

Written evidence submitted by Google (SMH0079)

X April 2025

Dear Chair,

Thank you again for inviting us to provide oral evidence to the Science, Innovation and Technology Committee on 25 February 2025. We are committed to responsible innovation, and grateful for the opportunity both to participate in this Committee's inquiry and to provide additional information in response to your follow-up questions.

Content Moderation at Google

Across all products and services, Google takes a comprehensive approach to online safety, involving large global teams and experts with a deep understanding of the key issues. In our experience, this is the best way to avoid over-reliance on any single approach and ensures we can address the different aspects that contribute to a safer online experience.

We support the goals of the Online Safety Act to keep people safe online and hope the act will raise the bar across industry. We are heavily engaged with Ofcom to show how we are meeting the duties set out in the act and codes of practice.

We have not waited to act on safety. Our approach has three elements: strong policies and guidelines, technological innovation and enforcement, and working in partnership with others. With respect to the removal of content from our platforms and services, we have two distinct, internal processes:

- (i) **Policy violations:** we remove content that violates the policy of a particular service. This process addresses a broad range of content, ranging from illegal material, to legal but potentially harmful content, depending on the product. We remove content for policy violations based on user reports and through our own content moderation processes.
- (ii) **Legal removals:** We remove content that our legal analysis has determined to be unlawful under applicable law, in response to a notification from a third party, such as a user or an authority. We measure the number of court and government legal removal requests biannually (across all products), and publish this information in [transparency reports](#).

Since Google launched the industry's first Transparency Report in 2010, we have been sharing detailed information related to content removal on a [quarterly basis](#). These reports break down data according to category of content, policies or laws implicated, Google services at issue, country-specific requests, and more. Additional reporting has documented that median time, in days, needed to take action on illegal content identified in accordance with Article 16 of the EU Digital Services Act. As shared in our February 25, 2025 [report](#), with respect to both Search and Ads, that figure was < 1 day. The OSA places a duty on categorised services to produce transparency reports, and for Ofcom to produce statutory guidance for such providers. We will have more to share through these transparency reports in due course.

Search and AI

Google Search relies on a combination of people and technology to enforce [its content policies](#), with machine learning playing a critical role in promoting content quality on the service. Google Search systems are built to identify and balance signals of authoritativeness so that people can find the most reliable and timely information available. Google Search algorithms look at many factors and signals to raise authoritative content and reduce low quality content. Google Search's publicly available website, [How Search Works](#), outlines the key factors that help determine which results are returned for a query.

Google Search works continuously to improve the effectiveness of automated systems to protect platforms and users from harmful content. To ensure our algorithms meet high standards of relevance and quality, Google Search has a [rigorous process](#) that involves both live tests and thousands of trained external Search Quality Raters from around the world. Raters do not determine the ranking of an individual, specific page or website, but they help to benchmark the quality of Google Search's results so that Google Search can meet a high bar for users globally. Under the Google Search [Quality Rater Guidelines](#), raters are instructed to assign the lowest rating to pages that are potentially harmful to users or specific groups or that are misleading, untrustworthy, or spammy. Between 2017 and 2024, Google Search launched multiple, new automated processes that detect webspam content. The typical precision of these processes is approximately 99 percent.

Given the volume of content on the web, the speed at which new content is created, and the need for human review to identify certain types of illegal content (in particular, where legal nuance, competing rights, and context are relevant), it is not possible to quantify the illegal content on the open web at any given time. Furthermore, as a search engine that does not, itself, host content, we cannot remove content from the web; this can only be done by webmasters themselves.

However, we can “delist” content, which is a process by which we “remove” links to certain web pages from the lists of displayed search results. This stops returning to users links to certain web pages (at times for all search queries, and at times following only certain search queries) and thereby prevents those web pages from being accessed through Search. Content that has been delisted is, however, still accessible via the open web, via direct navigation to hosting sites, social media platforms, and/or via searches on other search engines.

With respect to Google Search's share of the online search ecosystem, we believe there has never been more competition and innovation in information retrieval. The advent of AI is resulting in a new wave of competitive dynamics and Google competes fiercely to give consumers the best search experience. We help support businesses and consumers across the UK. Our products and services, like Google Search and Ads, contributed an estimated £118 billion in economic value to the UK in 2023, and we continue to deliver new innovations to UK users.

Advertising

We understand the importance of a safe and reliable ads ecosystem to protect our users and keep our ads platform safe, and to this end we require advertisers to comply with Google's robust [Ads policies](#).

Google uses a combination of automated and human evaluation to detect and remove ads that violate our policies and are harmful to users and the overall ecosystem.

