Discord Transparency Report: October - December 2023

At Discord, safety is one of our top priorities and commitments. Approximately 15% of our employees work on a cross-functional team that ensures users have a safe experience and make Discord the best place to hang out with friends.

This Transparency Report, our fourteenth since 2019, covers the fourth quarter of 2023, from October through December. We're proud of the work we do to keep people and communities safe on Discord. You can find an overview of our safety-related product partnerships and developments in 2023 here. We will also be expanding our safety content library on our Transparency Hub and Safety Center in 2024 to provide even greater insight into our work and resources for best practices for staying safe on Discord.

Community Guidelines Enforcement

Discord publishes and maintains a comprehensive set of <u>Community</u> <u>Guidelines</u> that explains what content and behavior is and is not allowed on the platform. Our <u>Policy Hub</u> expands on these Guidelines, providing more information about our policies.

We invest heavily in proactive efforts to detect and remove abuse before it is viewed or experienced by others. We do this by leveraging advanced tooling,

machine learning, and specialized teams, and partnering with external industry experts.

In addition to our proactive efforts, we encourage users, moderators, and our <u>Safety Reporting Network</u> to submit reports if they believe an account or server is violating our Community Guidelines. We analyze these reports to determine if the content or behavior violates our Guidelines and may take a number of enforcement actions including but not limited to issuing warnings; removing content; disabling or removing the accounts and/or servers responsible; and reporting them to law enforcement.

This report details the actions Discord has taken on accounts and servers that have violated our <u>Community Guidelines</u> and <u>Terms of Service</u>.

Actions Taken on Accounts and Servers

Account and Server Warnings

Warning type	Individual Accounts Warned	Servers Warned	Server Members Warned
Child Safety	2,018	99	67,862
Deceptive Practices	26,700	23	1
Exploitative and Unsolicited Content	22,943	25	1,255,533
Harassment and Bullying	2,141	66	32,072
Hateful Conduct	705	267	19,036
Identity and Authenticity	12	1	1
Misinformation	1	1	3,843
Platform Manipulation	17	5	33
Regulated or Illegal Activities	2	18	549,397
Self-Harm Concerns	133	117	4,046
Violent and Graphic Content	823	136	57,296
Violent Extremism	-	-	-
Total	55,495	758	1,989,120

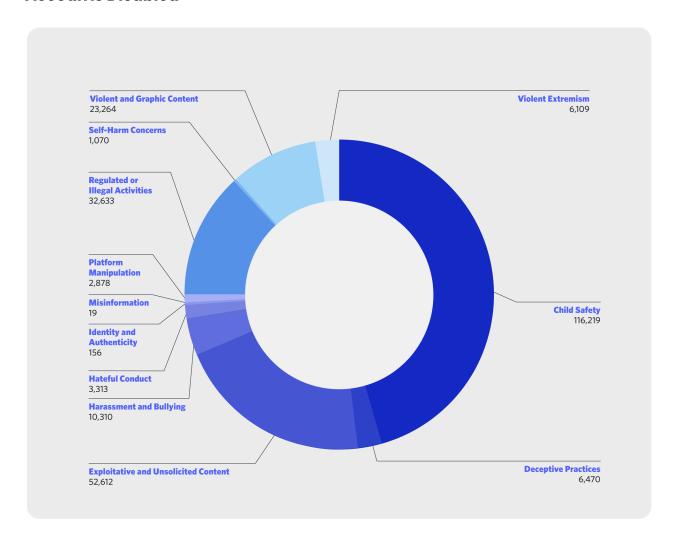
Discord issues warnings with the goal of preventing future violations of our Community Guidelines. For some high-harm issues such as Violent Extremism or Child Sexual Abuse Material (CSAM) - a subcategory of Child Safety - we don't rely on warnings and instead take immediate action by disabling

accounts and removing the content when we have identified this type of activity. In the case of CSAM and other types of high-harm Child Safety activity, we also report the account to the <u>National Center for Missing and Exploited Children</u>, who subsequently work with local law enforcement to take appropriate action.

