

User Safety

yubo



Some Metrics about Yubo



2016
Creation



Nationality



46
Staff members

Product Today

Find your Tribe



Socialize online

Discover new people from all over the world (languages, communities, interests, etc.)



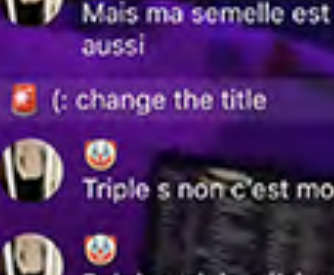
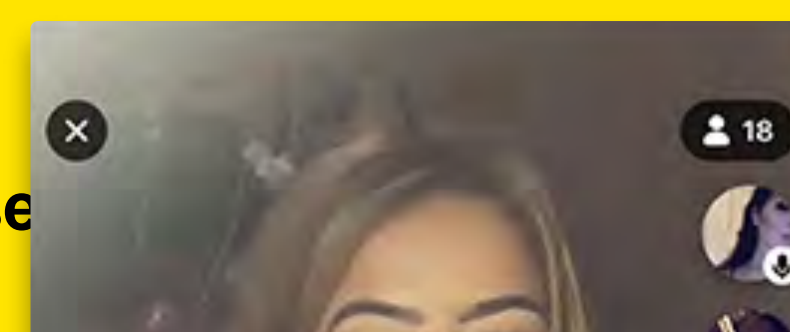
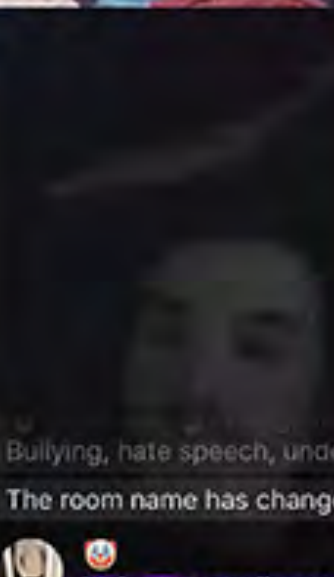
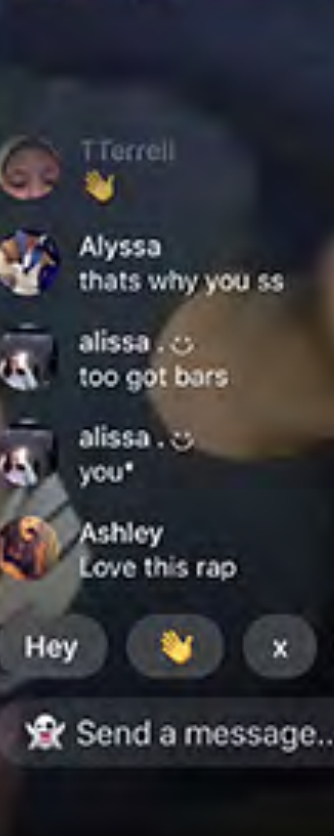
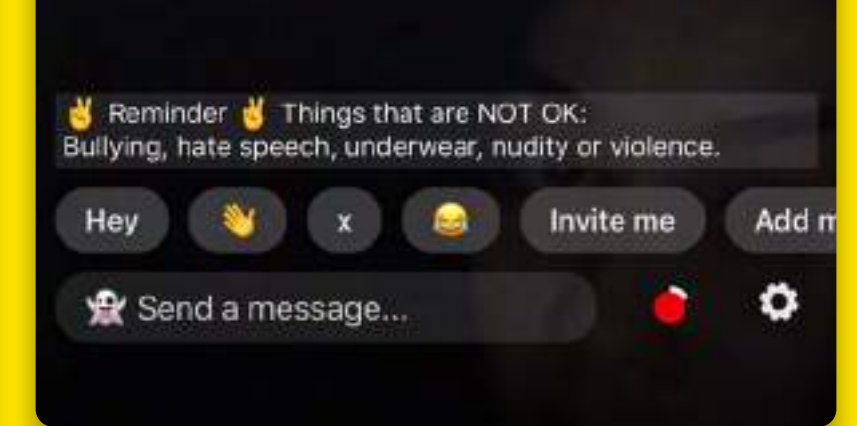
Group Live stream

Join groups of people based on the streamers' profiles

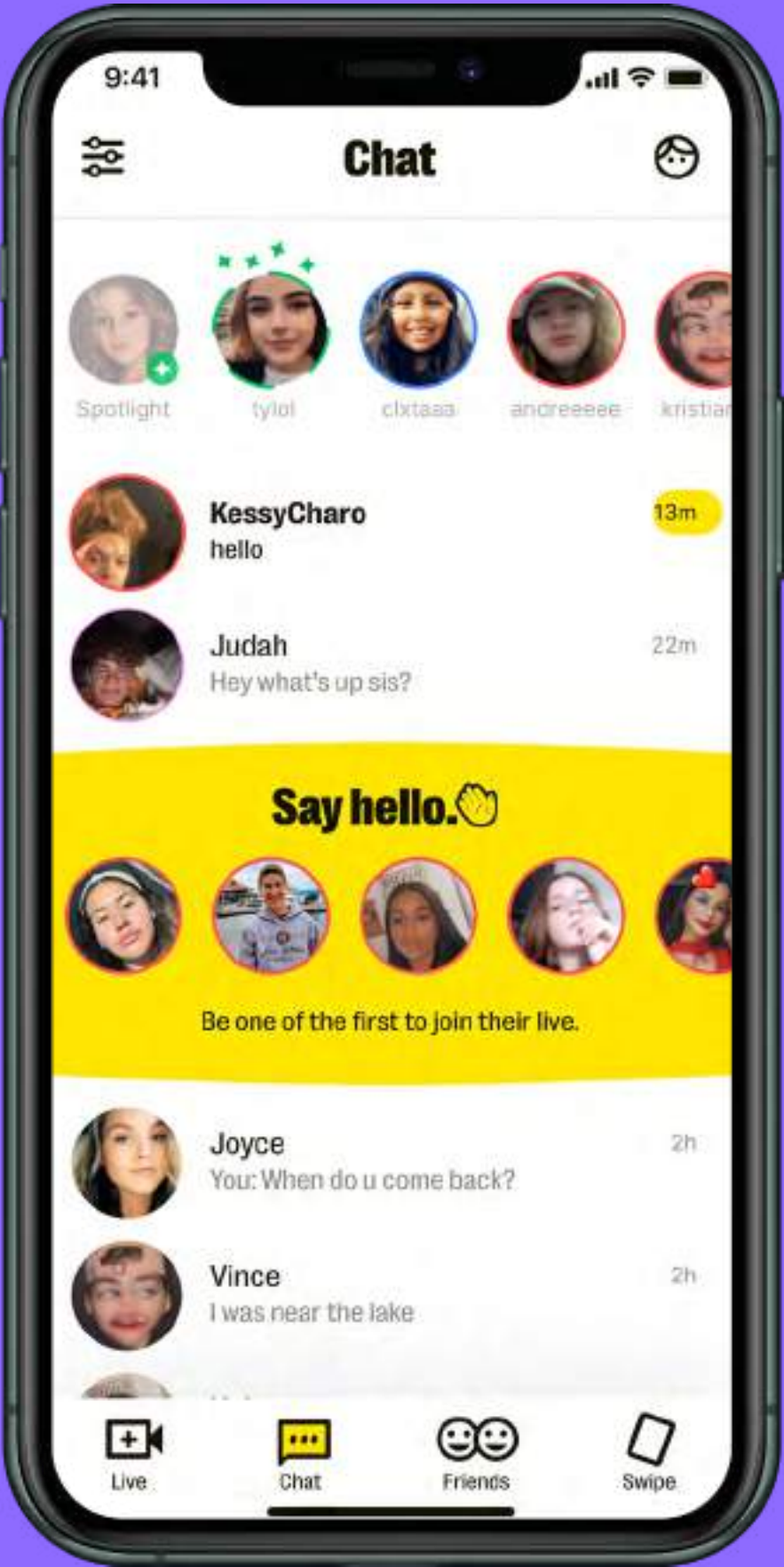
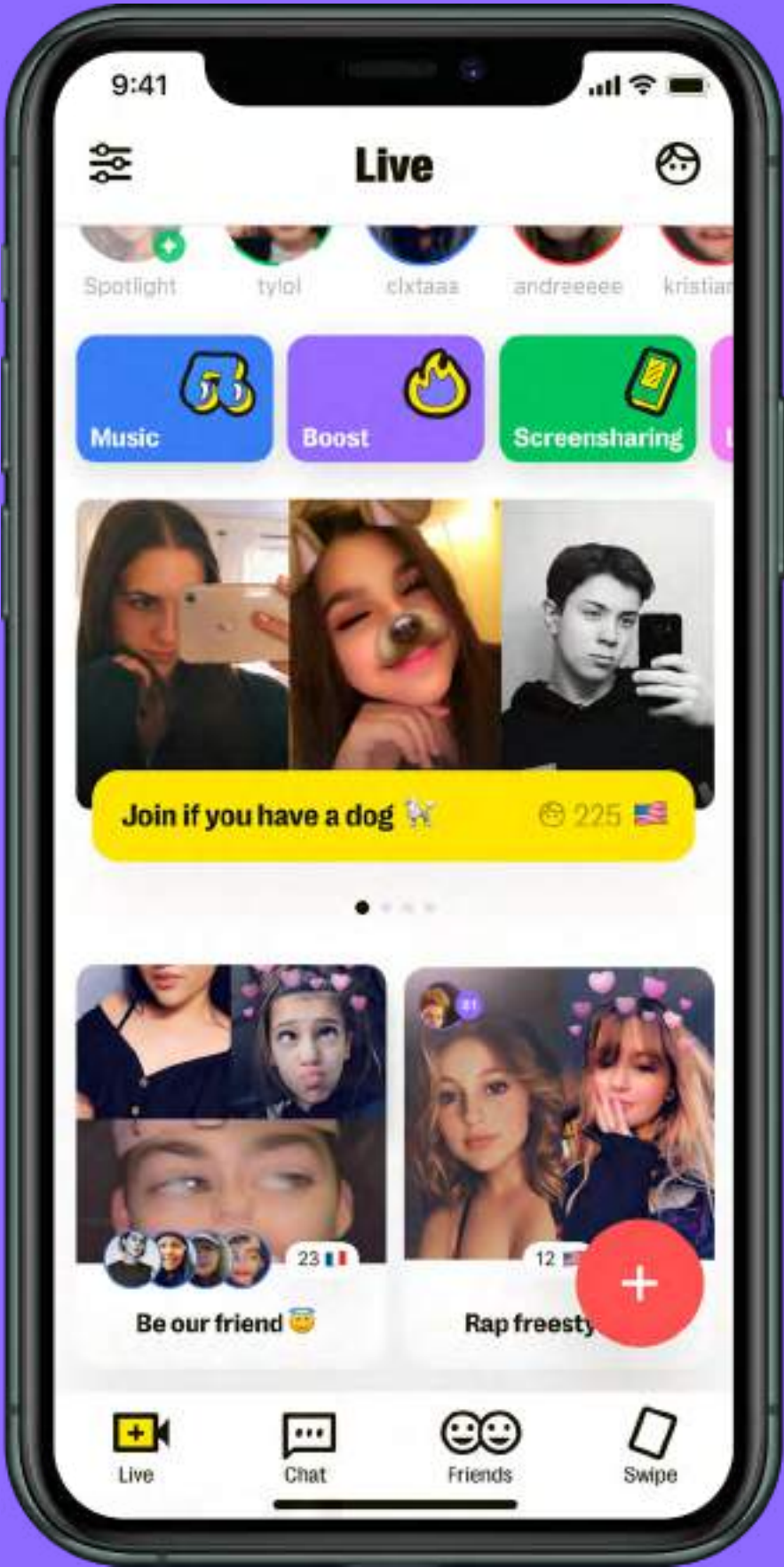


