HEARING BEFORE THE UNITED STATES SENATE
COMMITTEE ON THE JUDICIARY

"Big Tech and the Online Child Sexual Exploitation Crisis"

January 31, 2024

Testimony of Jason Citron,
Chief Executive Officer and Co-Founder,
Discord Inc.

Chair Durbin, Ranking Member Graham, and distinguished members of the Committee, thank you for inviting me to participate in today's hearing on this critically important topic. Ensuring child safety in the digital age demands a whole-of-society approach—parents, industry, civil society, and government all have essential roles to play.

At Discord, we believe that technology companies have an obligation to work hard to prevent our services from being exploited by bad actors who seek to harm children.

My name is Jason Citron and I am the co-founder and CEO of Discord, a voice, video, and text communications service that brings people together over shared experiences. I began my career in game programming before pursuing my own entrepreneurial ventures. Prior to co-founding Discord in 2015, I founded and served as the CEO of OpenFeint, a social mobile gaming platform. Both of my companies began as game studios with a mission to modernize the way people play games online. The success of Discord is tied inextricably to the safety of its users, and as the father of two children I understand instinctively the need to protect children's safety online. I am proud of the work that Discord is doing to prevent, detect, and respond forcefully to abuse on our service.

Allow me to start by stating clearly that Discord has a zero-tolerance policy for content or conduct that endangers or sexualizes children, and we are committed to continuing our work with parents, industry partners, and law enforcement to combat harmful content online and strengthen industry moderation practices and safety initiatives. We hope this hearing provides an opportunity for us to help the Committee understand both our overall approach to safety on Discord and the considerable work we do to prevent and address online harms targeting minors.

I.   **Introduction to Discord**

Discord is a real-time messaging service—a chat app—for people to talk and spend time with their friends and communities online. When Discord was first launched in 2015, its initial user

base drew primarily from friends who connected through Discord's voice chat feature while playing video games. Discord has since grown to include a broad range of people who use the service to talk about a variety of topics, from sports and travel to personal finance.

As a messaging service, Discord differs from traditional social media in several important respects.

First, unlike social media platforms, Discord users are in control of their experience—they decide with whom they interact and what communities they join. There is no news feed, no endless scrolling, no counting of likes, and no "going viral." Discord is not a service designed to maximize engagement by an algorithm picking and choosing the content users see; rather, Discord emphasizes real time interaction and connection among friends.

Second, Discord's users interact with the communities they choose to join via "servers," peer-to-peer spaces that enable conversation and sharing. Servers are made up of channels where users communicate by text, voice, and video—a more organized and powerful messaging experience. Text channels allow users to interact via text-based messages, as well as images, GIFs, emojis, and other files. Voice channels allow users to communicate by voice, video, and/or livestream sharing. Users are also able to chat one-on-one via direct messages ("DMs") or in small groups via group direct messages ("GDMs"). Messages are displayed in the order in which they are sent. The vast majority of Discord servers are invite-only spaces for small groups of friends and communities. Eighty percent of communications on Discord are in smaller group servers.

Third, Discord's business model differentiates our service from traditional social media platforms. In contrast to the advertising-based business models of social media companies, Discord does not sell advertising and instead builds premium features that users can choose to purchase. For example, users can buy subscriptions to products like Nitro, which gives users the ability to upload larger files, stream video in high definition, and personalize the user experience by using custom emoji and digital stickers in conversations with others. Discord users can also purchase Server Boosts, which allow users to share these perks with members of a server to which they belong. These features are designed to improve the experience of using Discord, not to maximize engagement by keeping users on the app.

II.   Discord's Users

Today, more than 150 million people log onto Discord each month. Over 75 percent of Discord's users are located outside the United States, and users communicate in more than 30 languages through the service.

