Contents

Behind the scenes at Yubo

03 Our company, purpose & values

Safety on Yubo

04 Safety strategy
05 Safety measures
05 Safety team
06 Trust & safety strategy
07 Promoting safety
08 Reporting concerns
08 Cooperating with law enforcement
09 Reports to law enforcement

Key metrics

10 Content and user moderation
10 User reports by category
10 Content removal
11 Suspension of user accounts
12 Appeals against permanent bans
13 Live moderation
Behind the scenes at Yubo

Yubo is a social live-streaming platform with more than 50 million users worldwide. Available to download for free on the App Store and Google Play, it is a space where anyone can belong, feel safe and make friends.

Our company: Created in France, growing worldwide

**FRENCH COMPANY**
Behind Yubo, there is a company called Twelve-App SAS, which is based in France. It was founded by three young French engineers in 2015.

**PASSIONATE TEAM**
The Yubo team includes creatives, engineers, data scientists, safety and privacy specialists, and marketers. Everyone is committed to creating a fun and safe space for the Yubo community.

**EXPANDING ABROAD**
Yubo/Twelve-App SAS has subsidiaries, offices and teams in Paris (France), London (UK) and New York (USA).

Our purpose: Empower youth to discover and belong

**POSITIVE CONNECTIONS**
Yubo is a social platform that celebrates the true essence of being young and makes it easy to connect with others wherever you are. It is our mission to inspire and enable a new generation to find out about the world and themselves.

**FRIENDS NOT FOLLOWERS**
Unlike many other social platforms, Yubo is about being yourself – it is not about collecting followers or ‘likes’ or engaging in social comparison. We empower our users to be themselves, with the support of others.

**AUTHENTIC COMMUNICATION**
Yubo provides a space where people from around the globe can communicate and learn in small or large groups in an authentic way and in real time. The app is a place to socialize rather than share content.

Our values: Openness, safety and privacy

**OPENNESS**
We want Yubo to be a place where anyone can belong and be themselves. We encourage members of the Yubo community to respect other people and their differences and we have created an environment where our users can build meaningful and trusted relationships.

**SAFETY**
We take our responsibility to building a safe digital space very seriously – this has been a priority for Yubo from day one. Our innovative approach, industry-leading tools, dedicated safety team, network of partners and independent Safety Board help to keep the Yubo community safe.

**PRIVACY**
Yubo’s products, features and processes are always developed with ‘privacy by design’. With no selling or sharing of personal data and no targeted advertising, our commitment to user privacy helps us to stand out from the rest of the industry.
Safety on Yubo

As the leading platform for young people to socialize online, we have a special responsibility to members of the Yubo community. Safety is at the core of our DNA and we have developed and implemented a range of safeguards in the app. We also work with the broader technology industry, NGOs and charities to create industry-leading child protection standards and run in-app awareness campaigns.

Safety strategy

PROTECT, SUPPORT, EDUCATE
We help our users to become more self-aware and responsible for their own safety and the safety of others on Yubo. We take an educational approach whereby we explain, support and prevent more than we penalize. For example, we send pop-up alerts that inform our users how they have broken our Community Guidelines and how to change their behavior.

PROACTIVE APPROACH
We endeavor to catch the damage before any harm is caused. For example, using AI technology, we have developed algorithms to proactively detect suspicious or inappropriate content or behavior in the Yubo app.
Safety measures

SAFETY TEAM

We have a team of over 150 safety experts who are dedicated to ensuring the safety of the Yubo community, including:

Frontend: Interacting with users

LIVE OPERATIONS
Monitors our Live feature (live-streams) 24/7 and, when it sees inappropriate content or behavior, intervenes to take action.

IDENTITY AND CONTENT OPERATIONS
Moderates the content in user profiles (e.g. photos) and also works on our age and identity verification operations.

INVESTIGATIONS
Monitors and analyses suspicious behavior patterns and related trends.

SUPPORT
Responds to any concerns or issues reported through our Help Center by users, parents, caretakers or educators.

