

Transparency Report 2022





### **Table of Contents**

Introduction Introduction Content Moderation in Social Audio	<b>3</b> 4 5
Community Guidelines Enforcement  2022 Overview  2022 Enforcement Data Child Endangerment Doxxing Graphic Violence, Threats, and Terrorism Harassment Hate Speech Nudity and Sexual Content Sexual Exploitation Spam, Bots, Platform Misuse, & Misinformation Suicide and Self-Injury Unauthorized Sales	6 7 9 10 11 12 13 14 15 16 17 18
Suspension Appeals Account Suspension Appeals	<b>20</b> 21
Proactive Detection Proactive Detection Child Safety Reporting	<b>22</b> 23 24
Government and Law Enforcement Requests 2022 Overview Data by Request Type Data by Country Data by Country and Request Type	25 26 27 28 29
Report Recap and Additional Resources	30



2022 Transparency Report

### Introduction

### Introduction

Founded in March of 2020, Clubhouse is like an always-on dinner party where you can hang out with friends, meet new ones and talk about anything. As we have built the company, we have also built strong trust and safety teams and practices to protect that sense of community. Prior to mid-2021, we had an invitation-only system to help grow our community and our systems—including our content moderation efforts—in a scalable way. This allowed us to build stable abuse-fighting systems and legal operations procedures as our user base grew.

As our app moved from an invitation-only to a general release, our team continued to think about how to provide the best environments to connect with friends. As such, in the second half of 2022 we introduced Houses to Clubhouse, a more intimate space to invite friends and their friends to talk, hang out, and have fun.

Our second transparency report includes our metrics for 2022, which was also our second 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, the introduction of our new Houses feature presents an exciting opportunity to build community.

Community, however, is a shared responsibility, and one that Clubhouse takes seriously. For the most egregious violations, including (but not limited to) hate speech, sexual exploitation, child safety, and violent extremism, we rely not only on user flagging but also our own investments in more proactive detection, such as conducting network analyses to identify accounts that require additional review. In instances where we discover a House has been created with the sole purpose to be dedicated to violating our policies, we will quickly intervene both at the user and House-level.

For other policy categories, we see our responsibility for maintaining a healthy community to be a collective task. Clubhouse staff will continue to enforce on content that creates genuine risk of harm to users, but we also expect and encourage House founders to play a role in cultivating a safe space for all the members of a House. Each House founder has the chance to decide on what types of House norms and rules should be upheld (in addition to our own Community Guidelines)—and is empowered with the opportunity to review complaints and remove users who are disrupting the communal space.

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.

Note: In our 2021 report, we were constrained by transitions to a new enforcement system that prevented us from being able to report data for the whole 2021 calendar year. This year's report does not face such limitations.



2022 Transparency Report

### Community Guidelines Enforcement

### **Community Guidelines Enforcement**

2022 Overview

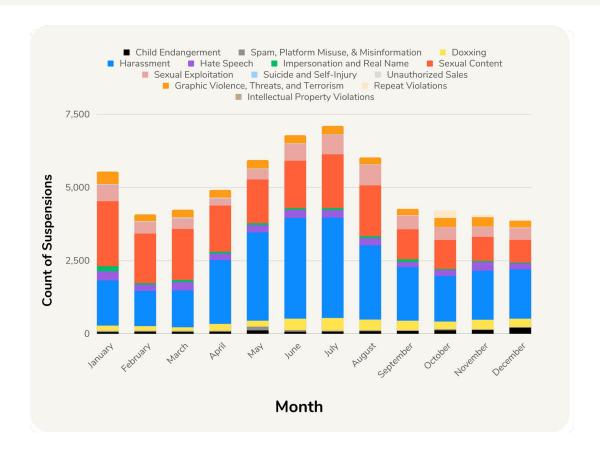
Our trust and safety team takes enforcement actions when our team investigates an incident and confirms a violation of the <u>Community Guidelines</u> that help keep the Clubhouse platform safe. Enforcement actions we may take range from warnings to indefinite suspensions. Suspensions vary in length based on the policy category and severity of the violation. We recommend reviewing the <u>Community Guidelines</u> for more details about how we curate a safe space for our users.