Our AI and machine learning tools help us detect and remove policy-violating ads and malicious actors rapidly, often even before they gain significant visibility. In some cases, potentially policy-violating content is tagged for review by trained operators and analysts who conduct content evaluations that might be difficult for algorithms to perform alone, such as when an understanding of the context of the ad is required. The results of these manual reviews are then used to help build training data to further improve our machine learning models. The vast majority of ads are reviewed within one business day and our classifier removals are working constantly.

In 2023, we blocked or removed over 5.5 billion ads, slightly up from the prior year, and suspended 12.7 million advertiser accounts, nearly double from the previous year. Similarly, we work to protect advertisers and people by removing our ads from publisher pages and sites that violate our policies, such as sexually explicit content or dangerous products. In 2023, we blocked or restricted ads from serving on more than 2.1 billion publisher pages, up slightly from 2022. We are also getting better at tackling pervasive or egregious violations. We took broader site-level enforcement action on more than 395,000 publisher sites, up markedly from 2022.

Eating Disorder Content, Advertising and Monetisation on YouTube

Mental health issues like eating disorders can be isolating and stigmatising for people around the world. We want to make sure YouTube creates space for community, recovery, and resources, while continuing to protect viewers.

As such, YouTube's [Community Guidelines](#) prohibit content that promotes, glorifies or provides instructions for disordered eating. We have developed a comprehensive framework that involves age-restricting certain videos, and surfacing crisis resource panels under videos discussing eating disorders or at the top of search results related to eating disorders.

YouTube's crisis resource panels provide resources and information from global mental health support organisations. We recently expanded these panels to serve as a new full-page experience that will help viewers pause for a moment when they search on YouTube for certain queries related to suicide, self-harm, and eating disorders. Viewers will now more prominently see resources for third-party crisis hotlines.

With respect to content related to eating disorders, specifically, our advertiser friendly guidelines do not allow content that includes explicit references to eating disorders accompanied by any of the following references or context: lowest BMI or weight, showing a thin or emaciated body, weight or body-based shaming or bullying, reference to bingeing, hiding, or hoarding food, exercising to reach a caloric deficit, vomiting or abusing laxatives, checking weight loss progress, reference to hiding any of the above behaviors. Our policies apply to all portions of the creator's content (e.g. video, Short, or live stream, thumbnail, title, description, and tags).

In addition to abiding by YouTube's community guidelines and policies, creators seeking to participate in the YouTube Partner Program and monetise on YouTube must also ensure that their channels adhere to the YouTube [monetisation policies](#). For creators who are in the YouTube Partner Program, content that is eligible to monetise with ads must also follow our [Advertiser-friendly Content Guidelines](#). Under these guidelines, [certain content](#) is prohibited from earning ad revenue on YouTube. This includes violent content, hateful and derogatory content, and content that is gratuitously incendiary, inflammatory, or demeaning, among other categories.

YouTube [enforces](#) its monetisation policies in a variety of ways, including by: withholding, adjusting, charging back, or offsetting earnings or payment; limiting ad revenues from videos; suspending or terminating participation in the YouTube Partner Program; or terminating a channel, account, or disabling a user's access to the Service.

Check My Ads Report and Publisher Monetisation at Google

Google helps to enable a free and open internet by helping publishers monetise their content and advertisers reach prospective customers with useful, relevant products and services. Maintaining trust in the ads ecosystem, however, requires setting limits on what we will monetise. To that end, we have strict [Publisher Policies](#) that govern what type of content can run ads. Ensuring the health of our ad network is key to supporting the ecosystem of creators, publishers and advertisers that rely on Google ads to build their businesses.

Policies and Enforcement

All publishers admitted to any Google ad network are required to adhere to our Publisher Policies. Our [content policies](#) specifically prohibit monetisation of illegal content, dangerous or derogatory content, or misrepresentative content, among other categories. We are constantly evolving these policies to keep pace with evolving needs and indicators of abuse.

Google requires all publishers wishing to monetise their content to undergo a thorough review process to determine their eligibility for the AdSense program, including both automated and manual reviews. As a result, only a subset of all accounts that apply are approved for the program. When publishers first apply to monetise with AdSense we run a series of checks to better understand who they are. These include checks to determine whether accounts are related to previously terminated accounts. Publishers who pass this review are, at that point, eligible to submit specific domains for approval. At this step we run additional automated checks to understand how well the domain and subdomains adhere to our [AdSense Program Policies](#). Before final approval for AdSense monetisation, publishers must undergo additional manual checkpoints designed to determine their adherence to our content and behavior policies.

Content found to violate our policies will not be served any ads from Google. Publishers that violate the policies face serious consequences including limited ad serving – which would limit advertising based on certain suitability controls – page-level demonetisation, site-level demonetisation, account suspension, or account termination, depending on the severity of the violation. In 2023, the last date for which figures are [available](#), we blocked or restricted ads from serving on more than 2.1 billion publisher pages and took

site-level enforcement action on more than 395,000 publisher sites.