Backend: Developing tools and processes

SAFETY POLICY
Develops and drafts all the safety policies and guidelines that are enforced on Yubo.

SAFETY ENGINEERING
Develops technical tools for all our safety teams so they can operate proactively and efficiently.

PRODUCT SAFETY
Designs the tools and processes that enable our safety teams to take action when our terms or Community Guidelines are broken.

LAW ENFORCEMENT
Works with law enforcement authorities on the reporting of illegal activities, responds to information requests and deals with emergency situations.

DATA SCIENCE
Develops analysis tools and analyses data in support of our detection and prevention safety operations.
Trust & safety strategy

**POLICIES**
Our policies determine the kind of content and behavior that is allowed on Yubo and ensure that our community behaves in a way that is aligned with our values (see our Community Guidelines). They provide the framework for all our moderation efforts to ensure that safety risks are treated in a consistent and coherent manner, in line with industry best practice.

**ACCURACY**
We conduct quality analysis to evaluate accuracy and continuously improve how we protect our users. In every case, automated detection triggers a report that a member of our safety team checks. Our global average accuracy is 95%.

**USER COMPLAINTS**
An internal Yubo complaint process is easily accessible for all users if they wish to challenge a moderation action taken because of content they shared or something they did on the platform.

**USER-GENERATED CONTENT MODERATION**
Our automated and innovative solutions combined with real-time interventions by our safety team enable us to detect and moderate inappropriate, harmful or illegal content and behavior in Lives and chat.

**PROACTIVE MODERATION**
We proactively block inappropriate content by default (SAFE Mode) and Yubo users can customize the filters by adding keywords to block particular content they do not want to see or receive (e.g. using our Muted Words feature).

**IDENTITY AND AGE VERIFICATION**
Our identity and age verification features help to build trust on Yubo and reduce the number of fake accounts. For example, we use face recognition algorithms to check the authenticity of profile pictures and users are able to provide e.g. their passport or driving license as ID.

**AGE GATES**
Yubo has a minimum age of 13 and we use age gates to create separate communities for ‘13 to 18-year-olds’ and ‘over 18s.’ We want to create a space where our users socialize with people of a similar age, as they would in real life.

**HUMAN MODERATION**
Our moderation processes that use automated solutions to detect inappropriate, harmful or illegal content and behavior include a review by a member of our safety team (a human moderator) of the detected or reported content or behavior.

**GROOMING DETECTION**
We have developed technologies and processes to detect suspicious behaviors through patterns. When grooming is suspected, we conduct detailed investigations and take specific action. This includes analyzing and removing inappropriate, harmful or illegal content and, where necessary, reporting it to law enforcement authorities.
Promoting safety

**PARTNERSHIP APPROACH**

Yubo has always worked closely with key stakeholders worldwide, including NGOs, charities and other relevant organisations. Our partners include NCMEC, Thorn, NSPCC, e-Enfance and two members of INHOPE (Internet Watch Foundation and Point de Contact). We are also an active member of the Tech Coalition and the ICT Coalition and we work with local partners to run in-app awareness campaigns on topics such as bullying and mental health.

**SAFETY HUB**

As we value education, prevention and transparency at Yubo, our Safety Hub provides detailed guidance on digital safety. It includes advice on specific issues (e.g. hate speech and sexual exploitation) and an overview of our safety tools. We have also created two guides to staying safe on Yubo – one for young people and one for parents, caretakers and educators – we also have a guide specifically for law enforcement.

**SAFETY BOARD**

The independent Yubo Safety Board provides invaluable advice and informs our proactive approach to safety.

**CHIEF SAFETY ADVISOR**

Annie Mullins OBE has been Yubo’s Chief Safety Advisor since 2016. With 15+ years of experience developing industry good practice guidance, Annie works with us to develop our safety strategy and build our relationships with Governments, NGOs and industry groups, such as the Tech Coalition and the ICT Coalition.