We issue two types of warnings for accounts: Individual Account Warnings, which are issued to individual users, and Server Member Warnings, which target members of a server and may be issued when warning or removing a server from Discord.

We issued **55,495** warnings to individual accounts, an increase of **2%** when compared to the previous quarter. The number of warnings issued to servers decreased by **23%** to **758**, and warnings issued to server members decreased by **18%** to **1,989,120** accounts warned.

Accounts Disabled

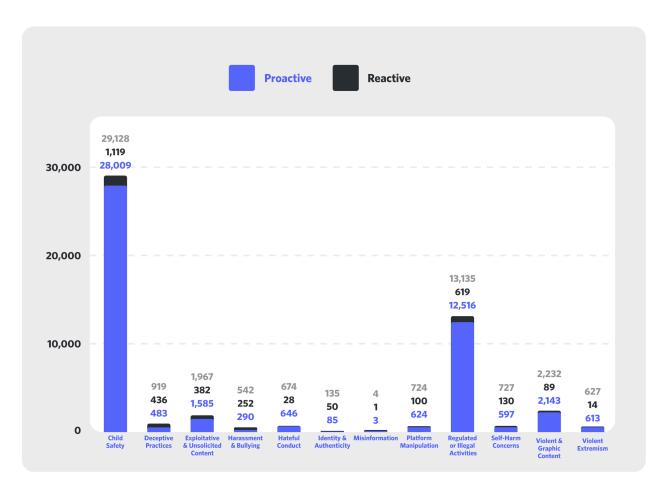


Accounts Disabled Category	Accounts
Child Safety	Disabled 116,219
Deceptive Practices	6,470
Exploitative and Unsolicited Content	52,612
Harassment and Bullying	10,320
Hateful Conduct	3,313
Identity and Authenticity	156
Misinformation	19
Platform Manipulation	2,878
Regulated or Illegal Activities	32,633
Self-Harm Concerns	1,070
Violent and Graphic Content	23,264
Violent Extremism	6,109
Total	255,063

We disabled **255,063** accounts for policy violations not including spam, a **2%** decrease when compared to the previous quarter. Accounts disabled for spam or spam-related offenses increased by **157%** to **32,973,473** when compared to the previous quarter.

An in-depth explanation of these trends can be found in the Enforcement Trend Analysis of this report.

Servers Removed



Servers Removed Category	Servers Removed Proactive	Servers Removed Reactive	Total Servers Removed
Child Safety	28,009	1,119	29,128
Deceptive Practices	483	436	919
Exploitative and Unsolicited Content	1,585	382	1,967
Harassment and Bullying	290	252	542
Hateful Conduct	646	28	674
Identity and Authenticity	85	50	135
Misinformation	3	1	4
Platform Manipulation	624	100	724
Regulated or Illegal Activities	12,516	619	13,135
Self-Harm Concerns	597	130	727
Violent and Graphic Content	2,143	89	2,232
Violent Extremism	613	14	627
Total	47,594	3,220	50,814

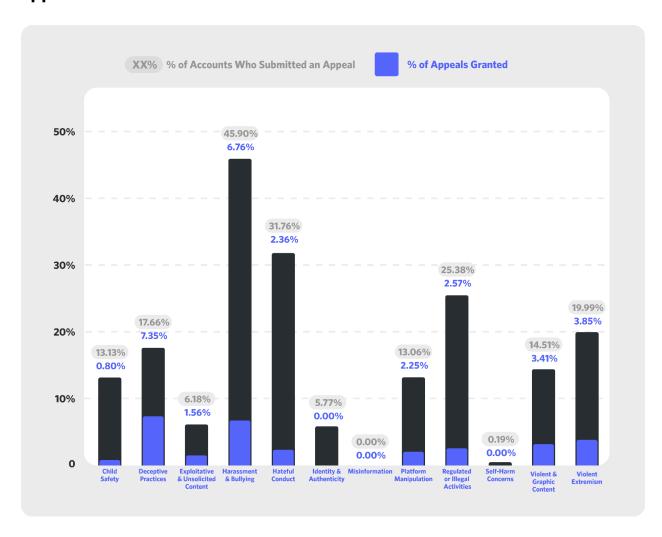
We removed **50,814** servers, an increase of **22%** from the previous quarter.

