Discord Transparency Report: January - June 2024

We built Discord to be the best communication tool for people to hang out and talk together around playing their favorite games. To achieve this mission, we know that Discord must be a safe space for our users. That's why safety is integrated into every aspect of our product and policies, and why about 15% of all Discord employees are dedicated to this work.

Community Guidelines Enforcement

Discord publishes and maintains a comprehensive set of <u>Community 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.

Combating harmful behavior requires industry-wide collaboration. Discord is deeply committed to making the internet and Discord safer, both through our own proactive interventions and with our various partners. We invest in proactive measures to detect and remove content that violates our policies by leveraging advanced tooling, machine learning, and specialized teams, and we partner with external industry experts. We also make it easy for people to report violative content and behavior, and we encourage users, moderators, and our Safety Reporting Network to submit reports directly to our team. We investigate these reports and may take a number of enforcement actions including, but not limited to: issuing warnings; temporarily suspending an account; temporarily limiting certain features of an account; removing content; disabling or removing the accounts and/or servers responsible; and reporting them to law enforcement when applicable.

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>.

Transparency Reporting at Discord

This Transparency Report is our fifteenth since 2019. During this time, we've evolved the structure of our reports in order to provide the most accurate and holistic report of the work we do. We're

making a number of changes in this report to reflect just that. First, we're shifting the cadence of publication from quarterly to semiannual to provide a more cohesive narrative of all of the work we do to keep people safe. This report covers the first half of 2024, from January through June.

This report also includes new data from our <u>Warning System</u>, a tool that provides more transparency about the interventions we take to help users who violate our rules learn how to do better in the future.

As a result of the introduction of the Warning System, we've also made changes to better reflect our updated, holistic approach to moderation. For example, this report no longer counts how many servers were proactively removed because that number is not effective at assessing the full impact of our safety operations.

As we learn more about the ways in which the Warning System changes the landscape of actions we take, our approach will evolve, and we'll continue providing the level of transparency that we strive to maintain.

These updates demonstrate our commitment to expanding the actions we take to enforce our Community Guidelines in an updated and fairer way. Some violations are more serious than others, and we'll take appropriate action depending on the severity of the violation. When appropriate, we focus less on disabling accounts and servers at the outset, and instead add more opportunities to learn our rules. We also let users know how their violations may impact their overall account standing and provide information on how to avoid further violations and more significant actions being taken on their accounts. Users can find this information in the Accounts and Server Interventions section, which provides information for warnings, temporary suspensions, and other forms of intervention that limit account features.

