

Google – supplementary written evidence (DAD0101)

Thank you very much for the invitation to appear in front of the Democracy and Digital Technologies Committee. As mentioned during the session, at Google, we believe the Internet and social media has had a transformational and positive impact on democratic debate and participation, but there is always more to be done as our technologies and society evolve. We are committed to playing a key part in this important work.

As requested, we have provided further information on a number of different areas relating to this discussion below.

Building trust in Google and our products

Our business relies on our users trusting us and the services we provide, and this is central to how we design and build our products. People continue to use platforms like Search or YouTube because they trust they will be able to find timely entertaining, informative and authoritative content. We try to build this trust through our approach to transparency, which is something we take very seriously. Google has made industry leading efforts in this area from our public How Search Works website which includes resources like our [Search Rater Guidelines](#) through to our [transparency reports](#) and our clear [community guidelines](#) which govern what is and isn't acceptable on YouTube.

We are keen to also improve transparency with policy makers, by providing written and oral evidence to Select Committees, such as this one, and we are pleased to have the opportunity to provide further information to you.

Making our platforms and products more understandable

At Google, we are constantly looking for ways to make all of our platforms easier to understand for the general public. This is shown through the products we provide to users, which are built with our core values of transparency, choice and control in mind. There are a number of ways that we are currently doing this across all of our platforms:

- On YouTube, our [community guidelines](#) clearly set out the rules for using the platform. This includes what content is and isn't acceptable, and how we protect the copyrights of our users. The guidelines are explained using straightforward language and illustrated using things like animations and videos to explain what they mean and how decisions are made. We have a dedicated [YouTube channel](#) called TeamYouTube, with over 7.75M subscribers, that keeps users updated about product changes, top tips and answers top questions about the platform. Through the [Creator Academy](#), we

offer additional support for the millions of people and businesses who use YouTube to upload videos and share them with global audiences.

- As we mentioned during our oral evidence session, we will continue to work on new ways to explain YouTube processes, such as content moderation in more detail, building on our efforts to explain how Search works (outlined below).
- We provide detailed information about our algorithms and ranking processes on our [How Search Works](#) site. The site includes information on our approach to algorithmic [ranking](#), including publication of our Search Rater Guidelines. We also offer [extensive resources to all webmasters](#) to help them succeed in having their content discovered online. We work hard to inform website owners in advance of significant, actionable changes to our Search algorithms and provide extensive tools and tips to empower webmasters to manage their Search presence - including interactive websites, videos, starter guides, frequent blog posts, users forums and live expert support.
- We keep users informed about our [privacy policies](#), using videos and illustrations to help our users understand them. We also embed the option for users to change their privacy settings directly through our privacy policies page, alongside these clear explanations. We regularly remind users to review their privacy settings and take their Privacy Checkup, which allows them to review their key settings in a few minutes, in an easy to understand format.
- We use easy to follow videos and illustrations to inform users about our privacy policies, so that they can better understand them. We have also embedded the option to allow users to change their privacy settings directly through our privacy policies page alongside these clear explanations. We regularly remind users to review their privacy settings and offer Privacy Checkups, which allows them to review their key settings in a few minutes, in an easy to understand format.
- We also look for ways to build transparency into our products directly, so that users can understand the privacy implications of their choices in situ. For example, if you add a Google Drive file to a shared folder, we'll check to make sure you intend to share that file with everyone who has access to that folder via an onscreen prompt.

This transparency is particularly important when it comes to the democratic process, and we demonstrate this via our work during election times in a number of ways:

- In 2019 we implemented a **[new process to verify advertisers for national elections and referenda in the UK](#)**. We also require that these verified election ads incorporate a clear **"paid for by" disclosure** and that all political ads are properly labelled.
- In May 2019, we expanded our portfolio of **[transparency reports](#)** to include a section on political advertising in Europe. This continues to provide voters with information about who purchases election ads on Google in Europe, and how much money is spent.

Digital literacy and education

We recognise there is always more to be done to ensure our users can understand how the online world works. The key to ensuring that users have a strong understanding of the platforms they use is by improving and expanding digital literacy and critical thinking, so that people have the skills to use and navigate new technologies safely. We believe digital literacy is an essential skill, and in 2019, we hosted the [Google Global Media Literacy Summit](#) in London to explore these important questions in more detail. We also sit on the advisory panel of Ofcom's *Making Sense of the Media* programme, which aims to improve online skills, knowledge and understanding of UK adults and children, and we look forward to engaging the Government on its upcoming Digital Literacy Strategy.

