A report from
cdt | Research

*Plain Language*

# Real Time Threats

## Livestreaming Platforms and how they try to stop Child Sexual Exploitation and Abuse (CSEA)

Robert Gorwa
Dhanaraj Thakur

March 2025

CENTER FOR
DEMOCRACY
& TECHNOLOGY

The Center for Democracy & Technology (CDT) is the leading nonpartisan, nonprofit organization fighting to advance civil rights and civil liberties in the digital age. We shape technology policy, governance, and design with a focus on equity and democratic values. Established in 1994, CDT has been a trusted advocate for digital rights since the earliest days of the internet. The organization is headquartered in Washington, D.C. and has a Europe Office in Brussels, Belgium.

**ROBERT GORWA**

Research Fellow at the WZB Berlin Social Science Center and Non-Resident Fellow at the Center for Democracy & Technology.

**DHANARAJ THAKUR**

Research Director at the Center for Democracy & Technology.

A report from

cdt | Research

# Real Time Threats

## Livestreaming Platforms and how they try to stop Child Sexual Exploitation and Abuse (CSEA)

**Written by Robert Gorwa and Dhanaraj Thakur**

*Plain Language Version by* ***Reid Caplan***

**OTHER PEOPLE WHO HELPED WITH THIS REPORT:**

Samir Jain, Gabriel Nicholas, Aliya Bhatia, Kate Ruane, Mallory Knodel, Silvia Lorenzo Perez, and Drew Courtney.

**THANK YOUS:**

We thank Riana Pfefferkorn, Jen Persson and David Thiel for looking over and giving feedback on this report. We also thank the participants in our June 2024 workshop who gave feedback on our research. Some of the people who helped us were Jamal Magby, DeVan L. Hankerson, Ozzie Oguine, Noor Waheed, and Saanvi Arora. Finally, thanks to everyone we interviewed for this report. You all had important experiences to share about this topic. This report comes from the Center for Democracy & Technology. What we say does not mean everyone we talked to for the report believes the same things we do.

Cover design and layout by Gabriel Hongsdusit.
Art direction by Timothy Hoagland.

We also want to thank Safe Online for covering the costs to write this report. What we talk about in this report comes from the Center for Democracy & Technology. That means Safe Online may not believe in the same things we do.

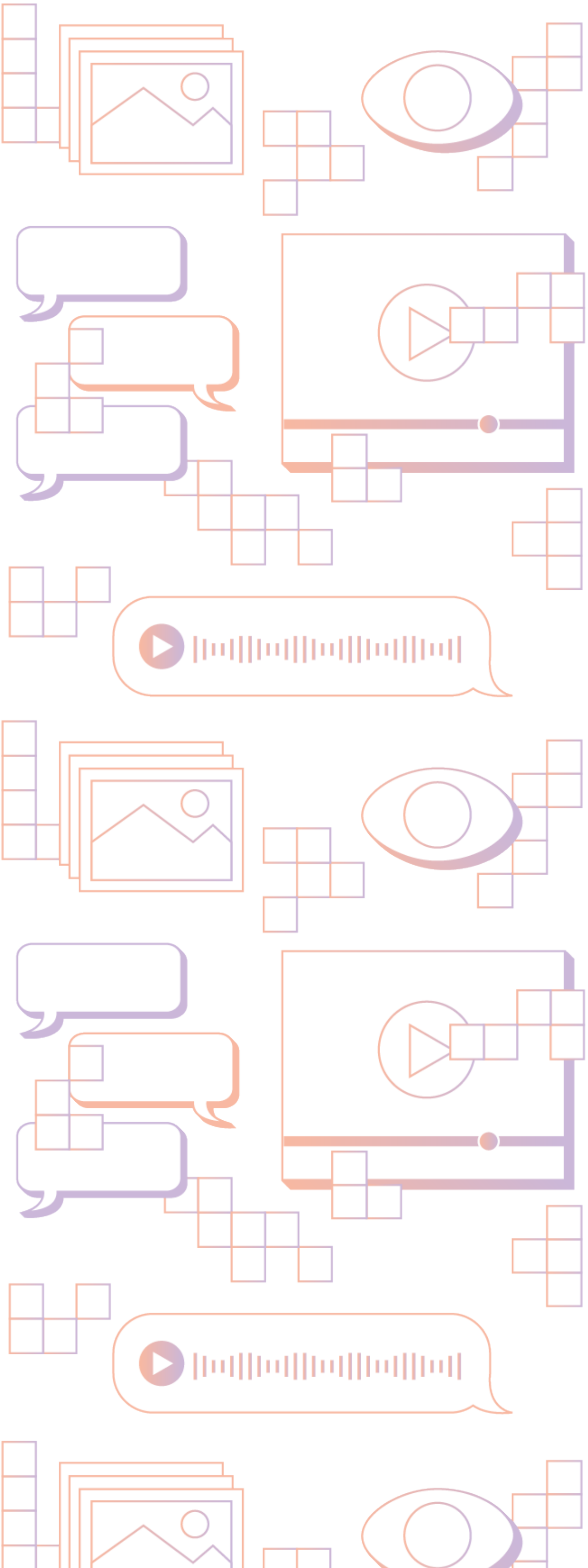
Safe Online

**When sharing the report, please use this information::** Gorwa, R. and Thakur, D. (2024) *Real Time Threats: Analysis of Trust and Safety Practices for Child Sexual Exploitation and Abuse (CSEA) Prevention on Livestreaming Platforms*. Center for Democracy & Technology. https://cdt.org/insights/real-time-threats-analysis-of-trust-and-safety-practices-for-child-sexual-exploitation-and-abuse-csea-prevention-on-livestreaming-platforms/

References in this report include original links as well as links archived and shortened by the Perma.cc service. The Perma.cc links also contain information on the date of retrieval and archive.

# Contents

# Words to Know



## Child sexual abuse material (CSAM)

Sexual pictures, text, and videos of children.

## Child sexual exploitation and abuse (CSEA)

When an adult hurts a child in a sexual way. This can also mean taking and sharing photos and videos of children doing sexual things.

## Grooming

When an adult tries to make a child trust them, with the goal of getting the child to do sexual things.

## Livestreaming services

Programs or websites that let people broadcast themselves live online.

## Predictive modeling

When a computer guesses what is in a picture or video based on what the computer has seen before.

## Prostitution

When someone pays to have sex with someone else. When this happens to children, an adult usually pays another adult to have sex with a child.

## Risk score

A way livestreaming services sort and figure out which accounts might be dangerous.

## Sextortion

When someone tries to blackmail a child to get the child to have sex with them, or threathens to share sexual pictures, text, and videos of that child.

## Sex trafficking

Sex trafficking is when someone makes a business out of forcing people to do sexual things.

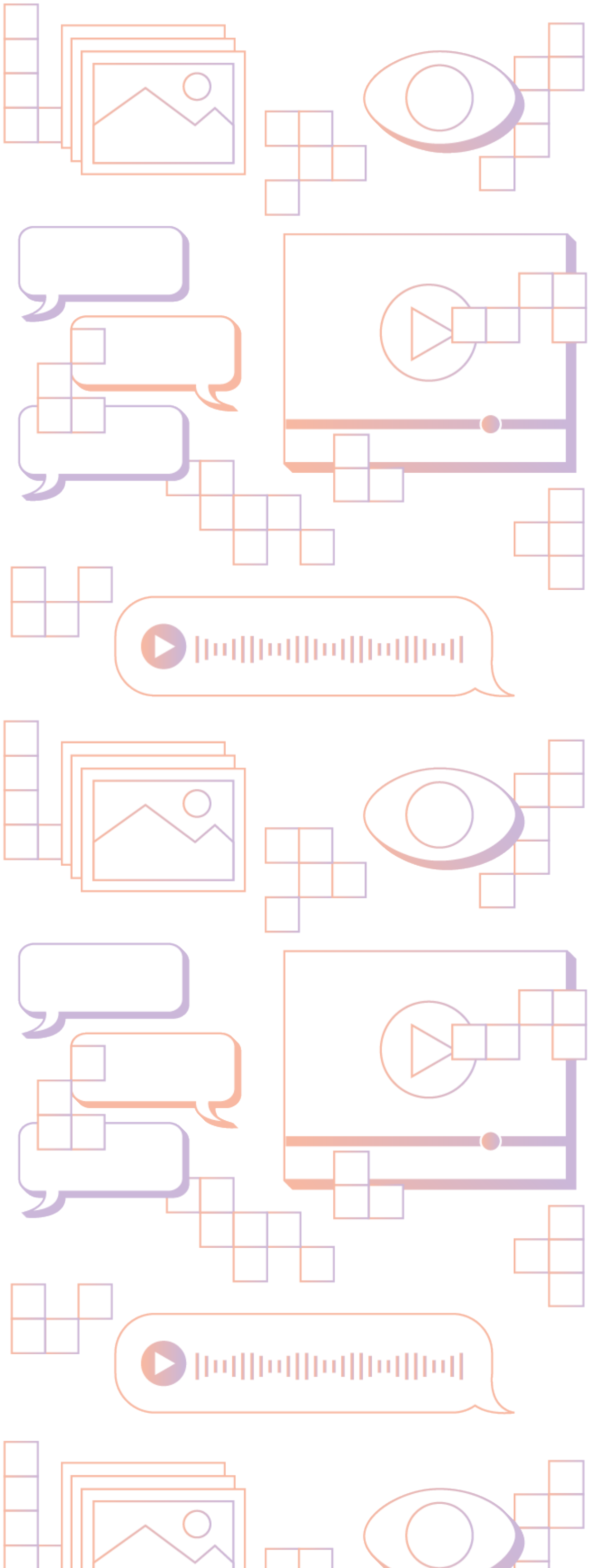
## Trust and safety practices

Online rules and tools that try to stop people from breaking rules or laws.

## Trust and safety tools

Online tools that try to make the internet safer.

# About this Report



Lately, there are lots of new kinds of websites and computer programs being made. One of these kinds of programs are called livestreaming services. **Livestreaming services** are programs or websites that let people broadcast themselves live online. Livestreaming can happen as a video, or as just sound, like a phone call. Livestreaming services let people show others around the world what they are doing.

For example: Someone who does livestreams is called a livestreamer. Suzie is a livestreamer. She does a “get ready with me” livestream. This is a popular kind of livestream. Suzie livestreams herself choosing an outfit and putting on makeup. While she does, Suzie talks to the people watching the livestream. People watching the stream can also type questions for Suzie. Then, Suzie can answer with her voice while she gets ready.

But livestreaming services can also be dangerous. Some people livestream things that are against the law or that hurt people. Some livestreamers tell people to hurt themselves or do violent things. In this report, we talk about a specific dangerous thing that happens on livestreaming services. We will talk about child sexual exploitation and abuse.

**Child sexual exploitation and abuse** is when an adult hurts a child in a sexual way. We call child sexual exploitation and abuse “**CSEA**” for short. Some kinds of CSEA are:

- Having or taking pictures or videos of children who are naked and/or having sex.
- Giving others pictures or videos of children naked and/or having sex.
- Forcing a child to have sex or do sexual things.
- Grooming a child. **Grooming** is when an adult tries to make a child trust them, with the goal of getting the child to do sexual things.
- Sextortion. **Sextortion** is when someone tries to blackmail a child to get the child to have sex with them. For example, an adult might threaten to tell a child’s secrets to their parents if they don’t have sex.
- Child prostitution. **Prostitution** is when someone pays to have sex with someone else. When this happens to children,

an adult usually pays another adult to have sex with a child.

To try and stop CSEA from happening in livestreaming, livestreaming businesses use certain tools. These tools are called **trust and safety tools**. This report looks at the kinds of trust and safety tools livestreaming businesses use. We look at how these tools work to make livestreaming safer and stop CSEA.

It is hard to stop bad things from happening in livestreams. Most online businesses have ways to find and get rid of dangerous pictures or videos. They have computer programs that can compare pictures or videos.

For example: Alton puts a video on Freevideos.com that shows people fighting. Fighting videos are not allowed on Freevideos.com. Freevideos.com deletes any videos that go against this rule.

Freevideos.com has a trust and safety tool to help with this. It is a computer program that checks everything that gets put on the website. The program can check if any new things on the website look like things that were deleted before. The program sees that Alton's video looks like another fighting video that got deleted for breaking the rules. So Freevideos.com deletes Alton's video from their website.

