

Mark Isakowitz
VP, Government Affairs &
Public Policy
Google LLC
25 Massachusetts Ave., NW 9th Floor
Washington, DC 20001

October 29, 2024

The Honorable Jeanne Shaheen United States Senate 506 Hart Senate Office Building Washington, DC 20510

The Honorable Rick Scott United States Senate 110 Hart Senate Office Building Washington, DC 20510

The Honorable Joe Manchin III United States Senate 306 Hart Senate Office Building Washington, DC 20510

The Honorable Mazie K. Hirono United States Senate 109 Hart Senate Office Building Washington, DC 20510

The Honorable Margaret Wood Hassan United States Senate 324 Hart Senate Office Building Washington, DC 20510

The Honorable Catherine Cortez Masto United States Senate 520 Hart Senate Office Building Washington, DC 20510

The Honorable Charles E. Grassley United States Senate 135 Hart Senate Office Building Washington, DC 20510 The Honorable Ron Wyden United States Senate 221 Dirksen Senate Office Building Washington, DC 20510

The Honorable Angus S. King, Jr. United States Senate 133 Hart Senate Office Building Washington, DC 20510

The Honorable Richard Blumenthal United States Senate 706 Hart Senate Office Building Washington, DC 20510

The Honorable Robert P. Casey, Jr. United States Senate 393 Russell Senate Office Building Washington, DC 20510

The Honorable Marco Rubio United States Senate 284 Russell Senate Office Building Washington, DC 20510

The Honorable Sherrod Brown United States Senate 503 Hart Senate Office Building Washington, DC 20510



Dear Senator Shaheen, Senator Scott, and Esteemed Senators:

Thank you for your letter inquiring about programs to locate and take down non-consensual intimate imagery (NCII) and involuntary synthetic pornographic imagery (ISPI). We condemn this abhorrent content and recognize the profound distress that individuals can experience when explicit or intimate imagery of them is discoverable online without their permission.

We set clear policies for the content and activity that is unacceptable on our platforms, and our Google Terms of Service include a provision requiring individuals that use our products and services to <u>respect others</u>, including their privacy rights. In addition, as described below, we take robust steps across Google Search, Google Play, and YouTube to help individuals affected by NCII and ISPI.

We continue to carefully evaluate how to better protect our users from this unacceptable material and enhance our efforts to combat this problematic content. Building upon our existing, robust information sharing with the National Center for Missing and Exploited Children, YouTube recently joined NCMEC's Take It Down initiative, which enables us to match hashes shared via their database that is specific to this type of content. Our teams are also engaged in ongoing conversations with the relevant teams at Stop NCII. In addition to working with organizations on a voluntary basis to tackle these important issues, we are proud to support S. 4569, the bipartisan Take it Down Act, that would give the force of law to takedown requirements.

To prevent abuse and harm on our products and services, we remain proactive and alert. We take special care to understand when users may be particularly vulnerable and take action when abuse is present. This is a large, dynamic challenge without a perfect answer, but it's a challenge we are dedicated to tackling — as we have been since Google's founding. In our actions to protect our users from harmful content like NCII and ISPI, we must consider factors such as the privacy rights of our users and the context and nuance around a piece of content. We are always aiming to increase the speed of our detection and response efforts, using both people and AI-driven technology.

## Efforts to Combat NCII and Other Types of Image-Based Sexual Abuse on Search

Google Search uses automated systems to crawl and index content from the web and other sources. These systems generate search results that provide useful and reliable responses to the billions of search requests that we process daily. Our systems are designed to prioritize what appears to be the most useful and helpful content on a given topic, as well as not to surface content that violates our content policies. In the limited and well-defined situations that warrant this, we may also take manual action, with humans reviewing cases where policy-violating content surfaces and taking steps to block it. More information is included below about our policies, reporting processes, and automated protections with respect to NCII and ISPI as they relate to Search. We continue to work on additional safeguards, including ranking improvements and tools to help people protect themselves at scale.

### Identification and Removal

Last year, we shared an update regarding an <u>expansion of our policies</u> that enable individuals to remove NCII from Search. Individuals can now also <u>request the removal of personal</u>, <u>explicit images</u> that they no longer wish to be visible in Search. For example, if a person identifies a personal, explicit image of themselves online that was published without their consent, they can



request the removal from Search. Additionally, if someone created and uploaded explicit content to a website, they can request its removal from Search if the content is being published elsewhere without approval. This policy complements our policies guarding against <u>ISPI</u>, as well as a reporting tool to facilitate removals of <u>irrelevant association with pornography</u>.

