Written evidence submitted by Google UK (OSB0175)

Executive Summary

Thank you for the opportunity to submit evidence to the Committee as part of the scrutiny process of the draft Online Safety Bill. This is an important, landmark, piece of legislation and Google is keen to work with parliamentarians, the Government, and Ofcom to make the new regulatory regime a success.

At Google, our mission is to organise the world's information and make it universally accessible and useful. We believe deeply in technology's ability to unlock creativity and engagement, but we also understand the responsibility we have to keep our users safe. Ensuring our platforms are used responsibly and people have the tools and knowledge they need to make responsible choices online is at the heart of everything we do—and our investment and innovation have often put us at the forefront of positive industry change in this area.

We strongly support the objectives of the draft Online Safety Bill to enhance the safety of UK citizens online and to protect users' fundamental rights of freedom of expression and privacy. We believe that the Bill is an opportunity to create an enduring regulatory system that gives platforms and users the clarity and flexibility that is needed to effectively keep people safe online. We hope that our experience in tackling harmful content online, over many years, can be useful to the Committee in its examination of the draft Bill.

We also recognise and welcome the Government's broader commitment "to using digital technologies and services to power economic growth across the entire UK".[1] The open internet has had a transformative and positive impact on society. Online services play a central role in the exchange of information and ideas. People use Google Search to help them study, find a job, and discover local businesses. They come to YouTube to express and entertain themselves. These services are open and anyone with an internet connection, no matter where they live or their background, can benefit from them.

It is this openness of the internet and the possibilities it has offered for innovation that have enabled the creation of these services and the huge social and economic value they have produced. In 2020, Google's search and advertising tools helped support an estimated £55 billion in economic activity for over 700,000 businesses in the UK. On an average day in Britain, 3.6 million people will use Google Search to look for a job, whilst YouTube makes a valuable contribution to the UK's thriving creative ecosystem, adding £1.4 billion to GDP and supporting 30,000 jobs in 2019.[2]

[1] Online Harms White Paper: Full response to the consultation, https://www.gov.uk/government/consultations/online-harms-white-paper/outcome/online-harms-white-paper-full-government-response.

Google

Our response to the draft Bill

The draft Bill represents an important step forward in the ongoing and multifaceted debate over how best to keep people safe online. This is an ambitious and complex piece of legislation, and it is important to get the details right. The right framework can deliver on the Government's objective to make the UK the safest place in the world to go online.

We agree with the Government that a risk-based approach is the best way to get to this outcome. A focus on the services and types of content that pose the greatest risk of harm to UK users, especially children, will concentrate efforts. The focus of regulatory attention on platform processes can promote innovative, service-specific steps to protect users. A systemic approach to platform accountability is the only way the regulatory framework can operate at the scale of today's internet.

The draft Bill's clear goal to robustly protect freedom of speech and privacy online reflects the importance of British users' fundamental rights. The legislation should make use of every available safeguard to minimise inadvertent restrictions on legitimate freedom of expression.

As a general principle, we believe platforms that host content should be free to set their own guidelines on the legal content they will and won't host, as we have done on YouTube. For example, YouTube guidelines do not allow users to post misleading content relating to COVID-19 that may cause egregious harm - even where such content is lawful.[3] Regulation should not seek to set the rules for what lawful content should remain online and what platforms should remove – what is legal offline should remain legal online. Platforms should have clear terms for harmful content, including how "harmful" is defined and determined, and this is the approach we will continue to maintain. If the Government believes that a category of content is sufficiently harmful that it should not be available online, we believe it should make that content illegal directly, through transparent, democratic processes.

We recognise the concerns that the Government and Parliament have on specific categories of lawful content. We haven't waited for legislation; we have developed our own guidelines and taken action. We understand the sensitivity and importance of these areas and have devoted careful attention to developing an approach that limits harm while protecting users' ability to express themselves online.

We think that it is vital that the Bill retains an approach that is based on encouraging platforms to clearly set out how they deal with harmful content, and that holds them accountable for delivering on their commitments.

---

[2] Public First, Google's Impact in the UK 2020, https://googleimpactreport.publicfirst.co.uk/uk/.
[3] See COVID-19 medical misinformation policy, https://support.google.com/youtube/answer/9891785?hl=en-GB.

Enshrining the principle that regulatory requirements should be proportionate is also important. Proportionate regulation supports a thriving digital economy and helps the UK's reputation as a global leader. The Government's aspiration to ensure that regulation does not negatively impact the value that digital services have brought to British citizens and the economy represents an important foundation for the debate.

Achieving the best outcomes for British society

We share the Government's objective of reducing both exposure to harmful content and levels of harm experienced by individuals online. We believe that the draft Bill can be strengthened by providing a more solid foundation for Ofcom, through simplification and greater clarity on how online services can keep users safe while simultaneously protecting their ability to access services and information. Without amendments to the bill that provide more clarity, there is a risk that the legislation does not fulfil its potential to create a safer online environment, and may inadvertently erode the ability of services to protect freedom of speech and privacy.

We look forward to working with parliamentarians and the Government to ensure that the final legislation will succeed in delivering our shared objectives. To do this, we believe the following points should be considered by the Committee:

a) Simplification of the new regime, providing greater legal clarity and flexibility for services to improve effectiveness

As it stands, the framework set out in the draft Bill will create arguably the most intensive regulatory environment for online services of any democratic country in the world.

The complexity of the draft Bill and the lack of legal clarity it provides to online services risks leading to uncertainty that detracts from quick and effective action to protect users. Some definitions are unclear and, with substantial details being left to Ofcom, a real operational burden will be placed on the regulator to implement this Bill in full, which could further delay a compliance regime being put in place. Further, many complexities — such as the offences that will constitute "priority illegal content" or the types of harm that will constitute "priority" and "primary priority" harmful content — will not be addressed by Ofcom and give considerable discretion to the Secretary of State.

We urge the Committee to explore how simplicity and clarity in the Bill can be achieved to help ensure the legislation is effective.

b) Targeted and better defined provisions that reduce the risk of over-removal

Despite its objective to protect free speech, without changes the Bill is likely to result in service providers over-removing content at scale because of the following issues:

- The range of content to which the Bill applies is broad, difficult to understand, and often depends on a complex assessment of context that cannot be applied at scale. This may lead to legitimate content being caught by the Bill in ways that were not intended.
- It is unclear how the overlapping matrix of competing duties are to be balanced. Faced with substantial penalties for failing to fulfil their safety duties, companies will be incentivised to simply remove legitimate content rather than to protect rights to free expression or privacy and risk enforcement action.
- Provisions in the Bill may amount to requirements for blanket monitoring and proactive removal of content, which can only be achieved at scale using automated tools. The Bill will lead companies to rely on imperfect automated tools even at the risk of blocking lawful content.

As a starting point, British citizens should be able to freely express and consume legal content as they communicate online. **The Committee can strengthen the draft Bill by pushing for further clarity on the content in scope and the obligations relating to it. This clarity will help guide effective and accurate decision-making from platforms.**

c) Clarifications on the duties to younger users to avoid blanket implementation of age-gating

Keeping children safe on our services is of the utmost importance to us. We have a comprehensive approach towards protecting children that has included progressively making substantive updates across our range of services.[4]

The Age Appropriate Design Code has come into force and is delivering against its remit to drive positive changes for children online. At its heart is a recognition that policy interventions must support and balance all the rights of children, both to protect them from harm but also to allow them access to the digital world and support their development.

This principle should also underpin the Bill's approach to stronger child protection. Yet in its current form, the draft Bill could inadvertently result in a majority of in-scope services being age-gated and making them unavailable to children, even if they do not pose a risk to children. Age gating may in some circumstances have a role to play in protecting children from content that is inappropriate for them. But the Bill's broad duties, combined with significant enforcement risks, may result in service providers concluding that compliance is most easily or reliably achieved by age gating in all cases, rather than working to improve their services to make them safe for children. This would not be a good outcome for British children, who would be unduly restricted in their ability to participate in the digital world and miss out on the personal and developmental benefits that this can bring.

[4] See, for example, our latest set of updates, Giving kids and teens a safer experience online, 10 August 2021, https://blog.google/technology/families/giving-kids-and-teens-safer-experience-online/.

Age-gating would also result in adult users having to verify their age in order to access online services, with serious consequences for privacy and the possibility of large numbers of British users being locked out of online services because they do not have the required documentation, or do not want to share documentation with the service. This is likely to have a disproportionate effect on vulnerable users, contributing to widening the digital gap.

To best protect children while retaining their access to online services, it is vital that the Committee carefully considers how the Bill's child protection duties can be simplified and aligned with the Age Appropriate Design Code and existing international standards like the EU's AVMS Directive.

d) Revised duties for search services that take into account the nature of the service

It is right that search services should be subject to expectations to protect users, but those expectations, both under the Bill itself and the Codes of Practice, must take into account the distinct nature of these services and be proportionate to the role they play in facilitating access to information.

While the structure of the Bill, and the Government's prior comments in its consultation responses, indicate an intention to adopt a different approach to search services, as compared to user-to-user services, we do not think there is a material difference of approach in the draft legislation.

Access to information is at the core of Google Search's mission, but we also have a strong commitment and responsibility to comply with local law and protect our users. When content is against local law, we remove it from being accessible in Google Search. We've made a number of investments over the history of Google Search to protect our users against harmful information online. Our approach is aimed at developing policies and tools that keep users safe while preserving their ability to access the full range of lawful information the internet has to offer.

Search engines are essentially indexes of the web and there are hundreds of new web pages published every second. Search engines do not host user-generated content like user-to-user services do and, given the scale of information that can be accessed through search results, it is particularly unworkable for search services to be expected to identify and evaluate whether individual pieces of content on trillions of web pages fall within the broad definitions of illegal and harmful content set out in the Bill.

The Committee could look at how the Bill can be strengthened so its approach to search fulfils the Government's stated policy intent to recognise the differences between user-to-user services and search services.

e) Ensuring ongoing industry and user engagement

The draft Bill provides considerable discretion for Ofcom in deciding how in-scope services should comply through Codes of Practice. We understand the intention is to ensure that the framework is future-proofed, and that Ofcom can adapt its approach as evidence of new harms emerge. While we have every confidence in Ofcom as a thoughtful, evidence-based and proportionate regulator, there is the risk that proliferating Codes of Practice create confusion for providers over how to prioritise the steps they should take and that they include requirements that are not technically workable. This risks making the framework less effective in strengthening user safety.