Messaging

You can talk to your friends in private text messages



How does it Work?



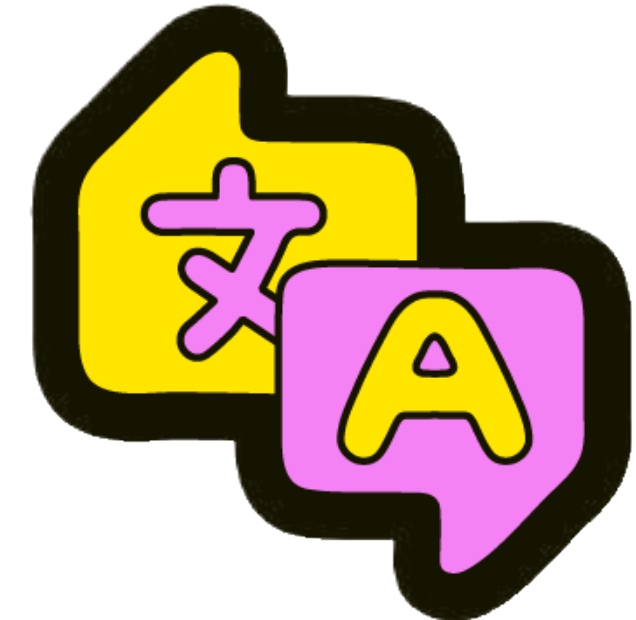
Some Metrics about Yubo



45m
users worldwide

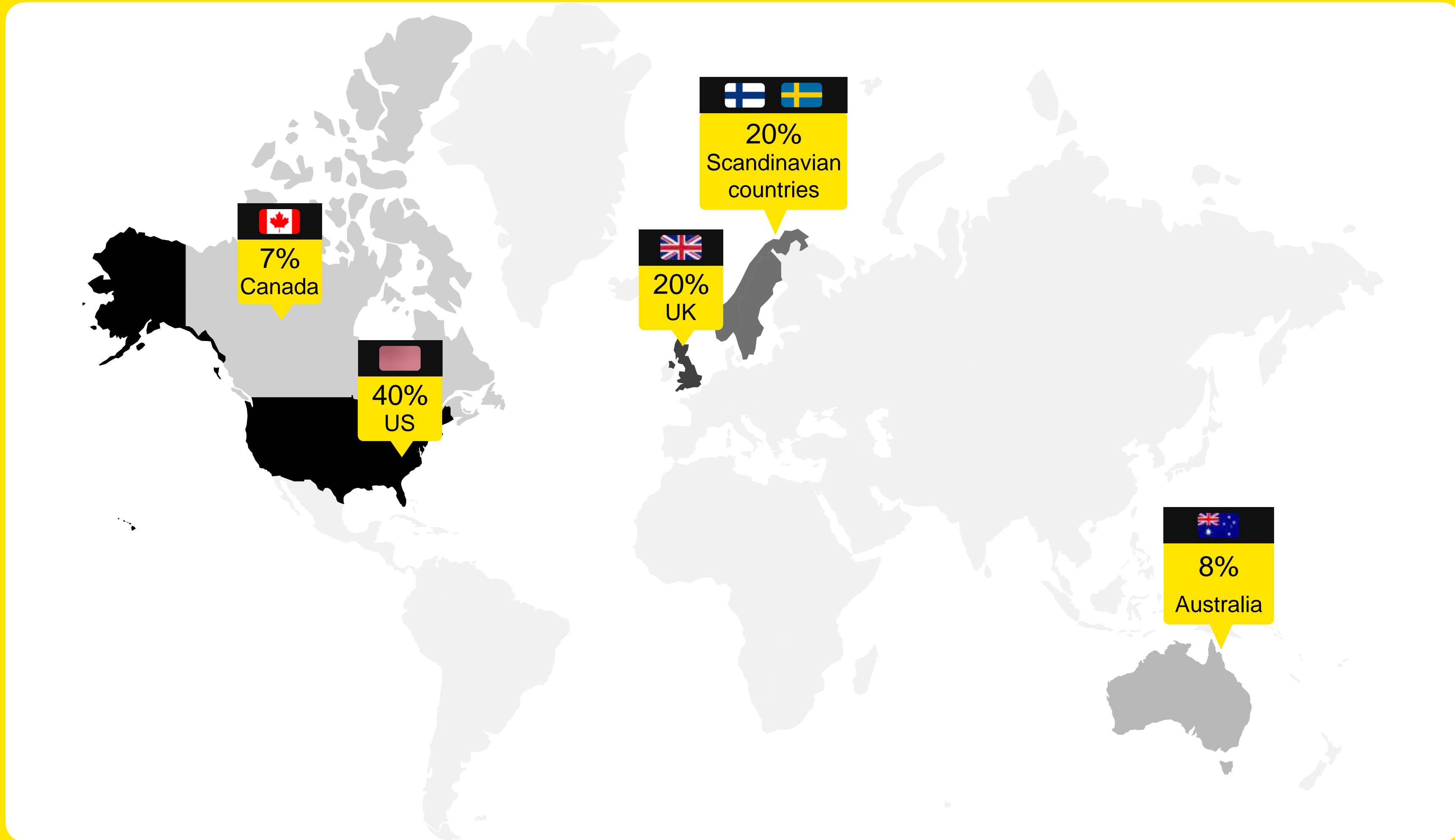


+100m
live videos

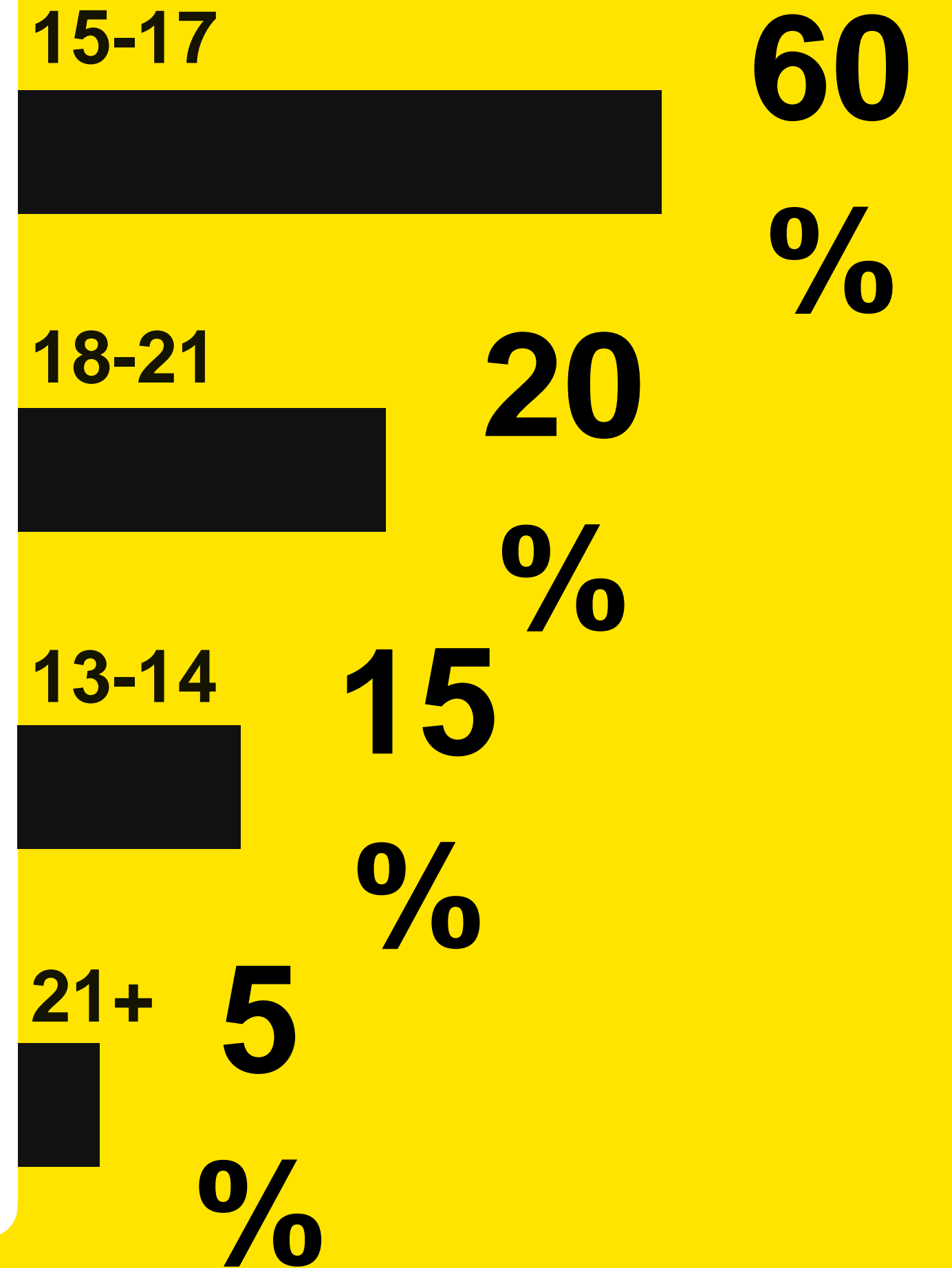


+10b
messages sent

Our User Base



AGE SPLIT



yubo

Safety

Some Difference with Other Platforms



STRICT GUIDELINES

We have put in place stricter guidelines and block content including shirtless pictures, underwear, weapons or drugs.