We have updated and simplified the forms individuals may use to submit removal requests for their personal sexual content on Search, as well as on websites containing their personal information or other content that may be removed under our Search product policies. Moreover, we allow people to upload multiple URLs so that any content a person identifies in search results can be reported through a single form. In addition to our user reporting forms, we also allow victims or authorized representatives to report NCII and other issues directly within the Image Search results page by electing an image result, clicking the three-dot menu in the image viewer, and selecting "Report this result." The goal is to make it easier for victims or authorized representatives to request removal of NCII they may find on Search.

### Automated protections

We know it can be difficult for survivors of NCII to navigate removal requests from multiple online services and platforms. When an image is removed from Search under our NCII or ISPI policies, we have systems in place to detect and remove duplicates of that image to reduce the need for victim-survivors to request removals one by one.

Using our own internal hashing technology, our systems detect and remove duplicates for the vast majority of NCII imagery reported and removed from Search. While Google makes best efforts to stop this content from appearing, we are aware that images can be modified to deliberately evade detection via current hash-matching technology. As a result, these "de-duplication" protections may not detect all manipulated (but visually similar) "near-duplicates". In <u>December 2023</u>, we further clarified how Google handles sites with a high proportion of non-consensual intimate imagery in our <u>spam policies</u> and <u>ranking systems guide</u>. If we process a high volume of such removals involving a particular site, we use that as a signal to improve our results. For sites that receive a high volume of NCII or ISPI removals, we <u>demote</u> other content from the site in our Search results.

## Additional protections

We provide an option for individuals to request that Search filter explicit results for Search queries similar to the one included in the NCIIremoval request. For example, if a user's removal request is related to the query (e.g., "name of person leaked nudes") and that request is approved, then we may filter explicit results for that query going forward. This mitigates the need for users to continually submit removal requests. Given the dynamic and ever-changing nature of the web, automated systems are not able to catch every explicit result that may appear.

### Open web content

Google Search shows information gathered from websites across the web. Even if content is removed from Google Search, it may still exist on the web. This means someone might still find the content on the page that hosts it, through social media, on other search engines, or in other ways.



To this end, it is important for such content to be removed also from where it is being hosted. We provide <u>information</u> to help users understand how to request removals from hosting websites and webmasters, if they feel comfortable doing so.

# Keeping Google Play a Safe and Respectful Platform

In addition to our policies on Google Search, we have standards defining and prohibiting content that is harmful or inappropriate for individuals that access applications through the Google Play Store. We enforce these policies and take action to keep the platform safe and protect the user experience. More information about our efforts across Google Play is below.

## **Applicable Policies**

Google Play's applicable <u>policies</u> prohibit apps that contain or promote non-consensual, pornographic, or otherwise illegal sexual content. We do not allow apps that contain or promote content associated with sexually predatory behavior or distribute non-consensual sexual content. These policies prohibit apps that degrade or objectify people, such as apps that claim to undress people or see through clothing, even if labeled as prank or entertainment apps. They also prohibit content or behavior that attempts to threaten or exploit people in a sexual manner, such as creepshots, hidden camera, non-consensual sexual content created via deepfake or similar technology, or assault content.

Our Play Store standards set clear expectations for developers, and we apply them to applications utilizing generative AI models as well. As these generative AI models become more widely available to developers, we <u>remind</u> developers that to ensure safety and in accordance with Google Play's Policy Coverage, apps that generate content using AI must comply with existing Google Play Developer Policies and contain in-app user reporting or flagging features that allow individuals to report or flag offensive content to developers without needing to exit the app. Pursuant to our <u>Policy Coverage</u>, these apps also agree to prohibit and prevent the generation of <u>restricted content</u>, which includes the generation of non-consensual sexual content and content that enables <u>deceptive behavior</u>.

#### **Review and Enforcement**

When app developers submit their apps to Google Play, Google utilizes a combination of automated processes and human review to assess these apps before they can be published for distribution on the Play Store. The automated processes—which include static and dynamic components—scan an app's code, app images, the developer profile, and the app description. More information about our analysis and review process is publicly available.

Google Play may take <u>enforcement actions</u> in response to a developer's violative app. If an app or developer account violates any of our policies, we may reject, remove, suspend, or limit the visibility of the app. Google Play may also restrict a developer's account or terminate the developer's account entirely, for violations. In 2023, we prevented 2.28 million policy-violating apps from being published on Google Play.



Individuals are also able to <u>report</u> an issue with an app on Google Play, including whether an app includes violative content, such as nonconsensual explicit imagery. If we find that an app violates any of our policies, we take appropriate action—up to and including termination of individual or related Google Play Developer accounts or the app.

