GOOGLE'S SUBMISSION TO THE OIREACHTAS COMMITTEE ON JUSTICE AND EQUALITY

HEARING INTO:

ONLINE HARASSMENT, HARMFUL COMMUNICATIONS AND RELATED OFFENCES

INTRODUCTION

Google welcomes the opportunity to contribute to the Committee's deliberations on the topic of online harassment, harmful communications and related offences.

A range of governments, tech platforms, and civil society are focused on how best to deal with illegal and problematic online content. There's broad agreement on letting people create, communicate, and find information online, while preventing people from misusing content-sharing platforms like social networks and video sharing sites. We recognise that there can be a troubling side of open platforms, and that bad actors have exploited this openness. We take the safety of our users very seriously, and we are committed to ensuring that inappropriate content that appears on our platforms is dealt with as quickly as possible.

Now 21 years old, Google has grown from a small start-up to a global company with legal obligations in each of the countries in which we operate. We work hard to protect our platforms from abuse and have been working on this challenge for years, using both computer science tools and human reviewers to identify and stop a range of online abuse, from "get rich quick" schemes to disinformation, to the utterly abhorrent, including child sexual abuse material (CSAM) online. We respond promptly to valid notices of specific illegal content, and we prohibit other types of content on various different services. A mix of people and technology helps us identify inappropriate content and enforce our policies, and we continue to develop and invest in smart technology to detect problematic content hosted on our platforms, which is driving progress.

As well as making significant investment in technology and human resources, we have engaged with policymakers in Ireland and around the world on the question of the appropriate oversight for online content sharing platforms. We are keen to work constructively with government to build on the existing legal framework and to build trust and confidence in the systems and procedures that ensure online safety. We support all efforts by legislators and government to engage with stakeholders to consider appropriate protection, remedies and forms of redress for individuals who are the victims of online harm.

Having considered the Committee's issues paper, our comments are directed towards those aspects that concern the role of internet service providers in preventing online harassment and certain harmful communications. In particular, we will address the Committee on:

- The role that regulation of internet service providers can play in combating online abuse (Issue 13).

- The legislative models adopted or proposed in Australia, New Zealand and the UK (Issue 2).
- The definition of "communication" contained within the Online Harassment, Harmful Communications and Related Offences Bill 2017 (“the Bill”) (Issue 2).
- The civil disclosure orders proposed by Section 14 of the Bill.

Our comments build on our response to the public consultation on the regulation of harmful content on online platforms and the implementation of the AVMS Directive that was conducted by the Department of Communications, Climate Action and the Environment earlier this year, and previous testimony to the Oireachtas Committee on Communications, Climate Action and the Environment on online safety and related matters.

PRINCIPLES FOR OVERSIGHT OF CONTENT-SHARING PLATFORMS

Google is supportive of carefully crafted and appropriately tailored regulation that continues to address the challenges of problematic content online. We believe that tackling this problem is a shared responsibility in which companies, governments, civil society, and users all have a role to play. We are keen to work constructively with legislators to build on the existing legal framework and to build trust and confidence in the systems and procedures that ensure online safety. We suggest a number of central principles that should be considered for approaching oversight of content-sharing platforms and problematic content online, including:

- **Clarity** - Content-sharing platforms are working to develop and enforce responsible content policies that establish baseline expectations for users and articulate a clear basis for removal of content as well as for suspension or closure of accounts. In this regard, it is also important for governments to draw clear lines between legal and illegal speech, based on evidence of harm and consistent with norms of democratic accountability and international human rights. Without clear definitions, there is a risk of arbitrary or opaque enforcement that limits access to legitimate information.
- **Suitability** - It’s important for oversight frameworks to recognize the different purposes and functions of different services. Rules that make sense for social networks, video-sharing platforms, and other services primarily designed to help people share content with a broad audience may not be appropriate for search engines, enterprise services, file storage, communication tools, or other online services, where users have fundamentally different expectations and applications. Different types of content may likewise call for different approaches.
- **Transparency** - Meaningful transparency promotes accountability. We launched our first Transparency Report more than eight years ago, and we continue to extend our transparency efforts over time. Done thoughtfully, transparency can promote best practices, facilitate research, and encourage innovation, without enabling abuse of processes.
- **Flexibility** - We and other tech companies have pushed the boundaries of computer science in identifying and removing problematic content at scale. These technical advances require

flexible legal frameworks, not static or one-size-fits-all mandates. Likewise, legal approaches should recognise the varying needs and capabilities of start-ups and smaller companies.

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

- **Cooperation** - International coordination should strive to align on broad principles and practices. While there is broad international consensus on issues like child sexual abuse material, in other areas individual countries will make their own choices about the limits of permissible speech.