Users must be at least 13 years old to create an account on Discord. We use a neutral age gate—an interface that requires new users to provide their date of birth upon creating an account. We further enforce our requirements through user reports. If we receive credible information that a given user is under 13 via a user-generated report, we ban that user from the platform until they submit a valid appeal proving that they are over 13 years old. Discord evaluates the credibility of each report and the available evidence, such as whether the user has made statements or posted information on Discord in which they admitted an age or date of birth that contradicts the information provided during account creation. If any of these factors are present, Discord will ban the user. We also welcome reports from parents or guardians who believe their child, who is under the age of 13, is active on the service.

III.   Discord's Approach to Safety

Discord's mission is to bring people together around shared experiences—to create a space where everyone has a place to belong. In order to fulfill that mission, we work hard to create a safe, welcoming, and inclusive space online, especially for teens. We believe that providing online spaces where people can connect authentically with individuals who share their interests improves people's lives. To fully realize that benefit, however, service providers must establish rules and practices to prevent, detect, and remove harmful content and bad actors. Nowhere is that work more critical than in tech companies' efforts to fight the spread of child sexual exploitation.

A. Discord's Moderation Practices

Discord takes a multi-pronged approach to combating harmful content, including child sexual exploitation. This includes: (1) our robust proactive and reactive efforts to enforce our rules and remove harmful content from our services; (2) our approach to product development, whereby we implement a rigorous "safety by design" practice when developing products; (3) our policies that make clear what is and what is not allowed on the service; (4) our work with industry partners and child exploitation experts, who help Discord evaluate and improve our practices; and (5) how Discord cooperates with law enforcement.

No matter how Discord identifies violative content, users who upload child sexual abuse material (CSAM) to Discord are permanently banned from the service and reported to the National Center for Missing & Exploited Children (NCMEC), which in turn refers cases to law enforcement agencies. For incidents related to high-harm content, such as CSAM, and for violations of Discord's Terms of Service or Community Guidelines, Discord examines account identifiers to identify and ban any additional accounts the user has on the service.

1. Discord's Enforcement Work: Proactive & Reactive Efforts

Discord prioritizes enforcement against violations that present harm to our users and our service, such as child sexual exploitation, violence, and sharing of illegal content. We enforce our Community Guidelines, described in detail below, through a mix of both proactive detection and reactive measures, supported by a range of human-powered and technical solutions. Our Safety engineering teams have developed sophisticated tools to help in this work. These tools include:

Image Detection and Machine Learning: Like many of our peers, and all of the companies before the Committee today, Discord uses PhotoDNA—a tool that uses a shared industry hash database of known child sexual abuse material (CSAM)—to proactively scan images uploaded to our service in order to detect and report CSAM content and perpetrators to the National Center for Missing and Exploited Children (NCMEC), which subsequently works with local law enforcement to take appropriate action. In addition to hash matching, Discord uses internally developed tools (including machine learning techniques) and works with industry partners (including peer companies, non-profits, and researchers) to detect CSAM distribution.

In addition to our use of industry standard tools to detect and remove known CSAM, Discord invests in developing innovative technologies to proactively identify *novel* CSAM—images that have not previously been detected and so do not appear in available hash databases. We also invest in new ways to detect the exploitation of minors (commonly referred to as grooming).

- **Novel CSAM:** Recently, Discord collaborated with industry peers to build and implement a visual safety technology to detect unknown CSAM and AI-generated CSAM with positive results. Effective tools that keep children safe online should be industry standards, not a means for companies to gain a competitive advantage. In recognition of that, as of October, we have made this technology open source, so that we can share our successes with other organizations—without cost—and contribute to the broader fight against CSAM online.

- **Exploitation of Minors:** Text-based sexual exploitation and extortion of minors is a persistent and growing problem. We worked collaboratively with Thorn, a leading non-profit in the minor safety space, to develop a machine learning model that will detect when a minor is in communication with someone actively trying to exploit or extort them. This technology continues to be developed for broader industry use and further demonstrates our ongoing commitment not just to make Discord—but also the broader communications ecosystem—safer for kids.