INTERVENTIONS IN REAL TIME

We send real time alerts to users when we feel they are not using the app safely (underwear, shirtless, etc.)



AGE GATES

We have created age gates to prevent 18+ users sending friend requests to younger users under 18 and vice versa

Our Safety Framework



MODERATION TECH

Building machine learning algorithms to scale moderation and help make good decisions.



REPORTS

Treat report in a very fast, efficient and smart way.



MODERATION TEAM

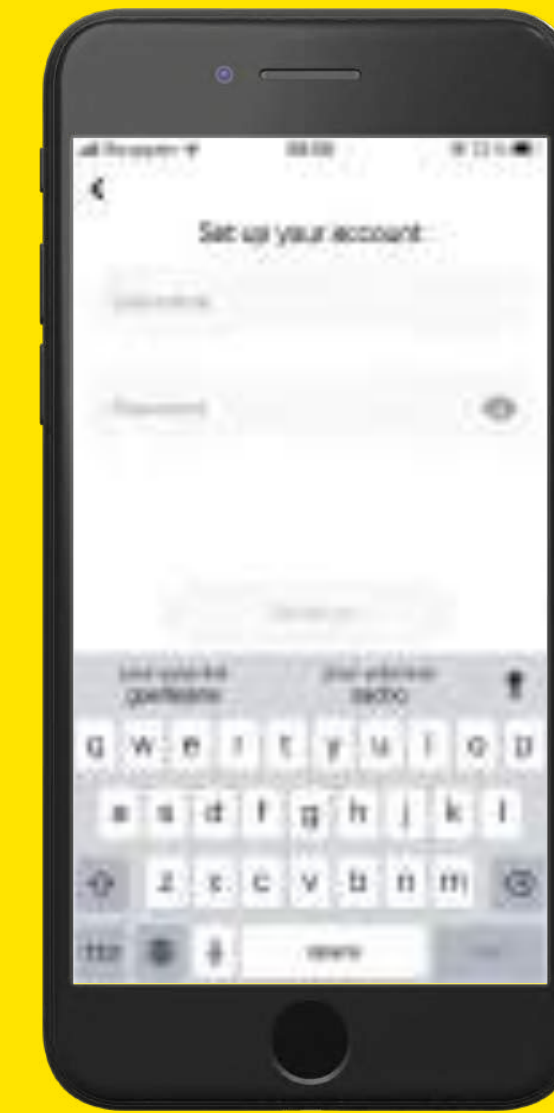
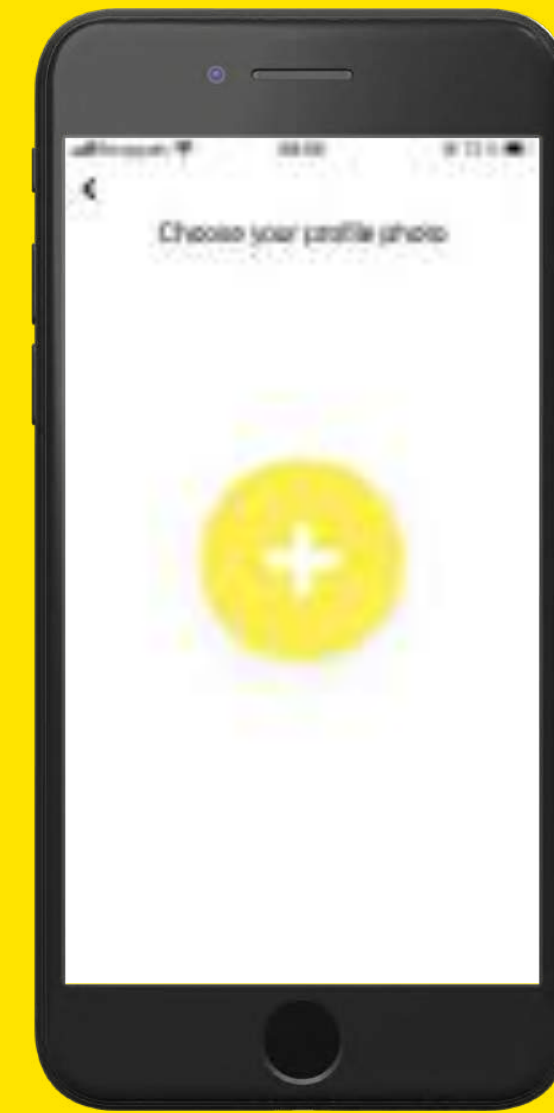
We have local moderation teams 24/7. They check profiles, content and live streams.



EDUCATION

Giving users the best tools to protect themselves before we intervene

What's Behind?



Username

We use our word filtering algorithm to keep users from using inappropriate words in their usernames

Multiple accounts

We check if you have already created multiples and spot it if there are age or gender variations.

Age

We analyze your profile picture and try to estimate your age. Under 13 users will have their devices blocked.

Phone numbers

We scrape everyday websites with fake phone number so you can't register with them.

Profile pictures

We analyze your pictures and video with our NSFW algorithm and reverse search it on Google Image.

Flagging

We can consider your profile as « unsafe » and double-check it. Before checking, it will be hidden from the app

Real Identity



Trust

To make new friends, you have to trust the people you are interacting with



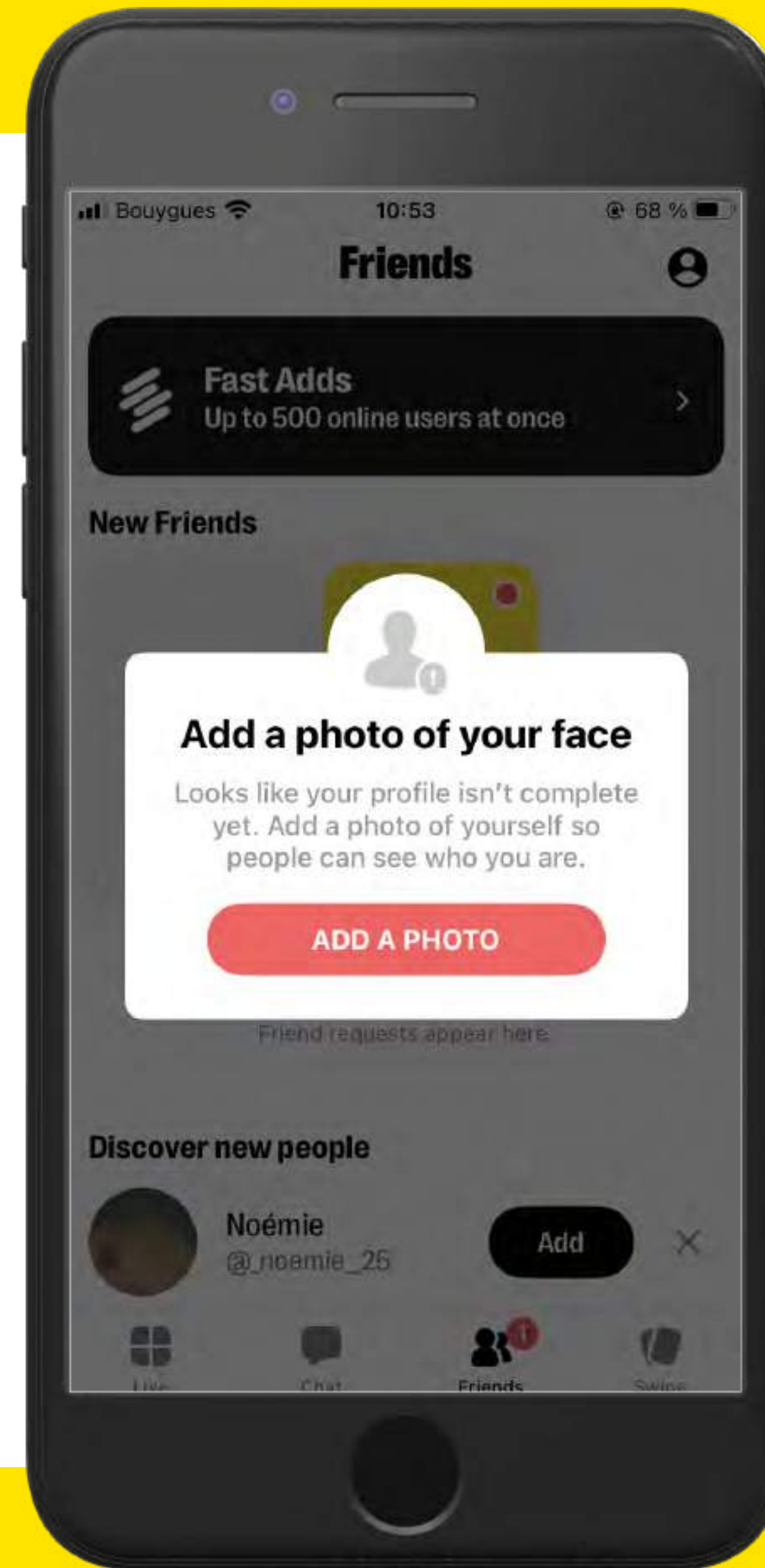
Real identity

We ask our users for real information and a photo when they set up their profile and we don't allow them to be anonymous.



Face detection

Our algorithms detect pictures with real faces. If you don't upload a real picture, you won't be able to access the App's features.



Profile and Video Detection



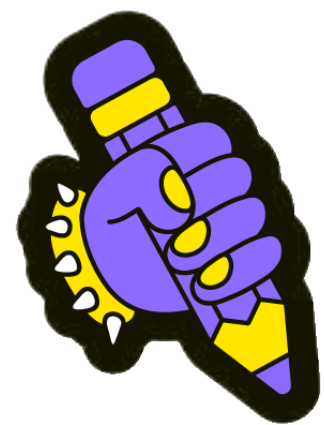
Inappropriate pics

You cannot upload nudity pictures when detected. If the picture is detected as suggestive (shirtless, underwear, etc.), you receive a popup alert.