But because livestreams happen live, this gets a lot harder to do. Everything a livestreamer makes is new and happens on-the-spot. That makes it harder to compare to other videos to see if there's anything dangerous in the livestream. That means livestreaming businesses need to use less helpful tools. They might need to listen to the livestream to hear what is happening. They might need to read a report of what got said in the live stream. Or, they might need to look at livestreamer's accounts to see who might do dangerous things.

We looked at everything we could find about livestreaming services. We also talked to people who work in livestreaming businesses or who study livestreaming. We found out that livestreaming businesses are trying to stop CSEA in 3 ways:

- **Looking at how the livestreaming service gets designed.** These are steps that happen before a livestreamer can start livestreaming. These tools make it harder for livestreamers to start before proving they are safe. For example, some services won't let new people livestream until they get enough followers on their account. This makes it so a new person can't keep making accounts to livestream CSEA.
- **Looking at what happens in livestreams.** People or computer programs can look at livestreams to find and stop CSEA. For example, people can take screenshots of a livestream and see if they match other livestreams that showed CSEA. Some computer programs can get "trained" to tell if a livestream has CSEA in it.

- **Looking at livestreamer accounts.** Livestreaming businesses can look at people's accounts to see who might livestream CSEA. They can put a "flag" on these account to look into them more. Livestreaming services can also spread the word to stop dangerous people from using a different service.

  For example: Paul tried to livestream CSEA on a livestreaming service called S. S found out and banned Paul's account. S also made sure nobody in Paul's home could sign up for S again. Lastly, S let other livestreaming services know about Paul. That way, Paul could not sign up for another livestreaming service.

It is still very hard to check for CSEA in livestreams. Right now, livestreaming businesses have to focus on finding and reporting CSEA. They do not have all the tools they need to stop CSEA before it happens. But livestreaming businesses hope that one day, they will be able to stop CSEA before it happens. Doing this will take a lot of new trust and safety tools. That's why the ways livestreaming businesses work to stop CSEA change all the time.

The way livestreaming businesses work on stopping CSEA has some problems. First, livestreaming business don't always explain how their trust and safety tools work. They do this to stop people from being able to break these tools and keep livestreaming CSEA. But it's important that people like policy-makers and victims of CSEA know how these tools work.

Second, it is almost impossible to tell how well trust and safety tools work. Some tools could end up banning livestreamers who were not doing anything wrong. We also don't know if there are certain kinds of CSEA that these tools might leave out.

Third, trust and safety tools can take away people's privacy. They can collect a lot of information about someone's life. Also, banning a livestreamer who was doing nothing wrong takes their right to free speech away. If the tools to stop CSEA make livestreaming too hard for everyday people to use, these websites could become a less safe place.

To help fix these problems, we talk about 4 ways to make things better:

1. **Livestreaming businesses need to show how their trust and safety tools work.** This will help make these tools better at stopping CSEA. Right now, there are no ways for businesses to test and compare how their tools work. Policy-makers and researchers need to know how well these tools work and how much these tools can do on their own.

2. **Livestreaming businesses should be honest about how well computer programs can find and stop CSEA**. Using computer programs to stop CSEA has become a popular topic recently. But there is only so much computer programs can do to stop CSEA. Livestreaming businesses should make sure people are always part of doing this work. People are better at understanding all the information needed to make a decision about dangerous things in livestreams.
3. **Livestreaming businesses should design livestreaming websites that help people protect themselves, especially children.** Livestreamers should have what they need to protect themselves against bad people online. For example, people on livestreaming websites should have better tools to report dangerous livestreams or accounts.
4. **Livestreaming businesses should work with many people and groups to stop CSEA on livestreaming.** Groups like the Tech Coalition are making guidelines for livestreaming businesses. But groups like child safety organizations and human rights groups should also be part of these conversations.

CSEA is a big problem that affects children, their families, and communities. It's very important that livestreaming businesses try to stop CSEA from happening on their services. Lots of different groups are working on trust and safety tools. But if these tools are not designed the right way, they won't be helpful. That will make everyday people and policy-makers trust livestreaming services less.

Trust and safety tools will get better if livestreaming businesses explain how they work. They will also get better by using the experiences of lots of people to make better tools. We hope this can help new tools get made that stop CSEA while being a good thing for livestreamers too.

# To Start

In the past few years, livestreaming services have gotten very popular. Some livestreaming services are:

- Discord
- Instagram Live
- Youtube Live
- Tiktok Live
- Twitch

Livestreaming services let people share videos as they make them instead of when they are finished.

For example: Johanne is making a Youtube video. They start by recording the whole video first. Then, they put the video on Youtube. After that, other people can watch Johanne's video. The people watching can leave comments, but they cannot talk directly to Johanne.

Later, Johanne does a livestream using Youtube Live. People can watch the stream in real time. They can talk directly to Johanne.

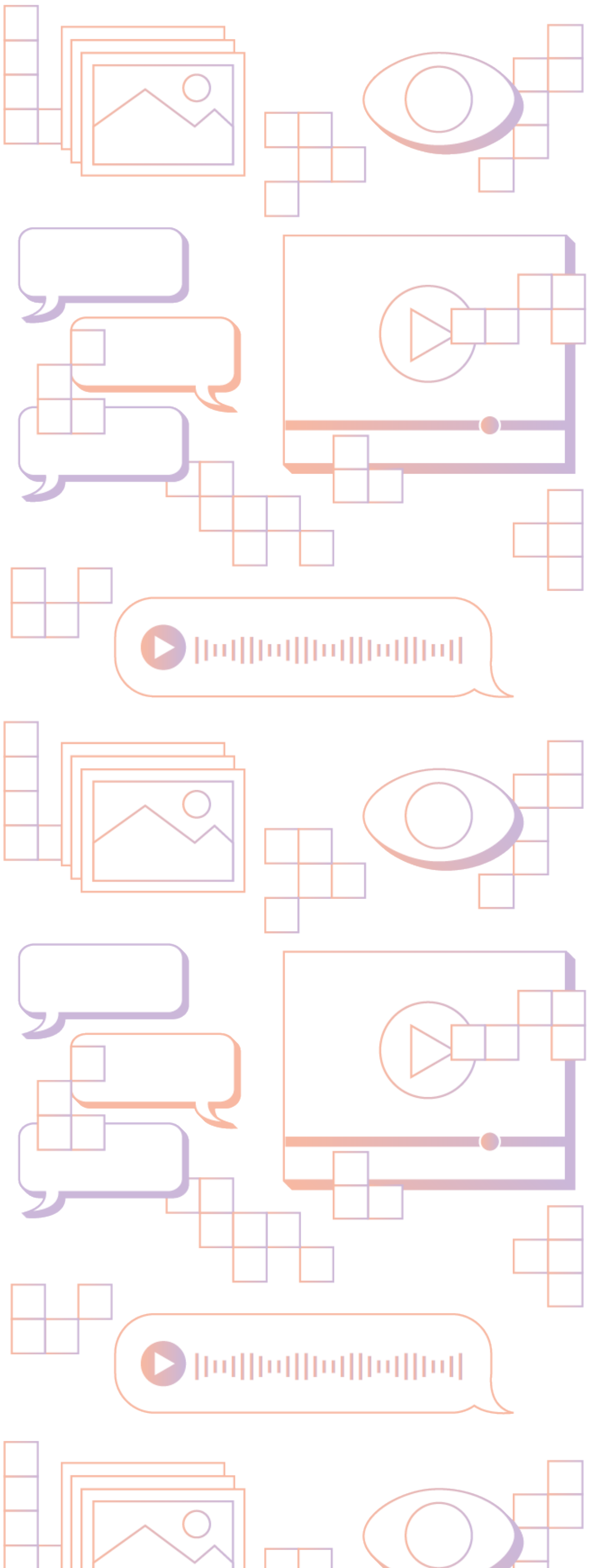
Livestreamers can share what they're doing to people around the world. Some livestreaming services are public, so anyone can watch. Others are private, so people can livestream to their friends or for work.

More and more young people are using livestreaming services. That's why people want to make sure that livestreaming services are safe. People like the police and parents are worried about CSEA happening on livestreams.

Livestreaming CSEA can look like:

- Showing pictures or videos of children naked and/or having sex
- Making a child undress or have sex on a livestream
- Child prostitution on a livestream
- Making a child do a sexual "performance" on a livestream

Sexual pictures and videos of children are called **child sexual abuse material**. We call this "**CSAM**" for short. Watching or making livestreams with CSAM is against the law. But many different livestreaming services get used to make or spread CSAM.



Websites that let people post pictures and videos already have guidelines to stop CSAM from getting spread. Groups like the WeProtect Global Alliance and the National Center for Missing and Exploited Children (NCMEC) helped make these guidelines. The biggest trust and safety tools these websites use are computer programs that can spot CSAM. These programs look at a database of CSAM from the government. They can check new pictures and videos to see if they "match" any in the database. If a new picture or video looks too much like one in the database, it gets marked as CSAM. Then, the website can delete the video and call the police if they need to. Matching lets businesses know if anything on their website looks like CSAM.

But livestreaming services make it harder to check for CSAM. Because livestreams happen in real time, they can't be "matched" in a database like regular videos can. This is a big problem because it stops the computer program businesses use to check for CSAM from working.

Livestreaming businesses have other tools they can use to find CSAM. But they are not as helpful as the matching tool. Instead, many businesses use a computer program called predictive modeling. **Predictive modeling** is when a computer guesses what is in a picture or video based on what the computer has seen before.

For example: Patty wants to make a predictive model to tell cats from dogs. She shows her computer program 1,000 pictures of cats and 1,000 pictures of dogs. This helps the program learn the difference between dogs and cats. Then, when Patty shows her computer program a new picture of a cat, the program can guess that the picture is a cat.

Livestreaming businesses can make predictive models to check for CSAM. But studies have shown that predictive models don't work well. They have even bigger problems when trying to sort through live video. And it would also take a lot of money and energy to run predictive models on a livestreaming services. Some livestreaming services have tens of thousands of people livestreaming at a time. That also slows down how fast the predictive modeling happens. The more people that livestream, the less well predictive modeling will work.

This report talks about how livestreaming businesses try to stop CSAE. We talk about what trust and safety tools businesses came up with to solve the new problems that livestreaming caused. We also talk about how these tools might affect everyday people and their privacy.

This report is written by the Center for Democracy & Technology (CDT). CDT studies how computer programs get used to check for dangerous things online. CDT looks at how these programs can help, but also might affect people's human rights and freedom.

This report will go through the trust and safety tools that livestreaming businesses made. There haven't been many studies about this topic. But lots of people use livestreaming services. And trust and safety tools affect how everyone does things online. There are more and more companies making trust and safety tools, too. So it's important to talk more about these trust and safety tools and what they do.

We will start this report by talking about research that has already been done about CSEA and livestreaming. Then, we will go through the ways livestreaming services try to stop CSEA. These ways are:

- Making it harder for people to livestream. This could make it harder for people to livestream CSEA.
- Looking at livestream videos and sound to find signs of CSEA.
- Looking for and stopping accounts that might livestream CSEA.

After that, the report will talk about how trust and safety tools might affect everyday people. We will also talk more about important things we've noticed about trust and safety tools. For example, more livestreaming businesses are focusing on livestreamer's accounts. They will look at an account's data to try and figure out if someone might livestream CSEA. We will talk about how these tools might affect people's privacy or human rights. We also talk about how livestreaming businesses could do a better job sharing information about their trust and safety tools.

## 1. Research on livestreaming and Child Sexual Exploitation and Abuse (CSEA)

The research on CSEA and livestreaming can be hard to understand. Some studies talk about different kinds of CSEA as if they were the same thing. Other studies talk about different websites, but not livestreaming services. Many everyday people know about livestreaming and that it might be dangerous for children. But there hasn't been a lot of research about livestreaming and CSEA. In 2024, there were only 8 peer-reviewed research papers about this topic. Peer-reviewed means the research was looked at by experts to make sure it was done right.

Most of these 8 research papers talked about how livestreaming gets used in sex trafficking. **Sex trafficking** is when someone makes a business out of forcing people to do sexual things. These papers talked about children from countries that have less money. Many of the papers talked about the Philippines.

These papers said that children end up forced to livestream sexual things to try and make money. These children want to make money to help their families, so they feel pressured to do sexual things online.

There is lots of research that's not peer reviewed about this kind of sex trafficking. Police have talked about how sexual livestreams of children get sold to people around the world. We also looked at 30 cases of online CSEA in the United Kingdom from 2013-2022. We found out that people figure out where livestreams showing CSEA are in many ways. These livestreams get posted on online dating websites, sexual websites for adults, and texting apps. Many non-profit groups work to try and stop this kind of sex trafficking.