User Controls and Transparency

Google provides advertisers with robust [brand suitability controls](#) that allow them to finetune where their ads appear, and to exclude specific [websites](#), URLs, apps, and entire categories that they wish to avoid. Contrary to recent reporting, advertisers can also generate placement reports to find out where their ads have served and how many impressions they have received, including down to specific webpages. More information about these reports is available [here](#) for most Google Ad campaign types, [here](#) for Performance Max campaigns, and [here](#) for DV360.

Channel3 NOW

With respect to Channel3 NOW, it was escalated to our teams based on the potential sharing of false information, including a false name about the identity of the suspect in the Southport knife attack. Following review, we removed the channel for violations of our spam and deceptive practices policies. We demonetised the domain on July 31 2024, several months before the January 2025 publication of the Check My Ads report.

Implementation of Intermediary and Platform Principles (IPP) Pilot

As noted above, we appreciate the importance of a safe and reliable ads ecosystem and have built robust systems to enforce its advertising policies efficiently. As part of our commitment, we agreed to participate in the Intermediary and Platform Principles (IPP) Pilot which ran from 1 June 2022 for one year, to support the Advertising Standards Authority (ASA)'s regulation online together with other participating companies.

In 2023, the ASA issued a "[final report](#)", assessing the merits of the Pilot's six principles and the extent of companies' voluntary compliance. With respect to the steps Google has taken, the independent findings of the report demonstrated, "without question," that Google had implemented the Principles applicable to its services and submitted evidence to substantiate its implementation.

Subsequently, Google has joined a working group - convened by the Interactive Advertising Bureau (IAB) and the ASA - to assess the future of the IPP. This objective constitutes Action 5 of the [Online Advertising Task Force's Action Plan](#).

Researcher Access

Our [Google Researcher Programme](#) equips qualified researchers with resources to advance the public's understanding of certain Google services and their impact. Qualifying researchers from diverse disciplines may apply to obtain access to data from Google Maps, Google Play, Google Search, Google Shopping, and YouTube to study those platforms from a wide range of perspectives and research frameworks. Eligibility requirements for the Researcher Programme, including affiliation requirements, are publicly available [here](#), and the Programme's Acceptable Use Policy terms are available [here](#).

Additionally, through the [YouTube Researcher Programme](#), eligible academic researchers in the United States and abroad, from diverse disciplines, can apply to use YouTube data. This data access enables individuals to study a variety of topics and gain access to YouTube's public data corpus, derive insights from global YouTube data, and utilise support and technical guidance from YouTube. Eligibility requirements, including affiliation requirements for the Programme are available [here](#), and the Programme's Terms & Conditions are available [here](#). We continue to improve both research programmes, and we welcome feedback from participating researchers.

For ads served on Google platforms, including Google Search and YouTube, Google makes available to all researchers a searchable repository of advertisers and the ads served over a certain time period via the [Ads Transparency Center](#). Researchers are able to search for an advertiser, view information about the advertiser and the ads run by that advertiser during that period. This enables researchers to learn more about an advertiser and report an ad if they believe it violates our policies. Additional transparency information is available for election ads in the section of our Ads Transparency Center on [political advertising on Google](#).

Generative AI Content

With more people using AI to create content, we are [building on the ways](#) in which we help our audiences identify AI-generated content through several new tools and policies. As we engage in this work, we are mindful that visibly labeling AI-generated content does not always provide the clarity users need to make informed choices, and in some cases can actually cause more confusion. For this reason, our focus has been on providing meaningful context beyond how content was created across our services. More information is available in our recent [white paper](#) on AI context and provenance.

- **Automated YouTube Ad disclosures:** We require advertisers to [disclose](#) all election ads that include synthetic content that does not authentically depict real or realistic-looking people or events. Recently, we added Google-generated disclosures for YouTube Election Ads for some YouTube formats.
- **Content labels on YouTube:** We seek to provide viewers as much context as possible about the content they watch. At time of upload, we require users to disclose content that is meaningfully altered or synthetically generated when it seems realistic. For most videos, a label will appear in the expanded description, but for videos that touch on more sensitive topics — like health, news, elections, or finance — we also show a more prominent label on the video itself.
- **Providing users with additional context:** The [About this Image](#) feature in Search helps people assess the credibility and context of images found online. Our [double-check](#) feature in Gemini evaluates whether there is content across the web to substantiate the responses it provides to user queries.
- **Digital watermarking:** We're continuing to bring [SynthID](#) — embedded invisible watermarking — to additional Google Gen AI tools for content creation and more forms of media including text, audio, visual and video. For instance, images generated in Gemini and with our most recent Imagen 3 model, are embedded with SynthID watermarks.
- **Visible watermarks:** We are experimenting with visible watermarks in some surfaces, such as Google AI Studio, with the launch of our most recent multi-modal model.