EXISTING LEGAL FRAMEWORK

In framing any measures in this area, it is important for legislators to have regard to and build on the existing legal framework. We operate in an environment where extensive regulation of online content and actions already exists and is being enforced. Many laws, covering everything from consumer protection to defamation to privacy, already govern online content. Much of the regulation for conduct online is equivalent to that which applies to offline conduct, with some additional protections applicable specifically to the online context. From consumer rights legislation to the new EU Audiovisual Media Services (AVMS) Directive, online behaviours come under the scope of a diverse and evolving set of legislation, multi-stakeholder initiatives and regulators.

Safe harbours and "Good Samaritan" laws for online platforms support the free flow of information, innovation, and economic growth, while giving platforms the legal certainty they need to combat problematic content. Over the internet's history, many countries have also developed codes of practice, for example those like the European Union's Code of Conduct On Countering Illegal Hate Speech and Code of Practice on Disinformation.

In addition to legal regulations, we have over the years developed extensive community guidelines and content policies that offer clear rules on what we don't allow on our platforms. These often go above and beyond the law and we employ thousands of staff around the world, working 24 hours a day to ensure violations are acted upon. Companies have also worked together to address these challenges, for example with the Global Internet Forum to Counter Terrorism, a coalition sharing information on curbing online terrorism.

We continue to improve on our processes and our technology to enforce those rules. We work diligently to protect our community across our other platforms by continuously reviewing and updating our policies based on new trends and investing in new machine learning (ML) technology to scale the efforts of our human moderators.

ML is helping us detect potentially violative content and surface it for human review. For example, YouTube removed 9 million videos during the second quarter of 2019, 7.9 million of which were first flagged through our automated flagging system. Of those videos, 81.5% had no views at the time of takedown.

In addition to enforcing our community guidelines, we have processes in place for the removal of content that violates a local law. Once we are on notice of such content, we are able to restrict access to it from the relevant jurisdiction.

REGULATION OF INTERNET SERVICE PROVIDERS AS A MEANS OF COMBATTING ONLINE ABUSE

In Google's role as an internet service provider, we provide tools and platforms for users to search for, create and share content online. Our role is distinct from the activities of authors, publishers and editors of content.

The distinctive role played by internet service providers is reflected in the EU legislation that underpins the regulation of electronic commerce in Europe, namely the E-Commerce Directive (Directive 2000/31/EC) as implemented in Ireland by the European Communities (Directive 2000/31/EC) Regulations 2003 (SI No. 68 of 2003).

The E-Commerce Directive provides that:

- In order to trigger a removal obligation on the part of the service provider, a user must properly notify the service provider of the content which is unlawful and provide it with the necessary information to identify the content and to determine whether or not it is unlawful.

- The service provider will only be exempt from legal liability where it has expeditiously removed properly notified unlawful content.

- Service providers cannot be placed under a general monitoring or active searching obligation.

The E-Commerce Directive framework provides strong incentives for service providers to establish and operate efficient notice and take down procedures. A service provider that does not operate such procedures will be exposed to potential legal liability for unlawful content hosted on its platform. Many service providers, including Google, have developed extensive infrastructures which provide efficient tools for the reporting and removal of illegal content.

The E-Commerce Directive has the advantage of setting out different requirements for different types of internet intermediaries, rather than being aimed at a particular business activity. It has led to the growth of a wide variety of services and business models, and is flexible enough to cover the multiplicity of activities and content types online. For example, an online news site can contain content authored by the news organisation, along with material licensed from third parties and also user-generated comments - the news site will be directly responsible for the editorial content it publishes, but will have different legal responsibilities with respect to user comments that the website is hosting as an intermediary. This online intermediary liability regime has fostered the huge economic and cultural benefits of the internet while ensuring platforms are taking appropriate and speedy actions in removing unlawful content.

The E-Commerce Directive also circumscribes the role of EU Member States in regulating service providers' notice and take down responsibilities. Within the context of the E-Commerce Directive, national regulatory bodies may be designated a role in determining what constitutes unlawful content in cases where there is disagreement between the service provider and a user. This adjudication function is typically carried out by courts, although in some instances specialist administrative bodies play a role.

APPROACH IN OTHER JURISDICTIONS

Google operates across many different jurisdictions, and we note that the Committee is interested in learning of lessons from other countries, including, Australia, New Zealand and the United Kingdom, where specific legislation for dealing with online harassment and harmful communications has been introduced or proposed.

As an initial point, we note that the Australian and New Zealand are not subject to overarching EU law requirements which apply in Ireland. In fact, in framing their laws, these countries started from a position where there was limited existing statutory rules applicable to notice and takedown procedures.

Australia

Australia has enacted the Enhancing Online Safety Act 2015 that has the aim of combating serious cyber-bullying involving children. A statutory body, the Office of the e-Safety Commissioner (**OESC**), has also been established for the purpose of enabling the reporting of serious cyber-bullying against children.