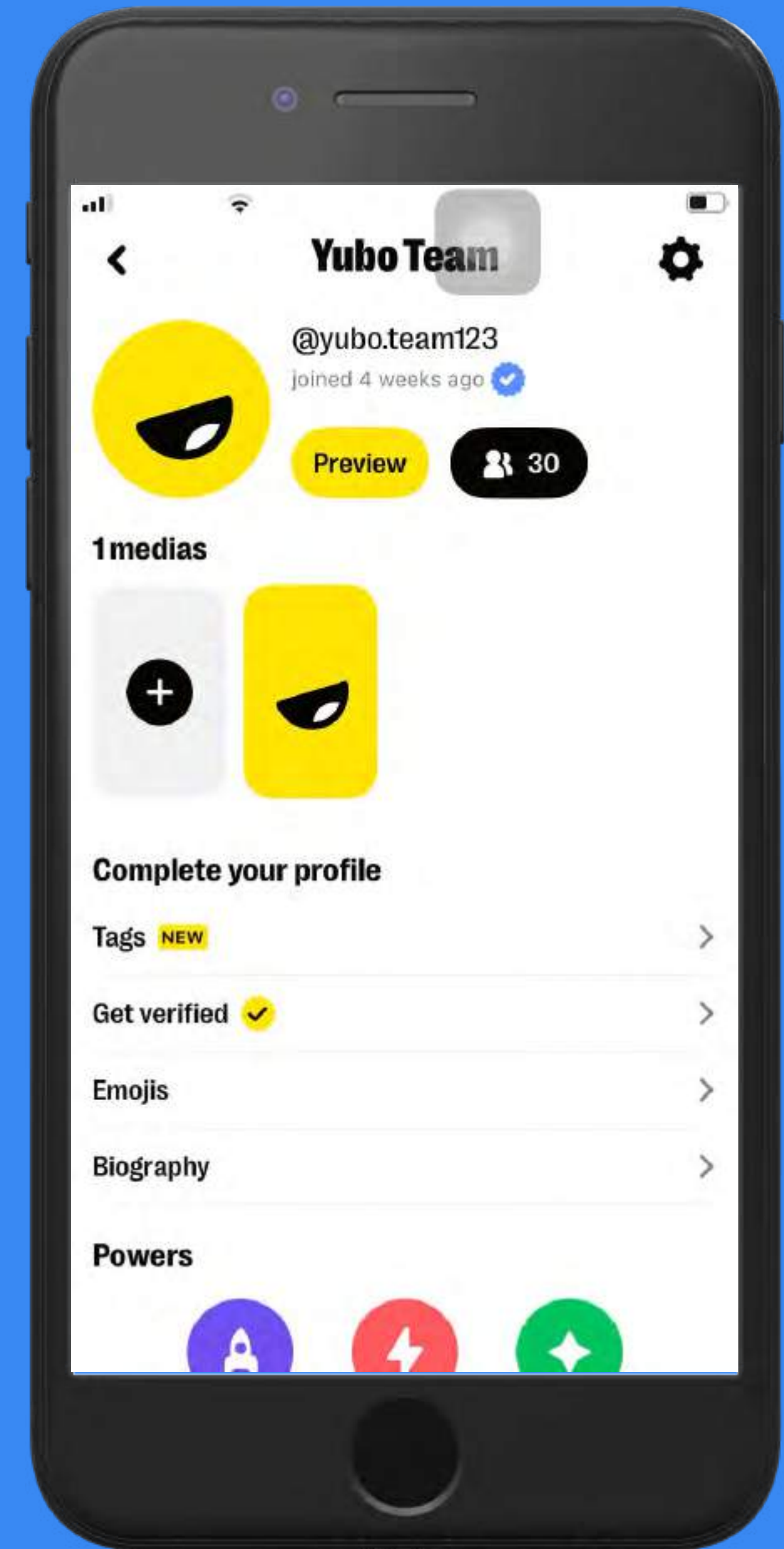
Prevention

25% cancel their content upload when they receive this alert.



Human Moderation (H24)

If the user uploads the picture, it will be then double-checked by our review team and the picture removed if it breaks our guidelines.



Identify and Age Verification are big Challenges

| Medium informations | |
|-------------------------------|---------------|
| Detected Faces | 2 |
| Snapcode | No |
| Identity | Not moderated |
| Moderation | Not moderated |
| Uploaded | Invalid date |
| | Invalid date |
| Hive | |
| Clean | 99.8 |
| Yoti Age | |
| Predicted Age | 10 |
| Minimum Predicted Age | 3 |
| Maximum Predicted Age | 17 |
| Prediction Uncertainty | 7 |

ello peepz Im 12, and im just chillen Im lookin for friends cuz in lonely forever 😂 Im bad at starting a conversation 😂 Im a little couch potato, im shy and I like animals and Anime/Yaoi This is my bio so BAI PEEEPZ 🤪

Device users for k6054:

Show total

Hector (@k6054)
Male - 21 years old
Dordrecht, Netherlands
4 minutes streams / an hour watch
Registered 3 days ago

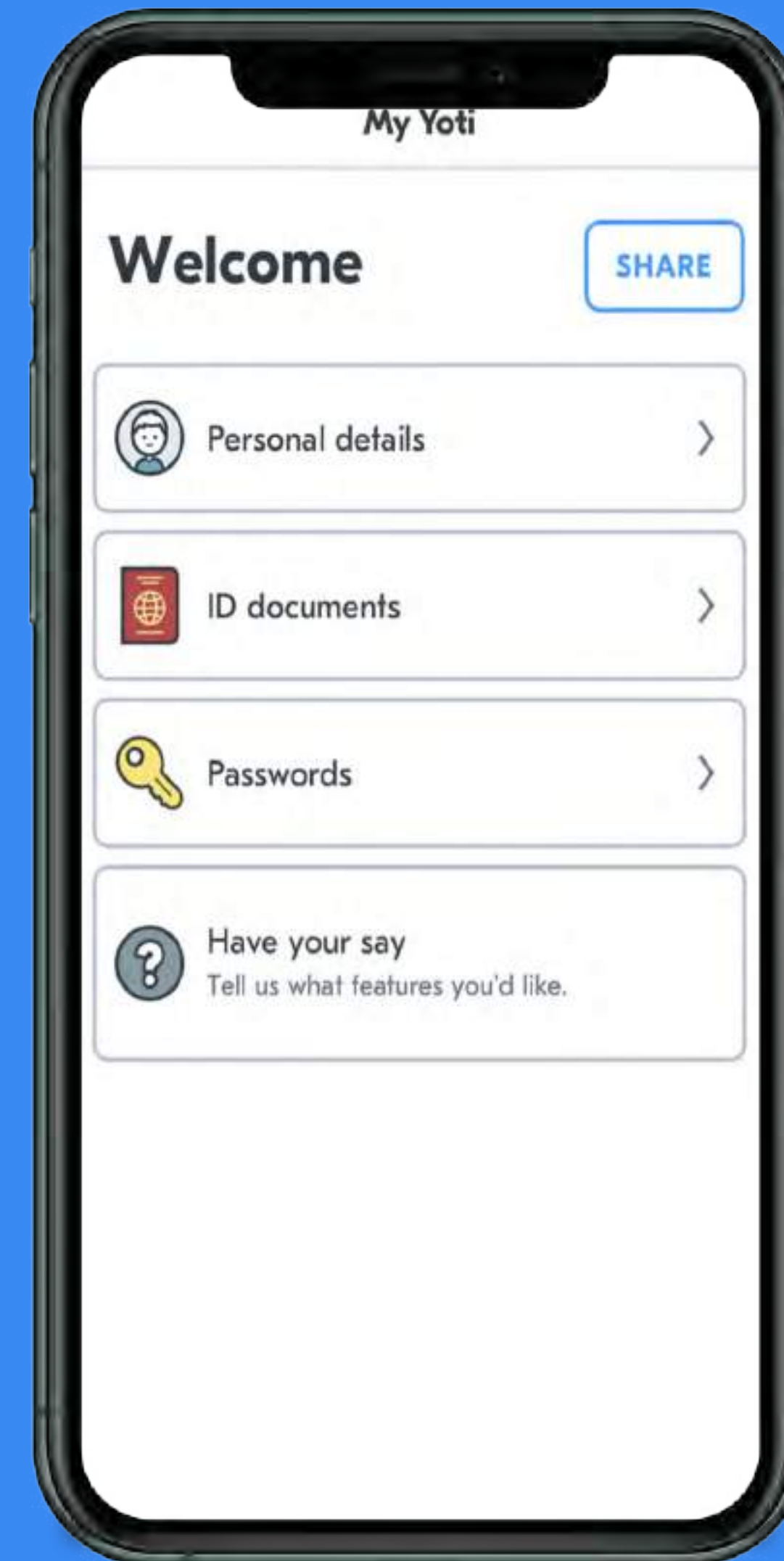
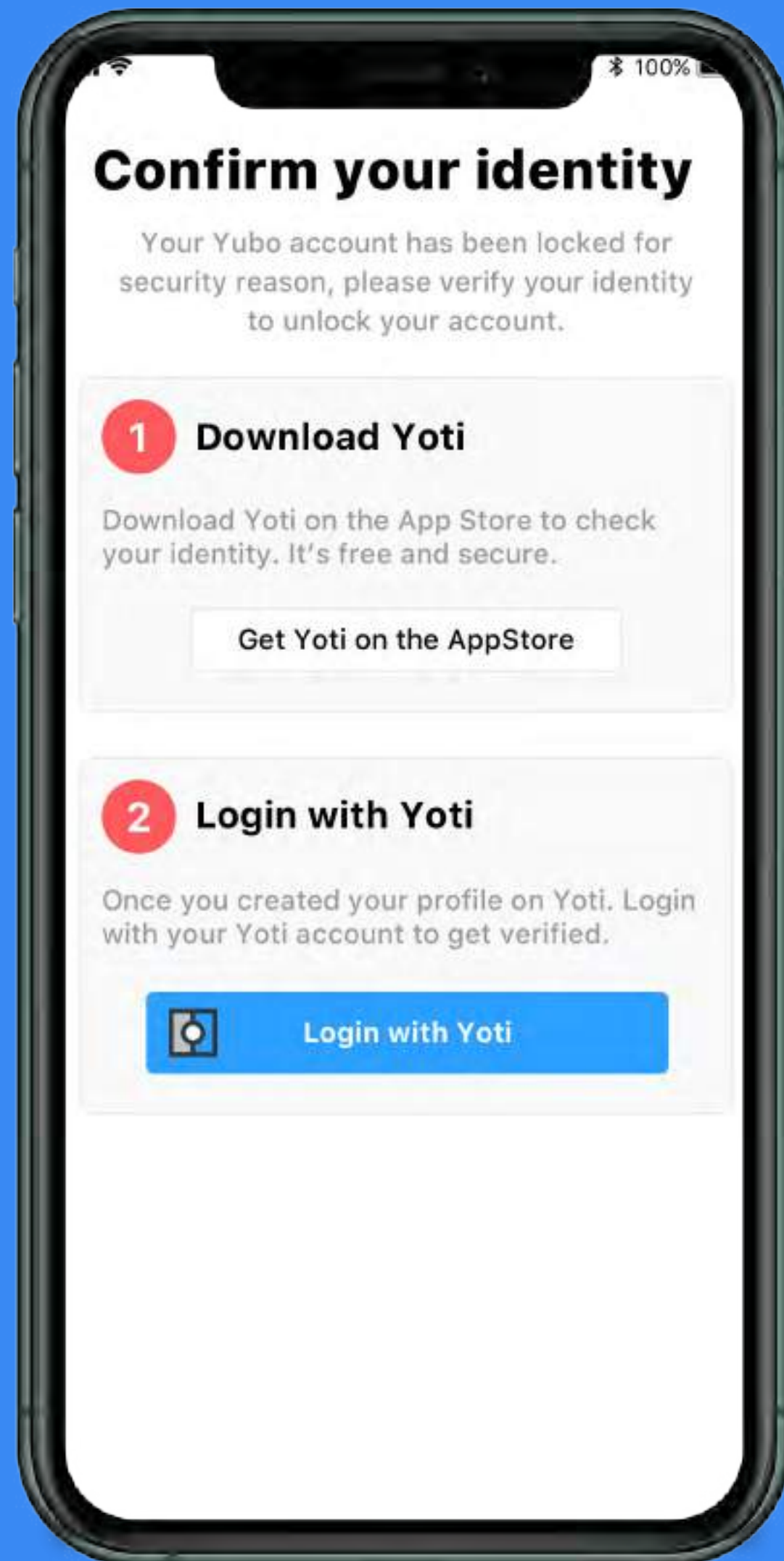
Show total