To ensure that the Codes of Practice are clear and effective in driving improvements from providers, the Bill should strengthen Ofcom's duty to involve all stakeholders impacted in the development and evolution of any Codes of Practice. This will ensure the Codes benefit from industry's technical expertise and experience in protecting users and that their requirements can be practically and swiftly implemented.

It is also important to consider the cumulative impact of different Codes of Practice on innovation in the digital economy. The Government's Plan for Digital Regulation rightly outlines the ambition to ensure that "new regulation minimises contradictions, undue burdens, or overlaps and gaps with existing frameworks".

We would also encourage the Committee to consider if Ofcom could be required to assess the regulatory burden of a new Code of practice on affected platforms and whether this is proportionate to the outcomes that will be achieved (which in turn should be evidence-based and carefully considered). This close dialogue between regulators and industry will help to achieve the Government's objective of ensuring that regulation is both effective in protecting UK users while remaining innovation-friendly.

We have provided a full response that addresses the Committee's Call for Evidence and develops and explains the critical points outlined above. We make a number of further specific suggestions to strengthen the Bill throughout the text of our response. Our full position is summarised in an Annex.

Detailed response

We have structured our response in line with the Committee's Call for Evidence, with sections aimed at addressing each part of the Committee's questions. We have underlined our main recommendations below and, for the sake of ease, have compiled them in an Annex at the end of the document. We have also started this response with a brief outline of how Google currently works to keep users safe online, meeting the duties that will be placed on our services by the new regulatory regime.

In each section, we include a number of sub-sections in which we aim to address specific questions asked by the Committee. We have reproduced those questions here for clarity.

1. Summary: How Google helps users stay safe online

The draft Online Safety Bill will place a number of duties on tech companies, including duties to protect users from content which may be illegal or harmful. At Google, our approach has never been to wait for new regulation before acting to keep our users safe and we have consistently made substantial progress in making our platforms hostile to those that try to abuse them. Below, you can find further details about the current approach we take to help our users stay safe online, across all our platforms.

1.1 Keeping our users safe

Google's approach to online safety has three elements: technological innovation, strong community guidelines and working in partnership with others. We are constantly looking to improve and evolve, introducing new policy changes, hiring thousands of new people dedicated to safety policy, and continuing to invest in technology to help us tackle illegal and harmful content at scale, such as Content Safety API and CSAI Match, which we make available to other companies to contribute to strengthen the ecosystem. We have also introduced a number of features to promote digital wellbeing and are working hard to ensure we understand how we can best protect our users by carrying out research and updating our products.

Across our platforms, products and services, we work to address harmful content based on four principles[5]:

- **Remove:** We set responsible rules for each of our products and services and take action against content and behaviors that infringe on them. We also comply with legal obligations requiring the removal of content.
- **Raise:** We elevate high-quality content and authoritative sources where it matters most.
- **Reduce:** We reduce the spread of potentially harmful information where we feature or recommend content.
- **Reward:** We set a high standard of quality and reliability for publishers and content creators who would like to monetise or advertise their content.

These levers allow us to be consistent in our methodology across products, while tailoring their implementation to fit the uses and needs of each product.

Across our platforms, we develop and maintain 'rules of the road,' which outline what types of content and behaviors are acceptable for each product or service. Known as 'content policies' or 'community guidelines,' we aim to make them clear and easily accessible to all users and content creators – whether those are video creators, webmasters, app developers, or advertisers. These 'rules of the road' articulate the purpose and intended use of a given product or service and represent a crucial part of what makes that product unique. They also explain what types of content and behaviors are not allowed, and the process by which a piece of content, or its creator, may be removed from the service. We regularly review and update those policies as new evidence emerges. On YouTube, for example, we have developed a robust COVID-19 misinformation policy over the course of the last year and a half. This policy evolved as scientific and medical consensus about the pandemic and related public health measures emerged.[6]

1.2 Protecting users on Search

Google processes billions of searches per day. In fact, every day, 15% of the searches that we process are ones that we've never seen before. Automation is how Google handles the immense scale of so many searches. Google uses automation to discover content from across the web and other sources. Automated systems – like our search algorithms – are used to surface what seems to be the most useful or reliable content in response to particular queries. Automation also helps power our SafeSearch feature, allowing those who wish to use it to help prevent explicit content from appearing in search results.

Automation is also generally Google's first line of defence in dealing with policy-violating content. Our systems are designed to prioritise what appears to be the most useful and helpful content on a given topic. Our systems are also designed not to surface content that violates our

[5] Susan Wojcicki, Preserving openness through responsibility, 27 August 2019, https://blog.youtube/inside-youtube/preserving-openness-through-responsibility/.
[6] For a more detailed explanation of our approach to promoting quality information and moderating content, see our Information Quality and Content Moderation White Paper, Google, https://blog.google/documents/83/information_quality_content_moderation_white_paper.pdf/

content policies

No system is 100% perfect. If our processes surface policy-violating content, we always look to resolve it by improving our automated systems. This allows us to better deal with both a particular issue that's been detected, and improve for related queries and other searches overall.

In some cases, we may also take manual action. This does not mean that Google uses human curation to rearrange the results on a page. Instead, humans are used to review cases where policy-violating content surfaces and take manual action to block this content, in the limited and well-defined situations that warrant this.

Our approach is aimed at developing policies and tools that keep users safe while preserving their ability to access the full range of lawful information the internet has to offer:

- **Our content policies for Google Search** specify that we:
  - Block search results that lead to child sexual abuse imagery or material that appears to victimise, endanger, or otherwise exploit children.
  - Remove certain personal information that creates significant risks of identity theft, financial fraud, or other specific harms.
  - Remove or demote spam, which we define as results designed to deceive users or game our search systems.

- **We have a separate set of policies for our Search features – which include, for example, auto-complete predictions and results spoken aloud.** We do not allow content in Search features that could directly facilitate serious and immediate harm to people, such as self-harm, eating disorders, or drug abuse.

- **Our SafeSearch feature helps filter explicit results from Google Search results**, even when they might be relevant for the query. While these algorithms will never be 100% accurate, turning on SafeSearch helps to filter explicit content, like pornography, from Google search results.

1.3 Protecting users on YouTube

Responsibility is our number one priority at YouTube. Our approach to addressing harmful content is based on the four principles set out above. We[7]:

- **Remove** content that violates our policies – our Community Guidelines provide clear, public-facing guidance on content that is not allowed on the platform. These include policies against spam, deceptive practices, scams, hate, harassment, identity

[7] Susan Wojcicki, Preserving openness through responsibility, 27 August 2019, https://blog.youtube/inside-youtube/preserving-openness-through-responsibility/.

misrepresentation and impersonation, and COVID-19 Medical Misinformation, amongst others. We remove content that violates our policies as quickly as possible, and removed videos represent a fraction of a percent of total views on YouTube. We work continuously to shrink this even further through improved detection and enforcement, relying on a combination of technology and people.

- **Raise up** authoritative voices – users searching for information about breaking news, science and historical events, and topics where accuracy and authoritativeness are key will be overwhelmingly directed towards authoritative sources such as the BBC and The Times. We've seen significant progress in our efforts to raise authoritative voices on YouTube. Globally, authoritative news watchtime grew by more than 85% from the first half of 2019 to the first half of 2020, with a 75% increase in watchtime of news in the first three months of 2020 alone.
- **Reward** trusted, eligible creators, artists, and media organisations – in addition to our community guidelines, creators need to meet an even higher bar to join the YouTube Partner Program and make money on YouTube. Since advertising has been at the core of creators' revenue, we need to ensure that advertisers have faith in our systems and feel comfortable with where their ads appear.
- **Reduce** the spread of content that brushes up against our policy line - while we have strong and comprehensive policies in place that set the rules for what we don't allow on YouTube, we also recognise that there's content that may be problematic but doesn't violate our policies. We use machine learning to reduce the recommendations of this type of content, including potentially harmful misinformation.

On YouTube, we operate a "strikes system" for channels flagged for violating our guidelines, meaning that if channels repeatedly publish content in violation of our policies, they will be terminated. For content that doesn't violate our policies but could be close to the removal line and be offensive to some viewers, we may decide to disable certain features, such as removing comments on videos, as well as placing videos behind a warning message. We have also worked to reduce the visibility of borderline content that remains by raising up more authoritative content in recommendations.

YouTube uses machine learning technology to help identify patterns in content that may violate our Community Guidelines or videos that may contain borderline content — content that comes close to violating our Community Guideline but doesn't quite cross the line. These systems scan content on our platform 24/7, enabling us to review hundreds of thousands of hours of video in a fraction of the time it would take a person to do the same. For example, more than 94% of the content we removed between October and December of 2020 was first flagged by our technology. This underscores just how critical machine learning is for content moderation.

The draft Bill proposes new legal duties on service providers - but we have not waited for regulation or legislation to take action to keep our users safe. Below, we set out in more detail

the steps we already take that align with the proposed duties being placed on companies as part of the new regulatory regime.

1.4 Removing illegal content

Illegal content has no place on our services, and we take a robust approach to identifying and addressing this material. For many issues, such as privacy or defamation, our legal obligations vary country by country, as different jurisdictions have come to different conclusions about how to deal with these complex topics.

We encourage people and authorities to alert us to content they believe violates the law. In fact, in most cases, this is necessary, because determining whether content is illegal is not always a determination that Google is equipped to make, especially without notice from those who are affected. For example, in the case of copyrighted material, we can't automatically confirm whether a given page hosting particular content has a licence to do so, so we need rightsholders to tell us. By contrast, the mere presence of child sex abuse material (CSAM) on a page is universally illegal and wrong, so we develop ways to automatically identify that content and prevent it from showing in our results.

In the case of all legal removals, we share information about government requests for removal in our Transparency Report.

We have invested in the best available methods for protecting users from illegal content:

- **Since our earliest days, we've been committed to fighting online child sexual exploitation and abuse both on our platforms and in the broader online ecosystem.**[8] We have invested in the teams, tools, and resources to deter, remove, and report this kind of content, and to help other companies do so. We use both hash-matching software like CSAI Match (a technology developed by YouTube engineers to identify re-uploads of previously identified child sexual abuse in videos) and machine learning classifiers that can identify never-before-seen CSAM imagery.[9] We share these tools with others in the sector, helping them identify illegal content at scale. Our Content Safety API helps partners classify around 2 billion pieces of content per month.