In 2022, our team actioned **61k** suspensions based on confirmed violations. The next few pages of this report further detail improvements we made in 2022 to our enforcement processes and systems, and the breakdown of enforcement actions by policy area.

While violation detection and enforcement are always top of mind to ensure a safe experience for our community, we also believe in learning from our mistakes and second chances. In 2022, we laid the groundwork to build less punitive enforcement actions (such as warnings) for less serious violations, which we began testing in early 2023. We also constantly assess our enforcement accuracy and gather insights to improve our trust & safety processes through a robust quality assurance effort, as well as reviewing our enforcement actions through suspension appeals (p. 21).

#### **Community Guidelines Enforcement**

2022 Overview



We undertook many safety initiatives over the course of the year which contributed to our improved detection of violating accounts and enforcement, including:

- **Proactive detection.** We implemented new tools and investigation procedures to identify more potential violations (p. 23-24).
- Duplicate accounts. Clubhouse has a one account, one user policy to ensure we're a
  community of real people, and that users who were suspended for serious harms to
  our community do not return. We continue to roll out improvements to detect and
  identify duplicate accounts, which may be actioned by our team, or automatically by
  our system.
- **Enforcement for repeat violations.** In the second half of 2022, we began testing and refining a new system to permanently suspend repeat violators who have been suspended at least three times for past violations.



### **Community Guidelines Enforcement Data**

**By Policy Category** 

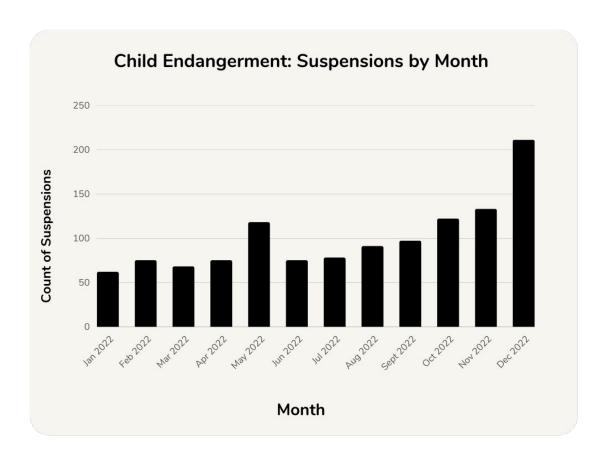
Policy Category	Suspensions
Child Endangerment	1,205
Doxxing	3,281
Graphic Violence, Threats, and Terrorism	3,353
Harassment	25,317
Hate Speech	2,921
Impersonation	830
Nudity and Sexual Content	17,506
Spam, Platform Misuse, & Misinformation	510
Sexual Exploitation	5,430
Suicide and Self-Injury	121
Unauthorized Sales	181
Repeat Violations	364
Intellectual Property Complaints	6
Total	61,025*

<sup>\*</sup> In 2022, we primarily used one enforcement system, which significantly improved the quality and accuracy of suspensions data. We continue to update and refine our data reporting methods. Data may not be fully inclusive of the "true" suspension count actioned due to duplicate investigations, overlapping suspensions, data deletions, unverified accounts, human error, and other factors. Suspension counts may also exclude certain types of suspensions which are actioned automatically by our security, anti-bot, and anti-abuse measures..

### Child Endangerment

Clubhouse strictly prohibits any content or behavior that endangers children. This includes any sexualization of or solicitation of minors, and any content that could put their safety at risk. We report all confirmed child endangerment cases to the National Center for Missing and Exploited Children (NCMEC), and/or to law enforcement.

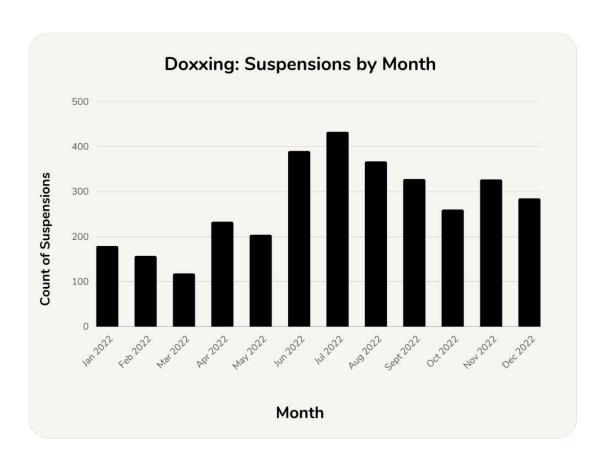
In 2022, our team actioned **1.2k** suspensions related to child safety violations. Child safety suspensions generally increased over the course of the year as we were able to improve our child safety team's proactive detection efforts over time. Learn more in the Proactive Detection (p. 23) and Child Safety Reporting (p. 24) sections.