In-App Reporting: Discord users are encouraged to report content they suspect to be CSAM or inappropriate sexual contact between adults and teens within DMs, GDMs, and within servers. In

order to better facilitate users' ability to report suspicious content for review, Discord streamlined the reporting process and eliminated extraneous steps between the submission of a report and review by our Trust & Safety teams. Reports concerning the highest harm material, including reports of CSAM or inappropriate contact with a minor, are prioritized for review and enforcement. In specific circumstances, reports are escalated to a dedicated investigations team, which is empowered not just to investigate specific infractions, but to conduct wider investigations and refer users to NCMEC and law enforcement agencies as appropriate.

Tools for Teens & Families: In addition to Discord's own enforcement, we provide users with sophisticated tools to help them tailor their experience and keep them safe. Fundamentally, each user has control over who they communicate with, what content they see, and what communities they join. But in recognition of the unique safety challenges that accompany teens' use of the internet, we have also created special features designed to give parents and teens even more control over their experience on Discord:

- **Family Center:** Discord provides parents with tools to better understand how their teens use our service. Our Family Center allows parents to get insights into the communities their teen children have joined, including which users their teens have recently added as friends, which users their teen has messaged or called in DMs or GDMs, and which servers their teen has joined or participated in. The Family Center also allows parents to get a weekly email with a high-level summary of the time their teen spends on Discord. The Family Center is designed not just to provide parents with greater insight into how their teen uses Discord, but to also help parents start conversations with their teen about online safety.

- **Teen Safety Assist:** In October 2023, Discord announced a series of new features designed to keep teens safe on our platform.

  - **Discord Safety Alerts** were designed in collaboration with the child safety non-profit Thorn to flag potentially unwanted conversations to teen users on Discord. This new feature helps determine when a DM sent to a teen from an unfamiliar user should trigger the delivery of a safety alert to the teen recipient. The alert encourages teens to think twice before replying to messages from strangers, allows teen users to block new and unfamiliar senders, and provides a link to safety tips for how teens can best protect themselves online.

  - **Sensitive Content Filters** automatically blur potentially sensitive media sent to teens in DMs, GDMs, and in servers. Blurring is enabled by default for teen users in DMs and GDMs with friends and in servers. In DMs and GDMs with non-friends, potentially sensitive media is blocked by default for teen users.

5

Discord is continuing to invest in this technology and we look forward to adding even more protective messaging features for teen users in the future.

Trust & Safety Specialists: Discord's Trust & Safety experts are trained to recognize harmful content, and we have teams of specialists dedicated to reviewing categories of content that we have designated as "high-harm:" exploitative content and minor safety (which includes child sexual exploitation), extremism, and cybercrime. Our minor safety experts are trained to identify and appropriately action content and behaviors that pose risks to teens on Discord and that violate our policies prohibiting CSAM, child sexualization, and inappropriate sexual conduct with children. These teams are a key component of our child safety work and use both machine learning and human review to proactively detect bad actors and networks of bad actors before they can proliferate.

Discord's commitment to safety is reflected in our staffing decisions: more than 15 percent of our employees work on safety. By way of context, that means our Safety team is larger than our Marketing team. In addition to the significant investment we have made in our full-time staff working on these issues, we also work closely with external partners to supplement our own work on critical safety issues.

Metrics: Discord's decision to invest in proactively detecting and removing servers before they are reported to us, especially for high-harm categories, has paid dividends in user safety. In 2023 alone, Discord scanned more than 47 billion images, or roughly 130 million images every single day. During the [fourth quarter of 2023](), Discord proactively removed servers for child safety 96 percent of the time; servers in which CSAM was posted were removed proactively 97 percent of the time. Overall, 94 percent of servers removed for policy violations across categories during this period were removed proactively. We continue to invest in our ability to proactively detect and remove servers before they are reported to us, especially for high-harm categories.

At the account level, in the fourth quarter of 2023 Discord reported 55,955 accounts to NCMEC through our use of PhotoDNA and hashing systems such as our [visual safety platform, PDQ, and CLIP](). 55,638 of those reports were media (images or videos). Additionally, 317 high-harm grooming or endangerment reports were delivered to NCMEC. Discord disabled 116,219 accounts and removed 29,128 servers for child safety during this same period.

