discussion paper suggests "*a new set of increased expectations around minimum standards for pre-emptive and preventative action.*" To the extent companies that take such action may incur liability for their failure to catch and take action on specific illegal content, the risk of liability creates a perverse incentive for companies to either refrain from taking reasonable preventive action, or to over-remove legitimate content in the course of moderating. 'Good Samaritan' protections would address that concern by giving protection for platforms to seek out and remove harmful content, without risking the loss of liability for occasional failures in that process. Any new law should ensure businesses can continue to invest in responsible proactive detection methods, without incurring an increased risk of legal liability in so doing.

We also agree that transparency is critical and could usefully feature in BOSE. Transparency reports can give the public and an oversight body a clear picture of what platforms are doing to tackle harms, and inform a regulator's judgments about systemic failures. They will also encourage companies to improve the measures taken to keep their users safe, long before fines or more significant sanctions are required.

Transparency has long been a priority at Google to help our users understand how our products work. We have a long history of producing transparency reports, notably concerning governments' removal requests. More recently, in 2018, YouTube began publishing a quarterly transparency report which provides aggregate data about the flags we receive and the actions we take to remove videos and comments that violate our content policies.

We recommend that any transparency reporting requirements are flexible, avoiding a rigid, templated approach and allowing enough room for each platform to report in a way that takes into account the nature of the services and the nuances that individual reports may need to cover as a result. Moreover, the benefits of transparency must be balanced with other interests and avoid negative unintended consequences. For instance, overly detailed transparency can allow bad actors to game a platform's systems through manipulation, spam, fraud and other forms of abuse. Transparency requirements must also be careful not to risk trade secrets or violate user privacy or data disclosure laws.

Further, if as proposed in the discussion paper, the eSafety Commissioner is enabled to impose sanctions for non publication of transparency reports, any such sanctions should focus on systemic failures to produce appropriate reports, rather than the detailed specifics of elements of reporting. Companies should also be given notice and an opportunity to rectify alleged failures.

Cyberbullying scheme

We support the recommendation made by the Independent Report of the Statutory Review of the *Enhancing Online Safety Act 2015* to do away with the existing tiering model for social media services. In practice most organisations within the scope of the legislation voluntarily comply with requests to review and remove cyber bullying material, regardless of whether they sit in tier 1 or 2.

Extension of the cyber abuse scheme to adults

The proposed BOSE already posit that online service providers would further strengthen their ongoing efforts to tackle abusive material. Extending the cyber abuse scheme to adults would seem to go further. If this is to occur, it is crucial that the definition of relevant content be tied to the Criminal Code, as the Consultation paper suggests. Tying the legal removal requirement to provisions of the Criminal Code is more helpful than amorphously asking companies to remove "abuse" or "harassment." In the absence of consistency with and adherence to the Criminal Code and relevant case law, the risk that companies will err on the side of caution and over-remove legitimate speech is very real.

When abusive material is classified as serious, illegal 'cyber abuse' of adults, then the well-known and largely successful method of notice and takedown for tackling illegal content can be used. We have detailed principles for combating illegal content and effective notice and takedown elsewhere (see: https://www.blog.google/perspectives/kent-walker/principles-evolving-technology-policy-2019/smart-regulation-combating-illegal-content/). In summary, notices of illegal cyber abuse material should be as specific as possible - such as referring to a precise URL, and detailing the purpose of the notice - and provide both the ability to appeal, and a penalty for people who abuse the system (see further below in relation to take-down periods).

The author of the Independent Report of the Statutory Review of the *Enhancing Online Safety Act 2015* noted that bullying is a broader social issue and is not confined to the Internet[2]. We remain concerned about the legislative separation of bullying that takes place online from bullying that takes place in an offline context. It is widely acknowledged that a very high proportion of cyberbullying is an extension of bullying behaviour taking place in an offline context and we urge the Government to consider how to adapt the legislative framework to tackle bullying behaviour in all of its manifestations.