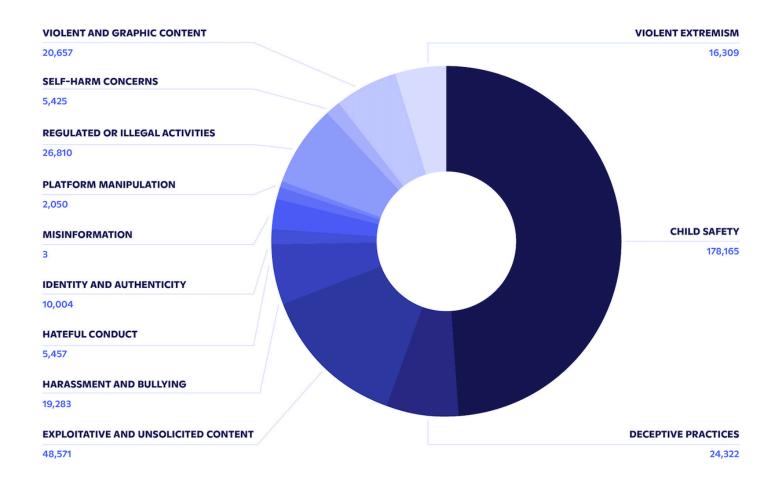
Actions Taken on Accounts and Servers

Accounts and Servers: Warnings & Temporary Interventions

POLICY CATEGORY	INDIVIDUAL ACCOUNTS	SERVERS	SERVER MEMBERS
CHILD SAFETY	202,028	121	225,175
DECEPTIVE PRACTICES	27,426	2	o
EXPLOITATIVE AND UNSOLICITED CONTENT	33,017	44	2,369,239
HARASSMENT AND BULLYING	56,042	67	37,231
HATEFUL CONDUCT	28,374	390	56,076
IDENTITY AND AUTHENTICITY	914	0	0
MISINFORMATION	7	1	16
PLATFORM MANIPULATION	6	5	44
REGULATED OR ILLEGAL ACTIVITIES	1,374	5	775,422
SELF-HARM CONCERNS	1,751	18	12,568
VIOLENT AND GRAPHIC CONTENT	12,158	97	102,285
VIOLENT EXTREMISM	1,842	0	39,459
TOTAL	364,939	750	3,617,515

Discord employs a variety of interventions against accounts and servers including warnings, temporary suspensions, and interventions that limit various features for an account. We took action on **364,939** individual accounts, **750** servers, and **3,617,515** server members. You can find more information about these interventions and our Warning System here.

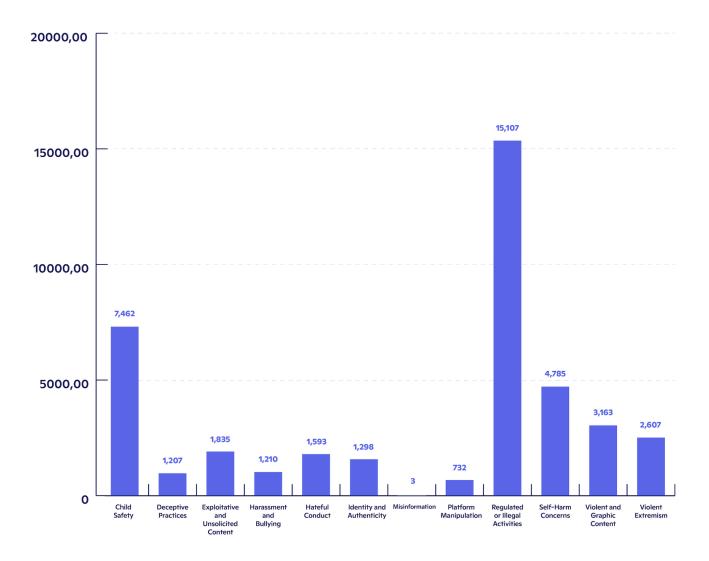
Accounts Disabled



POLICY CATEGORY	ACCOUNTS DISABLED
CHILD SAFETY	178,165
DECEPTIVE PRACTICES	24,322
EXPLOITATIVE AND UNSOLICITED CONTENT	48,571
HARASSMENT AND BULLYING	19,283
HATEFUL CONDUCT	5,457
IDENTITY AND AUTHENTICITY	10,004
MISINFORMATION	3
PLATFORM MANIPULATION	2,050
REGULATED OR ILLEGAL ACTIVITIES	26,810
SELF-HARM CONCERNS	5,425
VIOLENT AND GRAPHIC CONTENT	20,657
VIOLENT EXTREMISM	16,309
TOTAL	357,056

We disabled **357,056** accounts for policy violations not including spam during this period. We disabled an additional **35,456,553** accounts for spam or spam-related offenses.