2. Product Development: Safety by Design

Because safety is critical to our core mission, Discord takes a "safety by design" approach to our work—that means a company-wide commitment to building our service safely by assessing risks, adopting mitigating measures, and evaluating metrics *throughout* the product development

cycle, not waiting until *after* a product is launched. Safety by design means much more than merely adding a "report" button that allows users to flag content for review. It also means that our teams pursue responsible product design by studying human behavior and considering the impacts of our product on user safety—before we launch it into the market.

Our process includes a risk assessment during the design phase to identify and mitigate potential safety risks. Internal experts partner with product and engineering teams to think carefully about how product features might disproportionately impact teens, and whether the product facilitates more teen-to-adult interactions and/or any other unintended harm. Our teams identify and strategize ways to mitigate safety risks with internal safety technology solutions. We also draw on our relationships with outside experts in child safety and incorporate their insights and recommendations into the product design.

Discord's safety by design approach plays a critical role in our fight against child sexual exploitation and has not only allowed the service to proactively detect, remove, and report CSAM and inappropriate sexual contact between adults and teens, but has also improved users' ability to surface concerning or suspicious content for human review. This approach, leveraging technology and human expertise, allows us to better identify and remove spaces and users on Discord that we determine to have an increased likelihood of disseminating child sexual exploitation content.

3.   Discord's Policies: Community Guidelines

Discord's [Community Guidelines](#) contain the rules that everyone must follow on Discord. The Guidelines apply to all parts of the service, including content, behavior, servers, and apps. As I noted previously, Discord has a zero-tolerance policy for content or conduct that endangers or sexualizes children. This abhorrent material and behavior does not have a place on our service—or anywhere in society.

Discord's Community Guidelines prohibit CSAM and child sexualization. We do not allow CSAM on Discord, including AI-generated, photorealistic CSAM. When such imagery is found, it is removed and reported to NCMEC, which subsequently works with local law enforcement to take appropriate action. This policy is designed to ensure that the content does not proliferate and that the sexualization of children in any context is not normalized by bad actors.

Discord also has a zero-tolerance policy for inappropriate sexual conduct with children, meaning inappropriate sexual contact between adults and teens on the service, with special attention given to predatory behaviors such as online enticement and the sexual extortion of children, commonly referred to as "sextortion." This policy also applies to inappropriate sexual interactions—such as sextortion and other coercive behaviors—between a teen user and an older teen. When we

become aware of these types of incidents, we take appropriate action, including banning the accounts of offending adult users and reporting them to NCMEC, which in turn works with local law enforcement as appropriate.

Additionally, Discord has a policy on teen self-endangerment, an issue that we do not take lightly. We want teen Discord users to be able to express themselves freely while also taking steps to ensure that they don't engage in risky behaviors that could endanger their safety and well-being. In order to help teen users stay safe, our policies state that those under the age of 18 are not allowed to send or access any sexually explicit content, and users are only allowed to post explicit content in spaces that are behind an age-gate or in DMs or GDMs with users who are 18 and older. We also believe that teens dating online can result in self-endangerment. Under this policy, teen dating servers are prohibited on Discord, and we will take action against users who are engaging in this behavior.

Discord's Community Guidelines also prohibit a much broader range of harmful content and behavior, including content that glorifies, promotes, or seeks to normalize suicide and other acts of physical self-harm; illegal behavior; threats and harassment; hate speech and hateful conduct; violent extremism, including the glorification of violence; and other categories of harmful content and behavior that have no place on our service. Users seeking to better understand our approach to safety can learn more about the Community Guidelines in the [Discord Policy Hub](#), a resource for users that houses in-depth policy explainers across a range of topics, including [minor safety](#).

Our work in this space is constantly evolving to meet the changing threat landscape, and our Platform Policy team continues to monitor new developments in order to update and adjust our policies as needed.