Another kind of research about online CSEA focuses on grooming. This happens when a child does something live online, like playing a game or talking on a website. An adult will try to talk to the child online and gain the child's trust. Then, the adult will try to make the child do sexual things over video. The videos or pictures get made by the child, which is different than other kinds of CSEA. Online grooming has been talked about more in countries that have a lot of money. Grooming has not been talked about as much in countries that have less money.

When children make their own sexual pictures or videos, it is not always CSEA. Teens who are almost adults may choose to take these kinds of pictures or videos on their own. And it can be hard to know if a teenager makes a choice because of pressure or not. For example, one report found out that some teens sell sexual pictures on Instagram. Some teens might say it is their choice because they want to make more money. But if a teen's family doesn't have the money they need, they might feel forced to sell sexual photos.

We don't know how often children make their own sexual livestreams. But the news has started talking about CSEA and livestreaming more often. A news article from *Bloomberg* in 2022 looked at a video game livestreaming service called Twitch. They found out some young Twitch streamers were threatened and tricked into having sex. A chat app with livestreaming called Discord also got used in "sextortion" of children.

TikTok also looked into CSEA on its own app. They found out that children were taking their clothes off on TikTok Live in exchange for gifts.

The news also wrote about how some livestreaming services make it easier to share CSEA. Most people think livestreams disappear once they are done. But some livestreaming services let people save parts of a livestream. In 2024, *Bloomberg* reporters partnered with the Canadian Centre for Child Protection. They looked into Twitch's "Clips", short videos that livestream viewers chose of their favorite stream moments. Out of 1100 "clips", 83 had CSEA in them, which is a big portion of the clips. Sometimes, people will use these clips in new livestreams. This happens on other livestreaming services too.

*NBC* looked into Discord to find out more about CSEA happening there. They found out that many people worked together to try and make children do sexual things. For example:

- "Hunters" would find young girls and invite them into a Discord server.
- "Talkers" would chat with the girls and gain their trust.
- "Loopers" would pretend to be children by playing sexual videos of children over livestream. They would convince the real children to do sexual things with them.

Some news stories look at "trolling" livestreams. This is when someone, usually many people at once, go to a stream to cause trouble for a livestreamer. This happens a lot to livestreamers who are women or people of color. People can share CSAM in someone's livestream to try and get them in trouble or banned. Reporters at *404 Media* did a news story about "trolls" who share CSAM. They found out that this kind of trolling gets used by people to shut down Discord servers they don't like.

## GOALS OF THIS REPORT AND WHAT WE DID

For this report, we looked at research, livestreaming businesses, and the news. These different places used the word "livestreaming" to mean 7 different things:

- Social network websites with livestreaming built-in. These livestreams are meant for large groups to watch. For example, TikTok Live, Instagram Live, Facebook Live.
- Livestreaming services for video games that now also have other livestreams. For example, Twitch, Kick, Discord.
- Programs for livestreaming a big event, like a concert or show. For example, Clubhouse, Spotify Live.
- General video calling programs For example, Zoom, Teams, Skype, Jitsi, Webex.

- Text messaging apps that also let people make live group calls. For example, Facetime Video, WhatsApp, Signal, Telegram.
- "Random video chat" websites and apps. These match up two random people for a video call. For example, Shagle, ChatRandom, ChatHub.
- Websites or apps made for livestreaming adults doing sexual things. For example, StripChat, Chaturbate.

Research also talks about different kinds of CSEA that happen over livestream. These kinds of CSEA are:

- Sex trafficking over livestream. This research looks into people from countries with more money. People from these countries pay to see sexual livestreams from children. These children are usually from countries with less money.
- Showing non-live CSAM, like a picture or old video, over livestream. Or, sending CSAM to others through a live stream. For example, someone could link to CSAM in the chat of a livestream.
- Sexual pictures, videos, or livestreams that children make themselves. This usually happens in countries that have more money. These can get shared in livestreams, or in non-live places after a livestream. For example, a child might send a naked photo to someone they talked to in a livestream.

The 7 different types of livestreaming are very different. There isn't any research that talks about all of them at once. Also, there isn't good information about how much each different type of CSEA happens in livestreams. Some non-profit groups tried to look at how much CSEA happens in livestreams in the Phillipines. They found out that 1% of people under 18 could be part of sex trafficking over livestream. They also found out the people running the sex trafficking business were in Europe, the US, and the UK.

Sometimes livestreaming services will let everyday people know about CSEA on their livestreams. But they usually don't say if sexual livestreams were made by an adult or the child themselves. For example, Twitch said that from January to June 2023, they had 12,801 cases of CSEA that they took action to stop. But the way they measured also counted any livestreaming that could put a child in danger. So we can't know from Twitch's numbers how much actual CSEA happened there. Twitch also doesn't say how they "took action" in each of these cases.

Our goal in this report is not to find out how much CSEA happens in livestreams. We want to know what livestreaming businesses are doing to try and stop CSEA in livestreams. To do that, we started by looking at what these businesses have said about CSEA and livestreams. We looked at businesses who did all 7 different kinds of livestreaming. We looked the most at livestreaming services aimed at large groups of people. We also looked at how websites for adult doing sexual things keep children off their websites.

We looked at all kinds of materials from these businesses. We looked at their reports, white papers, new stories, and blogs. We also looked at work by businesses and groups like the Tech Coalition, who talk about online safety. We also did 15 interviews with leaders and experts in the livestreaming businesses. We asked these people to talk about trust and safety tools, and the problems they have keeping livestreaming safe. We used people's thoughts to help plan a half-day workshop hosted by CDT in June 2024. People from many different groups who care about livestreaming and CSEA came to the workshop.

This paper talks a lot about CSEA — Child Sexual Exploitation and Abuse. One part of CSEA is CSAM — child sexual abuse material/imagery. There are a lot of laws about CSAM. But CSEA talks about things that CSAM doesn't. For example, CSEA talks about grooming. Grooming itself is not against the law. But even though grooming isn't against the law, it is still dangerous. It can lead to an adult making a child doing things that are against the law. Our research showed that livestreaming businesses also try to stop grooming on their livestreaming services. They use trust and safety tools to try and stop grooming. We will talk about tools to try and stop grooming in this report.

## 2. All about trust and safety practices in livestreaming

Livestreaming businesses use many trust and safety tools to stop CSEA in livestreams. Most livestreaming services have rules that ban showing children doing sexual things. These rules get called things like "terms of service" or "community guidelines." These rules say that businesses can stop anyone trying to share CSEA.

Making rules like this is a part of online trust and safety practices. **Trust and safety practices** are the rules and tools websites use to stop people from breaking rules or laws. More and more technology businesses are making their own trust and safety practices.

Research about livestreaming services like Twitch show how hard it is for them to use trust and safety practices. One problem is that livestreams disappear as they get recorded. Unless someone saves a video of a livestream, people can't watch the livestream after it finishes. Sometimes livestreaming businesses save videos of each livestream. But other businesses don't have the money or tools they need to save that many livestream videos. This makes it harder for workers at livestreaming businesses to find and ban people who livestream CSEA. It also makes it harder to stop people who break the law on livestreams. Usually, the police need evidence that CSEA happened in a livestream before they arrest someone. If the video of the livestream is gone, this can make the police's job a lot harder.

Acting quickly when CSEA happens on livestreams is also hard to do. If CSEA happens on a livestream, it needs to get stopped before it gets spread to other places. But that is hard because everything on a livestream happens in real time. Livestreaming businesses may only have a few minutes or seconds to stop the livestream.

This can be a problem even in livestreams about normal topics. For example, big chess matches get livestreamed online a lot. If someone shared CSAM in the chat, thousands of people could see it. And some of those people could save the CSAM to share with others. It is harder to figure out who could have shared CSAM once the livestream ends. That's why livestreaming businesses need to act fast to stop this from happening.

There is lots of research about how livestreamers try to keep their own streams safe. They ask people they know to look over their livestream chat and ban people who break the rules. Livestreaming businesses should also think more about their own trust and safety practices. These practices need to happen quickly. Livestreaming businesses should also think about how the police or other groups can be a part of these practices.

One of the biggest problems with livestreams is that they are videos. It is very hard for computer programs to tell what happens in videos. It is much easier for them to understand words or pictures. There are some computer programs made to look at videos and tell what is going on. But these programs were made for regular videos, not for livestreams. That's why livestreaming businesses need to make new tools to keep livestreaming safe.

Right now, there isn't research that sums up trust and safety practices in livestreaming. We looked into many trust and safety tools that livestreaming businesses are making and using. Some of these tools are new, and got made just to find CSEA. Others are old tools used in a new way.

We split up these trust and safety tools into 3 types. We also list examples of trust and safety tools of each type. The 3 types are:

**Livestream design.** These tools affect who can livestream. They make it harder for people to start streaming. This could stop some people from sharing CSEA.

Some examples of livestream design tools are:

- Rules about what accounts can livestream. For example, someone might need 100 followers to livestream from a new account. Or, their account might need to be at least a week old to livestream.
- Making livestreamers prove their age. Some websites make people show photo ID to start livestreaming.
- Making livestreamers prove who they are. Some websites make people give out their phone number or credit card information before they can livestream.

**Checking what is in the livestream.** These computer programs look at livestreams and try to find CSEA.

Some examples of livestream checking tools are:

- Matching tools. These tools try to “match” pictures or videos with CSAM that already exists. The computer looks at someone’s profile picture, background, or other pictures or videos from an account. Matching tools can also look at livestreams by checking one short clip at a time. If anything “matches”, the computer knows it is CSAM and can take action to stop it.
- Predictive modeling. These computer programs look at different parts of a livestream. They check for clues that something might be CSEA. For example, a predictive model might check for how old the people in a livestream look. They might check if the people in the livestream are wearing clothes. They might also check what the people in the livestream are saying, and what the livestream chat says. The predictive model uses all this information to decide if a livestream has CSEA in it. If the predictive model finds CSEA, it can take action to stop it.

**Checking livestreamer’s accounts.** These tools check for signs an account might share CSEA.

Some examples of tools that check livestreamer accounts are:

- Account behavior indicator tools. These tools let a human or computer “flag” an account they think might share CSEA. This helps livestream businesses keep track of that account just in case. It also lets businesses share data with each other about bad accounts.
- User flags. This is when someone on a livestreaming service sees someone breaking a rule or law. They can “flag” the account for someone at the livestreaming businesses to check.

Next, we will talk in more detail about each of the 3 kinds of trust and safety tools.

## LIVESTREAM DESIGN: MAKING IT HARDER TO STREAM

There are already many ways businesses try and stop bad things happening online. Some websites use things like two factor authentication (2FA). 2FA means having to use both a password and your phone to log in to a website. 2FA helps stop people from making fake accounts to scam people. People who want to spread CSAM might do it less if they had to sign up for websites with their phone number. But people can also buy new phones and phone numbers to share CSAM.

Some livestreaming services have ways that livestreamers can protect their own livestream. For example, Twitch livestreamers can set "channel level verification requirements". This stops people from chatting who don't have their email or phone numbers on their accounts. Some livestreaming services let livestreamers block links in chat, or block certain words from being said.

Other livestreaming services check how popular someone is before they can stream. YouTube had a rule that livestreamer accounts needed to have at least 50 channel subscribers to livestream. This could stop someone from making a new account just to livestream CSEA.

Youtube also added extra rules about livestreaming. They said that livestreamers need to put in their phone number before they go live. Youtube also locked out certain parts of the website unless livestreamers shared more of their personal information. Tiktok has also done something like this. Anyone can use TikTok Live, but only accounts with more than 10,000 followers can use Live Studio.

### *Ways livestreaming services try to keep children safe*

Some livestreaming businesses are trying to stop children from signing up. TikTok's rules say that people under 18 can't livestream. Twitch's rules say livestreamers must be 13 or older. The U.S. has laws about children online, and the rules are different for children under 13. That is why many websites don't allow children under 13.

Some websites let people say what their age is without checking their ID. But people can lie about their age. Lately, more websites are using different ways to check people's ages. For example, Twitch uses a computer program to find and ban accounts of people under 13. Twitch also blocks the children who got banned from making new accounts. For accounts they are not sure about, they will make the account put in a phone number before livestreaming.

