



Digital, Culture, Media and Sport Committee

House of Commons, London SW1A 0AA
Tel 020 7219 6120 website www.parliament.uk/cms

Twitter UK
20 Air Street
Soho
London
W1B 5DL

By email

4 May 2020

Dear Ms. Minshall,

Follow-up questions from our session on the COVID-19 Infodemic

I am writing following your appearance before the DCMS Sub-Committee on Online Harms and Disinformation on Thursday morning.

We were deeply disappointed with the evidence given by all three social media companies, which fell short of the standards we expect in terms of clarity and openness. Witnesses seemed unwilling to answer direct questions, relying instead on pre-prepared answers, some of which seemed to be lifted from a press release rather than emerging from a statement of fact. This defensive positioning is deeply unhelpful, and is not a meaningful way of engaging with parliamentary scrutiny.

Ironically, the expert panel preceding your own evidence suggested that social media companies had made real efforts in some of their responses to the crisis. But we were not able to identify these positive trends due to oblique answers that did not deal with the question posed.

As a result, as well as asking follow-up questions now in writing, we need to leave open the option of recalling the three companies – represented this time by senior executives who have direct knowledge of, and are accountable for, company policy.

The Sub-Committee requests information on:

1. What proportion of identified accounts spreading misinformation and disinformation are bots or automated accounts or use automation in some way (irrespective of the number of overall/legitimate accounts using automated functions like scheduling)?
2. If verification is simply validation of identity, why does Twitter not roll out verification to all users who opt in (leaving the option of anonymity or pseudonymity for those who choose not to) to tackle disinformation, abuse and other harms?
3. Have you taken specific action against the accounts of politicians, celebrities, influencers, etc, such as by removing verified status or deleting accounts? What action is typically taken and are these accounts notified about why action is taken?
4. What defensive measures are used, such as how you identify specific tweets, hashtags and retweets, bots or automated accounts and hostile foreign actors that spread mis/disinformation, and how you disrupt or remove this activity? (Please provide confidential information in a supplementary response in the interest of transparency.)



Digital, Culture, Media and Sport Committee

House of Commons, London SW1A 0AA

Tel 020 7219 6120 website www.parliament.uk/cms

5. Facebook specifically allows users to flag posts specifically as 'false news'; Twitter allows users to report fake accounts but not false news specifically. Does your company have any plans to roll out functionality for users to specifically report posts as misleading, and if not why not?
6. Twitter's spokesperson, Katie Rosborough, has previously stated that the company "will not take enforcement action on every Tweet that contains incomplete or disputed information about COVID-19".¹ What is the reason for this approach, given that it implies that Twitter will not take action against disputed information even where you are aware of it?
7. Assuming the proposed duty of care set out in the Online Harms White Paper and initial consultation response has public and civil society buy-in, would Twitter support it in principle, given that the company has said that it will not take enforcement action on every tweet as per the above?
8. Ms. Rosborough's statement claims that "[w]e're prioritizing the removal of content when it has a call to action that could potentially cause harm". Can you be as specific and exhaustive as possible about what Twitter considers constituting a "call to action" and "potentially [causing] harm"? How is illegal or harmful content that does not specifically include a call to action prioritised?
9. Does Twitter today have the same number of full-time equivalent staff working directly in content moderation as before social distancing measures were in place (i.e. 1 February 2020), both for English-language users and globally?
10. How are you balancing moderation of different kinds of harmful material? Are there some categories that are being expedited?

Given the fast-moving nature of the crisis, and the urgency of the issue at hand, the Committee requests a response by 11 May 2020. I may also follow up with further questions should we require more information based on your response.

Yours sincerely,

Julian Knight MP
Chair, Digital, Culture, Media and Sport Committee

¹ *The Washington Post*, [On Twitter, almost 60 percent of false claims about coronavirus remain online – without a warning label](#) (8 April 2020).