- **As part of our membership in the Global Internet Forum to Counter Terrorism, we use technology to prevent re-uploads of known terrorist content before that content is available to the public.** In 2016, we created a hash-sharing database with industry partners where we share hashes (or 'digital fingerprints') of terrorist content to stop its

[8] Our efforts to fight child sexual abuse online, 24 February 2021, https://blog.google/technology/safety-security/our-efforts-fight-child-sexual-abuse-online/.
[9] See YouTube CSAI Match, https://www.youtube.com/csai-match/; Using AI to help organizations detect and report child sexual abuse material online, https://blog.google/around-the-globe/google-europe/using-ai-help-organizations-detect-and-report-child-sexual-abuse-material-online/.

spread. The shared database currently contains over 300,000 unique hashes, including both videos and images.

- **We also take a robust approach to removing content that incites hatred.** We have a network of over 180 academics, government partners – including the UK's Metropolitan Police – and hate speech NGOs through our YouTube Trusted Flagger programme. Participants in the Trusted Flagger programme receive training in enforcing YouTube's Community Guidelines. Because of this training and these partners' expertise in identifying hate speech, when they flag potential hate speech content to us, we prioritise it for review.

While we're proud of the steps we're making to better protect our users, we know we cannot do it alone. We partner with governments, industry leaders, and other experts to design better, safer products for our users and will continue to invest in and refine our approach.

1.5 Protecting children and young people

Over the years, we've been significantly investing in the policies, products and practices to help us protect kids and their privacy. This includes implementing additional protections for children and teens on our platforms to comply with the ICO Age Appropriate Design Code.

On 11 August, we committed to a series of global changes to our products for Google Account users under the age of 18.[10] This includes:

- The introduction of a new policy for users under the age of 18, or their parents or guardians, to request the removal of their images from Google Images results;
- New resources to help children and teens better understand our privacy practices, including a Family Link Privacy Guide and Teen Guide;
- Turning SafeSearch on for existing signed-in users under the age of 18 and users we believe to be under 18 (and making this the default setting for teens setting up new accounts);
- On YouTube, adjusting the default upload setting to the most private option available for people aged under 18, more prominently surfacing digital wellbeing features, and providing safeguards and education about commercial content;[11] and
- Expanding safeguards to prevent age-sensitive ad categories from being shown to teens, and we will block ad targeting based on the age, gender, or interests of people under 18.

[10] Giving kids and teens a safer experience online, 10 August 2021, https://blog.google/technology/families/giving-kids-and-teens-safer-experience-online/.
[11] New safety and digital wellbeing options for younger people on YouTube and YouTube Kids, https://blog.youtube/news-and-events/new-safety-and-digital-wellbeing-options-younger-people-youtube-and-youtube-kids/.

These announcements build on our existing dedicated product experiences tailored for users under 18, such as Family Link, our app which lets parents set digital ground rules for their children to help guide them as they learn, play, and explore online; YouTube Kids, an app that provides a separate experience designed especially for children that parents can customise; and Supervised Experiences on YouTube, which allows children to access YouTube through a parent-supervised Google Account. Other products include Teacher Approved Apps and Google Kids Space.

We have also established longstanding partnerships aimed at strengthening media literacy: we partner with Parent Zone on Be Internet Legends, the only PSHE-accredited online safety programme for 7-11 year olds in the UK, which has reached over 71% of primary schools in the UK since launching in March 2018.[12]

1.6 Transparency Reporting and the effectiveness of our efforts

Google is an industry leader in transparency. We have published data about government requests to remove illegal content on our services and government requests for user information since 2010. We've expanded our offering of transparency reports since then, and in 2018, we began publishing a dedicated YouTube Community Guidelines enforcement report. This quarterly report provides data about actions taken on violative comments, channels, and comments, as well as the volume of user flags and creator appeals of removal actions and reinstatements. Since 2018, YouTube has also published a dedicated community guidelines enforcement report.

From April to June 2021, YouTube:

- Removed 6,278,771 videos for violating our Community Guidelines. Almost 6 million of these videos were first flagged by machines rather than humans. Of those detected by machines, around 75% were viewed fewer than 10 times.
- Removed 1,874,729 videos and 40,383 channels for violations of our Child Safety Policies.
- Received 217,446 creator appeals, which led to just 52,696 videos being reinstated.
- Removed more than a billion comments, 99% of which were detected by our automated detection systems. This is a fraction of the billions of comments posted on YouTube each quarter, and the majority of actions we take on comments are for violating our guidelines against spam.

In April 2021 we introduced a new metric, called Violative View Rate, as part of our quarterly transparency reporting. This metric estimates that the proportion of views of YouTube videos that violate our Community Guidelines has fallen from c. 0.7% in Q4 2017 to c. 0.19-21% in Q2 2021. We recently worked with a professor of statistics at MIT Sloan to validate our statistical

[12] Be Internet Legends, https://beinternetawesome.withgoogle.com/en_uk/.

methodology and its application to estimating systemic efficacy of our enforcement and would welcome the opportunity to share that analysis with the Committee.

We publish a specific transparency report on our efforts to fight child sexual abuse material.[13] In the past 18 months, we made over 960,000 reports - containing over 7.8 million pieces of content - to the US National Center for Missing and Exploited Children, the global clearing-house for content relating to child exploitation. This transparency report demonstrates both our fight against this crime and our commitment to transparency.

1.7 Protecting journalistic content

Google shares the Government's aim to ensure appropriate protection for journalistic content and content of democratic importance and we already build these protections into a number of our review processes. For example, on YouTube, content that would otherwise violate the Community Guidelines but which we assess as having Educational, Documentary, Scientific or Artistic (EDSA) value will remain on the platform. To determine whether a video might qualify for an EDSA exception, we look at multiple factors, including the video title, descriptions and the context provided in the video's audio or imagery. These decisions are nuanced and context is important.

For some categories — like videos containing hate speech, graphic violence, content from violent criminal organisations, or COVID-19 medical misinformation — we have a higher bar, given the dangers they present to the public. For this type of material, we require clear context to be provided in the imagery or audio of the video itself and, for a video to benefit from an EDSA exception, it must be clear to the viewer that the creator's aim is not to promote or support the content that violates our policies.[14]

In evaluating content, YouTube makes decisions based on the content itself and surrounding context, not on the speaker or source of the content. Some content, such as a video promoting false cures for COVID-19, can be highly harmful to our users even if it is produced by an organisation with journalistic credentials.

This approach allows us to preserve content that has a distinct value to society while ensuring we provide consistently high levels of protection for our users. As an example, an *Economist* video on vaccine mistrust includes false claims on COVID-19 vaccines, but remains on the platform because the piece criticises and refutes false claims rather than propagating them.[15]

---

[13] Google's efforts to combat online child sexual abuse material, https://transparencyreport.google.com/child-sexual-abuse-material/reporting?hl=en_GB&lu=total_content_reported&total_content_reported=period:2020H1.
[14] A look at how we treat educational, documentary, scientific, and artistic content on YouTube, 17 September 2020, https://blog.youtube/inside-youtube/look-how-we-treat-educational-documentary-scientific-and-artistic-content-youtube/.
[15] The Economist, Covid-19: why vaccine mistrust is growing, 18 November 2020,

Our response to the Call for Evidence is below.

2. Objectives

Google supports the Government's objectives, as outlined in the Full Response to the Online Harms White Paper consultation, to keep people safe online, protect pluralism and freedom of expression, and to use digital technologies and services to power economic growth across the UK.[16] To ensure the Government's objectives are fully met, we believe there is an opportunity to provide greater legal clarity across a number of areas.

2.1. Objectives in the Bill and systemic approach

Relevant Committee questions addressed:

- *Is the "duty of care" approach in the draft Bill effective?*
- *Does the Bill deliver the intention to focus on systems and processes rather than content, and is this an effective approach for moderating content? What role do you see for e.g. safety by design, algorithmic recommendations, minimum standards, default settings?*

By making platforms accountable for the processes they have in place to protect users, the draft Bill has the potential to lead to continued improvements in user safety across online services in scope. The scale of online content – for example, 500 hours of content are uploaded to YouTube every minute — means that a systemic approach is the most effective way of focusing finite regulatory resources where it can have the biggest impact in improving online safety. As the Government has also outlined, "the focus on robust processes and systems rather than individual pieces of content means it will remain effective even as new harms emerge".[17] By focusing on the processes that platforms have in place, rather than individual pieces of content, the regulatory framework will be more impactful in driving sustained and continuous improvement in the way that platforms protect users. We encourage the Committee to maintain the Bill's focus on platforms' systems and processes.

At Google, we have dedicated extensive effort to developing processes that better protect users, and we are not waiting for regulation to take action. As outlined elsewhere in this submission, we have clear and robust policies in place so that our users understand what content we do and do not allow on our services, and we remove all content that violates these policies. We also know that removing content is, on its own, not enough to keep users safe

https://www.youtube.com/watch?v=3EK4VRmG3yM.

[16] Online Harms White Paper: Full Government Response to the consultation, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/944310/Online_Harms_White_Paper_Full_Government_Response_to_the_consultation_CP_354_CCS001_CCS1220695430-001__V2.pdf.

[17] Online Harms White Paper: Initial Government Response, February 2020, https://www.gov.uk/government/consultations/online-harms-white-paper/public-feedback/online-harms-white-paper-initial-consultation-response.

online. That's why our multi-faceted approach to responsibility also includes raising up authoritative content (such as content from the NHS in relation to COVID-19, or content from authoritative news sources) and reducing the spread of potentially harmful information where we recommend content. We also design our tools and products with safety in mind, including ongoing investments in policies, products, and practices to help keep children and younger users safe.

2.2. Improving online safety and protecting freedom of expression online

Relevant Committee questions addressed:

- *Will the proposed legislation effectively deliver the policy aim of making the UK the safest place to be online?*
- *Does the proposed legislation represent a threat to freedom of expression, or are the protections for freedom of expression provided in the draft Bill sufficient?*

We strongly support the Government's objectives to make the UK the safest place to be online and to protect freedom of expression. At Google, we take very seriously our responsibility to keep users safe and protect their rights to express themselves through our services. We have evolved an approach over many years that is focused on acting decisively to help users stay safe, without disproportionately limiting their ability to express themselves online. We share the Government and Parliament's aim to shape requirements that are both effective in setting clear rules for the removal of illegal content, and that protect users' freedom of expression by avoiding the mistaken removal of legitimate content.

The draft Bill indicates that Codes of Practice issued by Ofcom will specify how platforms can comply with their duties. This introduces the possibility of Codes of Practice later on being used to introduce further requirements that undermine freedom of expression. In this section, we invite the Committee to consider how guardrails can be added to ensure Ofcom's approach to implementation both strengthens user safety and protects freedom of expression online.