TikTok also uses a computer program to try and tell which accounts are under 18. If an account gets flagged by the program, that person must show their ID before they can livestream. They also need to take a photo of themselves with a code Tiktok gives them.

TikTok's said that from January-June 2024, they banned 21 million accounts of children under 13. They did not say how many of these accounts got banned because they tried to livestream.  We interviewed livestreaming businesses, and asked more about age-checking computer programs. We found out these programs look at what types of things each person on Tiktok watches, and who they talk to. This lets the program guess how old someone might be. But we still don't have a lot of details about how well this program actually works.

More and more livestreaming services are checking people's ages. This is similar to websites and apps for adults to do sexual things. U.S. law says that all actors in adult movies have to be over 18. Adult websites outside the U.S. also follow this law. For example, the website Stripchat is in Cyprus in Europe. They made a rule in summer 2024 that livestreamers needed to be over 18. Stripchat said that accounts had to show their ID before they could livestream.

We interviewed workers at one big adult website. They talked about how they make sure everyone on the website is an adult. They said humans are always in charge of checking people's IDs. Workers will also watch livestreams and look for dangerous things like violence and CSEA.

Checking people's ages and other information can help stop some CSAM from getting shared. But these tools can't stop all CSAM from being shared. Accounts can be hacked and taken over by dangerous people. People can make fake IDs, and computer programs that check people's information can make mistakes. Some people also worry that computer checking programs could take away people's privacy. If someone hacked into one of these programs, they could get information about a lot of people. Researchers once hacked into a program like this that TikTok used, just to show how dangerous it could be.

## CHECKING LIVESTREAMS TO STOP PEOPLE BREAKING RULES OR LAWS

It is hard for a computer program to tell what is happening in a livestream. But livestreaming businesses are still trying to make new programs to help with this. Some of these programs look for CSAM that already exists, but still gets shown in new livestreams. Other programs look at livestreams and try to guess whether or not they have CSAM. Both of these kinds of programs don't just get used on livestream video. They also get used on a livestream's chat box or the livestream's sound.

### *Finding CSAM that was already made: Video and Sound*

Livestreaming businesses haven't talked much about the tools they use to check for CSAM in livestreaming. For example, TikTok says it has its own tools, but also uses ones made by Microsoft and Google. Tiktok doesn't say what its own tools are or how it uses the tools from Microsoft and Google.

One tool that got talked about a lot was called "scene sensitive video hashing" (SSVH). SSVH is when a computer program scans through a livestream. The program guesses which parts of the video are most important. It turns those parts of the video into photos. Then, the program looks for CSAM just like it would for regular photos.

Another tool that livestreaming businesses use are computer programs that check sound. For example, there are apps that can tell what song is playing and let people know the title and artist. When the app hears a song, it checks a database of lots of music. When it finds a song that matches what it heard, it tells people the song.

Livestreaming businesses' tools work in the same way. Computer programs listen to the sound from a livestream. Then, the program checks a database of CSAM. If the sound from the livestream matches a sound from the CSAM, the program knows the livestream has CSAM. This tool can stop people from playing videos of CSAM on livestreams. These tools work fast and work well. But they cost a lot of money and energy. And the people we interviewed said that not many livestreamers play old CSAM on livestreams. So they felt like it might not be worth the cost for these tools.

The easiest thing for livestreaming businesses to do is "flag" pictures of videos of CSAM. People use pictures and videos in live streams all the time. Livestreaming businesses can put a "flag" on pictures or videos that break rules or laws. Computer programs that check livestreams will look for those flags. Then, if someone tries to share those pictures or videos again, they won't be able to. Or, they will get in trouble for doing so.

### *"New" Material: Video, Sound, Text*

When someone is livestreaming CSEA as it happens, that is called "new" material. It is much harder for livestreaming businesses to check for and stop new material of CSEA. Since the livestream is new video, it can't be compared in a database with old CSAM. That is why livestreaming businesses are trying to come up with new tools to solve this problem.

Many livestreaming businesses are using predictive modeling. This kind of computer program checks streams to guess if rules are being broken. For example, in 2023 Instagram talked about how they used predictive modeling to stop adults doing sexual things on livestreams. Predictive modeling could also get used to stop CSEA.

Some businesses make these kinds of tools just for livestreaming services. One big business that makes trust and safety tools explained their predictive modeling program to us. They said that the program will look at a part of a livestream. The program gives the livestream a "score" based on clues about how dangerous the livestream might be. If a livestream gets too high a score, the program can let someone know or do something about it.

Amazon has a trust and safety tool that works like this, called Rekognition. Rekognition is a predictive model that looks for naked or sexual pictures or videos. Businesses could combine this tool with tools that check people's ages. Then, they could figure out which pictures and videos are CSAM.

Some people are thinking about how to make predictive models specifically for CSEA. For example, a computer program could look at databases of CSAM from the government. Then, the program would be able to compare any new livestreams to the database. Any livestreams that looked close to what was in the database could get flagged as maybe having CSEA.

Some programs like this already get used online. They are called Thorn's 'Safer Predict' and SafetoNet's 'HarmBlock'. SafetoNet says that HarmBlock also works for livestreams. Groups working to protect kids like the idea of predictive models just for CSEA. They think it's a better idea than trying to guess how old people are in sexual videos or pictures. But we don't know if the CSAM databases show enough of a variety of children. That means the predictive model might miss certain children. We also don't know if it's fair to share these databases with private companies. Lastly, we still don't know how well these predictive models work.

Another idea for predictive models is to just check a livestream's sound instead of video. Some people we interviewed said that videos of CSAM had specific sounds in them. Computer programs could check the sound happening in livestreams. Livestreams that have the same kinds of sounds as CSAM could get flagged as maybe having CSEA.

But we still don't know how well these programs work, either. For example, many livestreaming services show people playing sexual video games. These games might use noises of people having sex. A predictive model might flag this kind of livestream as CSEA and ban the account, even if the game is allowed.

Instead of listening to a livestream's sound, many businesses use transcription. Transcription is when what someone says gets written into text. There are many computer programs that can transcribe speech from a livestream into text.

Transcription can be a helpful way to look for CSEA. Either workers or computer programs can read the transcription. They can look for key words that have to do with CSEA. If they find these words, they know to take a closer look at the account that said them. But not every livestreaming business uses transcription. Some businesses think it is too expensive to make and store all the transcripts.

Some problems might happen when a computer program looks at a transcript. Many livestreamers use slang or less common words. So the computer program might not understand how a word is being used. Or, people might speak in a language or accent the computer program doesn't understand. This could lead to computer programs banning accounts that weren't doing anything wrong. This is more likely to happen to people who speak different languages or are from different cultures. That is not fair. Livestreaming businesses need to make sure their tools are fair to everyone. They should think more about languages that not many people speak.

People on livestreaming services can "flag" people's accounts if they think the account broke a rule. People will flag accounts that do dangerous things like share CSEA. Using predictive models of sound or transcription can work well with the flagging that livestreamers already do. Livestreaming businesses can use predictive modeling on accounts that were flagged. That way, they can know which accounts might be the most dangerous. They can make sure to take care of the most important flags first.

One way livestreaming businesses could catch dangerous livestreamers is with "risk scores." A **risk score** is a way livestreaming services sort and figure out which accounts might be dangerous. Risk scores get made by looking at an account's:

- Messages in a livestream chat
- Video thumbnails (the picture shown before you click on a video)
- Video titles
- Video keywords
- Video and sound on a stream

If an account scores too high a risk score, it is a "problem account". That means a human worker has to look at the account and decide what to do. Right now, risk scores are not used in livestreaming services. But other online businesses say risk scores have worked well for them.

Some businesses are using transcript-checking tools to try and stop grooming. For example, Safer Predict says its program looks for CSEA in text. The program can stop people from using certain words, or ban people who use those words. Most businesses that make these kinds of computer programs don't explain how they work. It looks like these businesses come up with a list of keywords that might have to do with CSEA. But this could be a problem because people use words in lots of different ways. For example, adults might talk about ways they were hurt as children. It is not against the law to talk about this online. But computer programs might think they are talking about CSEA happening now. That could get these adults in trouble, even though they were doing nothing wrong.

One person we interviewed said computer programs that use keywords are still helpful. This person is a worker at a livestreaming business who looks at flags the computer program makes. They still have to deal with a lot of flags for accounts that did nothing wrong. But the computer program helps them find accounts that they would have missed otherwise.

## CHECKING LIVESTREAMER ACCOUNTS FOR SIGNALS

Livestreaming businesses have started looking at one more thing to try and stop CSEA. Instead of looking at specific videos or livestreams, they look at people's accounts. They look for "signals" that an account will do something dangerous or against the law. These kinds of trust and safety tools were not made for livestreams at first. But more and more livestreaming businesses use these tools in their trust and safety practices.

One way signals can be used is to share information about accounts that might share CSAM. Different livestreaming services can share this information with each other. This can stop dangerous people from making new livestreaming accounts to share CSAM. Another way signals get used is to guess which accounts might be sharing CSEA. Most big livestreaming businesses will use both of these kinds of signals.

Lantern is the name of a new project from the Tech Coalition. This project helps livestreaming services make signals to stop CSEA. Different livestreaming services can share signals with each other. Signals can have information like:

- Pictures or videos that shows an account has shared CSAM
- Someone's email address or username
- Keywords that show someone might be sharing CSAM

These signals can get taught to a computer program. Then, the program can flag or delete accounts that have signals on them.

For example, Meta has a signal program called ThreatExchange. This program lets people on Meta see signals and write their own. Meta has many different kinds of signals, like:

- Where someone lives;
- Their name;
- Their IP address (what computer or phone they are using);
- What someone looks at on the internet.

Livestreaming businesses don't have to share their signals with other businesses. But they can still use their own signals. A business could make a signal for anything they think is "strange."

For example: S makes its own signal program. It makes a signal for new accounts that get over 500 watchers in a livestream. That's because it is strange for a new account to have that many watchers. It could mean the account is sharing something dangerous like CSEA.

Livestreaming business can also make signals for things like:

- Livestreaming accounts that get many watchers coming from outside the website. This means they found the link to the livestream somewhere else.
- When a livestream only has new accounts talking in the chat.
- When a livestream account uses tools to hide who they are.

If an account has these signals, then someone who works for the livestreaming business looks at the account. That worker can decide if the account broke the rules, and what should happen next. Livestreaming businesses say signal tools already help workers make better choices about accounts. And signals can be used for more than just banning accounts. Livestreaming services can make it harder for accounts with a signal to livestream. They can make the livestream or chat slower, or kick people out of the livestream chat.

Using signals is still pretty new in livestreaming. There hasn't been a lot of research or news about it yet. But livestreaming businesses think signals are a good idea. They think signals are a good way to stop some people from livestreaming to share CSAM. They also think signals can be used more during streams to stop CSEA while it happens.

Today, big livestreaming businesses are thinking less about finding CSAM. They are thinking more about how to stop CSEA in livestreams before it happens. For example, instead of looking at pictures and videos of streams, livestreaming businesses will look at accounts. They will look at information from someone's account to check for connections to CSEA. If livestreaming businesses can figure out who will livestream CSEA, they can stop those people from sharing CSAM.

Some people we talked to thought signals could help protect some people's privacy. Certain trust and safety tools can take away people's privacy. For example, programs that collect someone's information, like their name and address. But signals can help livestreaming businesses know when to use these trust and safety tools. Livestreaming businesses can choose to only use these tools on accounts that have a signal. Signals also let a human worker know to check an account. This can help when computer programs make mistakes.

Because signals are so new, they have their own problems that need solving. People we interviewed said they were worried signals could get used without humans looking them over. That means computer programs would make choices about people's livestreaming accounts. This could be unfair to some livestreamers.

More research should happen about how signals affect people from different groups. For example, some people use tools online to protect their privacy. There is a chance livestreaming services might flag and ban these accounts. People might also get their accounts banned just because of where they live. Research should get done to see where this happens and how to stop it.

## 3. Problems with and effects of trust and safety practices

We talked about 3 ways livestreaming businesses try to stop CSEA. Each of these ways has its own problems. And each of these ways affect people's rights and lives on the internet.

### LIVESTREAM DESIGN CAN ONLY GO SO FAR

It is hard to tell how old someone is online. So it might be better for livestreaming businesses to check how popular someone is before they can livestream. This could stop a lot of people trying to make new accounts to spread CSAM.

But tools to check someone's popularity aren't perfect. Someone could hack into a popular account to livestream from it. Also, it can be hard to get enough people to follow a new livestreamer. That means new livestreamers might never actually get to livestream.