We continue to invest in our ability to proactively detect and remove servers before they are reported to us, especially for high-harm categories. Servers removed for Child Safety were removed proactively **96%** of the time, with CSAM servers removed proactively **97%** of the time. Servers removed for Regulated or Illegal Activities were removed proactively **95%** of the time, and

servers removed for Violent Extremism were removed proactively **98%** of the time. Overall, **94%** of servers were removed proactively.

An in-depth explanation of these trends can be found in the Enforcement Trend Analysis of this report.

Appeals



Appeal Category	Accounts Disabled	Appeals	% of Accounts Who Submitted an Appeal	Appeals Granted	% of Appeals Granted
Child Safety	116,219	14,334	12.33%	114	0.80%
Deceptive Practices	6,470	667	10.31%	49	7.35%
Exploitative and Unsolicited Content	52,612	2,429	4.62%	38	1.56%
Harassment and Bullying	10,320	4,039	39.14%	273	6.76%
Hateful Conduct	3,313	974	29.40%	23	2.36%
Identity and Authenticity	156	9	5.77%	0	0.00%
Misinformation	19	0	0.00%	0	0.00%
Platform Manipulation	2,878	311	10.81%	7	2.25%
Regulated or Illegal Activities	32,633	7,446	22.82%	191	2.57%
Self-Harm Concerns	1,070	2	0.19%	0	0.00%
Violent and Graphic Content	23,264	2,583	11.10%	88	3.41%
Violent Extremism	6,109	986	16.14%	38	3.85%
Total	255,063	33,780	13.24%	821	2.43%

Discord allows users to appeal actions taken on their accounts if they believe that the enforcement was incorrect.

We review appeals and reinstate accounts if we determine a mistake was made, or if we believe, based on the appeal, that the user has recognized the violation made for a low-harm issue and will abide by our Community Guidelines once back on Discord. We received appeals from **13.24%** of

disabled accounts. From these appeals, we reinstated **821** accounts which was **2.43%** of the accounts that submitted an appeal.

Reports

Reports Received By Category

Child Safety 416,036 Deceptive Practices 640,245 Exploitative and Unsolicited Content 895,182 Harassment and Bullying 2,530,225 Hateful Conduct 559,804 Identity and Authenticity 28,635 Misinformation 97,435 Platform Manipulation 6,094,106 Regulated or Illegal Activities 191,779 Self-Harm Concerns 135,128 Violent and Graphic Content 457,099 Violent Extremism 36,502 Total 12,082,176	Report Category	Reports Received
Exploitative and Unsolicited Content Harassment and Bullying 2,530,225 Hateful Conduct 559,804 Identity and Authenticity 28,635 Misinformation 97,435 Platform Manipulation 6,094,106 Regulated or Illegal Activities 191,779 Self-Harm Concerns 135,128 Violent and Graphic Content 457,099 Violent Extremism 36,502	Child Safety	416,036
Harassment and Bullying 2,530,225 Hateful Conduct 559,804 Identity and Authenticity 28,635 Misinformation 97,435 Platform Manipulation 6,094,106 Regulated or Illegal Activities 191,779 Self-Harm Concerns 135,128 Violent and Graphic Content 457,099 Violent Extremism 36,502	Deceptive Practices	640,245
Hateful Conduct Identity and Authenticity 28,635 Misinformation 97,435 Platform Manipulation 6,094,106 Regulated or Illegal Activities 191,779 Self-Harm Concerns 135,128 Violent and Graphic Content 457,099 Violent Extremism 36,502	Exploitative and Unsolicited Content	895,182
Identity and Authenticity 28,635 Misinformation 97,435 Platform Manipulation 6,094,106 Regulated or Illegal Activities 191,779 Self-Harm Concerns 135,128 Violent and Graphic Content 457,099 Violent Extremism 36,502	Harassment and Bullying	2,530,225
Misinformation 97,435 Platform Manipulation 6,094,106 Regulated or Illegal Activities 191,779 Self-Harm Concerns 135,128 Violent and Graphic Content 457,099 Violent Extremism 36,502	Hateful Conduct	559,804
Platform Manipulation 6,094,106 Regulated or Illegal Activities 191,779 Self-Harm Concerns 135,128 Violent and Graphic Content 457,099 Violent Extremism 36,502	Identity and Authenticity	28,635
Regulated or Illegal Activities 191,779 Self-Harm Concerns 135,128 Violent and Graphic Content 457,099 Violent Extremism 36,502	Misinformation	97,435
Self-Harm Concerns 135,128 Violent and Graphic Content 457,099 Violent Extremism 36,502	Platform Manipulation	6,094,106
Violent and Graphic Content 457,099 Violent Extremism 36,502	Regulated or Illegal Activities	191,779
Violent Extremism 36,502	Self-Harm Concerns	135,128
	Violent and Graphic Content	457,099
Total 12,082,176	Violent Extremism	36,502
	Total	12,082,176