### Doxxing 2022

Doxxing – disclosing an individual's personally identifiable information (including addresses, phone numbers, personal documents like IDs, and medical information), as well as threatening to share private information – is strictly prohibited and will lead to a permanent suspension.

In 2022, our team actioned 3.3k suspensions related to doxxing.

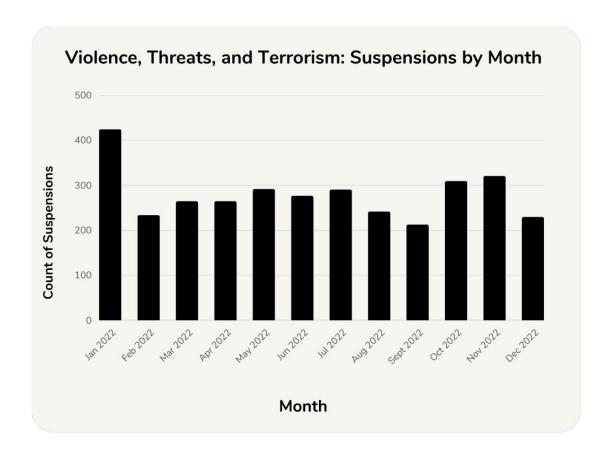


### **Graphic Violence, Threats, and Terrorism**

Clubhouse does not allow content and conduct related to graphic violence, threats, and terrorism, including (but not limited to):

- conduct and imagery supporting and encouraging the activities and ideologies of terrorist organizations
- threats of violence or harm against individuals or groups
- Images including excessive gore, violence, or harm toward people or animals, and images intended to be shocking or disgusting.

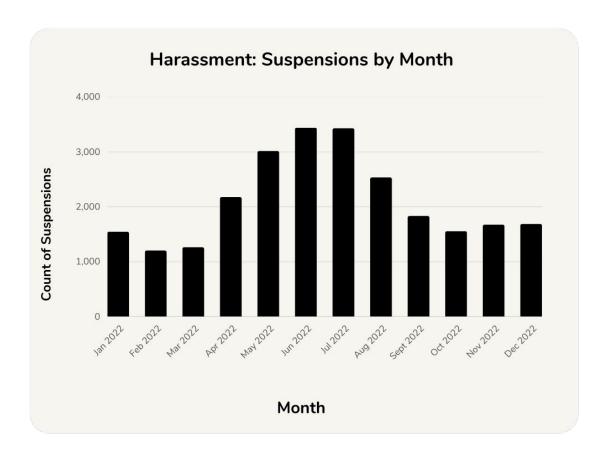
In 2022, our team actioned **3.4k** suspensions related to graphic violence, threats, and terrorism.



### Harassment

Clubhouse does not allow content intended to intimidate, threaten, or egregiously insult members of the Clubhouse community, or to incite harassment. Harassing conduct can include, but is not limited to, mocking the deceased or victims of tragedies, making lewd propositions at another user, and altering the photo of another user with intent to humiliate them.

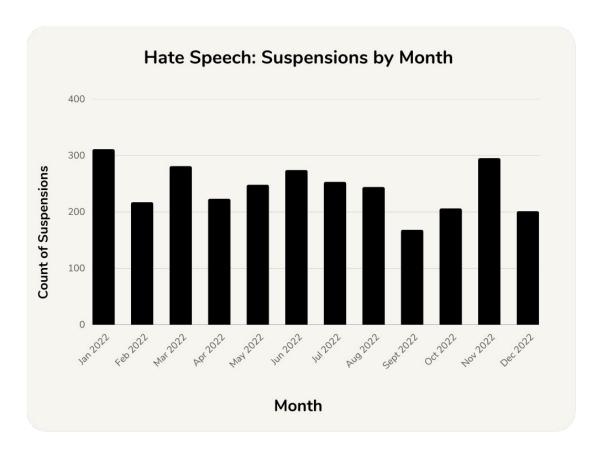
In 2022, our team actioned 25.3k suspensions related to harassment.