Another way livestream services try to stay safer is just asking people their age. Livestreaming websites could ask only livestreamers to share their age. Regular livestream watchers would not need to share their age. But people could still lie about their age.

Knowing for sure how old someone is can be hard to tell safely. Asking for someone's ID might not be enough. And asking for ID takes away people's privacy. Some trust and safety tools look at videos to try and tell people's ages. But these tools might not work for all kinds of children. For example, it might be harder for a computer program to figure out the age of someone with a disability.

Other ways that livestreaming services try to stop dangerous livestreamers might be unfair. For example, some livestreaming services make livestreamers give out their credit card information. This would leave out anyone who does not have a credit card. People in the U.S. also have the right to free speech. Trust and safety tools that ban large groups of people could be seen as stopping free speech.

## PROBLEMS WITH PREDICTIVE MODELING IN LIVESTREAMS

Right now, there are rules about reporting CSAM on livestreams. U.S. law says livestreaming businesses must tell NCMEC if they find CSAM. Livestreaming services can also use NCMEC's database to "match" CSAM they found on livestreams.

The problem is that most livestreams get made in real time. That means looking through CSAM databases won't always work. So livestreaming businesses need to find new ways to figure out what is CSEA on livestreams.

Making new trust and safety tools to find CSEA in livestreams comes with new problems. One of these tools we talked about earlier is predictive models. Predictive models may have "false positives". For example, the program might flag a livestream of kids swimming as CSEA, even though it isn't. This affects normal livestreamers who might get banned and not know why. In the worst case, the police could get called on a livestreamer who did nothing wrong.

False positives are a big problem since they happen a lot. They happen the most to people who are not men, and people of color. Computer programs also can have a hard time seeing what is happening in a livestream. The video might be dark or blurry, or the people in the video might be hard to see. Even predictive models that say they do a good job could have thousands of false positives a day.

CDT did research before about using computer programs to check things online, like predictive models. We found many problems with these programs. There was not a lot of information about how they worked. And they didn't work well in some situations.

We interviewed many people for this report. Most were honest that trust and safety tools are not perfect. They said computers will always be wrong sometimes. But people who sell these computer programs don't want to wait. They think they can make these programs better once they are already being used.

There are other things about trust and safety tools that livestreaming businesses need to think about. These things are important to how CSAM specifically gets handled in livestreams:

- Where do computer programs get the databases they use to learn about CSAM?
- Do people in these databases get asked if their information is okay to share?
- How can everyday people tell how well trust and safety tools work?
- How can we know how much is actually possible for predictive models to do?

A big problem is how to train computer programs to find "new" CSAM. Because CSAM is against the law, there aren't open databases to show a computer program. That makes it harder to train the program to do predictive modeling.

Some businesses work with the police to make tools to find CSAM. The police let the business use a database with CSAM that the police had. One example of this is a project called Project ARICA, which was paid for by the European Commission.

Computer programs using CSAM from the police usually only get used to help police. They use predictive modeling to try and find new cases of CSEA. But some of these computer programs get sold to livestreaming businesses. Businesses that sell these programs say their programs can tell how old someone else. They say the programs can tell when people are having sex on a livestream. They also say that their program does not have a lot of false positives.

But businesses selling these computer programs don't share information about these programs. They don't share how well these programs actually do in real life. They don't share how they made or tested their program. They don't compare their program to other programs. So we can't know if what businesses say about their programs is true or not.

The people we interviewed were worried about how well these kinds of computer programs could work. Even a predictive model that works 99% of the time will still have thousands of false positives. That could lead to people's accounts getting banned even if they did nothing wrong. And this would probably affect anything online that shows or talks about sex. People have the right to do sexual things online. Predictive modeling could make it harder for people to use that right.

Predictive modeling has a lot of problems. But CSEA is also a big problem. Some people we interviewed were worried businesses would stop trying to make tools to fight CSEA. If the tools are too hard to make, businesses might give up. Or, if the tools might affect people's rights or privacy, businesses might not want to deal with that. The people we talked to wanted businesses to keep trying.

But livestreaming businesses need to make sure they have human workers, too. Using computer programs to find and decide what is CSEA is not enough. People who make these computer problems might think the program never makes mistakes. But we know this is not true. That's why humans should always check the work computers do.

When training a computer program, it's important to use good information. For computer programs to do a good job finding CSAM, they need to understand what CSAM looks like. A good predictive model for CSAM would get trained on photos and videos from around the world. Then, the model would get trained on other photos or videos of people that were not CSAM. That way, the model could tell the difference between each kind of photo or video. But right now, we don't know of a computer program that does all this. So we can't say for sure whether or not predictive modeling can work well to find CSEA in livestreams.

A few people we interviewed talked about how hard it might be to get information about CSAM. Computer programs need CSAM information to work well. But there isn't an easy way to find the victims of CSEA in different databases. And victims of CSEA might not be okay with pictures or videos of them getting used for a computer program.

For example: a business made a computer program to predict CSAM. They used a database of 1,000 CSAM cases to train the program. Then, the business sells that program to others. But the business never asked the victims in the CSAM database if that was okay. And the victims don't get any of the money the business makes.

Livestreaming businesses talk about how hard it is to get databases of CSAM. Only some livestreaming businesses can work with CSAM databases from the police. Some businesses might get CSAM from databases that are against the law.

### PROBLEMS WITH SIGNALS: THEY TAKE AWAY PRIVACY AND PEOPLE LOSE THEIR RIGHTS

Experts on livestreaming and CSEA think signals are a good idea. Livestreaming businesses are just learning how to use signals. But signals can make it less likely someone's livestream is flagged as a "false positive" for CSEA. Signals also make it easier for human workers to know which accounts to look at.

But signals have their own problems. Signals might not get used in a way people can understand. It can also be hard to tell if they are working well or not.

For example, the Lantern project has a list of "keywords used to groom". But it can be hard to tell what is and isn't grooming. That means businesses using the keywords list might put a signal on livestreams that did nothing wrong.

Different livestreaming businesses might have trouble understanding how to use signals. Training a computer program to use signals might mean ending up with false positives. Training human workers or other livestreamers to use signals can also be difficult. That makes it harder for people to know if signals work well or not.

Using signals can affect someone's privacy. In Europe, human rights groups are already fighting to stop trust and safety tools that use signals. One member of government in Germany filed a lawsuit about this. He said that Facebook Messenger was using the "keywords used to groom" list to look at people's private messages.

The European Commission has talked a lot about stopping CSEA. They talked about tools they could use to stop grooming. But some people think it is too hard to stop grooming online. They think trying to stop grooming online would take away other people's rights and privacy. They don't want their private messages being read by a computer program.

Signals can be a problem in places besides livestreams. Signals gets used on many websites and apps that use accounts. Signals can lead to someone being kicked off a website or app. Signals can even lead to someone getting sued or arrested.

For example: In the US, there is a law called the REPORT Act. The REPORT act says that businesses that run websites and apps must report CSEA. The REPORT act says that an adult trying to get a child to have sex counts as CSEA.

But it's harder for a computer program to tell what counts as "an adult trying to get a child to have sex." Certain song lyrics talk about having sex and use the word "baby". Or, two teens could be talking to each other in a sexual way. This leads to more accounts getting signals that they might be sharing CSAM. If signals don't get used the right way, they will end up with many false positives like this.

Livestreaming businesses sometimes share signals with each other. That's why it's even more important that signals get made the right way. If one bad signal gets used in one place, many other places could start using it. And because everyone is using that signal, it gets harder to tell if the signal is doing anything wrong.

Signals could end up hurting people who didn't do anything wrong.

For example: Abel's S account gets hacked. Someone uses Abel's account to livestream CSAM. Then, S ban's Abel's account. They put a signal on Abel's account. The signal makes it so Abel can never sign up for another livestreaming service again.

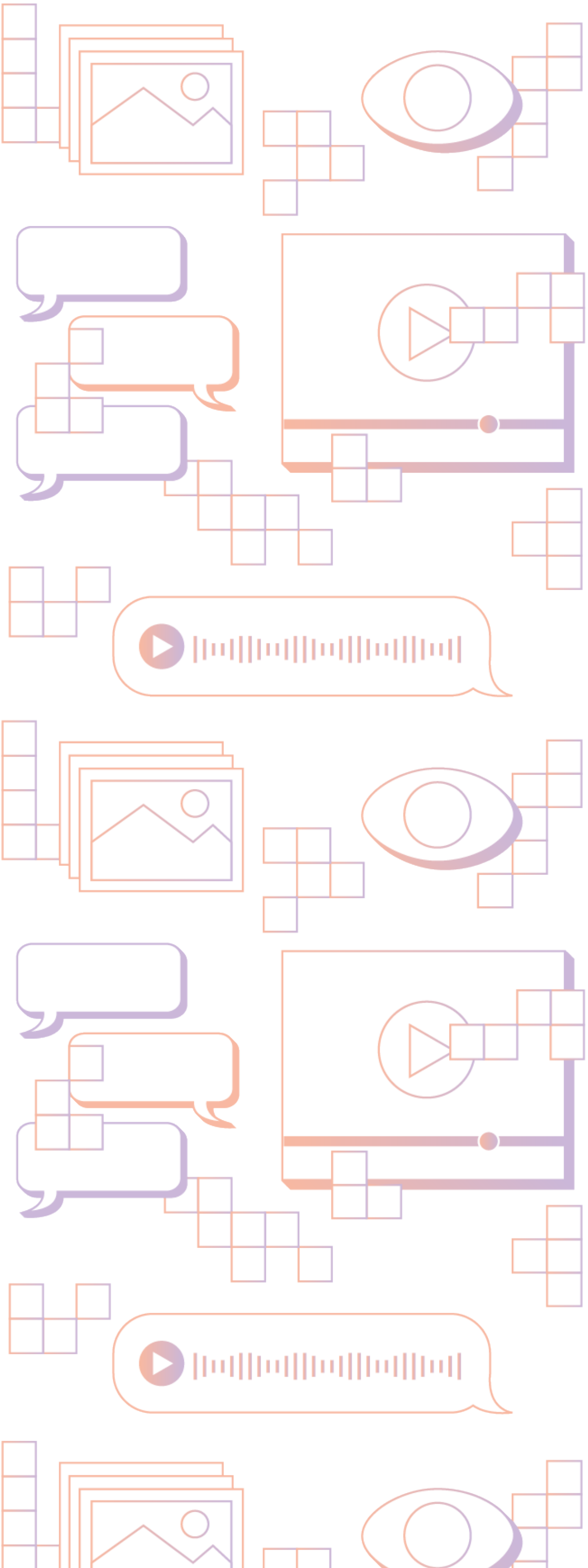
Livestreaming businesses need to think about what to do if things like this happen. They need to have ways for people to talk to a human worker about their livestreaming accounts. They need to have humans look at the accounts with signals and make sure everything is right.

When livestreaming businesses share signals, that can be an even bigger problem for people's privacy. There could be enough information in the signals for a stranger to figure out who someone is. This would be really bad if this information ended up in the wrong hands.

Livestreaming businesses use different ways to stop false positives from happening. But they do still happen. And someone might not even know they have a signal on their accounts until it becomes a problem.

Some livestreaming businesses let people try to get their account back if it was banned. These businesses talked to a lot of different accounts that got banned because of signals. They found out that the signals were flagging many accounts that did nothing wrong. The computer programs that made the signals were being too strict about what counted as dangerous. That's why it is very important livestreamers have a way to ask for their account back. Livestreaming businesses should also think about more ways to protect people's rights while using signals.

# Wrapping it Up



Livestreaming businesses have gotten a lot of pressure from people trying to stop CSEA. These businesses are working to stop CSEA happening on their livestreams. There are lots of different ways livestreaming services try to stop CSEA in livestreaming. The ways these services try to do this are still changing every day.

There hasn't been a lot of research about the best trust and safety practices to stop CSEA in livestreams. But this is a very important topic to help keep children safe around the world. And more and more businesses are making trust and safety tools to try and stop CSEA in livestreaming.

CSEA online in general is a big topic in the world right now. Lawmakers in different countries are trying to make more laws about online CSEA. Some of these laws try to stop children from going online. Other laws say that websites need to change to be a safer place for children.

Other groups have been trying to make laws about different websites. For example, some people want to make laws just for websites with adult sexual materials. These laws might make people show their ID before getting into an adult website. Or, they might make people do something else to show who they are. This shows that adults also get affected by laws about CSEA.