The OESC administers complaints in respect of communications on social media platforms but will respond differently depending on whether the complaint relates to a Tier 1 entity or a Tier 2 entity. To be regarded as Tier 1, a company must comply with a number of specified organisational conditions (e.g. ensuring there is a complaints scheme in place to allow users to request the removal of content); and apply to the OESC for Tier 1 status. The OESC also monitors on-going compliance with the Tier 1 conditions. Tier 2 companies are larger social media services formally designated by the Australia Government.

Where the OESC receives a complaint in relation to material on a Tier 1 company's platform, it will engage in cooperative dialogue with that company. If the OESC finds that the relevant material amounts to cyber-bullying, it will merely request the removal of that material.

If the OESC receives a complaint in respect of a Tier 2 company, and it finds that the relevant material amounts to cyber-bullying, it will send a formal notice to remove the harmful material. Failure to comply with such a notice within 48 hours will result in a fine. In practice, the OESC treats all partners (irrespective of whether they sit in Tier 1 or 2) the same and has not had to issue any formal notices.

Google works closely with the OESC and meets all of the basic safety requirements under the Act. We have never received a formal notice under the Act. We have only received a small number of

informal take down requests and have always complied. We agree with the recommendation made by the independent review of the legislation to do away with the tiering within the legislation - this should be a relationship fuelled by cooperation and underpinned by legislation.

New Zealand

New Zealand has enacted the Harmful Digital Communications Act 2015 (**HDCA**) as a means of combating "harmful digital communications".

The HDCA created a two-step system to provide for civil remedies for harmful digital communications, with a focus on engaging with alternative dispute resolution (**ADR**) processes before going to court. Complaints are initially made to a government approved agency which attempts to resolve the complaint through ADR. Should ADR fail an affected user may then seek remedies from the District Court.

The HDCA provides for a safe harbour period for online content providers whereby no civil or criminal proceedings may be brought if, after receiving a complaint from a user, the internet service provider engages with the author of the original material and the author consents to the material's removal. If the author does not respond or cannot be contacted within 48 hours of the complaint, the material must be removed or the internet service provider may be liable for the content. If the author objects to the removal, the internet service provider must leave the content online and inform the complainant of the objection. The safe harbour also supports the policy intent of the HCDA to place the primary focus on the actual protagonist or author, whose behaviour is at issue.

Netsafe, as the approved agency under the HDCA, provides a swift and effective service for people experiencing harmful digital communications and as a non-profit organisation with the government contract to be approved agency, provides confidence to complainants that their personal information (e.g. sextortion) is unlikely to be shared with wider government.

Another aspect of the approved agency worth noting is that it is not a regulator but is required “to establish and maintain relationships with domestic and foreign service providers, online content hosts, and agencies (as appropriate) to achieve the purpose of this Act” (section 8, HDCA). Netsafe’s focus on building effective relationships, rather than regulatory compliance, was critical to their ability to help remove violative content following the Christchurch shooting.

The HCDA also provides an appropriate escalation path - supporting parties to resolve an issue themselves, then with the help of Netsafe, and then ultimately to court. This supports swift solutions for people and reduces the cost of going to court.

United Kingdom

The UK Government published a White Paper on addressing online harms in 2019, and opened up a consultation process for feedback. The White Paper remains a proposal, and has not been enacted into law.

The White Paper proposes establishing in law a new "duty of care" towards internet users. This duty of care will be overseen by an independent regulator who will set out how companies can meet this

duty of care in codes of practice. The regulator will also have powers to take enforcement action against companies that have breached their statutory duty of care.

We believe the approach the UK Government has proposed within the White Paper has some advantages. In particular, we support the principle of oversight that looks at systemic risks to users, rather than individual pieces of harmful content. However, areas of legal uncertainty (especially around the 'duty of care' concept) and impact on user rights still need to be addressed.

We note that some of the proposals set out in the UK White Paper are inconsistent with the framework introduced by the E-Commerce Directive by requiring measures that would impose a general monitoring obligation on internet intermediaries. The White Paper broadly applies the same responsibilities on companies for both illegal content and "legal but harmful" content, which creates a conflict with rights to free expression and access to information. Where legislators believe a category of content is sufficiently harmful based on research and evidence, they may make that content illegal, through democratic processes, in a necessary and proportionate manner. Alternatively, the focus should be on companies enforcing their own community guidelines which often go above and beyond the legal requirements.

The proposed regulatory framework also applies equally to all companies across all sectors of the online economy. This broad and indiscriminate approach does not take into account the fundamentally different roles played by internet service providers. In addition, it is important that the obligations and expectations placed on companies must also reflect the resources available to them. For example, it would be unreasonable and overburdensome to develop a framework that requires the same machine learning techniques adopted by Google to also be required to be adopted by a new e-commerce start up.

