Karan Bhatia
VP, Government Affairs & Public Policy
Google LLC
25 Massachusetts Ave., NW - 9th Floor
Washington, DC 20001
bhatiakaran@google.com

July 31, 2019

The Honorable Adam B. Schiff
Chairman
House Permanent Select Committee on Intelligence
2269 Rayburn Building
Washington, D.C. 20515

Dear Chairman Schiff:

Thank you for your letter to Sundar Pichai and Susan Wojcicki dated July 15, 2019, and the opportunity to continue the discussion regarding disinformation and synthetic media, or deep fakes. As we have discussed with you and your staff, Google takes these issues of election security and election interference very seriously.

1. **How many YouTube users viewed the manually altered video of Speaker Pelosi before YouTube removed it? What triggered the review process before the video was taken down, and how long did the review take to complete?**

   We took action on these videos promptly under our Community Guidelines, which prohibit certain deceptive practices that aim to take advantage of the YouTube community. We can also share that these videos did not surface prominently on YouTube. In fact, search results and "watch next" recommendation panels related to Nancy Pelosi include videos from authoritative sources, usually at the top. Broadly speaking, our enforcement activities rely on a combination of machines and people. We do not disclose enforcement details publicly to prevent bad actors from manipulating our systems, but we would be happy to provide more detail in a closed-door briefing.

   2. Does Google have a written policy on deep fake content on YouTube or its other platforms, including use in advertising? If so, will you provide it in response to this letter? If not, are you developing such a policy and when will it be finalized?

We take the threat of manipulated media very seriously, whether they are AI-generated or low-tech edits. YouTube has clear policies that outline what content is not acceptable to post and we remove videos violating these policies when flagged to us. We are always working to invest in and improve on our processes and technology to enforce our guidelines, including against potential threats related to synthetic media and disinformation.

YouTube's Community Guidelines prohibit certain deceptive practices that aim to take advantage of the YouTube community, including in some contexts those involving the technical manipulation of content. Google has additional policies in place against misrepresentation, including for advertisers, which looks at the behavior of content creators. These advertiser policies also apply on YouTube.

We are always looking into new potential threats related to personal or societal harm arising from new technologies, including this one, and may further update our policies in the future if we identify gaps that are not currently covered by our existing rules or systems. For example, we recently updated our Google policy on involuntary pornographic imagery (colloquially referred to as 'revenge porn') to cover fake imagery in addition to real imagery.

3. Are fake images or videos that realistically portray individuals saying or doing something they never did allowed on YouTube, including use in advertising? Under what circumstances, if any, would Google remove such content and block its upload?

As described above, YouTube's Community Guidelines prohibit certain deceptive practices that aim to take advantage of the YouTube community, whether the content comes from a YouTube creator or an advertiser. YouTube removes content that violates its Community Guidelines.

4. Is Google conducting research into techniques for automatically detecting deep fakes and other forms of machine-manipulated media on its platform? To the extent machine-manipulated media is detected upon upload to a Google platform, will Google take specific steps to dampen the virality of such content, take it down completely, or require a human review for politically relevant content?

Across Google and YouTube, we are cognizant of the threat posed by these technological developments. We are exploring the field of detection, working closely with numerous academics and industry experts around the world. Deep fakes are a new form of media manipulation, but not the first time we've faced this type of challenge. Previous issues have included copyright infringing content (which we've

addressed with Content ID) and spam views (which we combat with anti-spam detection).

YouTube has been dealing with manipulated media ever since its early days, as simple edits in a video can alter its meaning and give a very different image of reality than originally intended. Our Deceptive Practices policies are geared to protect the YouTube community from being taken advantage of by those who would undertake such manipulations, whether AI-enabled or otherwise.

Beyond our policies, we actively engage in advancing research and best practices in defending against risks that may arise from 'deep fakes'. That's why we partner with many in academia to advance deep fake detectors. We are also exploring ways to help third parties build their own capabilities in this space. For instance, in January we released a dataset of audio synthetic media datasets to researchers as part of a global, open competition to develop and train new detection models. We've also been exploring resources for journalists and researchers to aid in identifying manipulated images or videos.

More broadly, we work to address disinformation by elevating content from authoritative sources, taking action against malicious and deceptive behaviors, and providing users with context about the content they watch. We continue to improve our ranking algorithms on Search, News, and YouTube so that we become better at identifying authoritative sources across the board, which are less likely to spread "deep fakes" or other forms of misinformation. We are also adding more context and indicators of trust, to make it easier for users themselves to determine whether a piece of content is trustworthy, including information panels and fact-check labels to give users more context. These features are all cues for users to look beyond a single headline or piece of content, and instead explore numerous sources on an issue to form their own opinions.

It's a responsibility we take seriously and will continue to invest in -- including integrating feedback from external specialists where relevant, to help improve our collective resilience and to further the public's awareness of the growing capabilities of malicious actors. We remain committed to make progress on this issue.

Thank you for the opportunity to respond and we look forward to continuing to work with you.

Sincerely,

Karan Bhatia
Vice President, Government Affairs and Public Policy