We provide resources to a wide array of external digital literacy programmes in the UK, including programmes which specifically encourage news literacy, such as, [NewsWise](#), [Student View](#), in addition to our own youth education programmes [Be Internet Citizens](#) and [Be Internet Legends](#).

Be Internet Legends teaches 8-11 year olds how to use the Internet safely and responsibly. Since launching in March 2018, the programme has already reached over 60% of UK primary schools, training more than 1.5m children via assemblies and/or teachers. **Be Internet Citizens** is aimed at teenagers and designed to teach critical thinking and digital citizenship; It helps teenagers spot fake news, recognise divisive 'us vs them' narratives, and respond effectively to online hate speech. So far, the programme has reached over 55,000 young people across the UK.

"How Search Works" initiative

Google handles trillions of searches each year and, every day, 15% of the queries we process are ones that we've never seen before. Building Search algorithms that can serve the most useful results for all these queries is a complex challenge that requires ongoing quality testing and investment. These ranking systems are made up of not one, but a whole series of algorithms. It is important to distinguish search engines, which index and rank webpages available on the Internet (reflecting the whole of the open Web), from video or social media platforms that actually host the content uploaded by the users.

To give users the most useful information, Search algorithms look at many factors, including the words of the query, relevance and usability of webpages, the expertise of sources and the user's location and settings. We have found that the query itself is by far the most powerful signal for which results are most relevant and useful. The weight applied to each factor varies depending on the nature of the query – for example, how new the content plays a bigger role in answering queries about current news topics vs. dictionary definitions. We explain how different signals are used on our [How Search Works](#) webpage.

To help ensure Search algorithms meet high standards of relevance and quality, we have a rigorous process that involves both live tests and thousands of trained external Search Quality Raters from around the world. These Quality Raters follow strict guidelines that define our goals for Search algorithms and are publicly available for anyone to see.

How News on Google works - providing authoritative content

The ranking and weighting of high quality content on Google Search and Google News

Trustworthy and timely information empowers people to better understand the world around them and helps them make educated decisions. Quality journalism provides that information when it matters the most, shaping our understanding of important issues and pushing us to learn more and seek the truth.

Google aims to make it easier to stay informed by using technology to organise what journalists are reporting about current issues and events. We don't have an editorial point of view. Instead, Google products are designed to connect users with a broad array of information and perspectives to help them develop their own points of view and make informed decisions. We're committed to fostering a healthy and diverse news ecosystem because we know journalism is vital to strong, well-functioning societies.

Across all our platforms we are committed to giving our users the best possible experience and this means serving authoritative, quality news and information, minimising exposure to poor quality content like misinformation, all the while ensuring a positive and safe experience. We also take additional steps to improve the trustworthiness of our results for contexts and topics that our users expect us to handle with particular care.

Google News' focus (coverage of current events) is narrower than that of Google Search. However, their goals are closely related. Both products present users with results that meet their information needs about the issues they care about. For that reason, both products have a lot in common when it comes to the way they operate. For instance, ranking in Google News is built on the same basis as that of Google Search ranking and they share the same defences against attempts at gaming our ranking systems. Similarly, both Google Search and Google News use algorithms to determine the ranking of the content they show to users.

Our algorithms are geared toward ensuring the usefulness of our services, as measured by user testing, not the ideological viewpoints of the individuals who build or audit them. The systems do not make subjective determinations about the truthfulness of webpages, but rather focus on measurable signals (such as expertise, trustworthiness, or authoritativeness) that correlate with how users and other websites value the reliability of a webpage on the topics it covers.

To better inform users on how we use technology to organise millions of pages of news content, in dozens of languages and make it discoverable to anyone, we have

launched a new website which explains [How News Works on Google](#). This provides an easy to understand guide for anyone interested in our mission, technology, content policies, and our efforts to support a healthy news ecosystem.

Partnership and information sharing with the research community

As we discussed during our appearance in front of the Committee, we are keen to understand the consequences of the use of our platforms, and encourage independent research into this topic.

