clubhouse

Transparency Report 2021

Introduction

Clubhouse was founded in March 2020 as a place to meet with friends and with new people and to learn, tell stories, ask questions, debate, have impromptu conversations, and find community. As we have built the company, we have also built strong trust and safety teams and practices to protect that sense of community. In 2020 and half of 2021 we had an invitation-only system to help grow our community and our systems–including our content moderation systems–in a scalable way. This allowed us to build stable abuse-fighting systems and legal operations procedures as our user base grew.

Here's a bit about our history:

- **March 2020:** The app was initially available in TestFlight for a small group of users
- **September 2020:** Clubhouse was first made publicly available for download on iOS
- **May 2021:** Clubhouse became available on Android
- **July 2021:** General Release - users no longer required an invite to Clubhouse

Making sure users have a great experience on the app remains our top priority. The vast majority of rooms on Clubhouse happen without incident and we want to keep it that way. That is why, since our earliest days, we have invested heavily in the three Ps: policies, product, and people to keep Clubhouse safe.

- **Policies:** We work to have strong, clear, and fair policies and we take action when users break these rules.
- **Product:** We work to build safety into our products before we launch them and to give users the tools they need to only interact with the people they want to interact with.
- **People:** We have grown from a full time team of just 9 people in January 2021 to near 100 people today, with a sizeable portion of our team working on trust and safety. Additionally, we have invested in an extended team to help us keep the platform safe in different languages and time zones.

We have also built systems to comply with valid law enforcement requests. We work with our legal teams and necessary outside counsel to make sure requests are consistent with internationally recognized standards on human rights, including due process, privacy, free expression and the rule of law.

Our first transparency report includes all of 2021, which was also our first full year of operation. We share data here on Community Guidelines enforcement (what we take action on), government requests to remove content or users, and law enforcement and government requests for user data. As we continue to grow and gather feedback, we hope to iterate on the content of our transparency reporting.

Content Moderation in Social Audio

Much of the public discourse on content moderation centers around removal of posts, whether that is text, video, or images. Social audio presents a different format and different opportunities for connection and discussion. The intonation, inflection, and emotion conveyed through voice allow you to pick up on nuance, have complex conversations, and form uniquely human connections with others. At the same time, social audio also presents a different set of moderation challenges because rooms are live, dynamic, and unpredictable.

For example, in a room that discusses medical topics, users may make points that are controversial or inaccurate. In a more traditional medium, such as text or video, companies may remove the content to prevent if from being broadcast further. In social audio, such a claim is usually a part of a conversation. We see cases where someone is misinformed about an important issue but a physician, a scientist, a researcher or other users challenge the false claim in real-time to correct certain claims and educate others. It is this dynamism and the opportunity for counterspeech and education that makes moderation of social audio unique.

Since we first started the company we've had strong Community Guidelines that clearly state what we do not allow. The world changes every day and our policies evolve with it. We design our policies to have zero tolerance for the most egregious and sometimes even illegal topics (e.g., violent extremism, hate speech, sexual exploitation). We also work to still allow space for important conversations, like political and social disagreements.

Any user can report something to us that they believe violates our rules. Despite being a small company in an emerging space and not having the same resources as much larger social networks, we also automated systems to flag potentially violations of our policies..

When a user is reported to us while a room is live or shortly thereafter, that report triggers our systems to keep a recording of the audio for our team to investigate. Our highly trained team works around the world in different languages to review flagged audio and user profiles.

The actions we take upon receiving flags (and the data reported here) are focused on suspensions against users. We can't predict what someone might say, but based on their speech and actions we can give a warning, require they take a break from our community, or indefinitely suspend them, and–in the worst cases–report their activities to law enforcement. In egregious cases or when a room is dedicated to violating our policies, we will also end the room.

Making sure the app is a great experience is essential to our success. We will continue to invest in a mix of people, products, and policies to enforce our rules, balance freedom of expression, and pioneer new ways to keep social audio communities safe.

A key caveat: As our app grew in 2021, we made big changes to our enforcement tools and systems to help scale. We had to quickly develop and deploy new systems—and when we made a switch to a new enforcement system, we lost granularity in our suspensions data. For example, after the switch to a new system, we no longer had clear counts of suspension reasons in a way that made is possible to report on suspension by type for the first half of 2021.

As a small, early stage company we are building our enforcement systems just like we are building our overall product. That means we have room for improvement. Another example: while we utilize some automated detection to flag severe violations of our guidelines, we currently do not have the ability to count user suspensions that result from user flags versus ones that originated using our automated tools.

Part of being transparent, also means being transparent about the challenges of building a new company with a small team and where we have areas to improve. As a result the data related to appeals and government and law enforcement requests reflects the whole year, but our suspensions by policy area data is only for August–December 2021.

Community Guidelines Enforcement Data

August–December 2021

Policy Area	Suspensions
Child Endangerment	309
Underage Accounts	363
False and misinformation	66
Harassment	3,964
Hate Speech	3,366
Impersonation	383
Personal Privacy	548
Sexual Content	1,682
Sexual Exploitation	813
Suicide and Self-Harm	120
Unauthorized Sales	176
Graphic Violence, Threats, and Terrorism	1,994

Account Suspension Appeals

2021

We believe in allowing all users to request an appeal, or an additional review of their accounts in instances when our team has suspended a user. Sometimes a user is suspended in situations where critical context was missed—when that happens, we work swiftly to reinstate the account. Our appeals data for the full year is as follows:

8,229
Appeals Submitted

6,597
Suspensions Upheld

1,632
Suspensions Overturned

20%
Overturn Rate

Government and Law Enforcement Requests

We receive and review a variety of requests from government bodies and law enforcement agencies. We make sure requests are consistent with internationally recognized standards on human rights, including due process, privacy, free expression and the rule of law. For government requests that do not meet these standards, we may ask governments to address any deficiencies and, where appropriate, we will challenge deficient requests

We will only review valid requests (in writing) and requests that are too broad or too vague are not processed by members of the team.

When we receive requests, we require government bodies to include the name of the issuing authority and agent and provide the pertinent legal documents behind their request.

Finally in cases involving imminent harm to someone or the safety of a child, we may voluntarily disclose information to law enforcement.

We report all child sexual exploitation to the National Center for Missing and Exploited Children (NCMEC). In 2021, we made **621** CyberTipline reports to NCMEC.

For the year 2021, Clubhouse received a total of **92** requests from governments and law enforcement. The following pages in our transparency report break down the types of requests as well as the number of requests by country.

Government and Law Enforcement Requests

By Request Type

Type of Request	Requests Received
Suspension Appeal	2
Block or Takedown Request	12
Emergency Disclosure	4
Preservation Request	5
Request for User Information	30
Request for User Information *and* Block or Takedown Request	24
Request for User Information *and* Preservation Request	3
Subpoena	12

Government and Law Enforcement Requests

By Country

Country	Total Requests
Canada	2
India	59
Nepal	3
Norway	1
Russia	2
Thailand	2
Turkey	2
United Arab Emirates	1
United States	20

Government and Law Enforcement Requests

By Country and Request Type

	Suspension Appeal	Block or Takedown Request	Emergency Disclosure	Preservation Request	Request for User Information	Request for User Information *and* Block/Takedown	Request for User Information *and* Preservation	Subpoena	Grand Total
Canada			2						2
India		8			24	24	3		59
Nepal					3				3
Norway					1				1
Russia	2								2
Thailand		1			1				2
Turkey		2							2
UAE		1							1
United States			2	5	1			12	20
Grand Total	**2**	**12**	**4**	**5**	**30**	**24**	**3**	**12**	**92**