Hector (@hectorrhorni)
Male - 16 years old
Dordrecht, Netherlands
no streams / 7 minutes watch
Registered 3 days ago

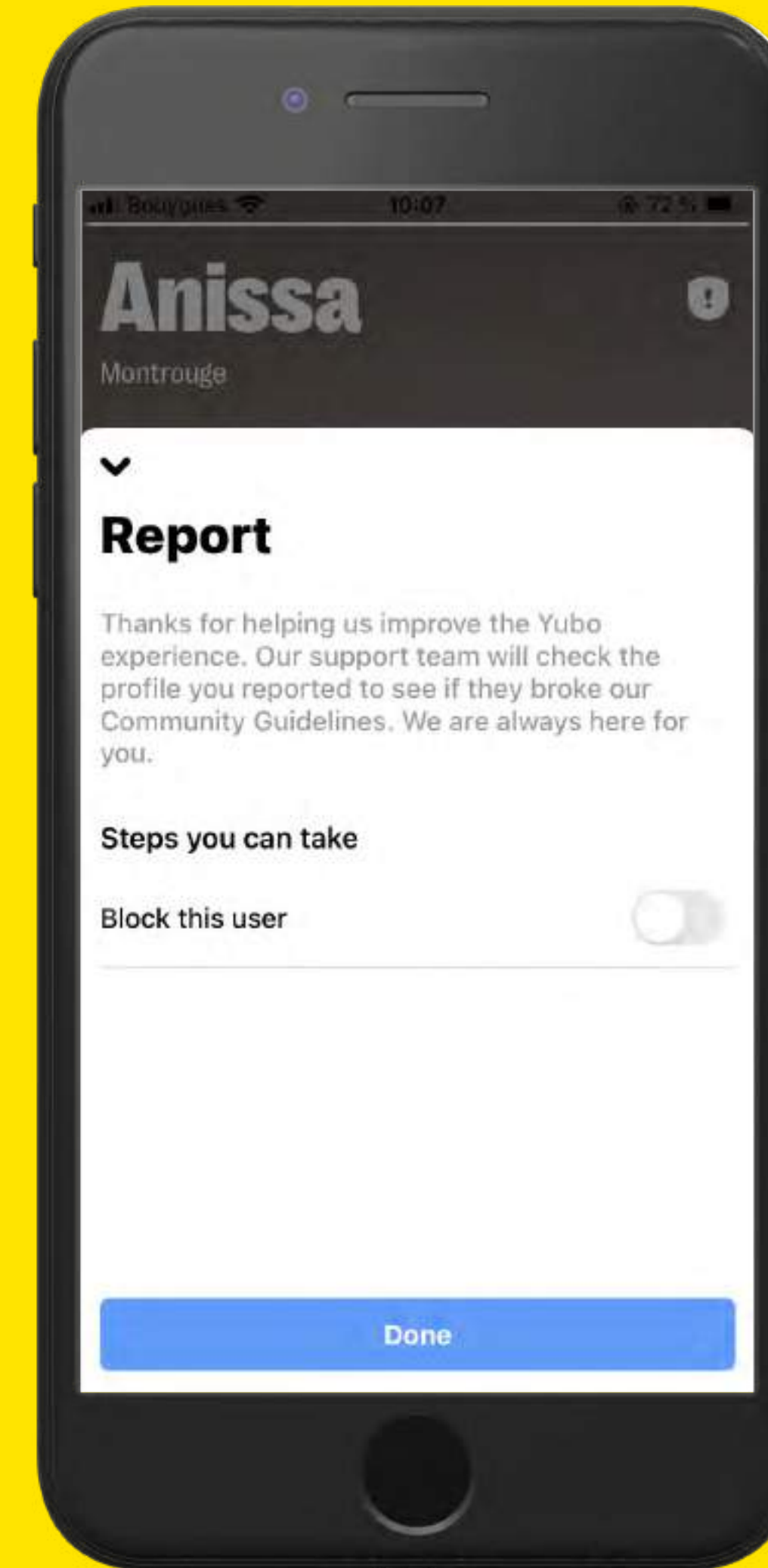
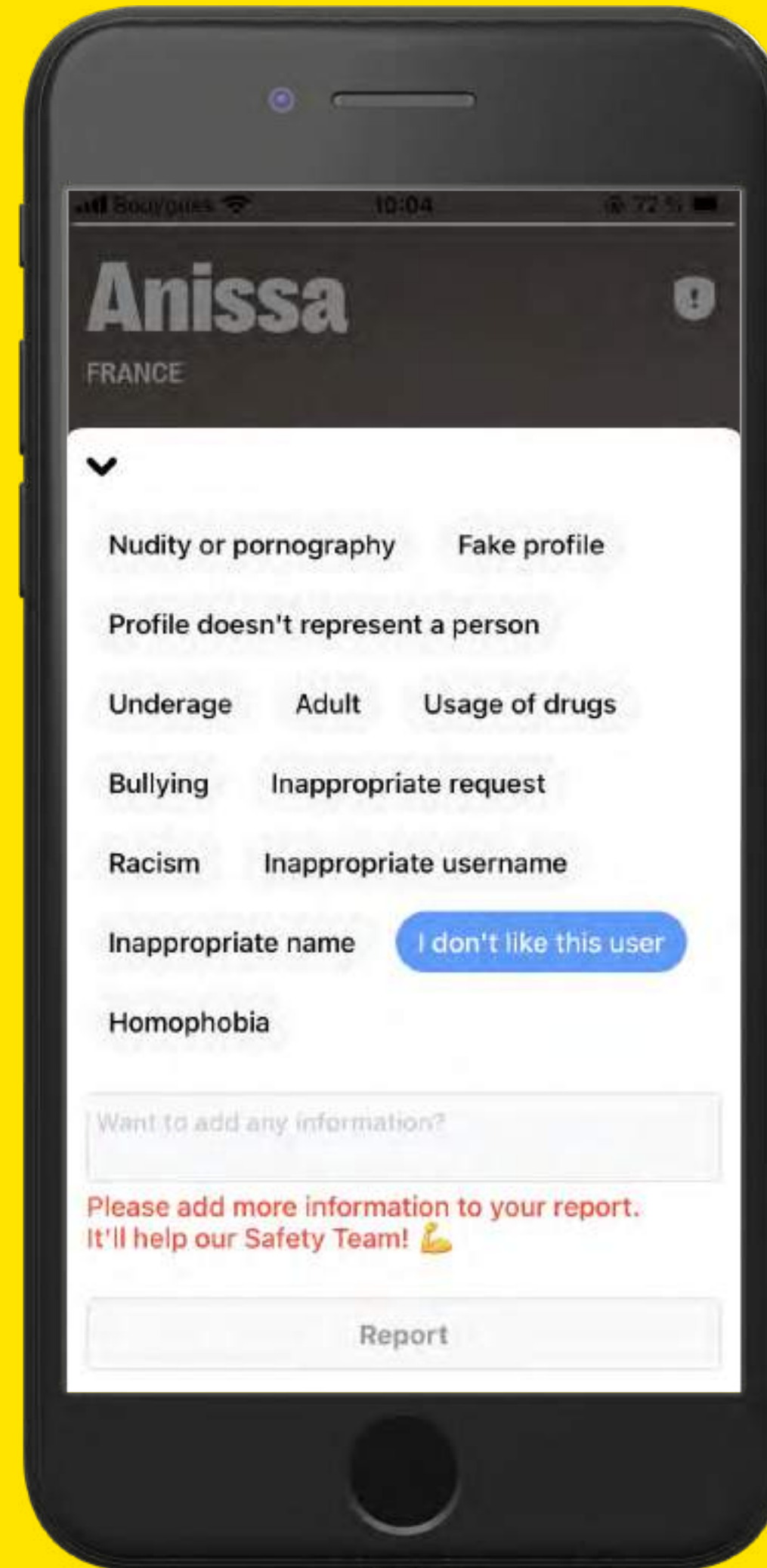
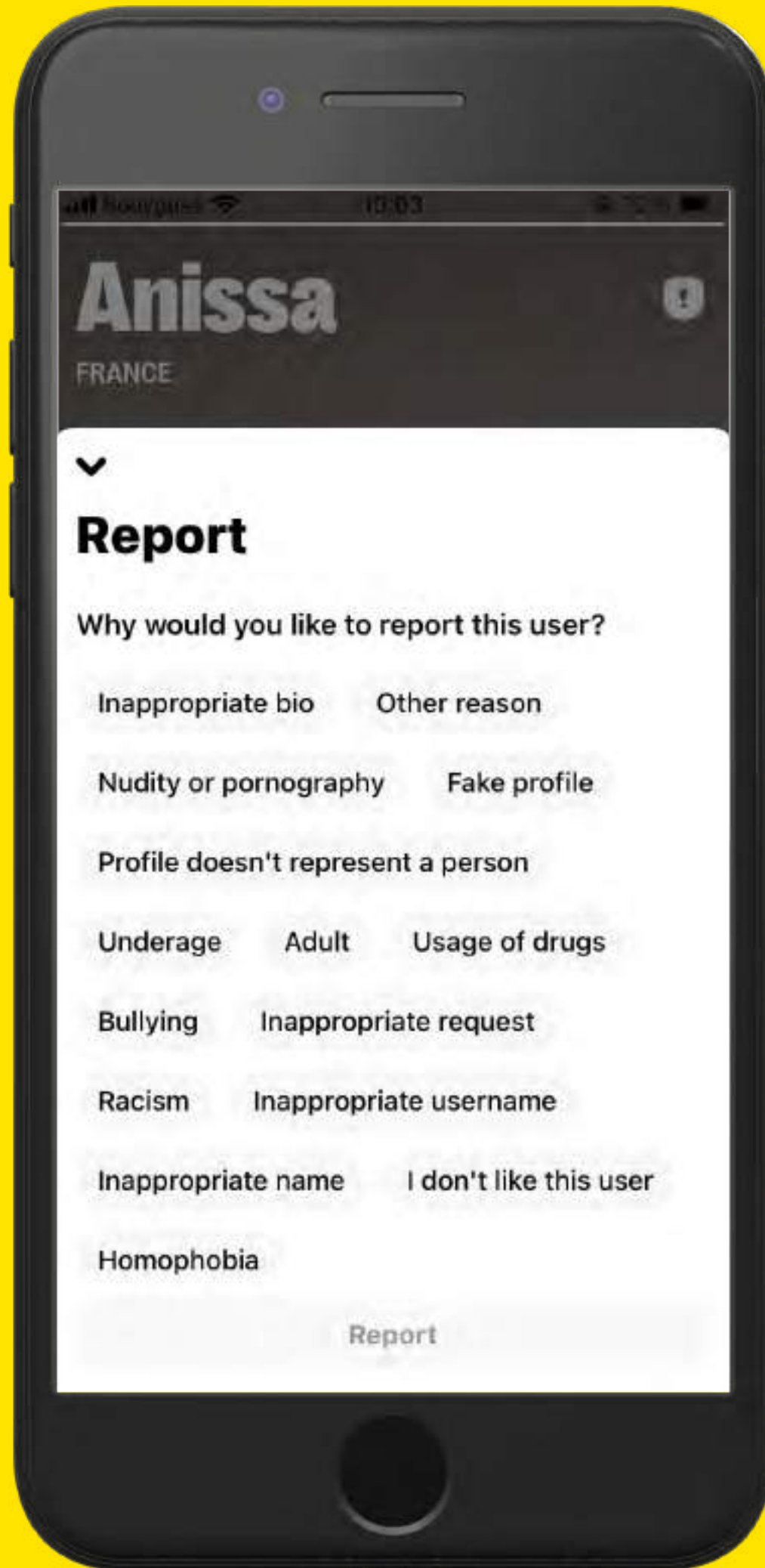
Show total