4. External Engagement & Partnerships

Discord's work to halt the spread of harmful content does not take place in a vacuum; we work closely with industry partners and outside experts to inform and enhance our efforts in this space. We develop and share best practices and information with peer companies, non-governmental organizations, academics, and subject matter experts.

To enhance our efforts at fighting child sexual exploitation specifically, Discord works with a range of groups to assist in the detection of harmful content and on the improvement of our internal policies and processes, including:
- NCMEC,
- INHOPE,
- Thorn,

- Family Online Safety Institute,
- Internet Watch Foundation, and
- Technology Coalition (a group of companies working together to end online child sexual exploitation and abuse), among others.

We also work closely with nearly 50 organizations around the world as part of our Safety Reporting Network, with special consideration for child safety. Trusted reporters are outside experts—civil society groups, non-governmental organizations—that help alert us to indicators of harmful content and assist in our efforts to detect and action such content on Discord.

In 2023, Discord was a founding member of Lantern, a cross-platform signal sharing program organized by the Tech Coalition that strengthens how companies enforce their child safety policies. Lantern allows tech companies to securely and responsibly share signals about activities and accounts that violate our child safety policies. As a result of Discord's participation in the program, we have been able to scale our enforcement against offending accounts and share data related to the highest-harm abuses with other participating companies. Discord has also acted on data shared with us through this program. The cross-company cooperation facilitated by Lantern helps prevent predatory actors from evading detection across services, and Discord is proud to participate in this program.

5. Cooperation with Law Enforcement

Discord provides law enforcement agencies with necessary information while respecting the privacy and rights of our users. When we receive a request from a law enforcement agency seeking information about a Discord user, we expeditiously review it to ensure legal compliance and its overall validity. Discord may also disclose user data to law enforcement in emergency situations when we possess a good faith belief that there is an imminent risk of serious physical injury or death. In order to ensure requests are handled appropriately and expeditiously, we adopted clear guidelines for law enforcement authorities seeking information or records about Discord users.

Discord publishes quarterly Transparency Reports that detail our interactions with law enforcement and provide greater insight into the work we do to keep users safe. Transparency is one of Discord's core safety values and we view the publication of these reports as a way to honor our commitment to that value.

We evaluate each law enforcement request that Discord receives to ensure its validity and provide responsive information. According to our most recent Transparency Report issued on January 16, 2024, Discord received 1,326 pieces of legal process from federal and state authorities in the United States during the fourth quarter of 2023, finding 1,073 (81 percent) both

legally valid and sufficiently specific for us to identify an account and produce relevant information. Discord received 1,125 preservation requests globally during the fourth quarter of 2023, finding 813 (72 percent) valid and specific enough for us to identify an account and preserve relevant information.

Discord received 203 valid emergency disclosure requests from law enforcement during this period. These requests originated from law enforcement agencies around the world. Discord reviews each emergency request to authenticate the requestor, confirm the request is specific enough for us to identify an account, and verify that the request meets the legal requirements necessary for providers to disclose information in emergency situations. Discord disclosed information in response to 159 (78 percent) emergency disclosure requests.

IV.   Conclusion

Discord's mission is to bring people together around shared experiences. We view user safety not only as our responsibility but absolutely essential to fulfilling our mission and creating a space online where everyone has a place to belong.

Discord is committed to identifying new and innovative ways to halt the spread of online child sexual exploitation. Although we are a mid-sized service, we have invested heavily in staffing, technological innovation, and partnerships to help ensure the safety of our users and the health of our service. As a result, since Discord was founded in 2015, we have banned more than 3 million accounts for child safety violations and reported more than 150,000 accounts to NCMEC. We have also worked closely with our partners in law enforcement as appropriate and made it a priority to engage efficiently and productively when agencies flag potential safety concerns.

The proliferation of child sexual exploitation content on the internet requires that industry work together with parents, civil society, and government to create new technologies and explore innovative methods to prevent, detect, and halt online harms to children. Discord is committed to continuing our work to combat harmful content online and strengthen industry moderation practices and safety initiatives. I am eager to learn more about the Committee's concerns and to find ways that Discord can be a partner in identifying solutions.