We received **12,082,176** reports during the fourth quarter. The largest category of reports was Platform Manipulation which includes all reports for spam. You can learn more about reporting abusive behavior to Discord <u>here</u>.

Enforcement Trend Analysis

Child Safety

Discord has a zero-tolerance policy for anyone who endangers or sexualizes children. Child-harm content is appalling, unacceptable, and has no place on Discord or in society.

We work with industry peers, civil society, and law enforcement to ensure that this effort extends beyond Discord. We invest heavily in resources and education so parents know how our platform works and understand the controls that can contribute to creating a positive and safe experience on Discord for their children.

This year we launched our <u>Family Center</u> which empowers parents and guardians to stay informed about their teens Discord activity, as well as our <u>Teen Safety Assist</u>, which is a series of features including safety alerts and sensitive content filters which are enabled by default for teens. We also launched our <u>Parent Hub</u> which is a resource on how to help teens stay safer on Discord.

We are an active supporter of the <u>National Center for Missing and Exploited</u>

<u>Children</u> (NCMEC) and their efforts to ensure that grooming and

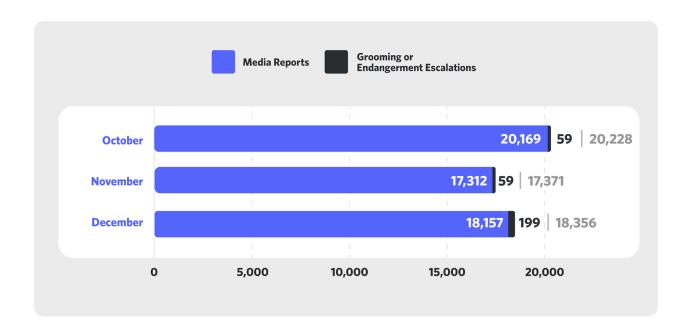
endangerment cases are quickly escalated to law enforcement. Users who upload abuse material of children to Discord or who engage in high-harm activity towards children are reported to NCMEC and removed from the service.

Discord is a member of the <u>Technology Coalition</u>, a group of companies working together to end online child sexual exploitation and abuse. We have partnered with the technology non-profit <u>Thorn</u> to enable us to build the right protections to empower teens with the tools and resources they need to have a safe online experience.

We are a proud sponsor of the <u>National Parent Teacher Association</u>, <u>ConnectSafely</u>, and partner with <u>The Digital Wellness Lab</u> to integrate their research on teen health and social media. Additionally, we are members of the <u>Family Online Safety Institute</u>, contributing to and learning from its important work.