Hector (@hectoramk)
Male - 16 years old
Dordrecht, Netherlands
a few seconds streams / 2 hours watch
Registered 5 days ago

Yubo x Yoti



Our Reporting Process



Real Time Intervention



Detection

Our algorithms automatically close live streams with a high certainty of nudity. We send push notification to our moderation team in other cases.



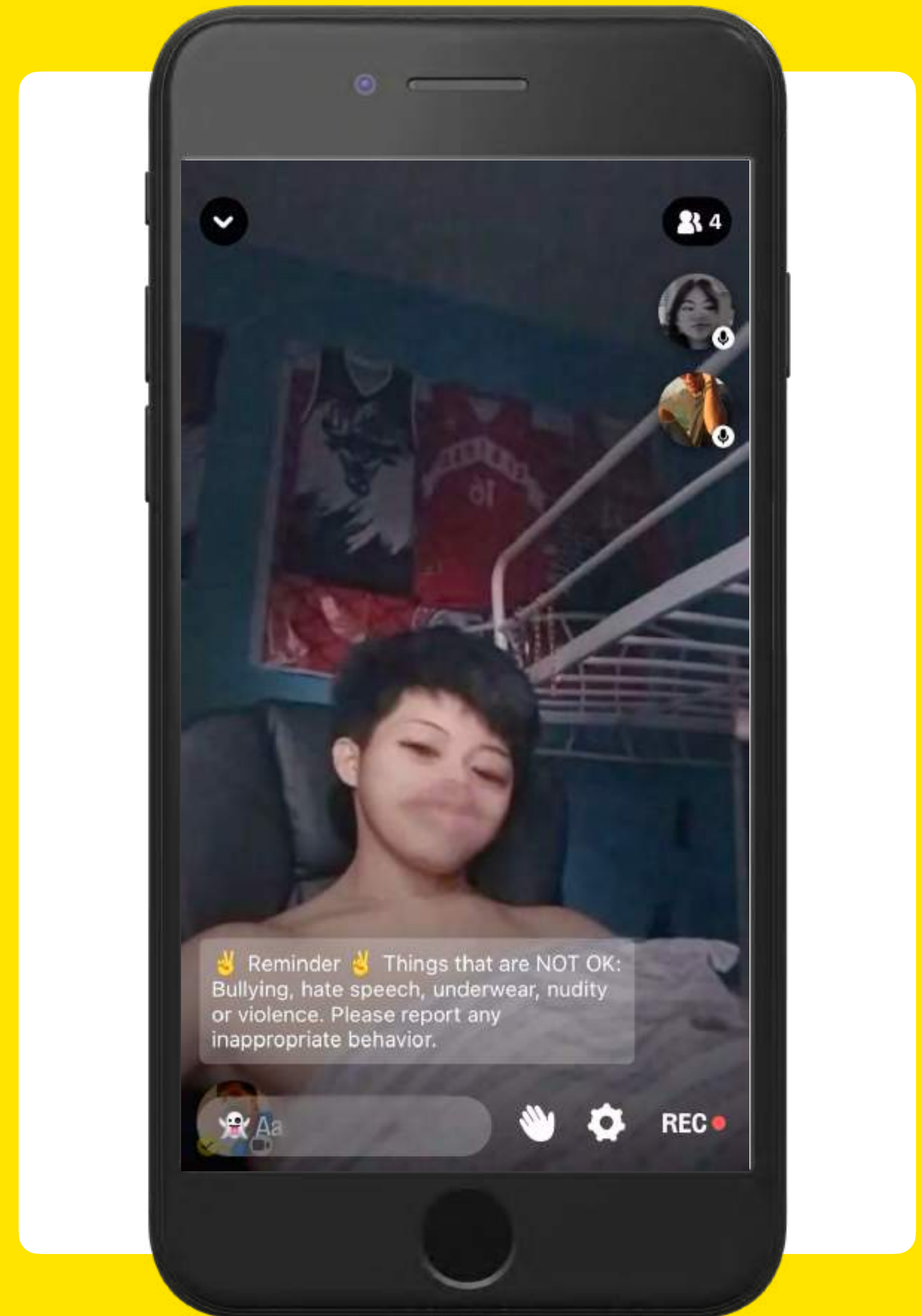
Intervention

We send alerts to users if they break our rules. More than 95% of users comply when getting alerts and learn our community guidelines.



Action

Our team works 24/7 to send alerts to users and close lives. We can block devices definitely or temporary.



Suicide Alerts



Detection

Our team check all « Suicide » reports from our community and our algorithms are able to detect blood (self-harm).