To make better tools to stop CSEA in livestreams, different groups will have to work together. Some of these groups are:

- Livestreaming businesses
- Businesses that make trust and safety tools
- Everyday people and livestreamers
- Experts about livestreaming, computer programs, human rights, and/or CSEA
- Researchers of livestreaming, computer programs, human rights, and/or CSEA

People from all these groups should get to test different trust and safety tools. This helps make sure these tools work well for everyone.

Here are some things we think livestreaming businesses should do. These things will help more people understand the trust and safety tools that can stop CSEA in livestreams. They will also help businesses make better tools to stop CSEA.

## Livestreaming businesses need to show how their trust and safety tools work.

Livestreaming businesses don’t have to share information about the trust and safety tools. That is why we don’t have as much research about this topic. And it is why we don’t know how well these tools work.

Livestreaming businesses don’t have a good reason to share their information. If they share that their trust and safety tool isn’t working well, it would make the business look bad. That is why we need to make rules about how livestreaming businesses should check their trust and safety tools. For example, a rule could say that businesses must show how they trained their computer programs. A rule like this would help people learn more about predictive modeling.

We think that anyone who wants to learn about these tools should be able to. But we also understand that it can be dangerous for everyone to know how these tools work. We think that the people who most need to know how these tools work are:

- People who work in livestreaming
- People who work in government
- People who do work to keep children safe

These groups should get the chance to research and learn about different trust and safety tools. A lot of research about trust and safety tools happens privately. The research doesn’t get shared with everyday people. We only learned about some of this private research while writing this report. That shows how hard it is for people to learn about trust and safety tools. It is important that this changes.

## Livestreaming businesses need to say when their trust and safety tools don’t work. Livestreaming businesses should always have humans look at what computers do.

All trust and safety tools have their limits. There are certain things they can’t do. Livestreaming businesses can’t rely just on trust and safety tools. The best work happens when humans work with trust and safety tools. For example, a predictive model can pick out certain livestreams that might have CSEA. But then a human would look over the livestream to make sure. This helps humans work faster, and helps predictive models get less wrong.

Livestreaming businesses should hire humans to make sure their trust and safety tools are fair. Humans can make sure these tools don't unfairly target one group of people, like people of color. Livestreaming businesses should always pay their workers well. Since these workers may end up seeing CSAM, it is a difficult job to do. These workers should get the training and support they need to do the job.

Livestreaming businesses should make trust and safety tools that work for everyone. They should let people, even children, protect themselves online.

Many trust and safety tools are just for livestreaming businesses to use. But trust and safety tools could get used by everyday people on livestreaming services. Giving livestreamers the power to use these tools could help make livestreaming safer.

For example, a livestreaming service could have ways for people to report other accounts. Adults could report accounts that they see sharing CSAM. This could even help children report if someone on a livestream tries to groom them.

If livestreaming businesses let people report accounts, they should make sure people can see what happens to the report. The person who made the report should know what the livestreaming business is doing about it.

## Many different groups of people should be a part of trying to stop CSEA in livestreaming.

A lot of work needs to get done to make sure trust and safety tools do the right things. And stopping CSEA in livestreaming is a big deal. Lots of groups need to work together to try and stop this problem.

Groups like the Tech Coalition are working on this topic. But other groups should be a part of the conversation, too. Livestreaming businesses should talk to child safety groups and the government. When only one group works on something, we miss out on important things.

We know that rules and tools come out better when groups work together. Different groups have already worked together to try and stop terrorist videos online. Different groups can work together to:

- Help design livestreaming services to keep people safer, such as by adding ways to check someone's age;
- Make predictive models;

- Help check trust and safety tools to make sure they work;
- Write up reports to show everyday people how trust and safety tools work;
- Look at different signals and come up with new or better.

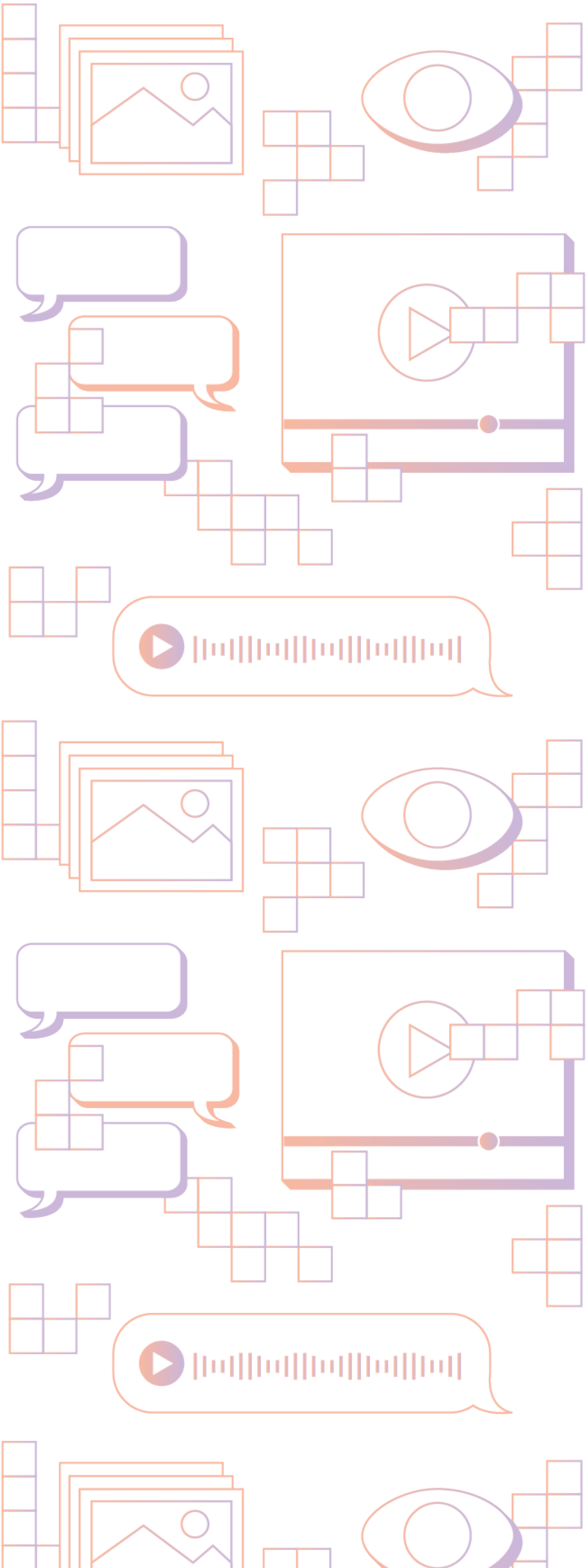
For example, The Tech Coalition’s Lantern project talked with human rights groups. They wrote a report about how their trust and safety tools might affect human rights. Livestreaming businesses should follow this example.

Livestreaming services are getting bigger and bigger. That’s why it is important to talk more about stopping CSEA in livestreams. This problem affects children, families, and communities. Many people in the government have also talked about this problem. So livestreaming businesses need to work on fixing this problem.

# Last Words

Lots of new trust and safety tools are being made to stop CSEA in livestreams. Businesses say these tools will stop CSEA in livestreaming. But it is hard to tell whether these tools are doing what they say they will do. If trust and safety tools don't work well, people might not want to use livestreaming services. The government might end up trying to shut these services down.

That's why it's important to make better trust and safety tools. And livestreaming businesses need to show how their trust and safety tools work. Different groups of people who care need to work together to make livestreaming safe for everyone. We can only stop CSEA in livestreaming with the right people and the right tools.



# To Learn More

**We used information from many other papers and websites to write our report.** The list of these websites and papers is below. You can click each link to learn more. Most of these papers and websites are not written in plain language.

ActiveFence. (2024). *Real-Time Video Content Moderation*. ActiveFence. https://www.activefence.com/video-content-moderation/ [perma.cc/LP6B-N2LH]

Allyn, B., Goodman, S., & Dara Kerr. (2024, October 13). Inside the TikTok documents: Stripping teens and boosting "attractive" people. NPR. https://www.npr.org/2024/10/12/g-s1-28040/teens-tiktok-addiction-lawsuit-investigation-documents [perma.cc/JBN9-4DVL]

Amazon Web Services. (2020, October 12). *Amazon Rekognition adds support for six new content moderation categories | AWS Machine Learning Blog*. https://aws.amazon.com/blogs/machine-learning/amazon-rekognition-adds-support-for-six-new-content-moderation-categories/ [perma.cc/A72N-T26W]

Angel, M. P., & Boyd, D. (2024). Techno-legal Solutionism: Regulating Children's Online Safety in the United States. *Proceedings of the Symposium on Computer Science and Law*, 86–97. https://doi.org/10.1145/3614407.3643705 [perma.cc/6AA8-L38R]

ARICA. (2023). *About*. ARICA. https://www.aricaproject.eu/about/ [perma.cc/DD34-5HM2]

Baines, V. (2019). Online child sexual exploitation: Towards an optimal international response. *Journal of Cyber Policy, 4*(2), 197–215. https://doi.org/10.1080/23738871.2019.1635178 [https://perma.cc/P6S9-SDM9]

Bhatia, A. (2024, September 11). The Future of the Christchurch Call Foundation and Lessons for Multistakeholder Initiatives. *Center for Democracy and Technology*. https://cdt.org/insights/the-future-of-the-christchurch-call-foundation-and-lessons-for-multistakeholder-initiatives/ [perma.cc/Q3JV-CQ54]

Bhatia, A., & Aboulafia, A. (2024, September 24). *Age Verification Technology Would Create New Barriers for Young Disabled People*. Teen Vogue. https://www.teenvogue.com/story/age-verification-technology-disabled-people [perma.cc/XP8C-ECNX]

Blake, P. (2019). Age verification for online porn: More harm than good? *Porn Studies*, 6(2), 228–237. https://doi.org/10.1080/23268743.2018.1555054 [https://perma.cc/86KF-LBC3]

Boburg, S., Verma, P., & Dehghanpoor, C. (2024, March 13). On popular online platforms, predatory groups coerce children into self-harm. *Washington Post*. https://www.washingtonpost.com/investigations/interactive/2024/764-predator-discord-telegram/ [https://perma.cc/S3ME-9UFC]

Brewer, J., Romine, M., & Taylor, T. L. (2020). Inclusion at Scale: Deploying a Community-Driven Moderation Intervention on Twitch. *Proceedings of the 2020 ACM Designing Interactive Systems Conference, 757–769*. https://doi.org/10.1145/3357236.3395514 [perma.cc/A9SB-K5K7]

Breyer, P. (2024, April 10). Pirate lawsuit: German Regional Court refuses to rule on legality of voluntary chat control scanning of private messages. *Patrick Breyer*. https://www.patrick-breyer.de/en/pirate-lawsuit-german-regional-court-refuses-to-rule-on-legality-of-voluntary-chat-control-scanning-of-private-messages/ [perma.cc/NB22-AN2Q]

Brunton, F. (2013). Spam: *A shadow history of the Internet*. MIT Press. https://doi.org/10.7551/mitpress/9384.001.0001 [https://perma.cc/Q7VQ-AWUJ]

Cai, J., Chowdhury, S., Zhou, H., & Wohn, D. Y. (2023). Hate Raids on Twitch: Understanding Real-Time Human-Bot Coordinated Attacks in Live Streaming Communities. *Proceedings of the ACM on Human-Computer Interaction*, 7(CSCW2), 1–28. https://doi.org/10.1145/3610191 [perma.cc/D2XG-4J8C]

Cai, J., & Wohn, D. Y. (2019). Categorizing Live Streaming Moderation Tools: An Analysis of Twitch. *International Journal of Interactive Communication Systems and Technologies (IJICST)*, 9(2), 36–50. https://doi.org/10.4018/IJICST.2019070103 [perma.cc/T6ET-9EU7]

Cai, J., & Wohn, D. Y. (2021). After Violation But Before Sanction: Understanding Volunteer Moderators' Profiling Processes Toward Violators in Live Streaming Communities. *Proceedings of the ACM on Human-Computer Interaction, 5*(CSCW2), 410:1-410:25. https://doi.org/10.1145/3479554 [perma.cc/CT9K-RHF6]

Caplan, R. (2023). Networked Platform Governance: The Construction of the Democratic Platform. *International Journal of Communication*, 17(22). Retrieved from https://ijoc.org/index.php/ijoc/article/view/20035 [perma.cc/RWX7-DNQX]