The relationship between duties

We acknowledge and share the Government's concern with ensuring that online regulation supports UK users' ability to access and express a range of perspectives online. The duties to protect users' freedom of expression and privacy are an important step in this direction, but there remains ambiguity over how providers should balance these duties with the safety duties. For example, providers lack clarity on how they should manage the risk that they inadvertently remove legitimate content as they seek to protect users. To provide more clarity, the Committee could consider how the Bill can be strengthened by clarifying that service providers must give equal weight to the duties to protect users' fundamental rights and the duties in relation to illegal and harmful content.

General monitoring and proactive content removal

We welcomed the Government's Online Harms White Paper's confirmation that the Bill would not compel companies to undertake general monitoring on their online services, but the breadth of content in scope of the Bill is such that it appears to establish requirements for service providers to monitor all content on their services, which would directly contradict established law in this area. General and proactive monitoring obligations encroach on fundamental rights and freedoms. They inevitably tilt the balance between protecting users from harm and protecting their freedom of expression toward more restriction of legitimate speech. **This could lead to large amounts of legitimate content being mistakenly removed, limiting UK users' ability to express themselves online.**

AI tools are an important part of our strategy for keeping users safe, and these tools require careful use. In particular, while AI tools are increasingly reliable in identifying and flagging CSAM content, other forms of content, for example hate speech, require an understanding of the context. This is why YouTube relies on a combination of humans and machines to review and moderate content. Machines can identify patterns and potential violations at scale; humans bring nuance and context to content review. **Without appropriate safeguards, AI tools can make errors that negatively impact on users' rights to express their opinion online and can entrench discrimination.** For instance, research has suggested that content featuring LGBTQ+ people has been erroneously blocked from some social media services because it was misidentified using AI. In one case, a newsletter aimed at women, trans, and non-binary people tried to publish advertising on a social media platform, but the advertising was rejected by the service's AI-based tools and erroneously marked as an "escort service".[18]

Our experience of deploying these tools has informed a number of steps we take to avoid their inadvertently impacting on our users' freedom of expression. In particular, when AI tools flag content for removal, we ensure there is an appropriate level of human oversight before removing the content, for most forms of content. In our response to the Committee's questions on Algorithms and User Agency, we explain further our approach to using algorithms in content moderation.

The particular emphasis in the draft Bill on reducing the amount of time content is on the platform will lead platforms to prioritise rapid removal using automation over getting the decision right. As an example of what can happen when automated removal is prioritised over careful human review, in Q2 2020, as COVID-19 lockdowns meant fewer human content moderators were able to work, YouTube depended more heavily on automated technology to remove content violating our policies. The number of appeals by users of content removal decisions doubled, compared to Q1: 50% of appeals resulted in reinstatement in Q2, compared with less than 25% in Q1.[19] This suggests a higher error rate than under normal conditions.

[18] See Heinrich Boll Stiftung, The state of content moderation for the LGBTIQA+ community and the role of the EU Digital Services Act, June 2021, p7, https://eu.boell.org/sites/default/files/2021-06/HBS-e-paper-state-platform-moderation-for-LGBTQI-200621_FINAL.pdf.
[19] YouTube Community Guidelines Enforcement Report, Q2 2020,

We consider that the Bill could be clarified to expressly state that the safety duties do not of themselves mandate general monitoring and proactive removal.

2.3 Protecting children online

Relevant Committee questions addressed:

- *Are children effectively protected from harmful activity and content under the measures proposed in the draft Bill?*

We strongly support the Government's objective of protecting children online. We take our responsibility for child safety and children's rights online very seriously. As outlined elsewhere in this submission, we have always implemented protective policies and standards for children and we continue to evolve our approach in order to build on our existing dedicated product experiences tailored for users under 18. We want to work with the Government and with Parliament to ensure the Bill protects children and preserves their rights to safely access online services.

Requirements to age-gate online services

We believe there is an opportunity for the Committee to ensure the Bill does not inadvertently block children, young people and vulnerable users from accessing the internet. The Information Commissioner's Office's Age Appropriate Design Code has only just come into force and is already fulfilling its objective to deliver meaningful change to the way that online services interact with children.

The Code acknowledges the benefits that children derive from online services. It takes a proportionate and considered approach focused both on enhancing children's safety and on protecting their right to participate online, delivering improvement "not by seeking to protect children from the digital world, but by protecting them from within it".[20]

Online services have had a powerful impact in enriching children's lives. For example, recent University of East London research outlined the different benefits young people derive from video-sharing platforms, including knowledge, connection, enjoyment and expression.[21] A recent GCHQ/DCMS/Home Office project highlighted a particular risk that an overly prescriptive approach risked excluding some users due to the need for specific documentation and encouraged children to circumvent the system to bypass the protections, making them

https://transparencyreport.google.com/youtube-policy/removals.

[20] Information Commissioner's Office, Age appropriate design code: a code of practice for online services, 2 September, 2020.

[21] University of East London, Research on Protection of Minors: A Literature Review and Interconnected Frameworks. Implications for VSP Regulation and Beyond, 24 March 2021, https://www.ofcom.org.uk/__data/assets/pdf_file/0023/216491/uel-report-protection-of-minors.pdf.

increasingly vulnerable.[22] A requirement to submit personal documents also has important implications for the privacy of all users; the Age Appropriate Design Code notes that "requiring hard identifiers [such as a passport] may also have a disproportionate impact on the privacy of adults".[23]

In relation to children, the Bill establishes three tiers of "content that is harmful to children", imposing subtly different requirements in each case, along with requirements to give separate consideration to how individual pieces of content in the different tiers might affect children in different age groups, taking into account the particular characteristics any such child may have. Such complex requirements detract from duties for service providers to consider more broadly how to best protect children and may not be workable in practice, undermining the aim to strengthen protection. A more practical and effective approach would focus on requiring providers to address content that would be inappropriate for persons under the age of 18, which is the approach taken in the EU's AVMS Directive.

Mandating age-specific experiences for different age groups can add complexity to the service, and technological solutions are not sufficiently developed to meet childrens' and families' differing and evolving needs. Further, a system that mandates specific experiences according to age does not take into account the differences in development and maturity that can take place during teenage years.

As drafted, the complexity of the draft Bill's requirements in relation to children mean that service providers are incentivised simply to prevent access to their services to under 18s, thereby denying them the right to use such services. Instead of mandating different experiences or content for different age groups, the Committee could explore how the use of safety by design tools such as flexible parental controls, e.g. Family Link, could help to better meet the goal of improving safety. Supported by Family Link, YouTube Supervised Experiences put parents in control of the content and experience that their children can access, giving them the flexibility to choose what is right for their children and their families. Similarly, Youtube Kids includes a set of parental controls to customize their child's individual experience allowing parents to limit what their children can watch, including only allowing videos, channels and/or collections that they've hand-picked and approved and selecting the appropriate content level based on their child's age.[24]

---

[22] GCHQ, DCMS, Home Office, ACE, Verification of Children Online, Phase 2 Report, November 2020, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/934131/November_VoCO_report_V4__pdf.pdf.

[23] Information Commissioner's Office, Age Appropriate Design Code, https://ico.org.uk/media/for-organisations/guide-to-data-protection/key-data-protection-themes/age-appropriate-design-a-code-of-practice-for-online-services-2-1.pdf.

[24] For further details, see Exploration starts here, https://www.youtube.com/myfamily/?gclid=Cj0KCQjws4aKBhDPARIsAIWH0JV6O27RFiRGgkN0SrCKUgRT5J8ETe-Zv0u6NmdKQeSHAuN_7yrhzhkaAiCaEALw_wcB&gclsrc=aw.ds.

To best protect children while retaining their access to online services, it is vital that the Committee carefully considers how the Bill's child protection duties can be simplified and aligned with the Age Appropriate Design Code and the EU's AVMS Directive, notably focusing on requiring providers to address well defined content that is inappropriate for persons under the age of 18.

2.4 Protecting vulnerable people online

Relevant Committee questions addressed:

- *Does the draft Bill make adequate provisions for people who are more likely to experience harm online or who may be more vulnerable to exploitation?*
- *What would be a suitable threshold for significant physical or psychological harm, and what would be a suitable way for service providers to determine whether this threshold had been met?*

We understand and share the Committee's ambition to protect users who are more vulnerable to harm online.

We take a zero-tolerance approach to all forms of abuse. On YouTube, we have significant protections in place, built up over time, to ensure that we are effective in protecting all of our users. This includes a comprehensive policy that prohibits content promoting violence or hatred against individuals or groups based on any of the following attributes: age, caste, disability, ethnicity, gender identity, nationality, race, immigration status, religion, sex/gender, sexual orientation, victims of a major violent event and their family members, and veteran status.

Simplifying the definitions of harmful content would make the framework more likely to be more effective in achieving the Government's aims. The definitions of "content that is harmful" may be overly prescriptive and complex, requiring service providers to determine whether there are "reasonable grounds to believe" that there is a "material risk" of the content having a "significant adverse physical or psychological impact" on an adult or child of "ordinary sensibilities", taking into account any "characteristic (or combination of characteristics)" or membership of "a certain group of people" which may "reasonably be assumed" to be particularly affected, and taking into account whether there is "a material risk of the fact of the content's dissemination having a significant adverse physical or psychological impact", and in all cases whether such impact is direct or "indirect", defined by reference to content that may not cause harm to a person but that may "cause" another person to cause harm to that person.

As the Lords Communications and Digital Committee has outlined, this definition could "lead to content which is legitimate and well-intentioned being censored due to speculation about the influence it might have on an unreasonable person's actions towards a third party", adding "no platform could reasonably be expected to enforce this provision without significant interference in their users' freedom of expression".[25]

We agree with the Lords Committee's recommendation that the Bill should at least refer to "the reasonable person of ordinary sensibilities", but we would encourage the Committee to consider more broadly how a simpler, less-prescriptive definition of harmful content would create a more flexible way for service providers to determine how best to protect their users.[26]

2.5 Supporting economic growth and innovation

Relevant Committee questions addressed:

- *Will the proposed legislation help to deliver the policy aim of using digital technologies and services to support the UK's economic growth? Will it support a more inclusive, competitive, and innovative future digital economy?*
- *Will the regulatory approach in the Bill affect competition between different sizes and types of services?*

Google supports the Government's ambition, as stated in its Plan for Digital Regulation, to encourage innovation by offering clarity and confidence to businesses and consumers.