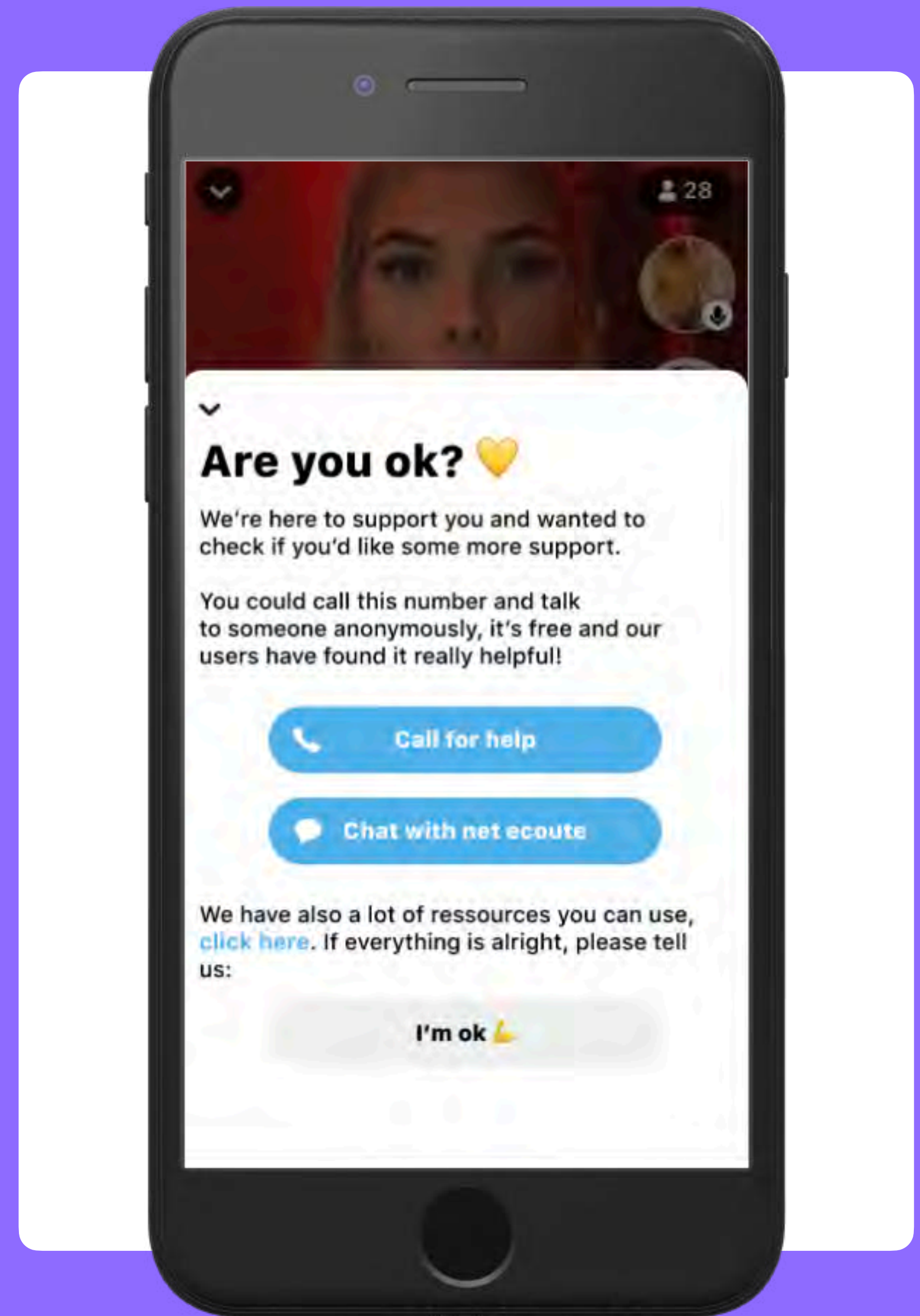
Prevention

When we detect a suicidal or depressed user, we send them a popup with a message and a local phone number to call



Action

We also send the user data to local enforcement so they can make sure the user is safe and that parents are aware of the situation



Private Conversations



Risks

Private messaging can be useful but also risky for communications as it is a less controlled area.



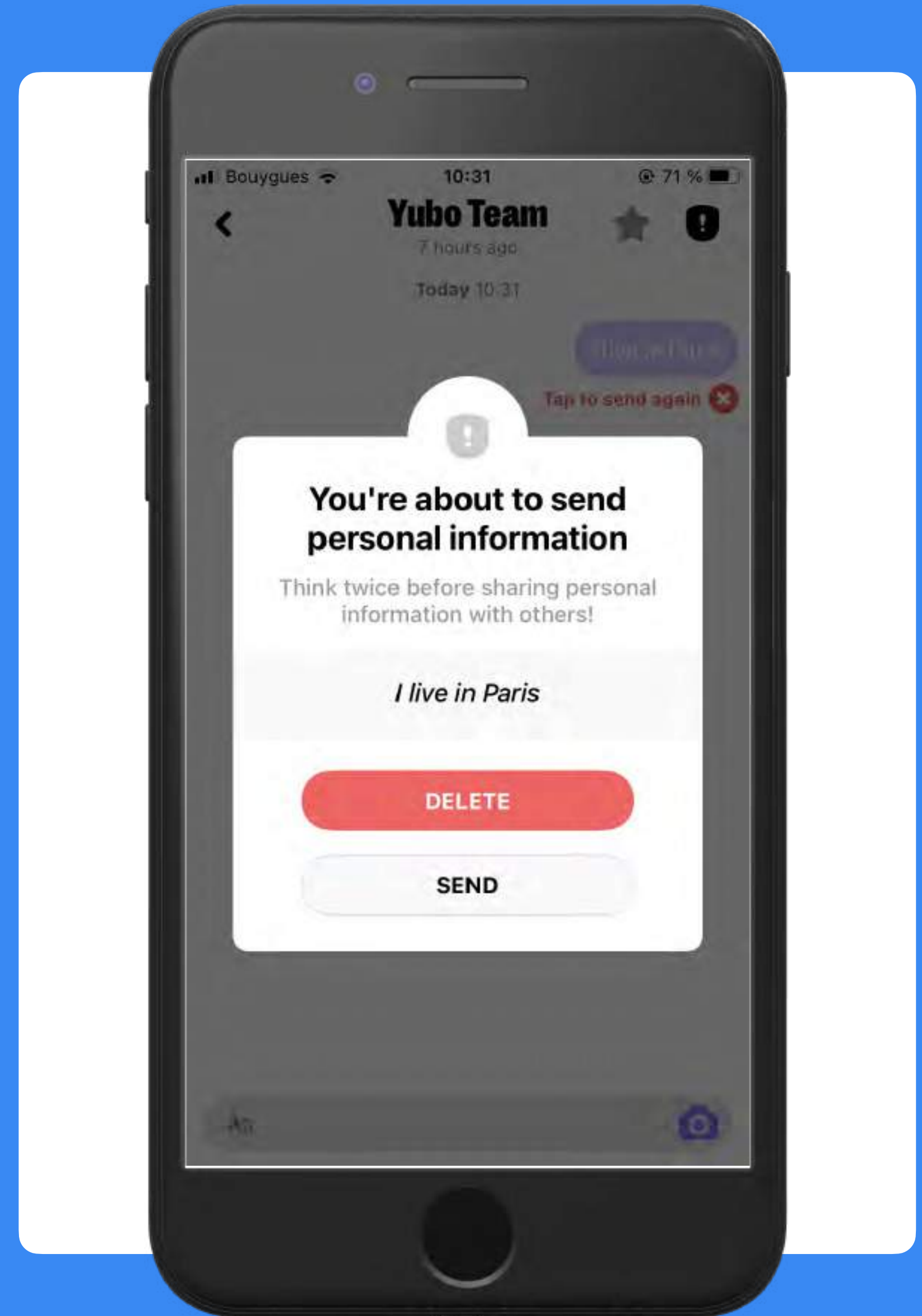
Compliance

We need to find a way to deal with user privacy and safety.



Personalization

We are thinking about solutions to help users personalize their messaging experience in a safe way.

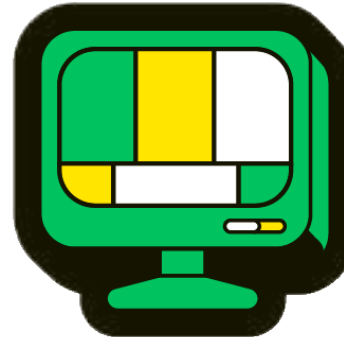


Law Enforcement



Process

We created a dedicated law enforcement guide available in the safety Center as part of our process. As we are based in France, you'll need to send request to French law enforcement.



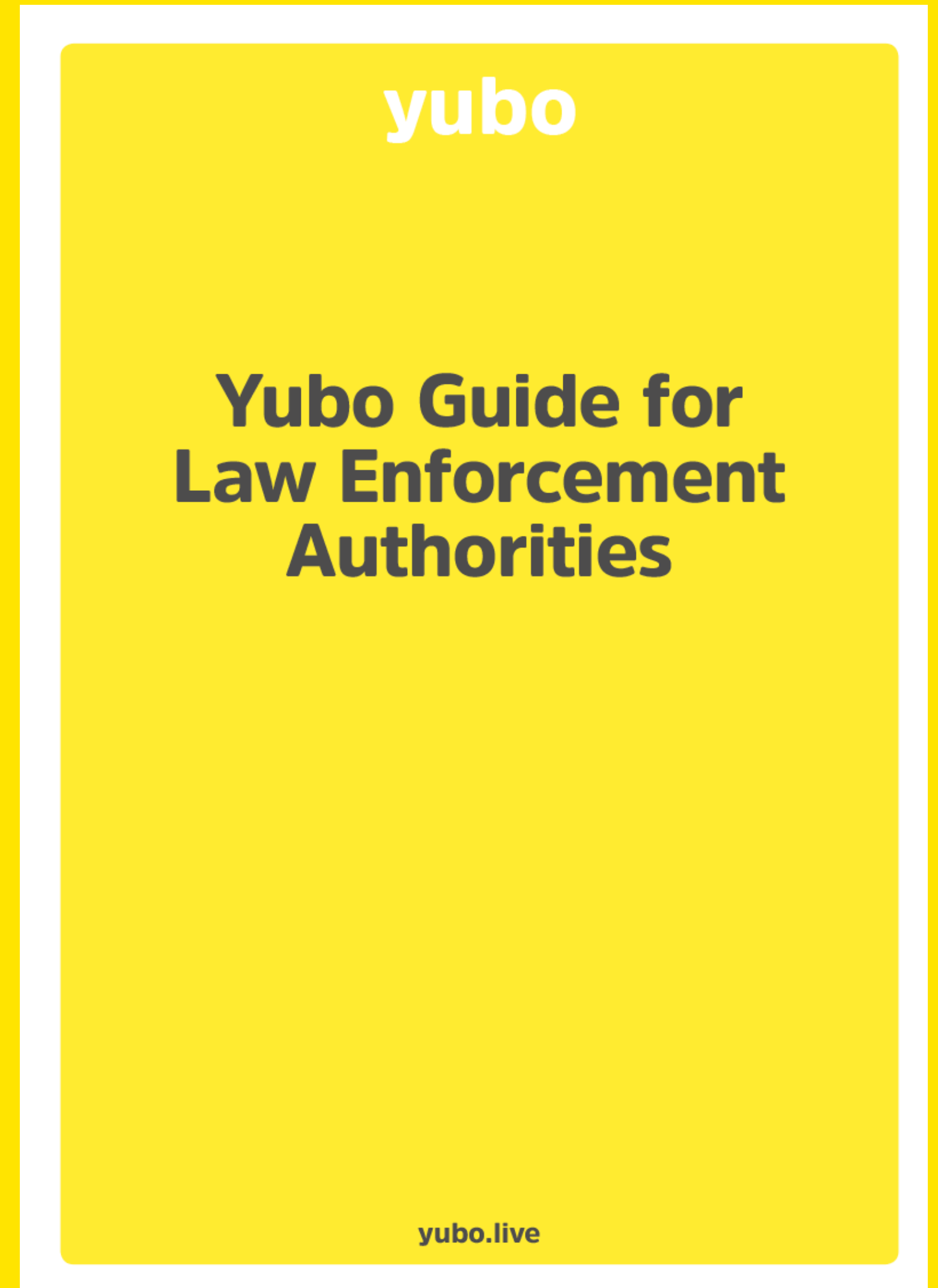
Data

We send a HTML page where you can navigate among profiles, messages and content created by the user. Despite requests, we also send regular reports.