Discord partners with INHOPE, the global network combatting online CSAM, and have become a member of the Internet Watch Foundation whose resources we will use to expand our ability to identify and prevent child sexual abuse imagery. We are a sponsor of events dedicated to increasing awareness of and action on child safety issues such as the annual Dallas Crimes Against Children Conference.

You can read our Child Safety policies, developed with the latest research, best practices, and expertise in mind, <u>here</u>.



Month	Media Reports	Grooming or Endangerment Escalations	Total
October	20,169	59	20,228
November	17,312	59	17,371
December	18,157	199	18,356
Total	55,638	317	55,955

Discord disabled **116,219** accounts and removed **29,128** servers for Child Safety during the fourth quarter of 2023. We removed servers for Child Safety concerns proactively **96%** of the time, and CSAM servers **97%** of the time. We reported **55,955** accounts to NCMEC through our use of PhotoDNA and hashing systems such as our visual safety platform, PDQ, and CLIP, which you can read more about here.

We have a longstanding team who solely focuses on child safety as well as a dedicated engineering team for our safety efforts. Discord is committed to continually exploring new and improved safeguards that help keep younger users safe on our platform and online.

Deceptive Practices

Using Discord for the purpose of distributing malware, sharing or selling game hacks or cheats, authentication token theft, or participating in either identity, investment, or financial scams is a violation of our Community Guidelines. You can read more about our Deceptive Practices policy here.

During the fourth quarter of 2023, **6,470** accounts and **919** servers were removed for Deceptive Practices. While actions taken decreased overall, accounts disabled for malicious links and malware increased.

Exploitative and Unsolicited Content

It is a violation of our Community Guidelines to share or promote sexually explicit content of other people without their consent. You can read more about our Exploitative and Unsolicited Content policies here and here and here.

During the fourth quarter of 2023, **52,612** accounts and **1,967** servers were removed for Exploitative and Unsolicited Content. This was an increase of **264%** accounts disabled and a decrease of **9%** for servers removed. We removed servers proactively **81%** of the time for this issue, an increase from 66% of the time in the prior quarter. While servers removed decreased by a moderate amount, accounts disabled increased substantially as the result of increased action against users seeking violative content.

Harassment and Bullying

Harassment and bullying have no place on Discord. Continuous, repetitive, or severe negative comments, circumventing bans, suggestive or overt threats, the sharing of someone's personally identifiable information (also known as doxxing), and server raiding are violations of our Community Guidelines. You can read more about our Harassment and Bullying policy here, and our Doxxing policy here.

During the fourth quarter of 2023, **10,320** accounts and **542** servers were removed for Harassment and Bullying.

Hateful Conduct

Hate or harm targeted at individuals or communities is not tolerated on Discord. Discord does not allow the organization, promotion, or participation in hate speech or hateful conduct. We define "hate speech" as any form of expression that denigrates, vilifies, or dehumanizes; promotes intense, irrational feelings of enmity or hatred; or incites harm against people on the

basis of protected characteristics. You can read more about our Hateful Conduct policy <u>here</u>.

During the fourth quarter of 2023, **3,313** accounts and **674** servers were removed for Hateful Conduct. The rate at which we removed servers proactively for Hateful Conduct was **96%**.

Identity and Authenticity

Using Discord for the purpose of coordinating and participating in malicious impersonation of individuals or organizations is a violation of our Community Guidelines. You can read more about our Identity and Authenticity policy here.

During the fourth quarter of 2023, **156** accounts and **135** servers were removed for Identity and Authenticity violations.

Misinformation

It is a violation of our Community Guidelines to share false or misleading information that may result in damage to physical infrastructure, injury to others, obstruction of participation in civic process, or the endangerment of public health. You can read more about our Misinformation policy here.

During the fourth quarter of 2023, **19** accounts and **4** servers were removed for Misinformation.

Platform Manipulation

Spam, fake accounts, and self-bots are examples of platform manipulation that damage the experience of our users and violate our Community Guidelines. You can read more about our Platform Manipulation policy here.