The Government's impact assessment estimates that the Bill will impact 24,000 businesses, including 17,100 micro businesses, defined as any business with ten employees or fewer. The complexity of the draft Bill means that businesses of this size will need to devote a considerable proportion of staff capacity and time to digest and understand what it means for how they should comply. In addition to these up-front costs, the draft Bill creates considerable uncertainty for businesses of all sizes, including over definitions of the content they will be expected to act against and on the precise expectations they will be subject to.

The possibility that providers will have to use technology to monitor content and age-gate their services could also have important implications for innovation in the digital economy. An approach that imposes a blanket mandate to use this technology will result in significant up-front costs for new businesses, for example in terms of access to skilled AI developers to implement algorithmic tools.

We believe the Committee could usefully intervene to ensure businesses are provided with the necessary clarity and confidence to continue to invest and build their businesses in the UK.

2.6 Global parallels

Relevant Committee questions addressed:

---

[25] Letter to Rt. Hon. Oliver Dowden MP, 25 May 2021, https://committees.parliament.uk/publications/6025/documents/68088/default/.

[26] Paragraph 15 of the House of Lords Communications and Digital Committee's report - "Free for all? Freedom of expression in the digital age", 22 July 2021

- *How does the draft Bill differ to online safety legislation in other countries (e.g. Australia, Canada, Germany, Ireland, and the EU Digital Services Act) and what lessons can be learnt?*

The UK's draft Bill differs from legislation being considered elsewhere in a number of important respects. The European Commission published a draft Digital Services Act (DSA) proposal in December 2020, which the European Parliament and European Council are now considering.[27] Google has a number of specific recommendations to improve the Commission's proposal and proposed amendments, but relative to the draft Online Safety Bill, it does provide a greater degree of legal clarity in several respects. In particular:

- It specifies that there should be no requirement to monitor all content;
- It does not require the age-gating of all services;
- It does not require the removal of legal content;
- It does not include a deferred power to introduce criminal personal liability for companies' staff;
- It does not propose a power for regulators to mandate uses of specific technology.

We would encourage the Committee to carefully weigh the potential consequences of the UK defining an approach that diverges so markedly from approaches being taken elsewhere. It could, for example, lead to online platforms prioritising rapid removal of content aimed at UK users, while retaining a more balanced approach with safeguards against over-removal in other jurisdictions. **This could disproportionately impact British citizens' freedom of speech compared to citizens in other countries**. Equally, the risk of mandating the use of technology to monitor content and age-gate services is that **innovative British companies will be put at a disadvantage relative to competitors abroad, and that future start-ups will choose to build their businesses elsewhere**.The bill could also set an unwelcome human-rights precedent. For example, more repressive states that wished to impose specific technology mandates would be able to point to the UK legislation as a sign that they are not out of step with the human rights consensus.

3. Content in Scope

Google supports the Government's aim to protect users from different types of harm online. Clearer definitions of illegal content and guidance about compliance requirements are vital to provide service providers with the certainty they need to effectively and consistently fulfil the objectives of the Bill. Similarly, more clarity over the expectations in relation to the duties on journalistic content and content of democratic importance will help online services to better understand how they should balance these duties with their duties to protect users' safety.

[27] European Commission, The Digital Services Act package, December 2020, https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package.

3.1 Protecting users from illegal content

Google agrees with the draft Bill's particular focus on ensuring there are effective processes in place for the removal of illegal content. We want the Bill to succeed in ensuring there is an effective, industry-wide approach to the illegal content that poses the greatest risk to users, without jeopardising users' fundamental rights.

Clear definition of illegal content

At present, the Bill's threshold for what constitutes illegal content covers a wide and technically challenging range of content. **If online platforms are held liable for content where they cannot be sure of its illegality, they will tend towards blocking legitimate but controversial content**. This could chill online discussion around contentious but important social and political issues. An over-broad definition will incentivise providers to calibrate content moderation tools towards removal, sweeping up both legitimate and illegal content.

The Bill will be more effective in protecting users from the most egregious forms of illegal activity if it provides platforms with the necessary clarity to take quick and decisive action. We would encourage the Committee to consider how the definition of illegal content can be amended to focus the industry's response to serious illegal activity.

3.2 Ensuring appropriate protection for democratic content and content of journalistic importance

Relevant Committee questions addressed:

- *The draft Bill specifically places a duty on providers to protect democratic content, and content of journalistic importance. What is your view of these measures and their likely effectiveness?*

Google shares the Government's aim to ensure appropriate treatment of journalistic content and content of democratic importance, and we have developed our own policies to achieve this goal.

This is a complex issue that requires careful consideration to avoid inadvertently impacting user safety. Under the draft Bill, the balance between a service provider's duties to protect journalistic content and content of democratic importance and the provider's duties to protect users is unclear. For content of democratic importance, for example, service providers must "take into account" the importance of free expression of this content (Section 13). It is unclear if this would extend to not removing, for example, a video of a candidate deploying racist language while discussing immigration policy or using antisemitic language while campaigning. The House of Lords Communications and Digital Committee has expressed concern about Ofcom and platforms' ability to assess which content falls under the definition of "citizen journalism".[28] For example, Tommy Robinson has previously described himself as a "citizen

journalist", and we are confident that any expedited review system would quickly be (wrongly) used by individual creators for whom it is not intended.

As a global company, we frequently face situations where content that arguably has democratic importance presents a risk to our users, and we would be concerned about obligations that would restrict our ability to keep our users safe. For example, in April 2021, we removed a video of an event featuring the governor of the US state of Florida, Ron DeSantis. The video violated our COVID-19 medical misinformation policy because it included content that contradicts the consensus of local and global health authorities regarding the efficacy of masks to prevent the spread of COVID-19, long after the WHO's guidance on this subject had become clear. In 2020, we removed two videos from the channel of Brazilian President Bolsonaro for violating our COVID misinfo policies, where he openly dismissed public health recommendations. In April, we removed another four videos where the President encouraged Brazilian citizens to use treatments for COVID that run counter to global medical consensus and have been shown to be unsafe in certain circumstances. We welcome the Government's approach whereby online platforms remain able to remove journalistic content and content of democratic importance where such content may be harmful and violates the platform's content policies.

3.3 Protecting users from the potential for harm caused by non-illegal content

Relevant Committee questions addressed:

- *Are the definitions in the draft Bill suitable for service providers to accurately identify and reduce the presence of legal but harmful content, whilst preserving the presence of legitimate content?*
- *Earlier proposals included content such as misinformation/disinformation that could lead to societal harm in scope of the Bill. These types of content have since been removed. What do you think of this decision?*

We recognise Parliament has concerns about specific categories of lawful content, such as disinformation and misinformation. This is a critical issue, particularly on social media platforms, and for video-sharing platforms such as YouTube, and **we have devoted careful attention to developing a multifaceted approach that seeks to limit harm while protecting freedom of expression**.

We think it is important that regulation holds platforms to account for the processes they have in place to protect users, without dictating which types of lawful content should remain online and which should be taken down. **It is important that there remains space online to express views that are contentious or which diverge from the mainstream and that private enterprises are able to set their own terms / rules of engagements for their members/users.**[29]

[28] House of Lords Communications and Digital Committee, Free for all? Freedom of expression in the digital age, 22 July 2021, p44, https://publications.parliament.uk/pa/ld5802/ldselect/ldcomuni/54/54.pdf.

Regulation that is not sufficiently thought through could leave British citizens less free to express themselves online than their peers in other parts of the world.

If the Government believes that a category of content is sufficiently harmful that it should not be available online, it may make that content illegal directly, through transparent, democratic processes. Indeed, the Law Commission has recently issued recommendations for the Government to introduce new criminal offences aimed at criminalising "genuinely harmful" online behaviour, including a new offence to target intentional encouragement or assistance of self-harm.[30]

Ultimately, an effective regulatory approach that balances user safety with freedom of expression would be to require platforms to clearly communicate their content moderation policies to users and to be consistent and fair in upholding those policies, without dictating which types of lawful content should remain online and which should be taken down.

4. Services in Scope

In relation to user-to-user services, Google supports the draft Bill's ambition to tailor obligations based on the risk posed by different types of service, such as "Category 1 services". Evidence shows that small services can still have the potential to provide a lucrative platform for illegal and harmful activity, and we agree that the conditions for designation as a "Category 1 service" should be based on both a service's functionality and the number of users.[31]

4.1 Keeping users safe on search services

Relevant Committee questions addressed:

- *The draft Bill applies to providers of user-to-user services and search services. Will this achieve the Government's policy aims?*

Search engines play a uniquely important role in facilitating access to the internet, enabling people to access, impart, and disseminate information. They provide enormous social and economic value to individual UK users and the country as a whole. They are vital both to the

[29] See Prager University v Google LLC (Judgement and BBC Report) - also see recent German Fedreal COurt of Justice case (vs Facebook) which confirmed that platforms may balance users' rights to freedom of expression vs their own rights to set guidelines: "The weighing of the conflicting fundamental rights and interests of the parties as well as the third-party interests to be included shows that Facebook is in principle entitled to require the users of its network to comply with certain communication standards in general terms and conditions that go beyond the requirements of criminal law".

[30] Law Commission, Reforms to protect victims of online abuse and safeguard freedom of expression announced, 21 July 2021, https://www.lawcom.gov.uk/reforms-to-protect-victims-of-online-abuse-and-safeguard-freedom-of-expression-announced/.

[31] As an example, Tech Against Terrorism, ISIS use of smaller platforms and the DWeb to share terrorist content, 29 April 2019, https://www.techagainstterrorism.org/2019/04/29/analysis-isis-use-of-smaller-platforms-and-the-dweb-to-share-terrorist-content-april-2019/.

day-to-day functioning of businesses but also in allowing consumers to find small businesses in their area, strengthening competition across a range of markets. In 2020 research, 40% of British workers said their job would be impossible or difficult without access to a search engine. In a separate poll, 76% of British businesses said online search was an important way that customers found their business, second only to word-of-mouth.[32]

Search engines also play an essential social function, in providing millions of UK users with instant access to information they can trust. The pandemic has illustrated more than ever the importance of search engines in times of need. Throughout 2020, Brits turned to Google Search to help keep them informed, up to date, and safe. According to research conducted last year, 42% of Brits have used Google Search to find out information on COVID-19 symptoms, and 51% of Brits have used Google Search to keep up to date with the latest government advice on COVID-19.[33]

Search technology and content policies are designed to help all users avoid harm, and connect them with resources in moments of need. Google processes billions of searches per day. Our automated systems enable Google to handle the immense scale of so many searches. Our systems are designed to prioritise what appears to be the most useful and helpful content on a given topic. Our systems are also designed not to surface content that violates our content policies.