Google shares data with academics and other platforms when it can be assured that the data is totally anonymous and cannot be re-identified. We consider ourselves to be best in class in releasing data and tools that help others. For example, we bring the benefits of AI to everyone by releasing over 5200 publications on computer science. Before opening up our datasets, we spend hundreds of hours standardising data and validating quality. We also ensure that datasets are shared in a machine-readable format that's easy for others to use, such as JSON rather than PDF. As data seekers might struggle to locate data scattered across repositories, we have also created a dataset search tool that helps people find data sources wherever they're hosted.

Some particular examples of how we share data with independent researchers include:

- We have also made more than 8 million YouTube datasets available publicly. [YouTube-8M](#) is a large-scale labeled video dataset that consists of millions of YouTube video IDs, with high-quality machine-generated annotations from a diverse vocabulary of 3,800+ visual entities. Our goal is to help accelerate research on large-scale video understanding, representation learning, noisy data modeling, transfer learning, and domain adaptation approaches for video.
- [CSAI Match](#) is our proprietary technology, developed by the YouTube team, for combating child sexual abuse imagery (CSAI) in video content online. It was the first technology to use hash-matching to identify known violative videos and allows us to identify this type of violative content amid a high volume of non-violative video content. When a match of violative content is found, it is then flagged to partners to responsibly report in accordance to local laws and regulations. Through YouTube, we make CSAI Match available for free to NGOs and industry partners like Adobe, Reddit, and Tumblr, who use it to counter the spread of online child exploitation videos on their platforms as well.
- The [What-If Tool](#) - a key feature of Google Cloud AI Explanations - makes research much easier for researchers. This tool allows users to visualise datasets, manually edit examples from a dataset and see the effect of those changes, and allows for automatic generation of partial dependence plots, which show how the model's predictions change when any single feature is changed. This allows developers to better understand their own models, and

to answer questions such as: How would changes to a datapoint affect my model's prediction? Does it perform differently for various groups – for example, historically marginalised people? How diverse is the dataset I am testing my model on?

- We also share significant amounts of information through our [Transparency Report](#). This includes regularly updated content about the information and data requests we receive from governments and, as of May 2019, a searchable and downloadable election ad library for EU election ads, including useful data such as which ads had the highest number of impressions, images of the latest election ads running on the Google platform, and information about how the ads are targeted in terms of age, gender, and location.

The 2019 UK General Election

How we helped to provide clear information during the 2019 general election

Google aims to make civic information more easily accessible and useful to people globally as they engage in the political process. We have been building products for over a decade that provide timely and authoritative information about elections around the world and help voters make informed decisions.

During the 2019 general election, we worked with the Government Digital Service to ensure that Search provided people with the information they needed when searching about the election. When users searched for instructions on how to register or how to vote, they could see those details directly on the results page.

Working to overcome challenges during elections

However, we also know that elections pose particular challenges that require all of our teams across Google and YouTube to work together. We have established dedicated teams across Europe who specialize in preventing abuse of our systems during elections, ranging from phishing schemes to attempts to hack Google products. These teams are trained to get ahead of abuse, clamp down on malicious activity, and react rapidly to breaking threats. Our teams work in partnership with other companies and law enforcement officials to identify malicious actors, disable their accounts, warn our users about them, and share intelligence.

We have also put in place additional requirements for political advertisers, meaning advertisers who wish to run ads that reference a political party, current elected officeholder or candidate for the UK Parliament have to be verified. This verification process allows us to incorporate a clear **“paid for by”** disclosure - a requirement which goes further than current UK law. We include these ads in our political ads Transparency Report and Ad Library, which we update regularly.

During the 2019 General Election, we identified that a human error caused 24 political adverts, out of a total of around 1,500 in our Ad Library, to temporarily not

appear in our Transparency Report. Once this issue was identified, we worked to update the ads in the report as quickly as possible, and they featured on the report approximately 48 hours later. We undertook this effort because we strongly support greater transparency in political advertising, and work hard to correct inaccuracies when we find them. We are constantly working to improve our reporting, and we appreciate feedback on how we can make it even better.

YouTube Content moderation

Community Guidelines on YouTube

Concern for the safety of our users is at the heart of everything we do. We have robust and effective policies and Community Guidelines that send a clear signal to users about what content is and is not acceptable on our platforms. For example, we have effective policies and strong enforcement against hate speech. We proactively remove hateful content from our platform and our policies extend beyond what is considered illegal hate speech. We also remove content promoting hate or violence against protected groups under the Equality Act, including attributes like age, disability, ethnicity, gender, race, religion, sex, and more.