Proposed take-down period of 24 hours

We are committed to tackling illegal content. Google invests millions of dollars in technology and people to combat illegal content in an effective and fair way. It's a complex task, and–just as in offline contexts—it's not a problem that can be solved by one silver bullet solution. Rather, it's a problem that must be managed in combination with other efforts, and we are constantly refining our practices. As a result, Google achieves generally expeditious removal, particularly of harmful content.

It would be helpful to better understand why there is a perceived need to reduce the turnaround time that exists under the Enhancing Online Safety Act 2015 from 48 hours to 24 hours, particularly when the eSafety Commissioner has made repeated references to the fact that most platforms remove content upon receiving a request from her Office very promptly. Respectfully, we do not believe that the discussion paper successfully makes the case for reducing the legislated turn around time for the removal of cyber bullying content to 24 hours. Our experience in implementing various frameworks elsewhere in the world that mandate a short and specific time for removal is that this inevitably leads to overblocking of legitimate speech.

Some take-down requests can be complex and necessarily take time to assess thoroughly. A complainant may not initially provide sufficient information; there may be questions as to the complainant's authority to make the complaint; the possible exception created by the material being shown for educational or documentary purposes; or simply the difficulty of assessing whether material has crossed the line of impropriety in the often-nuanced cases that we face nowadays, among other issues; each of which can take time to resolve and can only be accommodated by a flexible requirement.

Context often matters when determining whether content is illegal. Consider a video of military conflict. In one context the footage might be documentary evidence of atrocities in areas where journalists have great difficulty and danger accessing. In another context the footage could be promotional material for an illegal organisation. Even a highly trained reviewer could have a hard time telling the difference, and we need to get those decisions right across many different languages and cultures, and across the vast scale of audio, video,

[2] Page 28 of the Independent Report of the Statutory Review of the *Enhancing Online Safety Act 2015*

text, and images uploaded online.

Making context-sensitive judgments can be time-consuming when complainants are well-meaning. In other cases, complainants actively attempt to abuse our removal processes through falsified identities and misrepresentations. Google regularly encounters malicious and baseless attempts to remove legitimate content from its platforms using copyright removal processes, for example.[3] Indeed, YouTube recently commenced litigation in the United States, alleging that the named defendant repeatedly attempted to harass and extort money from YouTube content creators through bogus allegations of copyright infringement.[4]

A framework that sets specific timeframes for the removal of illegal content will encourage platforms to remove first and ask questions later (or, more likely, remove and not ask questions at all), curtailing the legitimate interests of individuals and organisations large and small who use digital platforms to express themselves and lawfully share content. It will frustrate careful, more considered human review. Specifying an exact turn-around-time, regardless of complexity of case, provides an incentive for companies to over-remove, thereby silencing political speech and user expression. In addition, quick and prescriptive turn around times and unexpected spikes in volume place a significant pressure on content reviewers / moderators (who are already looking at difficult content) to make quick decisions about content that in some cases are incredibly nuanced and complex. Indeed, focusing on the speed with which content is removed as a measurement of success may not actually reflect the public policy objective of minimising widespread exposure to a piece of inappropriate or harmful content. The metric that we prefer to use within the YouTube Community Guidelines Transparency Reports, for instance, is the proportion of videos removed without a single person viewing the content.

Looking at the case of Germany's NetzDG, the discussion paper points to the requirement for social media platforms, after receiving notice, to exercise a local take down of "obviously illegal" content (e.g. a video or a comment) within 24 hours after notification. Services have 7 days to remove content that is not "obviously illegal" and even longer if the content is referred to an accredited self-regulatory body for review. However, the NetzDG also demonstrates that the quality of takedown requests can vary wildly. As our Transparency Report notes, 76.62% of reported items were not removed or blocked because the content did not violate YouTube's Community Guidelines nor the criminal statutes referred to in NetzDG.[5] Spending time evaluating such a high volume of spurious complaints takes reviewers away from reviewing content that does violate YouTube's Community Guidelines or local law.