We've carefully developed content policies for Google Search to balance the real concerns about content safety, alongside the need for a search engine to provide access to information for users who are researching on almost any topic. For example, we are implementing a policy to permit anyone under the age of 18 to request to have images of themselves removed from search results. One instance where this could be useful is to reduce the risk of cyberbullying.

Google Search is also designed to provide high quality information in moments of need. For example when users search for information about sensitive topics like health or in emergency situations, they might see an informational panel from authoritative sources or get crisis prevention information in their search results. In the UK we've worked with Samaritans to make sure their helpline and information is available in this results box. The goal of this type of result is to connect vulnerable people in unsafe situations to reliable and free support as quickly as possible.

It is important to note that many of the ways in which Google Search protects users, do not depend on Search's ability to preemptively analyse content in results and make determinations about whether or in what circumstances it may cause harm. This reflects the fact that **Search engines are different from other types of online service, especially user-to-user services such as social media platforms,** which provide an environment for user interaction that give them a

[32] Public First, Google's Impact in the UK 2020, https://googleimpactreport.publicfirst.co.uk/uk/.
[33] Public First, Google's Impact in the UK 2020, https://googleimpactreport.publicfirst.co.uk/uk/.

responsibility and an opportunity to set and enforce the rules for that interaction. Search engines provide a map of the internet and are therefore, by their nature, a reflection of the information that is available elsewhere. It is therefore sensible to consider whether all of the obligations that are applicable to user-to-user services are applicable to search engines.

As an important platform through which UK users access information, we think it is right that search services should be subject to expectations to protect users. However, as it stands, the draft Bill fails to reflect the Government's intent to differentiate the framework's approach to search services**.** This approach is not appropriate as search services are fundamentally different to user-to-user services. Specifically:

- The duties in the draft bill for search services to prevent people from encountering content that may, in some contexts, be illegal or harmful, do not take into account the fact that such services are an index of the web and do not host content or control the context in which it appears. For services providers to meet the duty they may have to remove controversial content from results for all users in all contexts, which would severely and disproportionately impact the rights of UK users to receive and impart information. Rather than a focus on preventing users from accessing content they may be seeking and which is freely available on the internet, more appropriate duties for search services might focus on how users can be protected from inadvertently encountering content they do not wish to see and how they can be directed to help where their search queries indicate it may be appropriate.

- Google is already providing SafeSearch as a default for under 18 users and making it available for all users, to filter out explicit and offensive content. The draft bill includes duties to minimise the risk of children of different ages from encountering content that is harmful to children, which would have to be achieved through a granular understanding of individual users' ages and age-gating or the removal from search results of all content that might conceivably be harmful to a child of any age. Both of these requirements would equally impact child and adult users who would either face an age gate and need to verify their age or would have their access to lawful content curtailed.

- The draft bill also creates an unprecedented general monitoring provision for "priority illegal content," on Search services, which is as-yet undefined. Without a clear and focused scope it would be difficult for companies to comply and, however it is defined, is unlikely to be achievable at scale given the limitations on automated tools. While we set and enforce policies using a higher bar for Search features to prevent things like porn, hate speech, or violence from appearing in search panels or search prediction,[34] it

[34] Content policies for Google Search, https://docs.google.com/document/d/13yiHitXAd6W-2BTTg_IVW6Byn-OorHwyXN0WbbVngZk/edit#.

is not technically possible to do this for ranking results without restricting users access to sites that are legal.

To ensure the regulatory framework preserves the benefits that search services provide to users, we would encourage the Committee to consider meaningfully differentiated obligations for search engines that take into account their unique character and function.

4.2 A targeted approach on online advertising

Relevant Committee questions addressed:

- *Are there any types of content omitted from the scope of the Bill that you consider significant e.g. commercial pornography or the promotion of financial scams? How should they be covered if so?*
- *Should other types of services be included in the scope of the Bill?*

We are very aware of the damage that scams can cause to individuals and their families and the wider economy, so we are devoting sustained attention and resources to addressing this. We invest significantly in both cutting edge machine learning systems and human reviewers to enforce our policies, which we are continually updating against the latest threats. **Working closely with the Financial Conduct Authority, we have taken a proactive approach to address this issue.** We recently introduced new certification requirements for financial services advertisers targeting the UK. In order to show financial services ads to UK users advertisers must demonstrate that they have FCA authorisation or qualify for a very limited number of exemptions.

We believe that paid-for advertising requires an approach separate from user-generated content. We would recommend prioritising the DCMS Online Advertising Programme and the Home Office Fraud Action Plan as the appropriate vehicles to review online advertising. Advertising and financial fraud involve a complex and highly specialised set of issues and existing rules, and requires consideration of the implications of regulation for legitimate competition and innovation in the UK's dynamic fintech sector. **The Advertising Standards Authority and the Financial Conduct Authority are best placed to address these issues.**

We believe that a more targeted approach to paid-for advertising would not only be more effective than including it within scope of the Online Safety Bill, but would also be more proportionate in its impact on the wider economy**.** To add online advertising into scope would significantly add to the 24,000 companies the Government estimates will be affected by the Bill, including many of the UK's largest companies who advertise online. SLG Economics, jointly commissioned by the Advertising Association, ISBA, IAB and the IPA, has estimated the cost of the Government's proposed ban on HFSS advertising online to be £2.3 billion, but this could be a fraction of the cost of poorly designed regulation on online advertising that is wider in scope.[35]

[35] SLG Economics, Review of the Government's proposals for introducing a total ban on HFSS

4.3 Ensuring strong protection for privacy

Relevant Committee questions addressed:

- *Should other types of services be included in the scope of the Bill?*

As the Government's Full Consultation Response rightly pointed out, **users expect a greater degree of privacy in relation to private services — such as direct messaging and closed groups — and such services should be subject to different requirements**. However, the draft Bill does not currently provide for a different approach in respect of private communications. Providers are subject to a duty to protect users from "unwarranted infringements of privacy" (Clause 23), but there is a lack of clarity on how this relates to their online safety duties. **There is a risk that the current wording in the draft Bill leads to significant monitoring of private communications**, as providers seek to take action against content that is shared privately on the basis that it may, in theory, cause harm if shared publicly. Such requirements to monitor private communications are not subject to any of the safeguards provided for under existing UK and EU law (e.g. the need for a judge-approved warrant to monitor and intercept private communications under the Investigatory Powers Act), and it is not clear how the Bill's provisions could be reconciled with these laws, with the nature of encrypted services, nor with the commitments service providers may have given to users (including in their terms of service). We would recommend the Committee propose tightly specifying a differentiated set of obligations in respect to private services in the Bill.

5. Algorithms and user agency

Google has taken extensive steps to inform users and regulators about the functioning of the algorithms we use on Search, YouTube, and other Google products. Algorithms play an important role in helping us to provide value for users, both in directing them to content they will find valuable and in keeping them safe. No technology is perfect, and we look forward to engaging the Committee on this topic.

5.1 Algorithms in Google's products

Relevant Committee questions addressed:

- *What role do algorithms currently play in influencing the presence of certain types of content online and how it is disseminated? What role might they play in reducing the presence of illegal and/or harmful content?*

Algorithms are critical for Google Search's mission of sorting the world's information and making it useful. Our algorithms surface relevant and high quality sources and prevent poor

advertising online, December 2020, https://www.iabuk.com/sites/default/files/public_files/APPENDIX-A-SLG-Economics-Final-Report-on-online-HFSS-advertising-ban.pdf.

quality or harmful content from rising in search results. Our publicly available Search Quality Rater Guidelines clarify our aim to avoid surfacing pages which are "deliberately created with factually inaccurate content".[36] We are updating our ranking systems all the time to better serve users' desire for quality and trustworthy content. Since 2017, we've done over 1 million search quality tests, more than 1,000 per day.

Similarly, the algorithms used by YouTube to recommend videos are built to provide value to our users, by helping them find new content that appeals to their interests. Recommendations play a pivotal role across our entire community, introducing viewers to content they love and helping creators connect with new audiences. To provide such custom curation, our recommendation system doesn't operate off of a "recipe book" of what to do. A number of signals build on each other to help inform our system.[37]

Recommendations also play an important role in how we maintain a responsible platform. They connect viewers to high-quality information and minimise the chances they'll see problematic content. And they complement the work done by our robust Community Guidelines that define what is and isn't allowed on YouTube. The rise of misinformation in recent years led us to further expand the ways we use our recommendation system to include deprioritisation of problematic misinformation and borderline content—that is, content that comes close to, but doesn't quite violate our Community Guidelines. Because, while clicks, watchtime, user surveys, shares, likes and dislikes are important signals that inform our system, they are rightly overruled by our commitment to meeting our responsibility to the YouTube community and to society.

Both Google Search and YouTube also deploy algorithms to identify specific categories of content that may be illegal or in violation of our policies. In particular, we use machine learning to help us detect illegal child sexual abuse material (CSAM). **However, algorithms are not infallible.** A recent report by the Government's independent advisory body, the Centre for Data Ethics and Innovation (CDEI), noted that algorithms are "generally poor at contextual interpretation",[38] making their deployment to identify most forms of illegal content challenging. For example, algorithms would struggle to distinguish between content from a terrorist organisation glorifying violence and content from a journalistic or human rights organisation documenting such violence. We seek to mitigate this sort of risk by, among other things, regularly reviewing our machine learning systems to reduce the risk of unintended algorithmic

---

[36] Search Quality Evaluator Guidelines, https://static.googleusercontent.com/media/guidelines.raterhub.com/en//searchqualityevaluatorguidelines.pdf.

[37] "On YouTube's recommendation system," YouTube blog (15 September 2021), https://blog.youtube/inside-youtube/on-youtubes-recommendation-system/.

[38] Centre for Data Ethics and Innovation, The role of AI in addressing misinformation on social media platforms, p19, 5 August 2021, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1008700/Misinformation_forum_write_up__August_2021__-_web_accessible.pdf.

bias. On YouTube, we add a layer of human review before removing content that is identified by machines.