We are focused on combating spam and minimizing users' exposure to spammers and spam content on Discord. We have a dedicated cross-functional anti-spam team building sophisticated anti-spam measures. You can learn more about how Discord combats spam here, and here about Automod, a safety feature that enables server owners to automatically moderate certain abuse, including spam.

During the fourth quarter of 2023, **2,878** accounts and **724** servers were removed for non-spam related platform manipulation issues. An additional **32,973,473** accounts were disabled for spam or spam-related offenses. This represents an increase of **157%** in the number of accounts disabled for spam when compared to the previous quarter. This increase was the result of improvements to proactive scaled abuse detection of fake spam accounts. **99%** of accounts disabled for spam were disabled proactively, before we received any user report.

Regulated or Illegal Activities

Using Discord to organize, promote, or engage in any illegal behavior is a violation of our Community Guidelines. You can read more about our Regulated or Illegal Activities policies here and here.

During the fourth quarter of 2023, **32,633** accounts and **13,135** servers were removed for Regulated or Illegal Activities. This was a **31%** decrease in the number of accounts disabled, and a **42%** decrease in the number of servers removed. Our rate of proactively removing servers increased from 74% to **95%**, which we attribute to more fine tuned approaches to our work to detect and remove this behavior.

Self-Harm Concerns

For those experiencing mental health challenges, finding a community that is navigating similar challenges can be incredibly helpful for support. That said, platforms have a critical role to play in ensuring that these spaces do not normalize, promote, or encourage others to engage in acts of self-harm.

We may take action on content that seeks to normalize self-harming behaviors, as well as content that encourages self-harm behaviors or discourages individuals from seeking help for self-harm behaviors. These actions are only taken on accounts glorifying or promoting acts of self-harm, not on users seeking help or in need of medical attention. You may read more about this policy here.

We are proud to partner with <u>Crisis Text Line</u>, a nonprofit that provides 24/7 text-based mental health support and crisis intervention via trained volunteer crisis counselors. If a user reports a message for self-harm on Discord, they will be presented with information on how to connect with a volunteer Crisis Counselor. You can learn more <u>here</u>.

Crisis Text Line is currently available to those in the United States and is offered in both English and Spanish. You can read more about this partnership here. Since the launch of our partnership, there have been over 2,000 conversations started using Discord's keyword.

During the fourth quarter of 2023, **1,070** accounts and **727** servers were removed for Self-Harm Concerns.

Violent and Graphic Content

Real media depicting gore, excessive violence, the glorification of violence, or animal cruelty is not allowed on Discord. You can read about our Violent and Graphic Content policy <u>here</u>.

During the fourth quarter of 2023, **23,264** accounts and **2,232** servers were removed for Violent and Graphic Content. We proactively removed servers for these issues **96%** of the time, an increase from 92% in the previous quarter.

Violent Extremism

We consider violent extremism to be the support, encouragement, promotion, or organization of violent acts or ideologies that advocate for the destruction of society, often by blaming certain individuals or groups and calling for violence against them. You can read more about our Violent Extremism policy here.

By partnering and engaging in cross-industry work with the <u>Global Internet</u> <u>Forum To Counter Terrorism (GIFCT)</u>, <u>the European Union Internet Forum</u>, <u>the Christchurch Call</u>, and other organizations, we've made progress in our

tooling, policy, and subject matter expertise to ensure violent extremism does not have a home on Discord.

During the fourth quarter of 2023, **6,109** accounts and **627** servers were removed for Violent Extremism. We proactively removed servers for Violent Extremism **98%** of the time. Our sustained ability to remove violent extremist content proactively can be attributed to our continued cross-industry work to bolster our tooling, policy, and awareness of emerging trends.

Information Requests

When appropriate, Discord complies with information requests from law enforcement agencies while respecting the privacy and rights of our users.