[3] For a set of examples, see https://support.google.com/transparencyreport/answer/7347743?hl=en&ref_topic=7295796

[4] See https://www.theverge.com/2019/8/19/20812144/youtube-copyright-strike-lawsuit-alleged-extortion-minecraft

[5] https://transparencyreport.google.com/netzdg/youtube

Taking into account these considerations, a more workable standard would be one that instructed online platforms to remove content “with all due speed,” “without undue delay,” or “expeditiously” and without a fixed 24 hour turnaround, upon receipt of a clear and specific notice. The service provider would be afforded a reasonable period of time in which to assess the take down request once all the required information has been provided by the requesting individual. The exact time frame is not something that should be stipulated in legislation, as it will vary from case to case, depending on the complexities and volume of content under consideration.

Such a standard would allow platforms to provide the necessary human oversight, seek guidance, and consult legal doctrine before making a considered decision to remove content. The legislator could provide guidance in a recital or in the explanatory remarks of the law, for example, that “under normal circumstances it can be expected that a platform blocks illegal content 72 hours after obtaining knowledge.” This provides helpful guidance for platforms without creating an inflexible standard, legal uncertainty, or incentives to over-block content.

Expansion in scope of services subject to cyberbullying and cyber abuse schemes

The discussion paper proposes that both the existing cyberbullying and the cyber abuse schemes be expanded and that an increased range of services be subject to any new expanded obligations. We suggest that these schemes are limited in scope to content sharing services, like social media and video sharing services, which have the principal purpose of helping people to store and share content with the public or other broad audiences, over which the platform provider does not have editorial responsibility.

Section 474.17 of Australia’s Commonwealth Criminal Code already prohibits using a carriage service (defined as a service for carrying communications) to menace, harass or cause offence. Arguably this existing provision of the Criminal Code already prohibits using ‘electronic services’ and ‘designated internet services’ to bully and abuse others.

Addressing illegal and harmful content

There is a broad range of content contemplated within this section of the discussion paper, with illegal child sex abuse material at one end of the spectrum and legal adult oriented content at the other. The paper proposes extending the existing content regulation schemes codified in Schedules 5 and 7 of the Broadcasting Services Act 1992 through industry codes to cover a broader range of service providers irrespective of whether content was being hosted within Australia.

Google has longstanding escalation processes in place whereby any law maker or regulator can notify us that we are either hosting content that is illegal under Australian law. When it comes to removing web pages from ancillary services like Google Search, we are strongly guided by local law and decisions from the courts. Our approach is based on the belief that, when it comes to questions about what information should be stripped from public availability, those lines are better drawn by the rule of law than by Google. For content that Google hosts, we also have terms of use or community guidelines that we enforce robustly and that prohibit certain types of abusive content.

To the extent that Australia develops additional codes with respect to online content, we think it is prudent to focus on principles based codes that address categories of content and that such codes should be limited to content sharing services. The inclusion of private messaging services within such codes appears at odds with the public policy intentions to (a) limit the widespread distribution of content, and (b) prevent inadvertent exposure of harmful content to children and young people. We suspect that a strong justification would be needed in order to interfere with private communications between citizens.

Furthermore, it is important that any new regime recognise the clear distinction between content that is illegal under Australian law and content that is not illegal but may be considered harmful and we encourage the Government to consider plugging any gaps with new laws that prohibit the distribution of certain types of harmful content. We would prefer to make decisions to remove harmful content based on the certainty of codified Australian laws rather than the discretion of a regulator.

Finally, codes that require an assessment of individual pieces of legal content to determine whether or not it is harmful risks creating a significant burden on both the Office of the eSafety Commissioner and industry due to the sheer volume of cases that could foreseeably need to be considered.

By way of background, tackling illegal abuse material is a top priority for Google. We devote significant resources—technology, people, and time—to detecting, deterring, removing, and reporting child sexual exploitation content and behaviour. Google was a founding member of the Technology Coalition when it was founded in 2006, and our involvement in this Coalition of like minded organisations enables us to further scale our work in developing technical and operational solutions to prevent the distribution of child sex abuse material.
Since 2008, we've used "hashing" technology (unique digital fingerprints) to tag known child sexual abuse images, allowing us to identify copies of those images which may exist elsewhere. We also created a shared industry repository of video hashes which allows known child abuse videos to be identified and blocked, allowing other companies to remove known content from their platforms. Along with removing child sexual abuse material from Search, in 2013, we made changes to the Google Search algorithm to further prevent images, videos and links to child abuse material from appearing in our search results. We've implemented this change around the world in 40 languages.