DEFINITION OF "COMMUNICATION" AND INSCOPE COMMUNICATION CHANNELS

The issues list asks whether there is a need to expand the legislative definition of "communication" to cover online and digital means of communication.

The current legislation setting out the offence of harassment is the Non-Fatal Offences Against the Person Act 1997. Section 10 of that Act prohibits harassment "*by any means*" of communication, which is assumed to cover both electronic and non-electronic means of communication. It is not clear therefore that the revised definition of "communication" contained in Section 2 of the Online Harassment, Harmful Communications and Related Offences Bill 2017 is required to address any shortcoming in the existing legislation.

However, the definition of "communication" brings into focus an important issue that should be recognised in legislation that has the aim of creating offences to deal with online abuse or harmful communications. While offences such as harassment (Section 3) or distribution of intimate images (Section 4) will need to cover a wide and technologically neutral range of communications by potential offenders, care needs to be exercised in ensuring that such offences do not inadvertently criminalise communication service providers, be they traditional telecommunication providers or internet service providers.

For example, under Section 4, it is proposed that any person that "distributes" an intimate image may be guilty of an offence. It is assumed that the intention is not to make communication service

providers criminally liable for the acts of individuals that send intimate images. To avoid any ambiguity of this nature, new offences should be carefully and precisely framed.

PROPOSED CIVIL DISCLOSURE ORDERS

For data protection reasons, internet service providers will usually need to be served with a court order before they can disclose a user's subscriber information to a potential civil claimant (e.g. in a defamation claim concerning content posted to a social media platform by a user).

It is already well-established in case law that intended claimants can apply to the High Court to obtain a so called "Norwich Pharmacal Order" (**NPO**) for the purpose of requiring an internet service provider to disclose the identity of a user of its service where the provision of this information is necessary to aid the intended claimant to pursue a civil claim.

In light of the existence of this well-established common law remedy, the intent behind Section 14 of the Online Harassment, Harmful Communications and Related Offences Bill 2017 is unclear. The proposal to confer jurisdiction on the Circuit Court to grant NPOs may itself be welcome, however, the necessity for placing NPOs on a statutory footing is open to question. Furthermore, Section 14 as drafted raises a number of questions and concerns. In particular:

- S. 14(1) provides that court orders may be made in respect of "digital service undertakings". This term has not been defined and its exact parameters are unclear.

- S. 14(2) envisages that the service provider may be ordered to disclose the "name" and "address" (which may be a "digital address") of a user. Service providers typically are not required to verify the identity information that is supplied to them by users. Therefore it should be made clear that a service provider can only be ordered to disclose the subscriber identity information that has been provided to it by the specified user. The concept of a "digital address" should also be clarified.

- S. 14(2) provides that the NPO can be granted where there is a claim in respect of "communications" that are unlawful by virtue of being "*abusive, threatening, offensive, false, defamatory or an invasion of a person's privacy*". This could be interpreted as suggesting that all abusive and false communications are unlawful, and would thereby greatly expand the circumstances in which NPOs could be granted. It should be clarified that NPOs can only be granted for the pursuit of civil claims for alleged unlawful activity. It should be for the court to decide whether a proper claim of unlawfulness is made out by a claimant. The publication of mere offensive or abusive comments should not justify a NPO being granted - in fact the creation of such a remedy may pose a risk to user privacy rights and could have a chilling effect on free expression.

- S. 14(3) provides that the court may "order" the service provider to serve notice on a user to allow the user to appear and make representations in Court. It would be preferable to have it stated in legislation that a service provider may also at its own discretion notify the user about the order within a reasonable period prior to the date on which the disclosure under the order is required. This would remove any doubt as to whether court approval is required.

- S. 14(4) envisages the possibility that a NPO granted against one service provider can be served on, and be automatically binding on, another service provider. The rationale for this proposal is not clear, and it raises questions over due process safeguards for disclosing personal data.

- The legislation is silent on the issue of costs. Existing case law provides that the applicant should be required to pay the service provider's costs of complying with a NPO. Although service providers may not necessarily pursue their costs, this is an important economic protection and it can also deter unmeritorious applications, which prejudice user privacy rights. Consideration should therefore be made to include an "applicant pays" costs provision in the legislation.

CONCLUSION

We wish to thank the Committee for providing us with the opportunity to contribute to the Committee's deliberations on the topic of online harassment, harmful communications and related offences. Addressing problematic content is a shared responsibility across society, in which companies, governments, civil society, and users all have a role to play, and it is appropriate that the Committee is hearing from a variety of voices on this topic. We hope the Committee will give our suggestions for approaching oversight of content-sharing platforms due consideration and look forward to further engagement over the time ahead.