### Hate Speech

Inciting hatred or discriminating against other people based on their race, religion, national origin, caste, age, sex, gender identity, sexual orientation, or disability is strictly not allowed–including slurs, insults, conspiracies, and symbols of hate. Hate organizations and accounts dedicated to hateful conduct don't belong on Clubhouse. Our team consults with non-governmental organizations and expert partners on an ongoing basis to consistently update our hate speech policies according to new learnings and trends.

In 2022, our team actioned **2.9k** suspensions related to hateful or discriminatory speech or conduct.

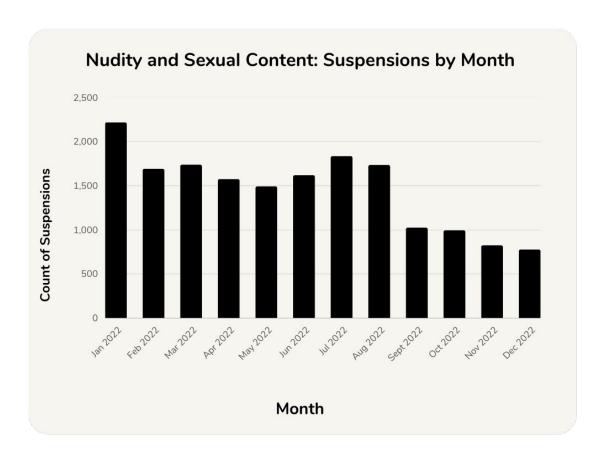




### Nudity and Sexual Content

We don't allow sexual imagery or overly graphic sexual content to be shared on the platform.

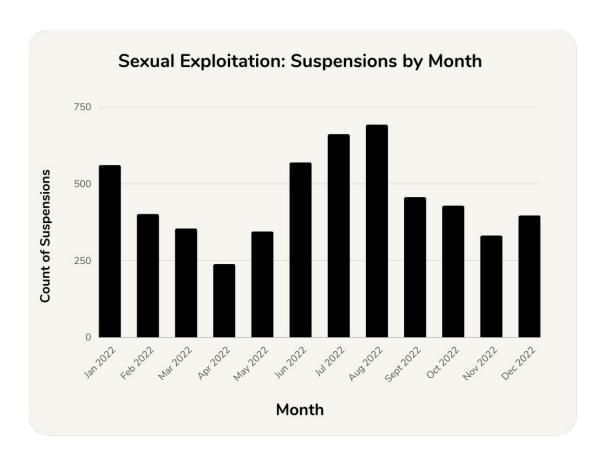
In 2022, our team actioned **17.5k** suspensions related to nudity and sexual content.



### Sexual Exploitation

In addition to child endangerment, which we detect, enforce, and report on as a distinct policy category (p. 10), the sexual exploitation of any person is strictly not allowed on Clubhouse. This includes, but is not limited to sexual images or other types of content being shared without consent, and content depicting or promoting sexual assault.

In 2022, our team actioned **5.4k** suspensions related to sexual exploitation.



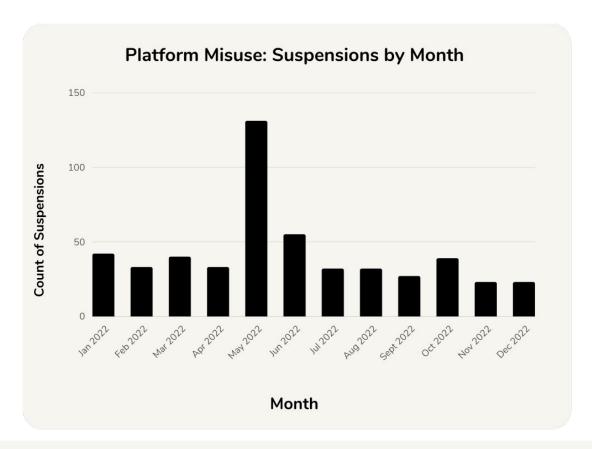


### Spam, Platform Misuse, and Misinformation

Clubhouse doesn't allow misuse of our platform for conduct such as:

- sharing irrelevant or unsolicited messages with other users, or otherwise behaving in a way that intentionally disrupts or ruins their experience.
- using scam tactics to solicit money or private financial information, or promoting investment and financial schemes.
- spreading harmful health misinformation, spreading misinformation for the purpose of making money, or spreading misinformation to manipulate or interfering in elections or other civic processes
- creating duplicate accounts by one individual.