As a technology leader, we understand we have a constructive role to play when it comes to this issue so we have made our new technology - Content Safety API - available for free to charities, industry partners, and other tech companies. This tool prevents further views of abusive material and protects content reviewers by prioritising and selecting content for review, meaning illegal material is reviewed faster and fewer people are exposed to it. We haven't waited for regulation to overcome these issues, we've created new technology, hired experts and specialists, and ensured our policies are fit for the evolving challenges we face online. Our work has the most impact when companies, government and communities work together.

Similarly, in the context of terrorist content, Google has made significant investments over the past few years on both technical and human resources to support our efforts to detect, review and remove illegal content from our platforms. We are also a founding member of the Global Internet Forum to Counter Terrorism (GIFCT) which, amongst other things, has established a hash sharing database of known terrorist content that is shared amongst GIFCT members to proactively detect and remove matching content. There are now over 200,000 distinct pieces of content within this database.

We also recognise that there are risks associated with inadvertent exposure of legal but inappropriate content to children and young people. We have built products for kids and families from the ground up to help parents and educators support safer experiences for their children and students. Most notably, Family Link is available by default on the latest Android operating system and helps parents stay in-the-loop as their child explores the internet on a compatible device. The app lets parents set digital ground rules for their family, like managing the apps their child can use, keeping an eye on screen time, or setting a bedtime and daily limit for their child's device. YouTube Kids also provides a separate YouTube experience designed especially for children, which parents can control. YouTube Kids, which is available as a standalone app, on living room devices and as a website, provides access to selected family-friendly YouTube videos, allowing children to explore a catalog of content in a more contained environment. Within YouTube Kids there are no social features like commenting on or uploading videos. In addition, parental control tools allow families to hand-select all of the content their children watch, or to choose content from third-party collections assembled by experts. For more information on our work, you can visit https://safety.google/families/.

Information on, and accreditation scheme for, Opt-in tools and services to restrict access to inappropriate content (filters)

This is a rapidly evolving area, and beside the considerable resources it would entail to set up and administer, we are concerned that an accreditation scheme might be too slow to accommodate users' varied needs and the diverse array of services they enjoy. In our experience, at the scale at which we operate, it is impossible to reliably identify, filter, and block all illegal content efficiently and without also blocking legitimate content. In many of these cases, context and external collaboration is essential to evaluating the legality of content, and our automatic tools are not as precise nor as adept at understanding context and nuance to filter content reliably.

If there is to be an accreditation scheme around filtering tools, it should follow a proportionate and reasonable approach. In particular, accreditation should maintain some flexibility, so as to let users benefit from any progress and innovative processes devised by providers in this space.

Blocking measures for terrorist and extreme violent material online

We are committed to ensuring we are doing all we can to fight the hatred and extremism that lead to terrorist violence. YouTube and Google hosted products have policies against violent extremism and prohibit terrorist groups from posting any content and we've been working along several key pillars to improve our efforts: (1) better detection and faster removal powered by machine learning; (2) collaborating with expert partners to help identify violative content; (3) counter-messaging; and (4) working with large and small platforms through the industry's Global Internet Forum to Counter Terrorism. Our significant investment in fighting this type of content is having an impact, ensuring it is removed before being widely viewed.

As part of these efforts, we work constructively with governments. This includes our work with Australia's Taskforce to Combat Terrorist and Extreme Violent Material Online, as well as our work as one of the founding signatories of the Christchurch Call to Action. We continue to make progress against its commitments. For instance, YouTube recently co-organised a crisis prevention workshop with the New Zealand Government in Wellington, NZ, attended by participants from around the globe, during which the Christchurch Call's principles and protocols were further refined. Also, the Global Internet Forum to Counter Terrorism (www.gifct.org) recently evolved into a full-fledged legal entity which will soon have dedicated staff, so as to further improve coordination and deliverables under the initiative. Google is and will remain one of the leading supporters of this important effort.