Servers Removed

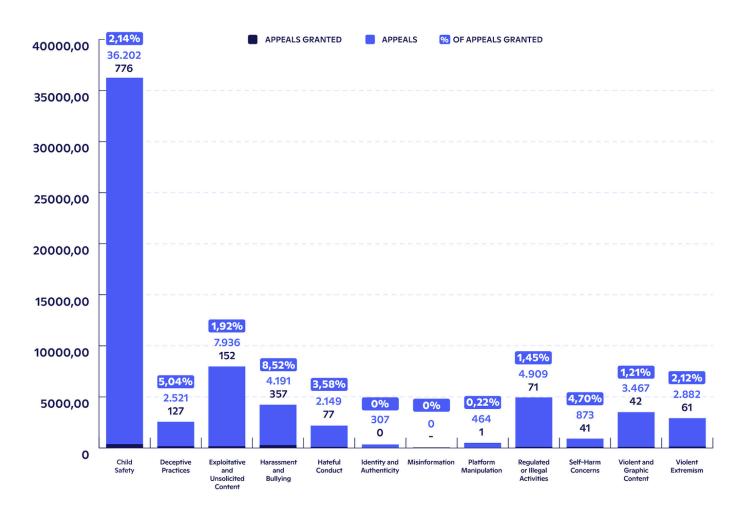


POLICY CATEGORY	SERVERS REMOVED
CHILD SAFETY	7,462
DECEPTIVE PRACTICES	1,207
EXPLOITATIVE AND UNSOLICITED CONTENT	1,835
HARASSMENT AND BULLYING	1,210
HATEFUL CONDUCT	1,593
IDENTITY AND AUTHENTICITY	1,298
MISINFORMATION	3
PLATFORM MANIPULATION	732
REGULATED OR ILLEGAL ACTIVITIES	15,107
SELF-HARM CONCERNS	4,785
VIOLENT AND GRAPHIC CONTENT	3,163
VIOLENT EXTREMISM	2,607
TOTAL	41,002

We removed **41,002** servers during this period.

Appeals

Appeals



POLICY CATEGORY	DISTINCT ACCOUNTS ACTIONED	APPEALS	% OF ACCOUNTS WHO APPEALED	APPEALS GRANTED	% OF APPEALS GRANTED
CHILD SAFETY	346,482	36,202	10.45%	776	2.14%
DECEPTIVE PRACTICES	33,533	2,521	7.52%	127	5.04%
EXPLOITATIVE AND UNSOLICITED CONTENT	2,091,583	7,936	0.38%	152	1.92%
HARASSMENT AND BULLYING	92,295	4,191	4.54%	357	8.52%
HATEFUL CONDUCT	67,969	2,149	3.16%	77	3.58%
IDENTITY AND AUTHENTICITY	10,891	307	2.82%	0	0.00%
MISINFORMATION	26	0	0.00%		
PLATFORM MANIPULATION	2,098	464	22.12%	1	0.22%
REGULATED OR ILLEGAL ACTIVITIES	795,569	4,909	0.62%	71	1.45%
SELF-HARM CONCERNS	18,781	873	4.65%	41	4.70%
VIOLENT AND GRAPHIC CONTENT	92,567	3,467	3.75%	42	1.21%
VIOLENT EXTREMISM	17,027	2,882	16.93%	61	2.12%
TOTAL	3,568,821	65,901	1.85%	1,705	2.59%

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

We review appeals to determine if a mistake was made or if we believe, based on the appeal, that the user has recognized the low-harm violation they committed and will abide by our Community Guidelines in the future. We received appeals from **1.85%** of unique accounts that had any enforcement action taken against them. From these appeals, we restored **1,705** accounts to their pre-intervention status, which was **2.59%** of the accounts that submitted an appeal.

Reports

Reports Received By Category

POLICY CATEGORY	REPORTS	% OF REPORTS
CHILD SAFETY	1,035,166	3.51%
DECEPTIVE PRACTICES	1,746,176	5.92%
EXPLOITATIVE AND UNSOLICITED CONTENT	2,504,685	8.49%
HARASSMENT AND BULLYING	5,594,669	18.97%
HATEFUL CONDUCT	1,248,125	4.23%
IDENTITY AND AUTHENTICITY	133,938	0.45%
MISINFORMATION	229,074	0.78%
PLATFORM MANIPULATION	15,180,176	51.47%
REGULATED OR ILLEGAL ACTIVITIES	433,874	1.47%
SELF-HARM CONCERNS	289,267	0.98%
VIOLENT AND GRAPHIC CONTENT	987,324	3.35%
VIOLENT EXTREMISM	112,839	0.38%
TOTAL	29,495,313	

We received **29,495,313** reports during this period. The largest category of reports related to 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 <u>policy</u> for anyone who endangers or sexualizes children. Child-harm content is appalling, unacceptable, and has no place on Discord or in society.

We have a long-established team focused solely on child safety and a dedicated safety engineering team. Discord is committed to continually exploring new and improved safeguards that help keep younger users safe on our platform and online.