We innovate to ensure that these policies are being effectively enforced. We use machine learning to find violative content at scale, and employ specialist teams to review the content flagged by those machines to make decisions based on context and to anticipate and understand the trends and themes that might produce harmful or violative content in the future.

Powerful machine learning is bolstered by people we've invested in, and we now have over 10,000 people working to address content that might violate our policies. Reviewers work around the world, 24/7, who speak many different languages and are highly skilled. We have three tiers of reviewers. Longer tenure, higher quality and deeper expertise is required to move through the tiers. From time to time, there are incidences where a further escalation is required - for example when a policy needs to be reviewed or where a very concerning case materialises. In these instances, a cross-company team is involved, often enlisting the help of outside academics and other independent experts to provide an external perspective.

We take our responsibility to keep hateful content off our platform seriously but we know we cannot do everything ourselves. That's why we partner with experts across the world, including government agencies, academics and NGOs through our Trusted Flagger programme, which allows us to understand issues with local nuances. In the UK, our partners include CTIRU, Henry Jackson Society, Parliament's Security Department, ICSR at King's College and many others.

Promoting Authoritative Content on YouTube

In 2017, we started to prioritise authoritative voices, including news sources like The Guardian, CNN, Jovem Pan and India Today, for news related queries both in Google News, our search results and "watch next" panels. For example, if a user is looking to learn more about a newsworthy event - for example, searching for

“Brexit”. While there will be slight variations, on average, 93% of the videos in global top 10 results come from high-authority channels. Authoritativeness is also important for evergreen topics prone to misinformation, such as videos about vaccines. In these cases, we aim to surface videos from experts, like public health institutions, in search results. Millions of search queries are getting this treatment today and we’re continually expanding to more topics and countries.

We know that reliable information becomes especially critical as news is breaking. But as events are unfolding, it can take time to produce high-quality videos containing verified facts. So we've started providing short previews of text-based news articles in search results on YouTube, along with a reminder that breaking and developing news can rapidly change. We’ve also introduced Top News and Breaking News sections to highlight quality journalism. In fact, this year alone, we’ve seen that consumption on authoritative news partners’ channels has grown by 60 percent. When an important event occurs, the Breaking News shelf is triggered automatically and populates content from partners such as those mentioned above.

For example, as we respond to the covid-19 crisis, we aim to surface videos from experts or media outlets, like public health institutions or news organisations, in search results and via information panels.

- We launched an information panel pointing users to an [article](#) from the World Health Organization about the Coronavirus, available in English, Arabic, French and Spanish.
- In order to reach more users with authoritative and up-to-date information about the Coronavirus, we have added localised information panels in 34 additional countries—including Korea, Japan, Italy—that link to local resources, such as Ministries of Health and Centers for Disease Prevention, and are translated into local languages.
- In addition, we are displaying a Top News Shelf on YouTube search results pages for certain Coronavirus queries, and a Breaking News Shelf may also appear on the home page for users who are regular consumers of news content on YouTube (see [Help Center article](#)). We are also displaying a COVID-19 news shelf across 20 countries, including the UK, Brazil and Indonesia.

Dealing with borderline content

The Committee also asked a number of questions about borderline content and how our guidelines manage this. Borderline content is content that could misinform users in harmful ways and content that comes close to but doesn’t quite cross the line of violating our Community Guidelines.

Our [Community Guidelines](#) set the rules for YouTube, and a combination of people and machines help us [remove more violative content than ever before](#). That said, there will always be content on YouTube that brushes up against our policies, but doesn’t quite cross the line. So over the past couple of years, we've been working

to raise authoritative voices on YouTube and reduce the spread of borderline content and harmful misinformation. This includes videos such as those promoting a phony miracle cure for a serious illness, claiming the earth is flat, or making blatantly false claims about historic events like 9/11.

Determining what is harmful misinformation or borderline, is tricky, especially given the wide variety of videos that are on YouTube. We rely on external evaluators located around the world to provide critical input on the quality of a video. And these evaluators use public guidelines to inform their work. Each evaluated video receives up to 9 different opinions and some critical areas require certified experts. For example, medical doctors provide guidance on the validity of videos about specific medical treatments to limit the spread of medical misinformation. Based on the consensus input from the evaluators, we use well-tested machine learning systems to build models. These models help review hundreds of thousands of hours of videos every day in order to find and limit the spread of borderline content. And over time, the accuracy of these systems will continue to improve.