Google has taken a number of steps to be transparent with our users on how algorithms work. For example, our How Search Works site provides extensive information about how we improve search quality and our approach to algorithmic ranking.[39] Similarly, our How YouTube Works site provides extensive information about how the product works and the controls we provide for users to manage their data.[40] We also regularly update our YouTube Blog with information designed to educate users about the platform - most recently we published a blog post dedicated to explaining how our recommendation system works.[41]

5.2 Algorithms and the Online Safety Bill

Relevant Committee questions addressed:

- *Are there any foreseeable problems that could arise if service providers increased their use of algorithms to fulfil their safety duties? How might the draft Bill address them?*

Algorithm-based content moderation can be a powerful tool in identifying certain forms of content, but there are clear risks that its use can lead to mistakes that seriously impact on users' freedom of expression. Algorithms and automated solutions can reliably identify specific types of content, such as CSAM, or patterns consistent with harmful content they've been trained to recognise. However, there are still limitations to how effective existing technology can be in enforcing our policies. Technology isn't a silver bullet and machine learning cannot replace human judgement, particularly for complex challenges where context is crucial - for example, differentiating between YouTube videos of hate speech at a rally and news reporting on that event. That's why our machine learning technology is supported by a highly trained and growing team of human reviewers with the ability to apply their own judgement for more complex cases that AI cannot tackle alone.

Requiring providers to use algorithms to monitor all content creates incentives to over-remove legitimate content in order to avoid penalties.

We recognise the draft Bill seeks to mitigate this risk, for example, through its duty about rights to freedom of expression and privacy. However, these mitigations are vague, specifying only

[39] https://www.google.com/search/howsearchworks/.
[40] https://www.youtube.com/intl/ALL_uk/howyoutubeworks/.
[41] On YouTube's Recommendation System, 15 September 2021, https://blog.youtube/inside-youtube/on-youtubes-recommendation-system/.

that providers should "have regard to the importance" of protecting users' rights. In practice, the draft Bill's emphasis elsewhere on minimising the presence of illegal content will lead platforms to prioritise using algorithms to monitor and remove content and to de-emphasise the human oversight required to prevent algorithms removing large amounts of lawful content. We believe the Committee should propose limiting the content which is in scope of the most onerous safety duties to that which can be reliably identified by algorithms.

5.3 Empowering users

Relevant Committee questions addressed:

- *Does the draft Bill give sufficient consideration to the role of user agency in promoting online safety?*

Google acknowledges the intention of the Online Safety Bill is to strengthen online services' accountability for the steps they take to protect users. **It is important to also recognise the role of user agency, the actions of individual users, and parental support in contributing to a safer online environment.** Media literacy and digital citizenship can help people make the most of online opportunities while empowering them to protect themselves and their families from the potential risks. This should not be a substitute for effective action from platforms, but can complement that action and further contribute to the UK's objectives to strengthen online safety.

Many of the tools we provide are aimed at empowering parents to supervise their children's experience online. For example, YouTube Kids includes a set of parental controls to customise their child's individual experience. Parents can decide to limit what their children can watch, including only allowing videos, channels and/or collections that they've hand-picked and approved and selecting the appropriate content level based on their child's age.

Our two PSHE-accredited programmes, Be Internet Legends and Be Internet Citizens, aim to equip young people with media literacy and digital citizenship skills so they can experience the internet in a safe and positive way. Be Internet Legends is Google and Parent Zone's online safety programme, supporting children, families and primary schools across the country.[42]

Be Internet Citizens is a programme launched by Google and YouTube in partnership with the Institute for Strategic Dialogue (ISD), delivering in-school workshops and practitioner training sessions across the UK.[43] The programme is primarily aimed at teenagers, empowering them to protect themselves online. In just three years the programme has reached an estimated 55,000 teenagers and 650 educators across England, Scotland and Wales. Based on feedback, 84% of young people having gone through the programme feel confident they would know what to do if they encountered hate speech online.

[42] Be Internet Legends, https://beinternetawesome.withgoogle.com/en_uk/.
[43] Be Internet Citizens, https://internetcitizens.withyoutube.com/.

Media literacy is integral to the UK Government and Parliament's objective of strengthening online safety. We would therefore encourage the Committee to assess how the Government can expand on its work to ensure citizens are provided with the information and tools they need to stay safe online as part of their scrutiny of this bill.

6. The role of Ofcom

6.1 Pursuing a transparent and proportionate regulatory approach

Relevant Committee questions addressed:

- *Is Ofcom suitable for and capable of undertaking the role proposed for it in the draft Bill?*
- *Are Ofcom's powers under the Bill proportionate, whilst remaining sufficient to allow it to carry out its regulatory role? Does Ofcom have sufficient resources to support these powers?*

Ofcom's expertise in regulating complex issues relating to content, harm, and freedom of expression means it is well positioned to meet the challenges of regulating online platforms. We have a long track record of working cooperatively with regulators across the UK, including Ofcom. We welcome regulatory approaches which are transparent and proportionate, and look forward to a constructive and impactful dialogue with Ofcom once the framework comes into force.

However, Google is concerned that some of the draft Bill's proposals in relation to enforcement may undermine the framework's efficacy in tackling illegal content.

Use of Technology notices

Specifically, we are concerned about the potential requirement for services, pursuant to a Use of Technology Notice, to use third-party technology. We are proud of the technology we have developed to proactively detect and remove Child Sexual Abuse Material (CSAM) content from services, and we make these technologies freely available for other platforms to use. We believe regulation should encourage use of such technologies but not mandate them.

Online threats are fast-moving: bad actors find new ways to circumvent existing technology, and the new power in this Bill risks locking providers into using tools that have been 'gamed' by bad actors and are therefore less effective at identifying illegal content. It also risks reducing the incentive for services to differentiate themselves and compete online. We are also concerned about how requirements to use such technology can be reconciled with users' rights to private and/or encrypted communications.

For example, whether an individual piece of content amounts to terrorism content — as defined by the wide range of offences set out in Schedule 2 to the Bill — will depend on the

context in which it has been created and made available. It would not be appropriate in our view for Use of Technology Notices to require the proactive monitoring and removal of terrorist content in circumstances where technological solutions are not capable of taking into account all relevant context. Equally, it would not be appropriate to mandate this for grey-area cases that require nuanced decisions around content removal. Instead, we would suggest to build on the progress made through the Christchurch Call to Action and our joint work in the Global Internet Forum to Counter Terrorism (GIFCT) to work with governments, industry and civil society to prevent terrorists and violent extremists from exploiting digital platforms, including by sharing best practices.

We would encourage the Committee to reassess the draft Bill's proposal for Ofcom to be able to mandate services to use third-party technology to proactively identify content.

Criminal liability for individual directors

We understand the Government and Parliament's effort to ensure there are tough penalties to incentivise compliance. **However, the threat of criminal liability will incentivise service operators to protect themselves by automating content removal at scale, rather than making careful judgements over content removal.** We believe that the heavy financial sanctions included in the draft Bill – up to 10% of global revenue — will be more effective in incentivising companies to comply, without chilling investment in the UK's digital economy. Criminal liability for company directors sets an extreme precedent and sends a message to entrepreneurs about the risks of investing and building digital businesses in the UK. It also risks making the UK an outlier compared to the approaches being considered in the EU and elsewhere.

We believe the Committee should assess the proportionality of the reserved power to make individual directors criminally liable, which could have highly negative consequences for online expression, international precedent, and the digital economy.

6.2 Identifying an appropriate level of ministerial intervention

Relevant Committee questions addressed:

- *How much influence will a) Parliament and b) The Secretary of State have on Ofcom, and is this appropriate?*

The Bill outlines a clear role for Ofcom in providing oversight of in-scope services, but it could be further improved by providing more clarity over the role of the Secretary of State. Much of the detail of the regime is left to regulations to be made by the Secretary of State, including:

- The power to amend the online safety objectives for all in-scope services (Section 30).
- The power to require Ofcom to modify a code of practice to reflect Government policy (Section 33).

- The power to define the meaning of "priority illegal content" (Section 41), "primary priority content that is harmful to children", "priority content that is harmful to children" (Section 45), and "priority content that is harmful to adults" (Section 46).
- The power to define minimum standards of accuracy for the purpose of technology notices (Section 66).
- The power to define the threshold conditions for different categories of services (Schedule 4).

These provisions give the Secretary of State significant discretionary power and could ultimately undermine the framework's effectiveness in protecting users. The Bill is rightly focused on empowering an independent regulator to make careful, evidence-based judgements on how to strengthen online safety and protect freedom of expression, including by setting out a process under which Ofcom should gather evidence and consult before issuing Codes of Practice that provide clear guidance to providers. The risk is that the Secretary of State can intervene to significantly reshape the expectations placed on platforms, prompting uncertainty and confusion that complicates their efforts to comply with the framework and keep users safe. **Over-broad powers for the Secretary of State risk detracting from the important principle of regulatory independence**.

We would recommend the Committee further considers which powers it is appropriate to assign to the Secretary of State, and which powers it is more appropriate to assign to Ofcom.

6.3 Ensuring the regulatory framework is itself accountable

Relevant Committee questions addressed:

- *Are there systems in place to promote transparency, accountability, and independence of the independent regulator?*

We would encourage the Committee to also assess the potential for the proliferation of numerous Codes of Practice, overwhelming in-scope providers and rendering the framework less effective. As the Codes will be developed separately from the Bill itself, it is also important to consider how to mitigate the risk they go further than the Government and Parliament's intent in their impact on freedom of expression. The Government's Plan for Digital Regulation outlines the aim to remove unnecessary regulations and burdens where possible. The Committee could usefully specify sensible steps that Ofcom should take to ensure accountability for the regulatory framework's impact on innovation and freedom of expression. This includes specifying that regulators need to fully consider and demonstrate that new burdens imposed are necessary, including undertaking an economic impact assessment.

We recommend the Committee suggests broadening the requirement on Ofcom under clause 58 to assess the impact of any new Code, to ensure the assessment covers the impact on all in-scope service providers, and considers whether achieving the specific objective of the Code is proportionate to this impact.

Since Codes of Practice will help shape how providers comply with the framework, it is also vital that the process for developing the Codes is fully consultative, deliberative, and thoughtful. The opportunity for industry and civil society to provide written input will improve the Codes, ensuring that requirements are technically workable and mitigating any inadvertent impacts on user safety and freedom of expression online.

We would recommend the Committee proposes specifying in the Bill that Ofcom must provide the opportunity for consultation to all companies impacted by changes to the Codes of Practice.

Additional transparency provisions for Ofcom would also increase its accountability for ensuring the online safety framework meets its objectives. The draft Bill includes a requirement for Ofcom to produce transparency reports based on the transparency reports produced by in-scope providers. We would recommend broadening this requirement, and that the Committee should propose requiring Ofcom to publish annual transparency reports on the impact of the framework on online safety, freedom of expression online, and innovation in the digital economy.