We collaborate with members of the industry, policymakers, law enforcement, and others to extend our efforts beyond Discord. We invest in <u>resources</u> and <u>education</u> to help parents and guardians understand how our platform works and how to create a safe and positive experience for their teens.

Discord actively supports the <u>National Center for Missing and Exploited Children</u> (NCMEC) and reports users who upload child abuse material or engage in other activity that seeks to exploit minors to NCMEC. As a member of the <u>Technology Coalition</u>, we collaborate with industry partners to drive critical advances in technology and adoption of best practices for keeping teen users safe online. Within the Technology Coalition, Discord helped pilot the <u>Lantern</u> initiative, a first-of-its-kind signal-sharing program for companies to enhance and strengthen how we detect attempts to sexually exploit and abuse children and teens online.

We also partner with <u>Thorn</u> and <u>The Digital Wellness Lab</u> to integrate their expertise and research on teen health and social media into our products and policies, including the development of our <u>Teen Charter</u>.

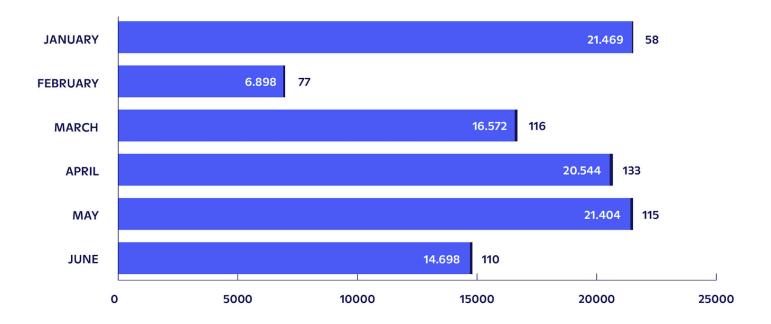
Discord is also a proud sponsor of the <u>National Parent Teacher Association</u> and <u>ConnectSafely</u> and is an active member of the <u>Family Online Safety Institute</u>.

We partner with <u>INHOPE</u>, the global network combatting online child sexual abuse material, and are a member of the <u>Internet Watch Foundation</u>, whose resources we use to identify and prevent child sexual abuse imagery. We also sponsor events dedicated to increasing awareness of, and action on, child safety issues such as the annual <u>Dallas Crimes Against Children Conference</u>.

At Discord, part of our holistic approach to safety also includes products and features for parents and teens. Our <u>Family Center</u> empowers parents and guardians to stay informed about their teen's Discord activity, and our <u>Teen Safety Assist</u> features, which are enabled by default for teens, include safety alerts and sensitive content filters. We also publish our <u>Guardians Guide</u> to help teens stay safer on Discord.

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

Discord took action on **346,482** distinct accounts for Child Safety during this period. This included disabling **178,165** accounts and removing **7,462** servers. We reported **101,585** accounts to NCMEC as a result of CSAM that was identified by our <u>hashing systems</u>, reactive reporting, and additional proactive investigations.



MONTH	MEDIA REPORTS	GROOMING OR ENDANGERMENT ESCALATIONS	TOTAL
JANUARY	21,469	58	21,527
FEBRUARY	6,898	77	6,975
MARCH	16,572	116	16,688
APRIL	20,544	133	20,677
MAY	21,404	115	21,519
JUNE	14,698	110	14,808
TOTAL	101,585	609	102,194

Deceptive Practices

Using Discord for the purpose of financial scams, malicious conduct, or fraud services is a violation of our Community Guidelines. You can read more about our Deceptive Practices policy here.

Discord took action on **33,482** distinct accounts for Deceptive Practices during this period, including disabling **24,322** accounts and removing **1,207** servers.

Exploitative and Unsolicited Content

It is a violation of our Community Guidelines to share, distribute, or create sexually explicit content depicting other people without their consent. You can read more about our Exploitative and Unsolicited Content policies <u>here</u> and <u>here</u>.

Discord took action on **2,091,583** distinct accounts for Exploitative and Unsolicited Content during this period, including disabling **48,571** accounts and removing **1,835** servers.