Since January 2019, we've launched over 30 different changes to reduce recommendations of borderline content and harmful misinformation. In the US, the result is a 70% average drop in watch time of this content coming from recommendations of channels users are not subscribed to.

Our work continues. We are exploring options to bring in external researchers to study our systems and we will continue to invest in more teams and new features. Nothing is more important to us than ensuring we are living up to our responsibility. We remain focused on maintaining that delicate balance which allows diverse voices to flourish on YouTube, including those that others will disagree with, while also protecting viewers, creators and the wider ecosystem from harmful content.

Informing users when we remove content

The Committee asked about how we engage with users when content has been removed. We are open with users and content creators about decisions to remove content, or decisions to keep content on our platform when it has been flagged as potentially being violative. YouTube notifies all users who flag a video for violation of community guidelines on its eventual decision to remove or retain that video. In 2018, we also launched a reporting history dashboard which each YouTube user can individually access to see the status of videos they've flagged to us for review.

In order to make this process more understandable, we have also produced a video called "The Life of a Flag" which is designed to help build awareness of our flagging system. We always clearly communicate changes to our community guidelines on YouTube so that our users are made aware of what is and what isn't allowed on the platform.

At YouTube, we work hard to maintain a safe and vibrant community. We have community guidelines that set the rules of the road for what we don't allow on YouTube. We give YouTube creators the option to appeal video removals or restrictions in case they think we made a mistake. If a creator chooses to submit an

appeal, it goes to human review, and the decision is either upheld or reversed. The appeal request is reviewed by a senior reviewer who did not make the original decision to remove the video. The creator receives a follow-up email with the result. Our YouTube Community Guidelines enforcement report now has a dedicated page on [transparency about Appeals](#).

Statistics on videos removed from YouTube

Our [Transparency Report](#) outlines in detail the number of videos removed for violation of our community guidelines, and what proportion of it was originally flagged by algorithms compared to human detection. It does not record content which is flagged and then not removed. We have broken down the most recent information (October-December 2019) below:

- From October-December 2019, we removed a total of 5,887,021 videos for violating our community guidelines
- Of the videos removed, 90.7% were first flagged by machines rather than by humans, and of those, 64.7% were removed before any views.
- 9.3% of videos taken down during this period were first flagged by humans
 - Of this, around 5.5% were flagged by users, and the remaining 3.8% by our Trusted Flaggers, NGOs and Government agencies.

A further breakdown of information, including the number of removed channels and comments as well as historical information, is available in our [Transparency Report](#). Similar reports are also available for many of our other platforms as well as for political advertising, as we outlined above.

Competition in digital markets

Digital markets are highly competitive, with relatively low barriers to entry, short innovation cycles, and a high degree consumer mobility. Across a wide range of digital services, there are typically multiple strong competitors. For example, there are multiple different cloud storage services (e.g. Microsoft OneDrive, Google Drive, Dropbox), music services (e.g. Spotify, Apple Music, YouTube Music, Deezer), and search engines (e.g. Google, Bing, Duckduckgo). It's also worth noting that online competitors aren't always symmetrical. Google and Bing provide a broader "general purpose" search service, but there are also specialist search services, including Amazon and EBay for retail, Expedia and Orbitz for travel, and many more across a range of categories. These specialist services, individually and in aggregate, compete fiercely with Google and Bing. A [recent study](#) showed that 66% of consumers start their product searches on Amazon.

Competition in these sectors is enhanced by the ease with which users can switch between services. Switching involves merely typing a new web address, opening a new app, or clicking on a browser bookmark. Users are willing and able to switch between services to find those that best meet their needs. In the case of search engines, a 2016 [survey by the European Commission](#) found that "nearly eight in ten Internet users would probably change search engines if the search results provided

were not useful (78%, vs. 17% who disagree). Four in ten totally agree that they would do so (40%), while almost as many tend to agree (38%)”.

The ease of switching services, combined with the fact that most digital services are provided to consumers free of charge, means competition for user attention takes place along parameters other than price, in particular, via innovation and improved quality. The frequency of product innovations are a useful way of understanding how competitive these sectors are. Large technology companies continue to spend more on R&D than companies in any other sector, driven by competitive pressures. As an example, Google spent \$26 billion on research and development in 2019. If Google didn't continuously invest in the quality of its services, it wouldn't be as useful to consumers. This imperative to improve and continue to attract users is a sure sign of healthy, competitive dynamics.