In 2022, our team actioned **510** suspensions related to spam, platform misuse, or misinformation.

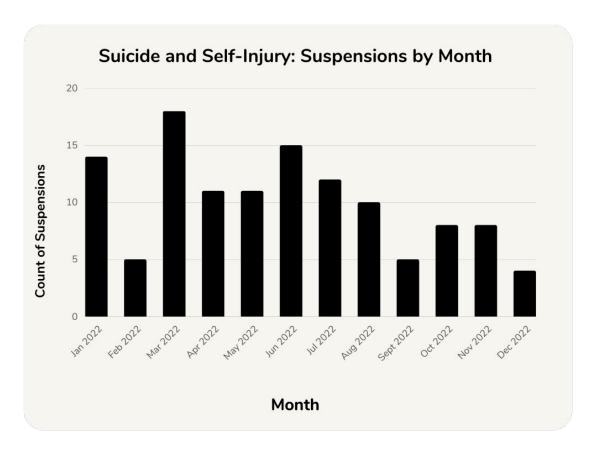


### Suicide and Self-Injury

We don't allow encouraging other people to engage in acts of self-harm or suicide.

In 2022, our team actioned **121** suspensions related to suicide and self-injury.

In addition to the enforcement of users who encourage acts of self-injury, we encourage our community to report users who have expressed that they may harm themselves. When our team encounters a user who may be at risk of self-harm, we send them an in-app message offering crisis support resources they can engage with, including how to locate region-specific support.



### Unauthorized Sales

Illegal activities and content are not allowed in Clubhouse, including (but not limited to):

- Promoting or selling illegal or highly regulated products
- Conducting any types of illegal or unauthorized sales or activities on the platform
- Encouraging others to take part in illegal activities
- Scams
- Soliciting funds or financial information
- High-risk investment and financial schemes
- Using the third party in-app payments system on Clubhouse for any of the above.

In 2022, our team actioned **181** suspensions related to unauthorized sales.





2022 Transparency Report

## Suspension Appeals

### Account Suspension Appeals

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:



Appeals Submitted\* 22.5k

Suspensions Upheld\*\*

1.9k

Suspensions Overturned 7%

Overturn Rate\*\*

<sup>\*</sup> The submitted appeal count will include duplicate appeals, where a complete re-review of the original enforcement decision will not occur, as the user has already received an appeal outcome previously. This figure may also not be fully inclusive of all appeals processed across 2022 due to cases where data was deleted by user request.

<sup>\*\*</sup> The count of "upheld" and "overturned" appeals is not fully reflective of cases where there is not a conclusive "upheld" or "overturned" outcome, such as repeat appeals, incomplete appeals that we couldn't process, and scenarios where an appealing user was not suspended.



2022 Transparency Report

## Proactive Detection

### **Proactive Detection**

In 2022, we invested further in our proactive detection safety efforts, using tools and investigation procedures to identify more potential violations. Some detection efforts involved keyword-based detection strategies, allowing us to identify profiles, channels, clubs, and houses that could be violating policies with higher safety risks, including violent extremism, child safety, hate speech, and self-injury or suicide. We then assessed these surfaces thoroughly for potential harms, and took actions accordingly for any confirmed violations of our Community Guidelines. We reviewed 14.4k total content surfaces (such as profiles, clubs, and houses) within this analysis across the four policy categories, and we took action on 38% of content surfaces overall.

14.4k

Content Surfaces Reviewed 5.7k

Content Surfaces
Actioned

38%

**Action Rate** 

Other noteworthy proactive detection projects we undertook in 2022 involved **social network analysis**, known to be an effective technique for identifying groups of users engaging in online harms, particularly engagement with child sexual abuse material (CSAM). One example was our **child safety** network analysis effort, through which we used several different detection techniques to conduct a thorough review of profiles closely connected to users suspended for child safety-related policy violations in the past.