Law enforcement must provide valid legal documentation for the requested information. We review each information request to ensure legal compliance. We work to limit disclosures of user information and content so they match the specific circumstances dictated by each request.

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. You can read more about how Discord works with law enforcement here.

United States Government Information Requests

Request Type	Requests	Information Produced*
Court Orders	46	34
Pen Register / Trap and Trace	0	0
Search Warrants	588	503
Subpoenas	692	536
Total	1,326	1,073

Discord received **1,326** legal requests from federal and state authorities in the United States during the fourth quarter of 2023, finding **1,073** (**81%**) both legally valid and specific enough for us to identify an account and produce relevant information.

*Please note, some requests are still under review at the time of publication. We may produce information for more requests in the coming weeks, and will update this report with those final numbers at a later date if necessary.

International Government Information Requests

Country	Requests	Information Produced
Australia	8	0
Belgium	2	0
Brazil	25	19
Canada	5	0
Czech Republic	8	0
Denmark	5	0
Finland	3	0
France	29	0
Germany	24	0
Hungary	1	0
Lithuania	2	0
Malta	2	0
Netherlands	1	0
Norway	2	0
Poland	21	0
Portugal	1	0
Serbia	2	0
Singapore	1	0
Slovenia	1	0
Spain	2	0
Sweden	5	0
Switzerland	2	0
Taiwan	1	0
United Kingdom	17	0
Total	170	19

Discord received **170** international legal requests during the fourth quarter of 2023, finding **19** (**11%**) both legally valid and specific enough for us to identify an account and produce relevant information. Note: Discord currently complies with international requests via the Mutual Legal Assistance Treaty (MLAT) or letters rogatory process and is evaluating disclosures per country on a case-by-case basis. During the fourth quarter 2023, **11%** of U.S. court orders received have been explicitly identified as having been issued as a result of MLAT requests, which originated in Brazil, Estonia, Ireland, Malta, and Spain.

Preservation Requests

Country	Requests	Complied
Argentina	0	0
Australia	3	1
Austria	1	0
Brazil	8	7
Canada	23	8
Croatia	1	0
Denmark	5	3
Finland	1	1
France	1	0
Germany	5	0
Ireland	15	9
Israel	1	0
Netherlands	8	4
Slovenia	2	1
Sweden	3	1
United Arab Emirates	2	0
United Kingdom	9	6
United States	1,037	772
Total	1,125	813

Discord received **1,125** preservation requests globally during the fourth quarter of 2023, finding **813** (**72%**) valid and specific enough for us to identify an account and preserve relevant information. We preserve information for 90 days and offer a one-time extension as needed.

Emergency Requests

Country	Requests	Information Produced
Australia	1	0
Brazil	13	13
Canada	10	8
Denmark	1	1
Finland	4	3
France	26	21
Germany	7	6
Greece	2	1
Hong Kong	1	0
Israel	7	7
Japan	2	2
Netherlands	2	2
New Zealand	3	3
Poland	3	3
Switzerland	1	1
United Kingdom	6	5
United States	114	83
Total	203	159

Discord received **203** valid emergency disclosure requests from law enforcement during this period. These requests originated from law

enforcement agencies around the world. We disclose user data to law enforcement absent legal process only when there is an imminent risk of serious physical injury or death. Our specialized team reviews each emergency request to authenticate the requestor and confirm the request meets our threshold for voluntary disclosure. Based on the information provided by law enforcement, Discord disclosed information in response to **159** (78%) of those requests.

Intellectual Property Removal Requests

Our Community Guidelines and Terms of Service prohibit the sharing of content that infringes third-party intellectual property rights. In accordance with the Digital Millennium Copyright Act (DMCA) and other applicable laws, Discord will remove content and/or accounts in connection with claims of copyright infringement on the platform. You can read about our unauthorized copyright access policy here.

We review each DMCA notice to ensure reports are valid, complete, and made in good faith.

Discord received **1,141** facially valid DMCA takedown notifications during this period, of which **837** (**73%**) provided information sufficient for content removal upon request.