We believe that appropriate legislative instruments already exist to address this type of content efficiently. However, if the proposal for additional legislative provisions were to go ahead, then it should be limited to blocking with respect to specific offending material, within a context that makes that material unlawful, and we would also urge the utmost caution, in the public interest, to ensure that blocking entire sites is not overly broad. After all, sites that host user generated content may feature a wide variety of material; blocking access to the site due to terrorist or extreme violent material may also render significant legitimate material inaccessible.

In turn, such powers should be narrowly tailored to address only those 'worst of the worst' platforms and services that willfully and systematically fail to respond to valid legal removal requests regarding specific items of identified content. The law should also focus on truly worst-of-the-worst content that would clearly constitute criminal offenses.

Ancillary service provider notice scheme

Any scheme that applies to “ancillary service providers” should specifically focus on notice-and-takedown of specific illegal material, as opposed to more vague concepts like ‘de-ranking’ harmful online content that is otherwise legal under Australian law. It is essential in a democratic society that people are able to access content that is otherwise lawfully available. Search engines play an important role in facilitating access on the internet, enabling people to access lawful information, impart and disseminate it.

Placing restrictions on the types of content that can be accessed through search engines would interfere with the right of freedom of expression, and the extent to which people can access and hear different views, as well as share their own. We believe a legal removals framework is the right way to approach content removals for ancillary services, including Search engines, and have a well ironed out process for taking action on legal removal requests. Today, upon request, we will review and remove links to specific illegal content in Search, and do so for millions of removal requests per year. Any additional notice scheme for search engines, app stores, or other ancillary services should build on this existing notice-and-takedown framework for illegal content.

Governance arrangements

In contemplation of extending the powers of the eSafety Commissioner, including by granting her quasi judicial and legislative powers to determine what constitutes a harm and issue sanctions, the Commissioner’s consultation of experts and stakeholders needs to be formalised, mainstreamed and ongoing.

In addition, oversight of the eSafety Commissioner’s Office should arguably be strengthened in accordance with any significant increase in powers, such as by setting up a multi-stakeholder body overseeing it and its decisions, especially to ensure the proper balance and respect of rights such as freedom of expression and opinion.

Irrespective of the Government’s preferred approach, we believe that there are a number of central principles which should be considered:

- True regulatory independence: we believe it is very important that any oversight body in this area is truly independent. It is important that, in circumstances where the Government is proposing to issue instructions over the content of new codes or practice (such as for terror content or CSAM), adequate protections are in place to ensure that this independence is not threatened. While an oversight body’s remit and powers should be clearly defined, it should be required to consult on the best ways of issuing guidance and codes of practice in order to ensure they are technology driven, platform agnostic, operationally sustainable and create a clear path to compliance for the platforms involved.

- Consultation with companies, experts, and other stakeholders: An oversight body would ensure that experts are consulted, and that any code or decision / determination is subject to an economic or human rights impact analysis. This would ensure that the requirements of the codes or decisions are technically viable, based on evidence of actual levels of harm, and economically and legally feasible.
- The Government could establish a multi-stakeholder forum involving representatives from companies and other relevant stakeholders, which could provide direct expertise from the field to make the regulator's decisions more effective and up-to-date with the existing social, legal and technological environment. The newly formed eSafety Advisory Council (formerly the Online Safety Consultative Working Group) could serve this purpose.
- Establish a formal industry board: To ensure industry is properly consulted, we propose the establishment of a Forum with representatives from industry, including companies of many different sizes. The Forum would provide input to appropriate codes of practice, and help set best practice for industry. To give the public confidence in the robustness and independence of this process, the minutes from meetings of the Forum could be published publicly. We understand that the Commissioner informally engages with many different industry organisations; perhaps this engagement can be formalised as a broader industry stakeholder forum?
- Reasonable expectation of ability to comply: Companies covered by the scope of the new framework also need a reasonable expectation that they can comply with any proposed regulation (for example, by avoiding mandating the use of technological solutions that would be inappropriate for some services or harms - as we explain earlier in this response).