Celiksoy, E., Schwarz, K., & Sawyer, L. (2023). *Legal and institutional responses to the online sexual exploitation of children* |. University of Nottingham Rights Lab. https://www.nottingham.ac.uk/research/beacons-of-excellence/rights-lab/resources/reports-and-briefings/2023/october/legal-and-institutional-responses-to-the-online-sexual-exploitation-of-children-the-philippines-country-case-study.pdf [perma.cc/8DXY-3C8C]

Child Rights International Network & defenddigitalme. (2023). *Privacy and Protection: A children's rights approach to encryption*. Child Rights International Network and defenddigitalme. https://home.crin.org/readlistenwatch/stories/privacy-and-protection [perma.cc/Y888-H7X8]

Christensen, L. S., & Woods, J. (2024). "It's Like POOF and It's Gone": The Live-Streaming of Child Sexual Abuse. *Sexuality & Culture, Online First*. https://doi.org/10.1007/s12119-023-10186-9 [perma.cc/35L8-HGLG]

Cloudflare. (n.d.). *What is live streaming? | How live streaming works*. Retrieved October 14, 2024, from https://www.cloudflare.com/learning/video/what-is-live-streaming/ [https://perma.cc/QN58-XKGW]

CNIL. (2022). *Online age verification: Balancing privacy and the protection of minors*. https://www.cnil.fr/en/online-age-verification-balancing-privacy-and-protection-minors [perma.cc/Z4H6-4BPJ]

Cobbe, J. (2021). Algorithmic Censorship by Social Platforms: Power and Resistance. *Philosophy & Technology, 34*(4), 739–766. https://doi.org/10.1007/s13347-020-00429-0 [perma.cc/U5RF-R2Y2]

Cooper, K., Quayle, E., Jonsson, L., & Svedin, C. G. (2016). Adolescents and self-taken sexual images: A review of the literature. *Computers in Human Behavior*, 55, 706–716. https://doi.org/10.1016/j.chb.2015.10.003 [perma.cc/83XT-59WL]

Cox, J. (2024a, March 28). Criminals Are Weaponizing Child Abuse Imagery to Ban Discord Servers. *404 Media*. https://www.404media.co/criminals-are-weaponizing-child-abuse-imagery-to-ban-discord-servers/ [perma.cc/S79A-V5PD]

Cox, J. (2024b, June 26). *ID Verification Service for TikTok, Uber, X Exposed Driver Licenses*. 404 Media. https://www.404media.co/id-verification-service-for-tiktok-uber-x-exposed-driver-licenses-au10tix/ [perma.cc/G4NZ-49Q9]

Crawford, K., & Gillespie, T. (2016). What is a flag for? Social media reporting tools and the vocabulary of complaint. *New Media & Society*, 18(3), 410–428. https://doi.org/10.1177/1461444814543163 [https://perma.cc/9HMD-FB2Z]

Crisan, A., Drouhard, M., Vig, J., & Rajani, N. (2022). Interactive Model Cards: A Human-Centered Approach to Model Documentation. *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, 427–439. https://doi.org/10.1145/3531146.3533108 [perma.cc/MP8M-QL7M]

D'Anastasio, C. (2022, September 21). Child Predators Use Amazon's Twitch to Systematically Track Kids Who Stream. *Bloomberg*. https://www.bloomberg.com/graphics/2022-twitch-problem-with-child-predators/?sref=P6Q0mxvj [perma.cc/BZJ9-UGAX]

D'Anastasio, C. (2024, January 5). Twitch "Clips" Feature Being Used to Exploit Minors. *Bloomberg*. https://www.bloomberg.com/news/articles/2024-01-05/twitch-clips-feature-being-used-to-exploit-minors [perma.cc/8EAK-832R]

Denyer Willis, G. (2023). 'Trust and safety': Exchange, protection and the digital market–fortress in platform capitalism. Socio-Economic Review, 21(4), 1877–1895. https://doi.org/10.1093/ser/mwad003 [perma.cc/FVD4-VTYJ]

Drejer, C., Riegler, M. A., Halvorsen, P., Johnson, M. S., & Baugerud, G. A. (2024). Livestreaming technology and online child sexual exploitation and abuse: A scoping review. *Trauma, Violence, & Abuse*, 25(1), 260–274. https://doi.org/10.1177/15248380221147564 [https://perma.cc/A4VH-NTPG]

Drejer, C., Sabet, S. S., Baugerud, G. A., & Riegler, M. A. (2024). *It's All in the Game—An Exploration of Extensive Communication on Gaming Platforms and the Risks of Online Sexual Grooming* (SSRN Scholarly Paper 4671140). https://doi.org/10.2139/ssrn.4671140 [https://perma.cc/7UQT-EDPT]

Duarte, N., Llanso, E., & Loup, A. (2017). *Mixed Messages? The Limits of Automated Social Media Content Analysis*. Center for Democracy & Technology. https://cdt.org/insights/mixed-messages-the-limits-of-automated-social-media-content-analysis/ [perma.cc/9DES-3EFJ]

EDRi. (2023, August 29). *Is this the most criticised draft EU law of all time?* European Digital Rights (EDRi). https://edri.org/our-work/most-criticised-eu-law-of-all-time/ [perma.cc/4MAJ-KN8M]

Europol. (2024, July 2). *Operational sprint generates 197 new leads on buyers of 'live distant child abuse.'* Europol. https://www.europol.europa.eu/media-press/newsroom/news/operational-sprint-generates-197-new-leads-buyers-of-live-distant-child-abuse [perma.cc/ETE3-HGMG]

Farid, H. (2022). Creating, Using, Misusing, and Detecting Deep Fakes. *Journal of Online Trust and Safety, 1*(4), Article 4. https://doi.org/10.54501/jots.v1i4.56 [perma.cc/X83Y-L9ZJ]

Forland, S., Meysenburg, N., & Solis, E. (2024). *Age Verification: The Complicated Effort to Protect Youth Online*. Open Technology Institute. http://newamerica.org/oti/reports/age-verification-the-complicated-effort-to-protect-youth-online/ [perma.cc/FRE2-NFJ4]

Goggin, B. (2023, June 21). Discord servers used in child abductions, crime rings, sextortion. *NBC News*. https://www.nbcnews.com/tech/social-media/discord-child-safety-social-platform-challenges-rcna89769 [perma.cc/SG7P-Q3QC]

Google. (2024a). *Create a live stream on mobile*. https://support.google.com/youtube/answer/9228390 [perma.cc/LKY4-2VNE]

Google. (2024b). *Verify your YouTube account—YouTube Help*. https://support.google.com/youtube/answer/171664?hl=en [perma.cc/3VSR-DX9B]

Gorwa, R., Binns, R., & Katzenbach, C. (2020). Algorithmic content moderation: Technical and political challenges in the automation of platform governance. Big Data & Society, 7(1), 205395171989794. https://doi.org/10.1177/2053951719897945 [https://perma.cc/9RWX-PKM6]

Gorwa, R., & Veale, M. (forthcoming). *Routine Resistant Platform Governance* (Working Paper).

Grossman, S., Pfefferkorn, R., Thiel, D., Shah, S., Stamos, A., DiResta, R., Perrino, J., Cryst, E., & Hancock, J. (2024). *The Strengths and Weaknesses of the Online Child Safety Ecosystem: Perspectives from Platforms, NCMEC, and Law Enforcement on the CyberTipline and How to Improve It*. https://doi.org/10.25740/pr592kc5483 [perma.cc/GA64-7LEY]

Han, C., Seering, J., Kumar, D., Hancock, J. T., & Durumeric, Z. (2023). Hate Raids on Twitch: Echoes of the Past, New Modalities, and Implications for Platform Governance. *Proceedings of the ACM on Human-Computer Interaction, 7*(CSCW1), 1–28. https://doi.org/10.1145/3579609 [perma.cc/5L6B-FLVF]

Horsman, G. (2018). A forensic examination of the technical and legal challenges surrounding the investigation of child abuse on live streaming platforms: A case study on Periscope. *Journal of Information Security and Applications, 42*, 107–117. https://doi.org/10.1016/j.jisa.2018.07.009 [perma.cc/S5J9-W5EE]

Insoll, T., Ovaska, A., & Vaaranen-Valkonen, N. (2021). *CSAM Users in the Dark Web: Protecting Children Through Prevention*. Suojellaan Lapsia/Protect Children. https://www.suojellaanlapsia.fi/en/post/csam-users-in-the-dark-web-protecting-children-through-prevention [perma.cc/JM6G-ALNN]

International Justice Mission & University of Nottingham Rights Lab. (2023). *Scale of Harm: Estimating the Prevalence of Trafficking to Produce Child Sexual Exploitation Material in the Philippines*. International Justice Mission. https://www.ijm.org/studies/scale-of-harm-estimating-the-prevalence-of-trafficking-to-produce-child-sexual-exploitation-material-in-the-philippines [perma.cc/8JB9-2JXB]

Jackson, G. (2019, October 14). Twitch Streamer Says She Was Banned For “Suggestive” Attire After Brigade From Racist Trolls. *Kotaku*. https://kotaku.com/twitch-streamer-says-she-was-banned-for-suggestive-atti-1839040894 [perma.cc/GQ7K-KP96]

Kamara, S., Knodel, M., Llansó, E., Nojeim, G., Qin, L., Thakur, D., & Vogus, C. (2021). *Outside Looking In: Approaches to Content Moderation in End-to-End Encrypted Systems* (p. 38). Center for Democracy & Technology. https://cdt.org/insights/outside-looking-in-approaches-to-content-moderation-in-end-to-end-encrypted-systems/ [perma.cc/97V3-M8H2]

Kennedy, Ü., Lala, G., Rajan, P., Sardarabady, S., & Tatam, L. (2024). *Protecting Children from Online Grooming: Cross-cultural, qualitative and child-centred data to guide grooming prevention and response*. Save the Children. https://resourcecentre.savethechildren.net/document/protecting-children-from-online-grooming-cross-cultural-qualitative-and-child-centred-data-to-guide-grooming-prevention-and-response/ [https://perma.cc/ED6M-LM4T]

Laranjeira da Silva, C., Macedo, J., Avila, S., & dos Santos, J. (2022). Seeing without Looking: Analysis Pipeline for Child Sexual Abuse Datasets. *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, 2189–2205. https://doi.org/10.1145/3531146.3534636 [perma.cc/2MLT-6Z9V]

Llansó, E. (2020, July 30). Human Rights NGOs in Coalition Letter to GIFCT. *Center for Democracy and Technology*. https://cdt.org/insights/human-rights-ngos-in-coalition-letter-to-gifct/ [perma.cc/NZZ6-XDPU]

Luria, M. (2023). *More Tools, More Control: Lessons from Young Users on Handling Unwanted Messages Online*. Center for Democracy & Technology. https://cdt.org/insights/more-tools-more-control-lessons-from-young-users-on-handling-unwanted-messages-online/ [perma.cc/L756-HP44]

Marwick, A., Smith, J., Caplan, R., & Wadhawan, M. (2024). Child Online Safety Legislation (COSL)—A Primer. *The Bulletin of Technology & Public Life*. https://doi.org/10.21428/bfcb0bff.de78f444 [perma.cc/A5LE-YVRL]

McAlinden, A.-M. (2006). 'Setting 'Em Up': Personal, Familial and Institutional Grooming in the Sexual Abuse of Children. *Social & Legal Studies*, 15(3), 339–362. https://doi.org/10.1177/0964663906066613 [https://perma.cc/YBW5-M7RW]

McKee, A., & Lumby, C. (2022). Pornhub, child sexual abuse materials and anti-pornography campaigning. *Porn Studies*, 9(4), 464–476. https://doi.org/10.1080/23268743.2022.2083662 [https://perma.cc/Q7ND-7FWN]

Meta. (2023, December 1). Our Work To Fight Online Predators. *Meta*. https://about.fb.com/news/2023/12/combating-online-predators/ [perma.cc/MBL8-9DUJ]

Meta. (2024a). *Child Sexual Exploitation, Abuse, and Nudity | Transparency Center*. https://transparency.meta.com/policies/community-standards/child-sexual-exploitation-abuse-nudity/ [https://perma.cc/A4ES-MERB]

Meta. (2024b). *IndicatorType—ThreatExchange—Documentation*. Meta for Developers. https://developers.facebook.com/docs/threat-exchange/reference/apis/indicator-type/v21.0/ [https://perma.cc/8RM5-T35N]

Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., & Gebru, T. (2019). Model Cards for Model Reporting. *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 220–229. https://doi.org/10.1145/3287560.3287596 [perma.cc/VEQ7-K2D8]

MSAB. (2023, May 31). *Safer Digital Spaces: The Vital Role of Technology in Combating CSAM*. MSAB. https://www.msab.com/blog/forensic-fix-tom-farrell-jesse-nicholson/ [perma.cc/DT3U-4UPE]

Nicholas, G., & Bhatia, A. (2023). *Lost in Translation: Large Language Models in Non-English Content Analysis*. Center for Democracy & Technology. https://cdt.org/insights/lost-in-translation-large-language-models-in-non-english-content-analysis/ [perma.cc/Y7JL-F5GW]

Payt, S. (2024, September 27). *Council Post: 3 Solutions To The Technology-Facilitated Crimes Against Children*. Forbes. https://www.forbes.com/councils/forbesnonprofitcouncil/2024/09/27/3-solutions-to-the-technology-facilitated-crimes-against-children/ [perma.cc/B8Q8-535U]

Peralta, D. (2023). AI and suicide risk prediction: Facebook live and its aftermath. *AI & SOCIETY*. https://doi.org/10.1007/s00146-023-01651-y [perma.cc/C3Z7-YNNN]

Pereira, M., Dodhia, R., Anderson, H., & Brown, R. (2023). Metadata-Based Detection of Child Sexual Abuse Material. IEEE Transactions on Dependable and Secure Computing, 1–13. *IEEE Transactions on Dependable and Secure Computing*. https://doi.org/10.1109/TDSC.2023.3324275 [perma.cc/HAJ3-ZAXK]

Persson, J. (2024). Age as a Gatekeeper in the UK Online Safety Agenda. In E. Setty, F. Gordon, & E. Nottingham (Eds.), *Children, Young People and Online Harms: Conceptualisations, Experiences and Responses* (pp. 169–181). Springer International Publishing. https://doi.org/10.1007/978-3-031-46053-1_7 [perma.cc/EYF9-G85L]

Pfefferkorn, R. (2023, December 19). *Child Safety-Focused REPORT Act Passes US Senate | TechPolicy.Press*. Tech Policy Press. https://techpolicy.press/child-safetyfocused-report-act-passes-us-senate [perma.cc/QCA5-EM4K]

Quayle, E. (2020). Prevention, disruption and deterrence of online child sexual exploitation and abuse. *ERA Forum*, 21(3), 429–447. https://doi.org/10.1007/s12027-020-00625-7 [perma.cc/7PDJ-L5U3]

Quayle, E. (2022). Self-produced images, sexting, coercion and children's rights. ERA Forum, 23(2), 237–251. https://doi.org/10.1007/s12027-022-00714-9 [perma.cc/PBE3-RVFD]

Reyes, I., Wijesekera, P., Reardon, J., Elazari Bar On, A., Razaghpanah, A., Vallina-Rodriguez, N., & Egelman, S. (2018, July 24). *"Won't Somebody Think of the Children?" Examining COPPA Compliance at Scale*. The 18th Privacy Enhancing Technologies Symposium (PETS 2018). https://dspace.networks.imdea.org/handle/20.500.12761/551 [perma.cc/YNW9-8CHG]

Ruane, K., Branum, B., Doty, N., & Jain, S. (2024, September 23). CDT Files Amicus Brief in Free Speech Coalition v. Paxton, Challenging TX Age Verification Law. *Center for Democracy and Technology*. https://cdt.org/insights/cdt-files-amicus-brief-in-free-speech-coalition-v-paxton-challenging-tx-age-verification-law/ [perma.cc/YG2W-263P]

Ruberg, B. (2021). "Obscene, pornographic, or otherwise objectionable": Biased definitions of sexual content in video game live streaming. *New Media & Society*, 23(6), 1681–1699. https://doi.org/10.1177/1461444820920759 [https://perma.cc/4NZP-CB8N]

Salter, M., & Sokolov, S. (2024). "Talk to strangers!" Omegle and the political economy of technology-facilitated child sexual exploitation. *Journal of Criminology*, 57(1), 121–137. https://doi.org/10.1177/26338076231194451 [https://perma.cc/SV28-5TWN]

Setter, C., Greene, N., Newman, N., & Perry, J. (2021). *Global Threat Assessment 2021*. WeProtect Global Alliance. https://www.weprotect.org/global-threat-assessment-21/#report [perma.cc/ZKY9-9VKK]

Shenkman, C., Thakur, D., & Llansó, E. (2021). *Do You See What I See? Capabilities and Limits of Automated Multimedia Content Analysis*. Center for Democracy and Technology. https://cdt.org/insights/do-you-see-what-i-see-capabilities-and-limits-of-automated-multimedia-content-analysis/ [perma.cc/5H78-DF8K]

Stardust, Z., Obeid, A., McKee, A., & Angus, D. (2024). Mandatory age verification for pornography access: Why it can't and won't 'save the children.' *Big Data & Society*, 11(2), 20539517241252129. https://doi.org/10.1177/20539517241252129 [https://perma.cc/EB9M-F7PX]

Stripchat. (2024, February 28). *What documents do I need to upload to create my account?* Stripchat FAQ. https://support.stripchat.com/hc/en-us/articles/4410734320785-What-documents-do-I-need-to-upload-to-create-my-account [https://perma.cc/ADZ4-ML7G]

Tech Coalition. (2022). *Tech Coalition | Trust: Voluntary Framework for Industry Transparency*. Tech Coalition. https://www.technologycoalition.org/knowledge-hub/trust-voluntary-framework-for-industry-transparency [perma.cc/ER2S-DH6U]

Tech Coalition. (2023, November 7). *Tech Coalition | Announcing Lantern: The First Child Safety Cross-Platform Signal Sharing Program*. Tech Coalition. https://www.technologycoalition.org/newsroom/announcing-lantern [perma.cc/NV2D-2PPW]

Teunissen, C., & Napier, S. (2023). The overlap between child sexual abuse live streaming, contact abuse and other forms of child exploitation. *Trends and Issues in Crime and Criminal Justice, 671*, 1–16. https://www.aic.gov.au/publications/tandi/tandi671 [https://perma.cc/34HA-8NQW]

Teunissen, C., Napier, S., & Boxall, H. (2021). Live streaming of child sexual abuse: An analysis of offender chat logs. *Trends and Issues in Crime and Criminal Justice, 639*, 1–15. https://doi.org/10.52922/ti78375 [https://perma.cc/J5AN-FM9T]

Thiel, D., DiResta, R., & Stamos, A. (2023). *Cross-Platform Dynamics of Self-Generated CSAM*. Stanford Internet Observatory. https://purl.stanford.edu/jd797tp7663 [perma.cc/CH99-A4YB]

Thorn. (n.d.). *Text Classifier for Child Safety | Safer Predict, Built by Thorn*. Retrieved October 15, 2024, from https://get.safer.io/text-classification-content-moderation [perma.cc/UVG7-YZ99]

Thorn. (2022, September 23). *How CSAM Detection Works | Safer by Thorn*. Safer: Proactive Solution for CSE and CSAM Detection. https://safer.io/how-it-works/ [perma.cc/Z5F9-3RDK]

Thorn. (2024, June 26). *CSAM Keyword Hub Application | Safer.i*o. Safer: Proactive Solution for CSE and CSAM Detection. https://safer.io/resources/csam-keyword-hub/ [perma.cc/5TUY-G7XU]

TikTok. (2024a). *Minimum age appeals on TikTok | TikTok Help Center*. https://support.tiktok.com/en/safety-hc/account-and-user-safety/minimum-age-appeals-on-tiktok [perma.cc/G84K-CNXB]

TikTok. (2024b). *Protecting teens online*. https://www.tiktok.com/transparency/en-us/protecting-teens/ [perma.cc/5TCH-PKWK]

TikTok. (2024c, January 19). *LIVE Center*. https://livecenter.tiktok.com/help_center/article/1023/tiktok-live-studio-operation-manual_en-US?lang=en [perma.cc/QQ8K-UMLF]

TikTok. (2024d, September 26). *Community Guidelines Enforcement Report—April 1—June 30, 2024*. https://www.tiktok.com/transparency/en/community-guidelines-enforcement-2024-9 [perma.cc/U6NN-KVKP]

TikTok. (2024e, October 25). *TikTok Creator Academy: Empowering Creators to Grow and Succeed on TikTok | TikTok For Creator*. https://www.tiktok.com/creator-academy/en/article/Going-LIVE?ref=kapwing-resources [perma.cc/45E2-2CF5]

Tirfe, D., & Anand, V. K. (2022). A Survey on Trends of Two-Factor Authentication. In H. K. D. Sarma, V. E. Balas, B. Bhuyan, & N. Dutta (Eds.), *Contemporary Issues in Communication, Cloud and Big Data Analytics* (pp. 285–296). Springer. https://doi.org/10.1007/978-981-16-4244-9_23 [perma.cc/UGF8-ALSA]

Tommy I. (2023, January 10). The Subscriber Requirements For Livestreaming On YouTube: How To Get Started | *TuBeast.com*. | TuBeast.Com. https://tubeast.com/the-subscriber-requirements-for-livestreaming-on-youtube-how-to-get-started [perma.cc/E8MF-QNUP]

Twitch. (n.d.). *Chat Verification Settings*. Retrieved October 14, 2024, from https://help.twitch.tv/s/article/chat-verification-settings?language=en_US [perma.cc/VUS7-LGN3]

Twitch. (2022, November 22). *Our Ongoing Work to Combat Online Grooming*. https://safety.twitch.tv/s/article/Our-Work-to-Combat-Online-Grooming?language=en_US [perma.cc/XU6Y-F9TJ]

Twitch. (2023). *H1 2023 Transparency Report*. Twitch. https://safety.twitch.tv/s/article/H1-2023-Transparency-Report?language=en_US [perma.cc/D7NR-C7UQ]

Twitch. (2024, March 26). *Twitch.tv—Terms of Service.* Twitch.Tv. https://www.twitch.tv/p/en/legal/terms-of-service/#2-use-of-twitch-by-minors-and-blocked-persons [perma.cc/8V7P-6YLU]

Vallance, C. (2024, April 22). *Three-year-olds groomed online, Internet Watch Foundation warns*. https://www.bbc.com/news/articles/cx9wezr1d1vo [perma.cc/6ZRW-3SD5]

Wang, A., Ramaswamy, V. V., & Russakovsky, O. (2022). Towards Intersectionality in Machine Learning: Including More Identities, Handling Underrepresentation, and Performing Evaluation. *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, 336–349. https://doi.org/10.1145/3531146.3533101 [perma.cc/G6E2-FGDD]

Winslow, L. (2024, January 5). Report: Predators Are Using Twitch "Clips" To Spread Child Abuse. *Kotaku*. https://kotaku.com/twitch-clips-feature-predators-child-abuse-tiktok-1851144631 [perma.cc/U56B-MQ5T]

Witting, S. K. (2019). Regulating bodies: The moral panic of child sexuality in the digital era. *Kritische Vierteljahresschrift Für Gesetzgebung Und Rechtswissenschaft, 102*(1), 5–38. https://doi.org/10.5771/2193-7869-2019-1-5 [perma.cc/67PT-WV94]

Xiao, F. (2024). Moderating for a friend of mine: Content moderation as affective reproduction in Chinese live-streaming. *Media, Culture & Society, 46*(1), 60–77. https://doi.org/10.1177/01634437231188465 [https://perma.cc/25KW-9SAH]

Zhao, D., Wang, A., & Russakovsky, O. (2021). Understanding and Evaluating Racial Biases in Image Captioning. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 14830–14840. https://openaccess.thecvf.com/content/ICCV2021/html/Zhao_Understanding_and_Evaluating_Racial_Biases_in_Image_Captioning_ICCV_2021_paper.html [perma.cc/8ZXT-CC2T]

Zornetta, A., & Pohland, I. (2022). Legal and technical trade-offs in the content moderation of terrorist live-streaming. *International Journal of Law and Information Technology, 30*(3), 302–320. https://doi.org/10.1093/ijlit/eaac020 [perma.cc/Y2CV-YJTJ]

cdt.org

cdt.org/contact

**Center for Democracy & Technology**
1401 K Street NW, Suite 200
Washington, D.C. 20005

202-637-9800

@CenDemTech

CENTER FOR
DEMOCRACY
& TECHNOLOGY