### Child Safety Reporting

Preventing online harms against children and taking swift enforcement actions when harms are detected are among our top safety priorities at Clubhouse. In 2022, we continued to invest significantly in our efforts to protect minors and enforce on confirmed violations, particularly violations related to child sexual abuse material (CSAM). We grew our trust & safety teams focused on child safety, and implemented new tools and processes for proactive detection and enforcement.

959

PhotoDNA matches\*

1.02k

Reports to NCMEC\*\*

We use tools including Microsoft's PhotoDNA to detect, review, and report on known illegal images of child sexual abuse. In addition to taking enforcement actions, such as suspensions, as required by law, we report all illegal content related to harms against minors to the U.S. National Center for Missing and Exploited Children (NCMEC). NCMEC assesses reports received from online platforms and contacts law enforcement teams around the world to address them accordingly. In 2022, we submitted **1.02k** reports to NCMEC.

<sup>\*</sup> PhotoDNA data reflects the time period of Feb 3, 2022 - December 31, 2022.

<sup>\*\*</sup> Reports to NCMEC count may exclude some duplicate reports sent in relation to the same user, or in relation to multiple copies of the same image.



2022 Transparency Report

**Government and Law Enforcement Requests** 

### Government and Law Enforcement Requests

For the year 2022, Clubhouse received a total of **147** 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.

147

Requests Processed 11

Number of Countries

121

Requests for Information

16

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

Law enforcement and government agencies can find more information about submitting legal requests on the <u>Clubhouse Knowledge Center</u>.



### Government and Law Enforcement Requests Data By Request Type

Type of Request	Requests Received
Block/takedown	2
Emergency disclosure request	9
Fraud	7
Non-disclosure order	2
Preservation	14
Request for information	98
Request for information AND preservation	1
Request for information AND block/takedown	14
Grand Total	147

<sup>\*</sup> In addition to the request types shown above, we also received **9** regulatory notices from government agencies and **2** trust & safety reports from law enforcement. We also confirmed **9** Community Guidelines violations and took enforcement actions in response to requests where we did not provide information to law enforcement.



#### Government and Law Enforcement Request Data By Country

Country	Total Requests
Canada	2
Germany	8
Hong Kong	1
India	62
Iran	1
Japan	1
Nepal	1
Norway	1
Sweden	1
United Kingdom	5
United States	64

In 2022, we received requests from **11** total countries. We received the greatest numbers of requests from the **United States**, **India**, and **Germany**.



### Government and Law Enforcement Requests By Country and Request Type

	Emergency Disclosure	Takedown	Fraud	Non- Disclosure	Preservation	Request for Information	Information + Preservation	Information +	Total
	Disclosure	Takedown	Trada	Discosure	Treservation	mormacion	riescivation	Tukcuowii	Total
Canada	2								2
Germany						8			8
Hong Kong		1							1
India	1		7			40		14	62
Iran		1							1
Japan	1								1
Nepal						1			1
Norway						1			1
Sweden						1			1
United									
Kingdom						5			5
United									
States	5			2	14	42	1		64
Grand Total	9	2	7	2	14	98	1	14	147

### Transparency Report Recap

In 2022, our team at Clubhouse was able to build significantly upon the trust and safety foundations we established in 2020 and 2021, rolling out countless improvements to our community policies, detection practices, and enforcement tools. Throughout 2023, we will continue to update aspects of our safety ecosystem with the aim to continuously improve our detection and enforcement efforts, while also promoting transparency and fairness in our community members' experiences. Improving our proactive detection and child safety efforts, increasing transparency around enforcement, and shaping our enforcement actions to be less punitive where appropriate will continue to be among our key initiatives for the year ahead. Notably, we also plan to continue to improve our data collection and transparency reporting practices in 2023, in an effort to share even more data insights with our community and align further with emerging global legislation and industry best practices. We look forward to sharing our progress with the Clubhouse community as we continue to roll out new improvements. **Thank** you for reading!

#### **Additional Resources**

- Community Guidelines
- Trust & Safety Center
- Terms of Service
- Privacy Policy