**SAFETY BOARD**

Our Safety Board is composed of world leading experts on digital safety and other related topics. Its members are: Dr. Richard Graham (Adolescent Psychiatrist), Anne Collier (Net Family News), Michael Moran (ex Interpol), Alex Holmes (The Diana Award), Travis Bright (Thorn) and John Shehan (NCMEC).
Cooperating with law enforcement

- **DEDICATED TEAMS**: We are committed to dealing with all reports from users as quickly as possible. Our safety teams review all reports and respond to the user within two hours depending on the kind of report and its complexity. If the report is about a Live, we aim to respond within one minute; if it is about a user profile, we aim to respond within one hour.

- **EASY REPORTING**: A reporting button is embedded in every Yubo profile and feature in the app so that users can easily send us reports about inappropriate, harmful or illegal content or behavior, whether it concerns them or another user. Our users are diligent in reporting profiles and content on the app that goes against our Community Guidelines.

- **PARENT CONCERNS**: Parents, caretakers and educators can contact us through our Help Center at any time, whether it is about a specific concern or a general question. We aim to respond to these reports within 24 hours. We have also published a guide to staying safe on Yubo for parents, caretakers and educators.

---

Cooperating with law enforcement

- **POLICE REQUESTS**: Within the applicable legal framework, and following a specific process, our dedicated Law Enforcement outreach team processes all requests from law enforcement authorities. 99% of the Law Enforcement requests we receive are processed within 24 hours.

- **STRONG NETWORK**: We are building relationships with law enforcement authorities and other relevant organizations in order to gain input about, and increase visibility of our law enforcement outreach processes.

- **PROACTIVE REPORTING**: When necessary, we proactively report illegal activities and situations presenting an imminent safety risk for our users to law enforcement authorities and other relevant organizations (e.g. NCMEC).
Reports to law enforcement

Reporting illegal content and behavior on the Yubo platform is the logical extension of our safety strategy. Analysis of the related data also presents an important opportunity to improve our safety policies and processes.

2,420 reports to PHAROS
We report all illegal content and behavior on Yubo to the French police platform for reporting illegal online content, PHAROS. When the situation involves users outside of France, PHAROS transfers the reports and information to the relevant local or international law enforcement authorities.

885 reports to NCMEC
We report all child sexual exploitation and abuse cases in the USA and Canada to the National Center for Missing and Exploited Children (NCMEC), a leading organization in the fight against child abduction, abuse and exploitation.

30 countries
In 2021, we made proactive reports to law enforcement authorities in 30 different countries, including:
- THE UNITED STATES (927 REPORTS)
- UNITED KINGDOM (412 REPORTS)
- AUSTRALIA (141 REPORTS)
- CANADA (105 REPORTS)
- BRAZIL (96 REPORTS)

Information requests
We receive data and data preservation requests from law enforcement authorities. In accordance with the applicable legal framework, our dedicated law enforcement outreach team replies to these requests within 24 hours in 99% of cases.

382 requests
In 2021, we received 382 data and data preservation requests from law enforcement authorities around the world.
Content and user moderation

1.6 MIL reports per month
Our safety team processed an average of 1.6 million reports per month in 2021. These were either sent by our users or generated by our algorithms.

80% from proactive detection
80% of the reports processed by our safety team and by our preventive moderation tools are the result of our proactive detection strategy. In 2021, 75% of these reports resulted in action being taken (e.g. temporary suspension of a user’s account).

20% from user reports
20% of the reports processed by our safety team came from the Yubo community. The actions taken by our safety team are guided by our Community Guidelines. 30% of user reports in 2021 resulted in action being taken (e.g. a Live being closed down).

By comparing the figures between our proactive detection systems and the user reports it appears that, after human review by our safety team, our proactive systems leads to ten times more actionable reports than user reports.

User reports by category
The metrics below display the main reasons selected in user’s reports throughout 2021. Hate speech and bullying constitute around 50%. As detailed above, those user reports represent 20% of the safety reports and only 30% of them were actionable.

