PRIVILEGED ATTORNEY WORK-PRODUCT DRAFT—FOR DISCUSSION PURPOSES ONLY

MEMORANDUM

To:	File
From:	Jacob Glick
Date:	August 15, 2022
Subject:	Briefing with Discord on July 29, 2022

Overview

On May 11, 2022, Committee staff met with Discord for a briefing to address outstanding questions following several document productions by Discord covering aspects of its content moderation operation, counterextremism measures, and response to the attack on the U.S. Capitol on January 6, 2021. Bri Riggio, Platform Policy Manager, was the main briefer on the call. Also present was Elena DiMuzio, Senior Director, Litigation and IP at Discord, and Jeff Landis and Marc Zwillinger, outside counsel for Discord.

Riggio provided an overview of Discord's operation and focused in particular on its content moderation strategy, which includes both proactive and reactive components. She also went into greater detail to respond to the Committee's inquiries about the changes that Discord observed following the President's tweet on December 19th telling his supporters to come to D.C. on January 6th. She explained that one server was banned after content following the tweet became very specific, planning-focused posts about an operation in D.C. In addition, she discussed how Discord saw concerning signals from off-platform sites, such as TheDonald.win, which were of great concern after January 6th.

Key Takeaways

Discord's moderation team combines different methods of review to detect extremist content. Riggio explained that Discord has a three-tier moderation process, with the lowest tier being users' ability to control which communities they join. Above that, there is platform moderation by Discord employees and contractors, which includes the efforts of the Trust & Safety Team that are most of interest to the Committee's investigation. Finally, there is community moderation that can be set by users, aided by automated tools, that go beyond the platform requirements.

Riggio told Committee staff that the Trust & Safety team consists of approximately 15% of all Discord employees and is divided into specialists working on detecting high-harm content such as sexual exploitation, violent extremism, and cybercrime. These high-harm areas are the subject of proactive investigations as well, including quotas about certain servers they must review. Riggio said that there is not a "timebound component" for proactive measures, such as monitoring a server for high-harm content, and that employees can continue to look at servers while a potential threat exists. Riggio said that Discord is able to look at closed servers. When asked by the Committee, she said that most of these servers are quite small.

PRIVILEGED ATTORNEY WORK-PRODUCT DRAFT—FOR DISCUSSION PURPOSES ONLY

In response to the Committee's questions, Riggio said that proactive efforts are the bulk of the investigative measures for these high-harm areas, since Discord "is aware" of the gaps in user moderation when the userbase may not have an interest in reporting problematic content. She also emphasized that Discord has worked with third-party organizations to identify areas of high-harm content where they should initiate investigations. At the time of January 6th, according to Riggio, third party signals were incorporated into Discord's workstreams by receiving third party reports on specific content and also by using research on expert trends.

When asked by the Committee about documents Discord provided about investigations into militia servers, Riggio said that they were based on proactive investigations that looked at metadata signals to see connections by region, and discussion of firearms, firearms training, and election-related misinformation.

Riggio went on to explain that Discord employees choose the level of punishment depending on the level of harmful activity that is seen. If the decision is made to ban an entire server, it likely means that most users on the server *know* why the server exists and are actively discussing prohibited content, as was the case with the January 6th-related examples provided to the Committee.

Discord incorporates signals from "off-platform" sites to monitor high-harm content. Riggio went on to discuss off-platform signals that are key to an analysis of risks such as violent extremism. This "off-platform evidence," which can include posts on other platforms such as TheDonald.win or Parler, can be used to verify that there is problematic content occurring on Discord, and would be information Discord would consider when determining whether to remove a serve. Since January 6th, there is a new "off-platform policy" that specifically codifies the need to refer to signals off of Discord itself to determine the level of risk.

The promotion of a Discord server on TheDonald.win was an alarm bell for Discord. Riggio said that the promotion of The Donald Discord server on TheDonald.win was concerning to staff at Discord, and they began to look for other indications that TheDonald.win was using Discord for implicit or overt calls to violence. Riggio said that she did not see that sort of behavior on The Donald server prior to January 6th, but after the attack on the Capitol there was an explosion of activity and Discord became concerned that it would be used for future threats of violence. Following the attack, The Donald server appeared to be using membership on TheDonald.win as a vetting process to enter the server. Prior to that point, Riggio said, it appeared that the server was acting independently of the forum.

Following Trump's December 19th tweet, Discord banned DonaldsArmy.US for specific planning posts. Riggio explained that the DonaldsArmy.US server was created a few weeks prior to December 19th, shortly after the election. Discord discovered the server by monitoring TheDonald.win. While the server was not violating rules prior to the 19th, Riggio said that thee appeared to be "some connection between individuals viewing that tweet as a call to action." Discord observed users on TheDonald.win directing individuals to the server as a place to form a response to the President's tweet.

PRIVILEGED ATTORNEY WORK-PRODUCT DRAFT—FOR DISCUSSION PURPOSES ONLY

At that point, Discord saw very clear organization on the server, with users organizing by region and trying to connect with one another and make plans to travel to D.C. Riggio said that the most concerning part of what they saw was the discussion of how and whether to bring firearms to the President's rally, a weighing of violence versus non-violence, and how to evade D.C.'s gun control laws. Riggio reiterated that this was all seen right after the tweet. Discord removed the server within hours of the tweet.

When as about other examples, Riggio said that DonaldsArmy.US was the main example of organization after the tweet, but that there was a shift in other forums related to the sharing of election misinformation. While Trust & Safety employees are not scanning activity across Discord as a whole, they did see increased conversation about the tweet across the platform.

Discord saw more organizing during the lead-up to January 6th than on the day of the attack itself. Riggio said that Trust & Safety "extremism squad" members were on a joint call to review content on January 6th and align around enforcement actions. While Discord set up a filter to identify the most relevant content, and proactively reviewed servers that had previously been flagged for misinformation related to the election, they did not see much organizing beyond what had already been shared with the Committee. Riggio said that there were only isolated reports of people allegedly organizing, and some individuals "glorifying violence," but generally the bulk of planning activity occurred prior to the attack itself.

After January 6, Discord attempted to expand its violent extremism efforts. Riggio explained that Discord's short-term response to January 6th included a review of relevant servers, like The Donald, as well as prioritizing user content flagged pertaining to the inauguration. In the longer term, she explained that Discord decided to form a counterextremism team that is only meant to perform investigations on violent extremism. She also stressed that Discord has increased cooperation with third party organizations, and that they are currently updating their violent extremism policy to better explain identification criteria and enforcement actions.

Riggio also said that it increased the quota of how many servers the Trust & Safety Team must review each week and implemented additional counterextremism trainings.