

United States Senate

WASHINGTON, DC 20510-4606

COMMITTEES:
FINANCE

BANKING, HOUSING, AND
URBAN AFFAIRS

BUDGET

INTELLIGENCE

RULES AND ADMINISTRATION

August 12, 2024

Mr. Jason Citron
Chief Executive Officer
Discord Inc.
444 De Haro Street
Suite 200
San Francisco, CA 94107

Dear Mr. Citron:

I write today regarding disturbing reports that Discord is being used by violent predatory groups to coerce children into self-harm. The failure of your company to stop this activity is deeply troubling, and the lack of adequate safeguards to protect vulnerable individuals, especially teens and children, from this degrading and violent form of abuse is of grave concern. I urge you to quickly take steps to remove malicious actors from your platform, prevent their future access, and collaborate with law enforcement officials to bring safety and justice to the victims.

On September 12, 2023, the FBI's Internet Crime Complaint Center (IC3) issued a warning to the public that violent online groups are deliberately targeting minor victims on messaging platforms to extort them into recording or live-streaming acts of self-harm and producing child sexual abuse material (CSAM).¹ IC3 noted that these groups are targeting minors between the ages of 8 and 17 years old, especially LGBTQ+ youth, racial and ethnic minorities, and those who struggle with a variety of mental health issues. The warning further noted that these groups often control their victims through inflicting extreme fear, extorting them through threats of sharing sexually explicit videos or photos of the minor victims with their friends and families, and many have an end-goal of forcing these minors into completing suicide on live-stream to view and record for their own entertainment or sense of fame.

I am extremely concerned about this abuse, and I am profoundly saddened that it has affected Virginia families, including the daughter of a military family who was coerced into self-harm and to attempt suicide. The severe harm that the family's daughter faced from a predatory user going by the name "King" closely mirrored a story published in the Washington Post.² This report detailed how one of these violent online groups misused your platform, engaging in pervasive harassing conduct that resulted in the deaths of several minors. It further described how Discord's Trust & Safety team has struggled to keep this specific group off the platform despite knowing of its existence. "King" ultimately coerced the Virginia minor into attempting suicide. Fortunately, first responders were able to reach her in time to save her life.

¹ FBI Internet Crime Complaint Center. "Violent Online Groups Extort Minors to Self-Harm and Produce Child Sexual Abuse Material", 12 September 2023. <https://www.ic3.gov/Media/Y2023/PSA230912>.

² Boburg, Shawn, Verma, Pranshu and Dehghanpoor, Chris. "On popular online platforms, predatory groups coerce children into self-harm" The Washington Post, 13 March 2024. <https://www.washingtonpost.com/investigations/interactive/2024/764-predator-discord-telegram/>.

I recognize that Discord's Trust & Safety team is aware of this type of activity and has taken some actions to detect and remove some of these violent groups from their platforms. However, despite increased moderation, predators continue to target minors on your platform. I urge you to devote more resources to this problem, including dedicating a greater number of content moderators, investigators, engineers, and legal professionals to it. It is my understanding that Discord currently enforces its policies through actions like suspending policy-violating users' accounts and servers, as well as banning their Internet Protocol (IP) addresses and email addresses. I also understand that there are far more sophisticated measures, such as device-based or cookie-based bans, that could be taken to prevent identified malign users from returning to your platform. Further, I am aware of measures that could be used to proactively detect harmful activity and initiate an early intervention to prevent harm and loss of life.

On July 22, 2024, the Biden-Harris Administration's Kids Online Health and Safety Task Force issued a report providing guidance to address the online health and safety for children and youth with specific recommendations made to industry.³ Several recommendations that address the harm detailed in this letter were made; including developing and deploying mechanisms and strategies to counter child sexual exploitation and abuse, using data-driven methods to detect and prevent online harassment and abuse, and providing age-appropriate parental control tools. The findings and recommendations of this task force underscore the need for platforms like Discord to act on the self-harm extortion of minors. I urge Discord to review the detailed recommendations made in the report and to take them seriously.

I respectfully request that you respond to this letter with detailed answers to the following questions:

1. What processes, procedures, plans, or other organizational policies are in place to identify, review, and remove content involving activity that violates Discord's Terms of Service and other user agreements with respect to harassing, manipulative, abusive, harmful, or dangerous user activity?

Your response should address content and behavior relating to coerced self-harm, to grooming, to CSAM production, to user-to-user extortion of a sexual and of a non-sexual nature, to physical, mental, or sexual abuse, and any other category of behavior that is responsive to this question (e.g. animal cruelty extortion and abuse).

2. What enforcement actions may Discord utilize in response to the harmful activities noted in Question 1? How were these enforcement action options developed, and how does Discord determine the appropriate enforcement action for a given violation?
 - a. How many violations and of what type (grooming, sharing of CSAM, extortion, etc.) are identified before each enforcement action is made?
 - b. For a given enforcement action, what is the lowest employee position of authority (e.g. manager, director, vice president, etc.) at which that given action may be approved and carried out? Is there a process for internally reviewing and redetermining a given enforcement action? If so, please describe that process.
3. What types of detection mechanisms (e.g. technical indicators, content, behavior, social network, server membership composition, etc.) does Discord employ for activities noted in Question 1? Does Discord utilize machine learning technologies for detecting violations of company policy?

³ NTIA. "Kids Online Health and Safety Task Force Announces Recommendations and Best Practices for Safe Internet Use", 22 July 2024. <https://www.ntia.gov/press-release/2024/kids-online-health-and-safety-task-force-announces-recommendations-and-best-practices-safe-internet>.

- a. Does Discord employ user identification methods, including device-specific or cookie-based detection methods, that enables identification of returning violators who take simple evasive measures like changing their username, email address, and IP address?
 - b. Please describe policies, processes, or procedures used by Discord to ensure that violators are consistently tracked and information is shared across the security and trust and safety officials.
 - c. What is the mean time to detection (from content creation to identification by Discord's detection tools) for this activity?
4. Once Discord has removed a violating account or server, does Discord collect and store technical indicators to detect the return of the malicious actor(s) and creator of the server?
- a. What is the mean time to live (from account creation to account suspension) for the accounts engaging in this activity? What about the servers?
5. How many accounts have been removed over the last four years (provide a breakdown by year for each violation category that resulted in account removal of activities noted in Question 1)?
- a. How many of these removed accounts were initially identified via a reporting mechanism vs. a detection mechanism?
 - b. How many accounts in total were flagged for removal by a detection mechanism? For those accounts flagged by that mechanism, describe the review process for determining if the account violates policy.
6. How many unique images or videos have been shared in these servers depicting or ideating suicide or that could be reasonably interpreted as depicting or ideating suicide?
7. Have you identified activities of the types noted in Question 1 from a user going by the name of "King" (or any successor, related, or otherwise affiliated account or accounts) and what actions has Discord carried out in order to prevent ongoing and future malicious activity from this user?
8. Please describe any actions, communications, or deliberations that Discord has taken with respect to the violent groups identified in the September 2023 FBI warning: 676; 764; CVLT; Court; Kaskar; Harm Nation; Leak Society; and H3ll?

Thank you for your prompt attention to this letter, and I look forward to reviewing your response.

Sincerely,



Mark R. Warner
U.S. Senator