

Attachment—Additional Questions for the Record

**Subcommittee on Communications and Technology and
Subcommittee on Consumer Protection and Commerce
Joint Hearing on
“Fostering a Healthier Internet to Protect Consumers”
October 16, 2019**

Dr. Hany Farid, Professor, University of California, Berkeley

The Honorable Anna G. Eshoo (D-CA)

- 1. Platforms have been criticized for doing too much moderation (particularly with respect to political biases) as well as doing too little moderation (particularly with respect to streaming of mass shootings). Which criticism, if any, is more accurate?**

Response: Although these criticisms are not mutually exclusive, I don't believe that the data supports both claims.

There is significant evidence that online platforms are not doing enough to remove child sexual abuse material, terrorism and extremism material, dangerous and deadly conspiracy theories, dis- and mis-information campaigns, and the sale and distribution of deadly drugs, weapons, and the illegal animal trade. At the same time, there is little evidence – beyond anecdotal – that online platforms are moderating with a systematic political bias.

The Honorable Kathy Castor (D-FL)

- 1. On June 19, 2019, The Verge published an investigation into one of Facebook's content moderation sites in Tampa, FL, which is operated by the firm Cognizant. The article details allegations of appalling working conditions including sexual harassment, verbal and physical fights, theft, and general filthiness in addition to adverse mental health effects associated with the nature of their work.**
 - a. Operationally, how should tech platforms moderate their content? What role should human content moderators play? What role should technology play?**

Response: The issue of horrific working conditions for Facebook moderators is not new, and dates back at least two years – see, for example:

“The Impossible Job: Inside Facebook’s Struggle to Moderate Two Billion People” by Jason Koebler and Joseph Cox, published in *Vice* on August 23, 2018.

“Scouring Facebook for disturbing content: How risk to moderators is raising concerns” by Davider Mever, published in *ZDNet* in on March 28, 2017.

Because the technology sector has not invested enough in developing automatic algorithms for content moderation (except in the case of copyright infringement, as mandated by U.S. law, and legal adult pornography, as needed to appease advertisers), human moderation remains necessary. Despite rapid advances in artificial intelligence and machine learning, it remains likely that even with significant investments in automated technology, human moderators will remain necessary for years to come.

Facebook is fond of citing their 30,000 moderators as evidence of their commitment to moderating their platform. While this number may sound impressive, it is almost well below the number of moderators that Facebook needs. According to [1], across major US cities, there are an average of 50 employees and officers per 10,000 citizens, yielding a law enforcement to citizen ratio of 1:200. There are approximately 2.3 billion active monthly Facebook users and according to Facebook they have 30,000 moderators. Let’s be generous and say that they have another 10,000 Facebook employees who work as part of the safety/integrity team. This yields a safety to user ratio of around 1:60,000. This is a factor of 300 times less than how we police our streets. It is clear from these numbers and from reports on how overwhelmed moderators are, that Facebook has not committed nearly enough resources to manage their services.

Facebook, and the other technology giants, must invest significantly more into technology to automatically moderate their platforms with the goal of reducing the burden placed on human moderators that are now viewing some of the most horrific and vile content for upwards of 10 hours per day. At the same time, the industry must establish more humane working conditions for these workers or have them mandated by Congress. And, the industry must stop out-sourcing this work and treat these moderators the same way that they do their cherished engineers.

[1] <https://www.governing.com/gov-data/safety-justice/law-enforcement-police-department-employee-totals-for-cities.html>

- b. What standard should a private company use to evaluate content? “Quasi constitutional”, a “community standard” established by the company along the lines of other private media, other?**

Response: The standard is and should be terms of service / community standards. Facebook and YouTube routinely ban constitutionally protected speech – most notably legal adult pornography – as is their right.

- c. Given that private companies are not governed by standards that government would be when it decides not to post content, why do content moderators have to spend so much time reviewing and in such great detail evaluating explicit, violent, or hateful content? What value is there to society and the site owner to work to ensure that such explicit, violent, or hateful content is given every opportunity to be posted?**

Response: There is no inherent value. This is not an issue of freedom of speech or expression. The issue is that taking down content is bad for business and so these companies create absurdly complex, contradictory, and constantly changing rules in attempt to take down as little content as possible while feigning concern for freedom of speech. This is not a speech issue (as seen by their aggressive and effective of adult pornography) – it is simply an economic issue. Let's stop pretending otherwise.

- d. This explicit, violent, or hateful content often is known to be inconsistent with the tech platform's content bylaws. Why do tech platforms, like Facebook, force content moderators to not only look at but also evaluate in great detail explicit, violent, or hateful content that is often inconsistent with the tech platform's bylaws?**

Response: Please see response to part (c).

- e. Should content moderators have more leeway to ban harmful content so they don't have to look at it over such lengthy time periods and evaluate the content in such detail?**

Response: Yes. Facebook and the like could create much simpler and easier to enforce guidelines which would in turn make content moderation easier to implement. This, however, runs against their core business model and so I am not optimistic that they will change their practice without significant public or regulatory pressure.

- f. What should industry best practices be for treating content moderators? Should Congress play a role in ensuring worker rights in this unique industry? If so, how?**

Response: I am not an expert in worker's rights, but this is an important question. I believe that Congress should act to protect workers because the technology sector has shown over the past two decades that they are not able to self-regulate and that they routinely put profit and growth ahead of all else.