Next steps

We hope that the evidence we have provided to the Committee is useful, as it carefully considers the important questions before it. Our recommendations are aimed at strengthening the Bill so that it provides the necessary clarity to deliver substantial improvements in UK users' online safety, while protecting freedom of expression and promoting innovation and competition between online services. The Bill represents an opportunity to set a positive example to the rest of the world. We welcome any questions on the evidence we have provided and look forward to discussing in further detail with the Committee.

28 September 2021

Annex - Summary of our response

2. Objectives

2.1. Objectives in the Bill and systemic approach

- By making platforms accountable for the processes they have in place to protect users, the draft Bill has the potential to lead to continued improvements in user safety across online services in scope.
- We encourage the Committee to maintain the Bill's focus on platforms' systems and processes.

2.2. Improving online safety and protecting freedom of expression online

- We strongly support the Government's objectives to make the UK the safest place to be online and to protect freedom of expression.
- We acknowledge and share this Government's concern with ensuring that online regulation supports UK users' ability to access and express a range of perspectives online.
- The Committee could consider how the Bill can be strengthened by clarifying that service providers must give equal weight to the duties to protect users' fundamental rights and the duties in relation to illegal and harmful content.
- Provisions in the draft Bill may amount to requirements for blanket monitoring and proactive removal of content, which can only be achieved at scale using automated tools. This could lead to large amounts of legitimate content being mistakenly removed, limiting UK users' ability to express themselves online.
- Without appropriate safeguards, AI tools can make errors that negatively impact on users' rights to express their opinion online and can entrench discrimination.
- We consider that the Bill could be clarified to expressly state that the safety duties do not of themselves mandate general monitoring and proactive removal.

2.3 Protecting children online

- We strongly support the Government's objective of protecting children online. We believe there is an opportunity for the Committee to ensure the Bill does not inadvertently block children and young people from accessing the internet.
- The ambiguity of the draft Bill's requirements in relation to children mean that service providers are incentivised simply to prevent access to their services to under 18s, thereby denying them the right to use such services.
- Instead of mandating different experiences or content for different age groups, the Committee could explore how the use of safety by design tools such as flexible parental controls, e.g. Family Link, could help to better meet the goal of improving safety.

- To best protect children while retaining their access to online services, it is vital that the Committee carefully considers how the Bill's child protection duties can be aligned with the Age Appropriate Design Code, notably focusing on requiring providers to address content that is inappropriate for persons under the age of 18.
- We would encourage the Committee to further consider how simplifying the definition of content that is harmful to children and the associated duties will make the framework more effective in driving improvements in child safety.

2.4 Protecting vulnerable people online

- We understand and share the Committee's ambition to protect users who are more vulnerable to harm online. We take a zero-tolerance approach to all forms of abuse.
- Simplifying the definitions of harmful content would make the framework be more likely to be more effective in achieving the Government's aims.
- We would encourage the Committee to consider how a simpler, less-prescriptive definition of harmful content would create a more flexible way for service providers to determine how best to protect their users.

2.5 Supporting economic growth and innovation

- Google supports the Government's ambition, as stated in its Plan for Digital Regulation, to encourage innovation by offering clarity and confidence to businesses and consumers.
- We believe the Committee could usefully shape the Bill to ensure businesses are provided with the necessary clarity and confidence to continue to invest and build their businesses in the UK.

2.6 Global parallels

- We would encourage the Committee to carefully weigh the potential consequences of the UK defining an approach that diverges so markedly from approaches being taken elsewhere.
- The Bill could lead to online platforms prioritising rapid removal of content aimed at UK users, which would disproportionately impact British citizens' freedom of speech compared to citizens in other countries.
- The risk of mandating the use of technology to monitor content is that innovative British companies will be put at a disadvantage relative to competitors abroad, and that future start-ups will choose to build their businesses elsewhere.
- Other legislation like the proposed DSA provides a greater degree of legal clarity in several respects. Notably it specifies that there should be no requirement to monitor all content, it does not require the age-gating of all services, it does not require the removal of legal content, it does not propose a power for regulators to mandate uses of

specific technology, and it does not include a deferred power to introduce criminal personal liability for companies' staff.

3. Content in Scope

3.1 Protecting users from illegal content

- The Bill's threshold for what constitutes illegal content covers a very broad and technically challenging range of content. If online platforms are held liable for content where they cannot be sure of its illegality, they will tend towards blocking legitimate but controversial content.
- We would encourage the Committee to consider how the definition of illegal content can be focussed on serious illegal activity.

3.2 Ensuring appropriate protection for democratic content and content of journalistic importance

- Google shares the Government's aim to ensure appropriate treatment of journalistic content and content of democratic importance. This is a complex issue that requires careful consideration to avoid inadvertently impacting user safety.
- We welcome the Government's approach whereby online platforms remain able to remove journalistic content and content of democratic importance where such content may be harmful and violates the platform's content policies.

3.3 Protecting users from the potential for harm caused by non-illegal content

- We recognise Parliament has concerns about specific categories of lawful content, such as disinformation and misinformation. This is a critical issue, and we have devoted careful attention to developing a multifaceted approach that seeks to limit harm while protecting freedom of expression.
- An effective regulatory approach that balances user safety with freedom of expression would be to require services to clearly communicate their content moderation policies to users and to be consistent and fair in upholding those policies, without dictating which types of lawful content should remain online and which should be taken down.

4. Services in Scope

4.1 Keeping users safe on search services

- Search engines play a uniquely important role in facilitating access to the internet, enabling people to access, impart, and disseminate information.
- We think it is right that search services should be subject to expectations to protect users. We've carefully developed content policies for Google Search to balance the real concerns about content safety, alongside the need for a search engine to provide access to information for users who are researching on almost any topic.
- The draft Bill implies the Government's intent to differentiate the framework's approach to search services. However, in reality the framework subjects search services to most of the same duties as user-to-user services. This approach is not appropriate as search services are fundamentally different to user-to-user services.
- To ensure the regulatory framework preserves the benefits that search services provide to users, we would encourage the Committee to consider meaningfully differentiated obligations for search engines that take into account their unique character and function.

4.2 A targeted approach on online advertising

- We have taken a very proactive and industry-leading approach to addressing the problem of paid-for advertising directing consumers to fraudulent products, working closely with the Financial Conduct Authority.
- We would recommend prioritising the DCMS Online Advertising Programme and the Home Office Fraud Action Plan as the appropriate vehicles to review online advertising.
- As action is already being taken against fraudulent advertisements, we believe that a more targeted approach to paid-for advertising would not only be more effective than including it within scope of the Online Safety Bill, but would also be more proportionate in its impact on the wider economy.

4.3 Ensuring strong protection for privacy

- Users expect a greater degree of privacy in relation to private services - such as direct messaging and closed groups - and such services should be subject to different requirements.
- There is a risk that the current wording in the draft Bill leads to significant monitoring of private communications, as providers seek to take action against content that is shared privately on the basis that it may, in theory, cause harm if shared publicly.
- We would recommend the Committee propose tightly specifying a differentiated set of obligations in respect to private services in the Bill.

5. Algorithms and user agency

5.1 Algorithms in Google's products

- Algorithms are critical for Google Search's mission of sorting the world's information and making it useful. Our algorithms surface relevant and high quality sources and prevent poor quality or harmful content from rising in search results.
- The algorithms used by YouTube to recommend videos are built to provide value to our users, by helping them find new content that appeals to their interests.
- Both Google Search and YouTube deploy algorithms to identify specific categories of content that may be illegal or in violation of our policies. However, algorithms are not infallible. We deploy safeguards to mitigate the risk that algorithmic tools mistakenly identify lawful content as illegal.

5.2 Algorithms and the Online Safety Bill

- Algorithm-based content moderation can be a powerful tool in identifying certain forms of content, but there are clear risks that its use can lead to mistakes that seriously impact on users' freedom of expression.
- We believe the Committee should propose limiting the content which is in scope of the most onerous safety duties to that which can be reliably identified by algorithms.

5.3 Empowering users

- It is important to recognise the role of user agency, the actions of individual users, and parental support in contributing to a safer online environment. Media literacy and digital citizenship can help people make the most of online opportunities while empowering them to protect themselves and their families from the potential risks.
- We would encourage the Committee to assess how the Government can expand on its work to ensure citizens are provided with the information and tools they need to stay safe online as part of their scrutiny of this bill and would welcome sharing our experiences from our Be Internet Legends and Be Internet Citizens programmes.

6. The role of Ofcom

6.1 Pursuing a transparent and proportionate regulatory approach

- Ofcom's expertise in regulating complex issues relating to content, harm, and freedom of expression means it is well positioned to meet the challenges of regulating online platforms.
- Online threats are fast-moving, as bad actors find new ways to circumvent existing technology, and the new power to mandate providers to use third-party technology risks locking providers into using tools that are less effective at identifying illegal content.

- We would encourage the Committee to reassess the draft Bill's proposal for Ofcom to be able to mandate services to use third-party technology to identify content.
- The threat of criminal liability will incentivise service operators to protect themselves by automating content removal at scale, rather than making careful judgements over content removal.
- We believe the Committee should assess the proportionality of the reserved power to make individual directors criminally liable, which could have highly negative consequences for online expression, international precedent and the digital economy.

6.2 Identifying an appropriate level of ministerial intervention

- The draft Bill outlines a clear role for Ofcom in providing oversight of in-scope services, but it could be further improved by providing more clarity over the role of the Secretary of State.
- Provisions in the draft Bill give the Secretary of State significant discretionary power and could ultimately undermine the framework's effectiveness in protecting users. We would recommend the Committee further considers which powers it is appropriate to assign to the Secretary of State, and which powers it is more appropriate to assign to Ofcom.

6.3 Ensuring the regulatory framework is itself accountable

- We would encourage the Committee to also assess the potential for the proliferation of numerous Codes of Practice, overwhelming in-scope providers and rendering the framework less effective.
- We would recommend the Committee suggests broadening the requirement on Ofcom under clause 58 to assess the impact of any new Code, to ensure the assessment covers the impact on all in-scope service providers and considers whether achieving the specific objective of the Code is proportionate to this impact.
- Since Codes of Practice will help shape how providers comply with the framework, it is vital that the process for developing the Codes is fully consultative, deliberative, and thoughtful.
- We would recommend the Committee proposes specifying in the Bill that Ofcom must provide the opportunity for consultation to all companies impacted by changes to the Codes of Practice.
- Additional transparency provisions for Ofcom would also increase its accountability for ensuring the online safety framework meets its objectives.
- We would recommend requiring Ofcom to publish annual transparency reports on the impact of the framework on online safety, freedom of expression online, and innovation in the digital economy.