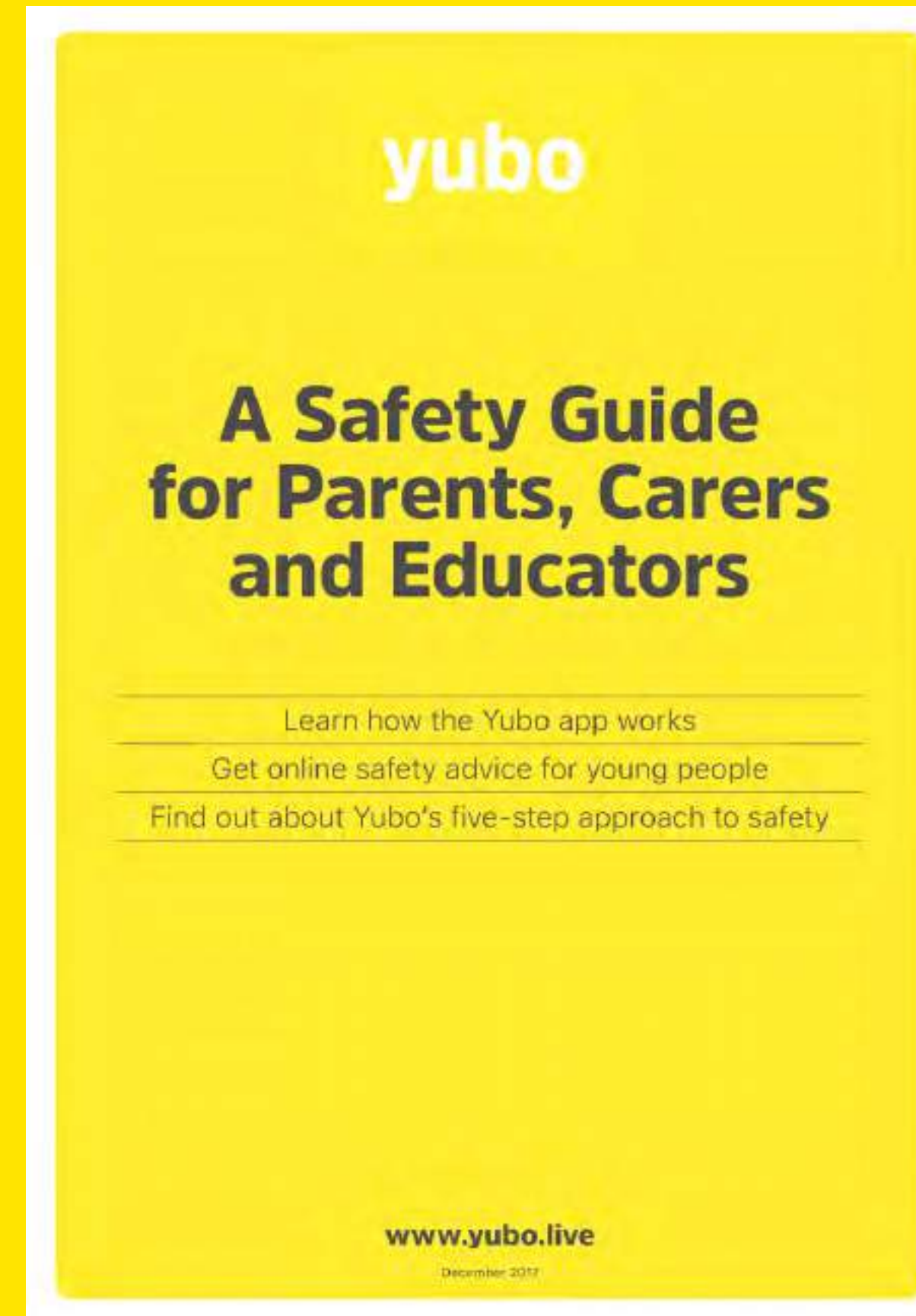
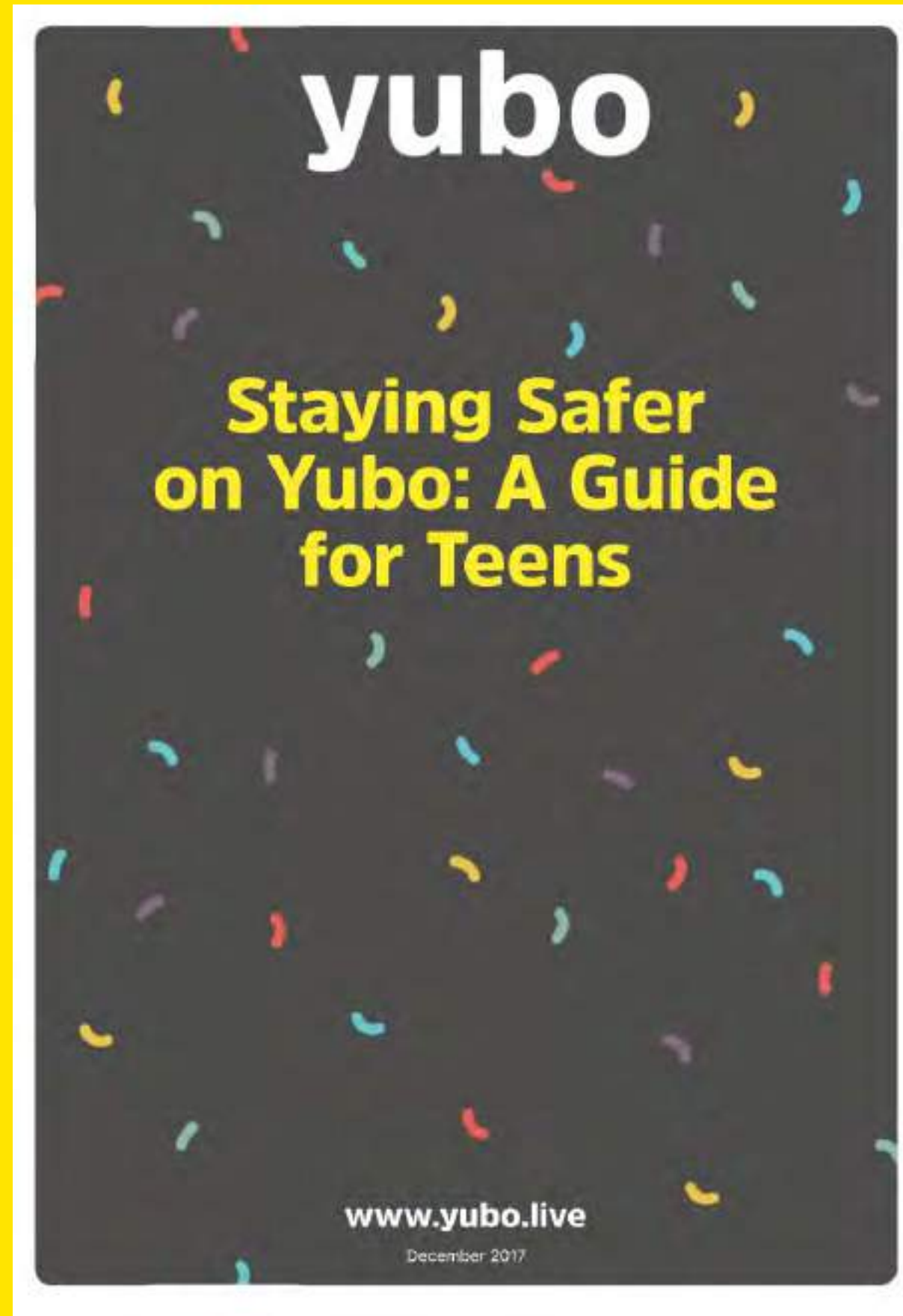


Emergency situations

We can directly send data to law enforcement if necessary.



Education and Resources



A Dedicated Safety Center

Advice

Overview

Yubo is a great place for young people to connect with friends but, as with any social network, you might have some concerns about your child or students using it.

Our community guidelines outline what is acceptable and what is not on Yubo and we encourage our users, their families and other concerned adults to contact us about anything inappropriate or illegal they see on the app so that we can take action.

We respond quickly when we are aware of any rules being broken and we report potentially illegal activity, such as Fake profiles & scams, to law enforcement.

In this section, we look at some of the main concerns you might have, provide online safety advice and explain what we're doing to help protect young people on Yubo.

| | | |
|---|--|--|
| Bullying Bullying can be really hurtful, whether it's online or in the real world. | Inappropriate language Young people often test boundaries with the language they use on the internet – do you understand their online code? | Sexual exploitation Unfortunately, some people target children and teenagers online for sexual purposes. |
| Nudity Flirting, exploring sexual feelings and having relationships – a natural part of growing up – increasingly happens online. | Fake profiles & scams On social networks, people sometimes pretend to be someone they are not – how can you spot a fake and avoid a catfish? | Self-harm & Suicide Sadly, the pressures of growing up sometimes lead to depression and suicidal thoughts. |
| Harmful & illegal content Violence, drugs, racism, homophobia, violent extremism... | Self-esteem Some teenagers experience body shaming, bullying and other forms | Peer pressure From just-for-fun pranks to more serious dares, teens often find |

We Work with the Best Safety Experts

In addition to the technical resources that guarantee our users' safety, Yubo is supported by a **Safety board made up of the best international experts.** The Board meets 3 times a year and goes through the program of product features and reviews and the safeguards in place.



Annie Mullins
Head of Trust & Safety
Ex Yahoo, Vodafone

Aol. YAHOO! vodafone



Anne Collier
Journalist & Safety expert
Safety advisor board @Snapchat, @Twitter, @Facebook

Twitter Instagram Facebook




Alex Holmes
Safety advisor @Twitter
Deputy CEO Diana Awards

Twitter



Michael Moran
Ex-interpol
Crimes against children unit expert in cybercrime

INTERPOL



Travis Bright
Safety Engineer
Director of product @Thorn. Formerly safety engineer @facebook and @google

THORN



John Shehan
Expert
VP of the Exploited Children Division @NCMEC

MISSING & EXPLOITED CHILDREN



Dr Richard Graham
Teen Psychiatrist
Nightingale Hospital

Nightingale Hospital

Comments and feedback are always welcome!

lucien.grandval@yubo.live