- g. Is it common practice among tech platforms to use contractors to conduct content moderation for their sites? Why do some tech platforms use contractors to conduct content moderation for their sites? Should tech platforms do this?**

Response: I am only aware that Facebook out-sources most of their content moderation. I am not aware of how Google/YouTube or Twitter employ their moderators. I cannot say for sure why Facebook out-sources content moderation, but it is reasonable to assume that this is the most cost-effective approach and it allows Facebook to wash their hands of the ugly business of content moderation.

The Honorable Lisa Blunt Rochester (D-DE)

1. What can the federal government do to improve the capacity and ability to effectively moderate online content, including technological research?

Response: As we discussed at the hearing, modest changes to Section 230 would go a long way to forcing the technology sector to invest in more effective technological and human moderation. Despite years of public outcry and bad press, profits at Google and Facebook are up. Change will only happen when these companies are held financially responsible for their failure to create safe products that don't lead to the disruption of our democratic elections, don't lead to horrific violence against our citizens, don't lead to allowing child predators to freely exploit children, and don't lead to the daily abuse and marginalization of women and under-represented groups. Like every other industry, the technology sector should be held responsible when their products are unsafe and lead to real and measurable harm.

The Honorable Tom O'Halleran (D-AZ)

1. Dr. Farid, in the testimony of Dr. McSherry on behalf of the Electronic Frontier Foundation, she states how changes to Section 230 could increase liability risks for platforms and force some to over-censor due to a lack of resources to review content. Dr. McSherry used Craigslist's decision to remove its personal ads section as an example.

a. With the increasing amount of user-generated content being published on platforms daily, what do you believe to be the correct balance between using algorithms and human reviewers for platforms moderating content?

Response: I was part of the team that in 2008 developed photoDNA, a technology designed to find, remove, and report child sexual abuse material. At the time we heard from the EFF and others that deployment of this technology would lead to the over moderation of other content. This slippery slope argument is constantly trotted out anytime there is a discussion of content moderation. I find this argument lazy and not backed by the evidence: photoDNA did not, as predicted, lead to over removal of material.

Today, automatic removal of *re-uploaded* content is highly efficient and effective once content has been identified by human moderators as being illegal or a violation of terms of service. Human moderators, however, still need to make the initial determination of what material should be removed. Moving forward, new technologies can and should be developed to automatically flag problematic content, thus reducing the burden on human moderators. I do not, however, foresee the ability to completely remove human moderation in the coming years.

The Honorable Greg Walden (R-OR)

1. In a letter submitted for the record by TechFreedom, the author states,

“[the Republican staff memo] then claims that “platforms” have failed to meet their end of the bargain: “Internet platforms have, in many instances, benefitted from the ‘shield’ without using the ‘sword’ as intended.” Both claims are false: the first misrepresents the legislative history of Section 230 and the second fails to acknowledge how much interactive computer service providers, both large and small, wield the ‘sword’ of content moderation—and why they do so, without a legal mandate to.

Is the author of that letter correct: have interactive computer service providers met their end of the bargain?

Response: I do not believe they have. From the earliest days of the modern web, the technology giants have followed a similar pattern when it comes to dealing with everything from child sexual abuse material, terrorism and extremism, dangerous and deadly conspiracy theories, dis- and mis-information campaigns, and the sale and distribution of deadly drugs, weapons, and the illegal animal trade: Deny the problem exists, minimize the extent of the problem, concede that the problem is real but deny that a solution is possible, and once there is sufficient public or regulatory pressure, respond as anemically as possible.

Nearly everyone agrees that the technology sector is not doing enough to reign in the abuses and misuses on their services, including web inventor Tim Berners-Lee (see, Mr. Berners-Lee’s *New York Times* op-ed “I Invented the World Wide Web. Here’s How We Can Fix It.” Published on November 24, 2019). The technology sector has consistently put growth and profit over all else and has consistently hid behind Section 230 when called out on their failures.

2. In an October 15, 2019 letter to the Energy and Commerce Committee, TechFreedom states, “The [Republican staff] memo appears to suggest that the shift towards an ‘advertising-centric business models [sic] built upon user-generated content’ has made websites less willing to wield the sword of content moderation. In fact, just the opposite is true: relying on advertising generally gives platforms more of an incentive to monitor and remove objectionable user content.”

That may be true for monitoring and removing objectionable content placed next to ads, but is that the case for affiliated products or platforms that are not as proximate to ads?

Response: The data simply does not support TechFreedom’s argument that an advertising-based economy encourages removal of harmful or illegal material. In fact, all evidence is to the contrary. YouTube, for example, has for years allowed child predators to linger on their services with impunity. It wasn’t until more than three separate Disney-led advertising boycotts, spanning

several years, that YouTube responded (although with only a limited effort) to reduce the exposure of children to predators and inappropriate content. Facebook has publicly (and almost proudly) claimed that they will allow anyone to explicitly lie in political ads and target advertise these lies with laser focus. Twitter has for years been unable or unwilling to reign in daily abuse, often directed at women and under-represented groups. And, for years, Google has repeatedly failed to remove illegal and dangerous content from their search engine.

The simple fact is that when services like Google/YouTube, Facebook, and Twitter are free, these companies are not answerable to us the public. They need only appease the advertisers which in turn need only see – with few exceptions – a return on their advertising dollars. The advertising driven technology sector – once thought to be a boom for the public – is the underlying poison of the today's internet. This business model puts engagement, views, likes, and shares, ahead of all else.