Harassment and Bullying

Harassment and bullying have no place on Discord. Promoting, coordinating, or engaging in harassment, threatening to harm another individual or group of people, or sharing or threatening to share someone's personally identifiable information (also known as doxxing) are violations of our Community Guidelines. You can read more about our Harassment and Bullying policy here and our Doxxing policy here.

Discord took action on **92,295** distinct accounts for Harassment and Bullying during this period, including disabling **19,283** accounts and removing **1,210** servers.

Hateful Conduct

Hate or threats of harm targeted at individuals or communities is not tolerated on Discord. Using hate speech or engaging in other hateful conduct are violations of our Community Guidelines. We define hate speech as any form of expression that either attacks other people or promotes hatred

or violence against them based on their protected characteristics. You can read more about our Hateful Conduct policy <u>here</u>.

Discord took action on **67,969** distinct accounts for Hateful Conduct during this period, including disabling **5,457** accounts and removing **1,593** servers.

Identity and Authenticity

Using Discord to misrepresent your identity on Discord in a deceptive or harmful way or evade permanent Discord-level enforcement action is a violation of our Community Guidelines. You can read more about our Identity and Authenticity policy <u>here</u>.

Discord took action on **10,891** distinct accounts for Identity and Authenticity during this period, including disabling **10,004** accounts and removing **1,298** servers.

Misinformation

It is a violation of our Community Guidelines to share false or misleading information. You can read more about our Misinformation policy <u>here</u>.

Discord took action on **26** distinct accounts for Misinformation during this period, including disabling **3** accounts, and removed **3** servers.

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 <u>here</u>.

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 about Automod, a safety feature that enables server owners to automatically moderate certain abuse, including spam.

Discord took action on **2,098** distinct accounts for non-spam-related Platform Manipulation issues during this period, including disabling **2,050** accounts, and removed **732** servers. An additional **35,456,553** accounts were disabled for spam or spam-related offenses.

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 <u>here</u> and <u>here</u>.

Discord took action on **795,569** distinct accounts for Regulated or Illegal Activities during this period, including disabling **26,810** accounts and removing **15,107** servers.

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 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 can 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 in English and Spanish via trained volunteer crisis counselors in the US. 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 about Crisis Text Line <u>here</u> and about our partnership <u>here</u>. Since the launch of our partnership, there have been over 3,000 CrisisTextLine conversations that were started by texting "DISCORD".

As part of our efforts to provide resources for all users, we're excited to partner with ThroughLine Care, a global network of vetted helplines. Non-US Discord users can find helplines in their country by going to <u>discord.findahelpline.com</u>. Learn more by checking out <u>the Family Center</u>.

Discord took action on **18,781** distinct accounts for Self-Harm Concerns during this period, including disabling **5,425** accounts and removing **4,785** servers.

Violent and Graphic Content

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

Discord took action on **92,567** distinct accounts for Violent and Graphic Content during this period. This included disabling **20,657** accounts and removing **3,163** servers.

Violent Extremism

Discord has a zero-tolerance <u>policy</u> against violent extremism, and it is against our Community Guidelines to organize, promote, or support violent extremist activities or beliefs. The term violent extremist organization describes groups that promote a political, ideological, or religious agenda and tolerate, advocate for, or use violence to achieve their goals.. We also take reported <u>off-platform harmful activities</u> into account when evaluating potential violations of this policy.

We partner with industry-leading organizations, such as the <u>Global Internet Forum To Counter Terrorism (GIFCT)</u>, <u>the European Union Internet Forum</u>, <u>the Christchurch Call</u>, and other organizations across the world to combat extremism and understand global trends.

Discord also leads its own <u>Safety Reporting Network</u> for direct communication with expert third parties, including researchers, industry peers, and journalists, for intelligence sharing about potential extremist and hateful content or activity on the platform. We also take appropriate action based on wider platform trends and intelligence we receive.

Discord took action on **17,567** distinct accounts for Violent Extremism during this period, including disabling **16,309** account, and removing **2,607** servers.

Information Requests

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

When law enforcement seeks to compel Discord to disclose user information, it must provide valid legal process for the requested information, and we review each information request to ensure legal compliance. Discord determines what data to disclose based on the type of legal process received and the data requested.