# Maintaining a Safe and Responsible Environment on YouTube

Responsibility remains at the center of all that we do on YouTube. More information on YouTube's policies and their enforcement is below.

### Applicable Policies

We believe it is in everyone's interest to maintain a healthy ecosystem on YouTube. To this end, all content uploaded to YouTube is subject to YouTube's <u>Community Guidelines</u>—regardless of how it is generated.

Explicit content meant to be sexually gratifying is <u>not allowed on YouTube</u>. This policy applies to real-world, dramatized, illustrated, and animated content, including sex scenes, video games, and music. Prohibitions under this policy cover nonconsensual sex acts or the promotion or glorification of nonconsensual sex acts, as well as nonconsensual zooming in or prolonged focus or emphasis on private areas for the purpose of sexual gratification.

YouTube also <u>prohibits harassment and cyberbullying</u>, including content that contains unwanted sexualization of an identifiable individual. This prohibition includes content that describes someone in a lewd, degrading, and sexually explicit manner; content that fantasizes about, threatens, or supports sexual assault; as well as content that shares, requests, or shows how to distribute nonconsensual intimate or explicit imagery.

We also have long-standing <u>policies</u> that specifically prohibit technically manipulated content that misleads viewers and may pose a serious risk of egregious harm. These policies encompass certain types of misinformation that can cause real-world harm and certain types of technically manipulated content.

#### **Enforcement**

If content, including <u>external links</u> in the content, violates our Community Guidelines, <u>we remove the content</u>. We may terminate a channel or account for repeated violations of the Community Guidelines or Terms of Service. We may also terminate a channel or account after a single case of severe abuse, or when the channel is dedicated to a policy violation. This can include clickable URLs, verbally directing users to other sites in video, and other forms.

We also ask individuals to <u>report</u> inappropriate or abusive content. We offer a separate complaint process for individuals to <u>report Al-generated or other synthetic content</u> that looks or sounds like them, and ask that the content be removed. Reporting content is anonymous, so other users cannot tell who made the report.



We work continuously to improve our efforts through enhanced detection and enforcement, relying on a combination of technology and people. Our automated detection systems are an area of significant investment for YouTube, and our engineering teams continuously evaluate their efficacy and make improvements. These systems help our human review teams remove content at scale, with the speed and volume that could not be achieved with people alone.

To measure our progress on removing violative videos, we have developed a metric called <u>Violative View Rate</u> (VVR). This metric estimates the percentage of views on violative videos. We started tracking VVR in 2017, and share the rate in our quarterly <u>Community Guidelines</u> <u>Enforcement Report</u>. In Q2 2024, the VVR was 0.09%-0.11%, which means that out of every 10,000 views on YouTube, only 9-11 come from violative content.

#### Generative AI and CSAM

Al-generated CSAM, or computer-generated imagery depicting child sexual abuse, is a threat that Google takes very seriously. Our work to detect, remove and report CSAM has always included violative content involving actual minors, modified imagery of an identifiable minor engaging in sexually explicit conduct and computer-generated imagery that is indistinguishable from an actual minor engaging in such conduct.

Google places a heavy emphasis on child safety when creating our own GenAl models and products. We follow Google's responsible generative Al principles in protecting all of Google's publicly available models and the services built on top of these models.

We deploy a variety of child safety protections for our GenAl models and products. This can include protections against the presence of CSAM in the training data underlying our models, against CSAM-seeking and -producing prompts, and against violative outputs. We also conduct robust child-safety testing on our models prior to public launch to understand and mitigate the possibility of CSAM being generated.

We work with others in the child-safety ecosystem – including the Technology Coalition and child-safety NGOs – to share and understand best practices as this technology continues to evolve. We also signed on to the <u>Safety by Design Generative Al principles</u> — developed by Thorn and All Tech is Human – that complement our existing work to prevent the creation, dissemination and promotion of Al generated child sexual abuse and exploitation.

Additionally, Google <u>removes images</u> of anyone below the age of 18 from search results at the request of the individual under 18 or their parent or guardian, with the exception of cases of compelling public interest or newsworthiness. This means these images won't appear in the Images tab or as thumbnails in any feature in Google Search.

\* \* \*

We are deeply committed to addressing non-consensual intimate imagery, and our work to combat this malicious content persists. We look forward to continuing to collaborate with legislators,

Male W/ salonety



non-governmental organizations, and affected individuals to develop innovative solutions that address these complex challenges.

Thank you, again, for the opportunity to respond to your inquiry.

Sincerely,

Mark Isakowitz

Vice President, Government Affairs and Public Policy