- 27% Bullying and harassment
- 26% Hate speech (e.g. racism, homophobia)
- 14% Sexual and nude content
- 14% Other
- 8% Harmful or dangerous behaviors
- 7% Inappropriate title of a Live (live stream)
- 4% Self-harm

Content removal

15,000 user biographies removed per month
In 2021, our safety team reviewed an average of 30,000 biographies submitted by Yubo users per month for their profile, of which 50% were removed for breaking our terms or Community Guidelines.

Over 107,000 media removed per month
In 2021, our safety team reviewed an average of 153,837 photos or videos submitted by Yubo users per month, of which 70% were removed for breaking our terms or Community Guidelines.
Suspension of user accounts

Education is one of the pillars of our safety strategy. In order to empower our users to manage their own safety, we have developed a range of moderation policies and processes. Depending on the severity of each situation, a user may receive a pop-up alert, be temporarily suspended from Yubo or permanently banned from the app.

437,688 temporary suspensions in 2021 on average per month
Our suspension system temporarily restricts a user’s usage of Yubo when they have broken our terms or Community Guidelines. The length of suspension can vary from two hours to one month depending on the severity of the violation. This system gives the user the opportunity to use the app appropriately once their suspension has been lifted.

72,001 permanent bans in 2021
When a violation is severe, we can permanently ban the user from the Yubo app. Where the content or behavior could be illegal, our 24/7 safety team reports it to the relevant law enforcement authorities on the same day the offense is detected.

Number of permanent bans by reason in 2021 (rounded up)

- **38,000** Illegal commercial activities
- **12,000** Propositioning
- **8,000** Graphic nudity
- **5,000** Grooming
- **3,700** Child sexual exploitation
- **3,000** Child sexual abuse material
- **2,000** Scam
Appeals against permanent bans

Users who have been banned permanently from the Yubo app can appeal our decision. We have a dedicated team that reviews these cases and allows some of the users to have their ban lifted.

7,000 cases investigated
In 2021, our safety team investigated 6,792 cases of permanent bans that were appealed by users.

19% of accounts reinstated
As a result of our investigations following appeals in 2021, we reinstated almost a fifth (19%) of user accounts that had faced permanent bans.

Identity and age verification

Identity and age verification are key assets to ensure trust and safety on a social media platform. For this reason, Yubo develops and deploys identity verification processes, age verification tools and age gate solutions.

67,000 accounts per month suspended following age verification
Where there is any doubt about the age provided by a user at sign-up or their activity on Yubo is inconsistent with the age provided, we require the user to verify their age through an ID check process. The user's account is temporarily blocked until verification is completed and, if age verification does not prove they are the required age, their account is suspended.

30,000 accounts per month suspended following identity verification
When there is any doubt about the authenticity of a user account, we ask the user to verify their identity with a relevant proof of identity (e.g. passport, driver's license, school ID). The account is temporarily blocked until verification is completed.

7,000 verifications processed per day
Age verification and identity verification requests result in a similar process of ID document verification. On average, we processed 6,500 responses to our verification requests per day in 2021.
As a live-streaming platform, real-time moderation is a key focus for Yubo. In line with our ‘protect, support and educate’ safety strategy, our safety team developed a real-time moderation process several years ago, which we continue to improve.

880,239 automated detection alerts on average per month
When a violation is detected by our AI technology, an automated detection alert is sent to our safety team for them to review and take action.

118,462 user report alerts on average per month
Users can report content or behavior that they consider inappropriate or offensive in a live-stream. Our safety team reviews the report and takes the appropriate action.

Reasons for automated detection alerts

- **88%** Partial nudity
- **6%** Weapon on display
- **4%** Graphic nudity
- **<2%** Self-harm
- **1%** Drug use
- **<1%** Other
- **<1%** Hate speech
Thanks!

www.yubo.live @yubo_app

To learn more, please visit our Safety Center at www.yubo.live/safety