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	122	110
PEN REGISTER / TRAP AND TRACE	9	9
SEARCH WARRANTS	1,586	1495
SUBPOENAS	2,065	1,609
TOTAL	3,782	3,223

Discord received **3,782** legal requests from federal and state authorities in the United States during this period, finding **3,223** (**85%**) both legally valid and specific enough for us to identify an account and produce relevant information.

International Government Information Requests

COUNTRY	REQUESTS	INFORMATION PRODUCED
ARGENTINA	2	0
AUSTRALIA	29	1
AUSTRIA	13	1
BELGIUM	7	0
BRAZIL	106	73
BULGARIA	1	0
CANADA	27	15
CHILE	1	0
COSTA RICA	1	0
CROATIA	1	0
CZECH REPUBLIC	3	0
DENMARK	6	0
ESTONIA	1	0
FRANCE	58	0
GERMANY	90	0
GREECE	6	0
HUNGARY	10	О
INDIA	1	0
ITALY	1	0
KAZAKHSTAN	2	0
KUWAIT	1	o
LITHUANIA	5	0
MALTA	2	0
MONGOLIA	1	0
NETHERLANDS	12	7
NEW ZEALAND	5	0
NORWAY	3	o
POLAND	42	0
PORTUGAL	3	0
ROMANIA	4	0
SERBIA	2	0
SINGAPORE	1	0
SPAIN	12	0
SWEDEN	5	0
SWITZERLAND	3	o
TAIWAN	3	0
THAILAND	4	0
TURKEY	6	0
UNITED ARAB EMIRATES	1	0
UNITED KINGDOM	64	29
TOTAL	545	126

Discord received **545** international legal requests during this period, finding **126** (**23%**) 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 evaluates requests to disclose from foreign governments on a case-by-case basis. During this period, **22** U.S. court order requests were identified as having been issued as a result of MLAT requests that originated in: Brazil; Canada; Croatia; Czech Republic; France; Germany; Greece; Ireland; Latvia; Netherlands; Slovakia; Spain; and Switzerland.

Preservation Requests

COUNTRY	REQUESTS	COMPLIED
AUSTRALIA	14	6
AUSTRIA	5	3
BRAZIL	27	19
CANADA	68	35
CHILE	1	1
DENMARK	8	4
FRANCE	15	1
GERMANY	29	3
ICELAND	2	0
INDIA	1	0
IRELAND	28	12
ISRAEL	1	0
NEW ZEALAND	1	0
POLAND	3	0
PORTUGAL	1	0
SLOVENIA	1	1
SPAIN	1	0
SWITZERLAND	2	2
TAIWAN	1	0
UNITED KINGDOM	11	5
UNITED STATES	3,015	1,989
TOTAL	3,235	2,081

Discord received **3,235** preservation requests globally during this period, finding **2,081** (**64%**) 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	COMPLIED
AUSTRALIA	9	5
AUSTRIA	2	1
BELGIUM	1	1
BRAZIL	40	25
CANADA	56	40
CZECH REPUBLIC	1	1
DENMARK	2	1
ESTONIA	1	1
FRANCE	28	17
GERMANY	33	15
GREECE	1	1
INDIA	4	1
ISRAEL	9	6
JAPAN	10	4
NETHERLANDS	6	5
NORWAY	1	0
NEW ZEALAND	1	1
NORWAY	4	3
POLAND	14	5
PORTUGAL	1	1
ROMANIA	1	0
SPAIN	1	0
SWITZERLAND	3	2
TURKEY	1	0
UNITED KINGDOM	30	14
UNITED STATES	280	164
TOTAL	539	314

Discord received **539** 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 we possess a good-faith belief that there is an imminent risk of serious physical injury or death. Our specialized team reviews each emergency request to authenticate the requester and confirm the request meets our threshold for voluntary disclosure. Based on the information provided by law enforcement, Discord disclosed information in response to **314** (**58%**) 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 **2,927** facially valid DMCA takedown notifications during this period, of which **2,040** (**70%**) provided information sufficient for content removal. Upon review, 400 of these notifications flagged content that had already been removed by users.