

# TRUST & SAFETY

FACT SHEET | MAY 2024

Aylo is committed to working with partners to prevent and eliminate illegal content, including non-consensual material and child sexual abuse material (CSAM).

Over the years, Aylo has put in place robust measures to protect their platforms from abuse. Trust and safety policies and practices are always evolving to better identify, flag, remove, review, prevent and report illegal or abusive material.

Aylo views it as its responsibility to ensure its platforms are not used for sharing, consuming, or searching for CSAM. We take all forms of online abuse very seriously and are fully committed to the trust and safety of the public, including minors – as well as of our content-creators, users, partners, and employees and their families.

As part of Aylo's efforts and dedication to trust and safety, it is important to provide information about initiatives, technologies, moderation practices and how its platforms respond to harmful content. Aylo does this via Pornhub's Transparency Reports.

Ethical Capital Partners (ECP) is confident that the Aylo team and all Aylo platforms operate with trust and safety at the forefront of everything they do.

## PROGRAMS

Every online platform has the responsibility to join the fight to prevent and eliminate illegal content, including CSAM, and it requires collective action and constant vigilance. Aylo works with government, law enforcement agencies and NGOs on the following programs:

**Stop it Now Program** Aylo partnered with the Lucy Faithfull Foundation (LFF), a widely respected non-profit organization at the forefront of the fight against child sex abuse whose crucial efforts include a campaign designed to deter searches seeking or associated with underage content. Aylo actively works with LFF to display deterrence messaging for certain user searches on our platform and rolled out localised messaging in partnership with regional organisations such as Talking for Change, and Safe to Talk. In partnership with the Cyber Civil Rights Initiative (CCRI) and other regional NGOs, Aylo also added deterrence messaging for non-consensual material in multiple jurisdictions.

**reThink Chatbot** Internet Watch Foundation, Stop It Now and Pornhub launched a first of its kind (deterrence) chatbot to prevent child sexual abuse and educate users. The reThink chatbot engages with adult pornography users attempting to search for sexual imagery of children online. The chatbot, triggers based on search words, engages users in a conversation and signposts them to Stop It Now UK and Ireland where they can receive help and support to address their behaviour.

**IWF/Aylo Partnership Project** Following a two-year pilot project, Aylo and the **IWF** introduced the world's first **standards for adult pornography websites** to prevent the upload and distribution of online child sexual abuse. Under the standards, adult sites would be required to introduce a series of measures to show they are committed to making the internet safer, including using AI to scan videos being uploaded, and to take clear steps to make sure users are over 18. As part of the two-year pilot, Aylo supplemented its existing Trust and Safety program by taking all relevant IWF services and deployed them across its platforms including the URL List, Keyword List and Hash Lists for CSAM, and Non-photographic Imagery. The project is being overseen by an external advisory board of experts who will guide the outputs of the partnership. Key learnings from the project will be shared publicly to support others' work in this area.

**NCMEC Take It Down** This tool allows young people to proactively prevent or subsequently remove nude or sexually explicit images and videos (including AI-generated content (source: **USA Today**)) of themselves that may be or have been posted to the participating platforms. Aylo deployed the use of this service for all content uploaded to its platforms.

**STOPNCII.org** A global initiative (developed by Meta & **SWGfL**) that prevents the spread of non-consensual intimate images (NCII) online. If any adult (18+) is concerned about their intimate images (or videos) being shared online without consent, they can create a digital fingerprint of their own material and prevent it from being shared across participating platforms. Aylo volunteered for and is currently using this program for all content uploaded to its platforms.

## TRUSTED FLAGGER PROGRAM

Launched in 2020, through their **Trusted Flagger Program**, members are empowered to alert Aylo to content they believe may violate their terms of service. The Trusted Flagger Program consists of nearly 60 leading non-profit internet and child-safety organizations. Members have a direct line of access to Aylo's moderation team, and any content identified by a Trusted Flagger is immediately and automatically disabled, with no human intervention.

### Some examples of our Trusted Flaggers:

- Internet Watch Foundation - UK
- Child Focus - Belgium
- Centre for Expertise on Online Sexual Child Abuse (Offlimits) - The Netherlands
- Point de Contact - France
- National Center for Missing & Exploited Children (NCMEC) - USA

## DETERRENCE PARTNERS

Aylo actively works with organizations to display deterrence messaging for certain user searches on our platforms. As a result, attempts to search for certain words and phrases\* now yield a clear deterrence message about the illegality of non-consensual material or of underage content, while offering a free, confidential, and anonymous support resource for those interested in seeking help for controlling their urges to engage in illegal behavior, including:

- Lucy Faithfull Foundation
- Stop It Now US, NL, BE, AUS
- Safe to Talk
- Centre for Addiction and Mental Health (CAMH)
- Revenge Porn Helpline
- Cyber Civil Rights Initiative

\*the banned word list contains over 34,000 terms in over a dozen different languages.

We look forward to further consulting with staff and other stakeholders as we continue to build on Aylo's industry-leading commitment to quality adult entertainment made by and for consenting adults on platforms that set the standard for trust and safety.

For more information please consult the **[Pornhub Transparency Report](#)** page.

## AUTOMATED DETECTION TECHNOLOGY

Aylo remains vigilant when it comes to researching, adopting, and using the latest and best available detection technology to help make their platforms safe, while remaining an inclusive, sex positive online space for adults. Aylo deploys several different automated detection technologies to help moderate content before it can be published, including:

**CSAI Match** – YouTube's proprietary technology for combating child sexual abuse imagery. All video content appearing on Pornhub has been scanned against **YouTube's CSAI Match**. All new video content is scanned at the time of upload and before it can be published.

**PhotoDNA** – Microsoft's technology that aids in finding and removing known images of child sexual abuse material. All photo content appearing on Pornhub has been scanned against **Microsoft's PhotoDNA**. All new photo content is scanned at the time of upload and before it can be published.

**Google's Content Safety API** – **Google's Content Safety API** uses Artificial Intelligence to evaluate content as a possible CSAM to help companies better prioritize illegal/abuse material for human review thus helping reviewers find and report content faster.

**MediaWise** – **Vobile's** cyber "fingerprinting" software that scans all new user uploads to help prevent previously identified content from being re-uploaded.

**Safer** – Aylo became the first adult content platform to partner with **Thorn** (2020), **Safer** gives platforms an additional layer of protection to their compliance and content moderation process and helps protect visitors from unwanted or illegal material.

**Instant Image Identifier** – The Centre for Expertise on Online Sexual Child Abuse (Offlimits, previously known as **EOKM**) tool, commissioned by the European Commission, detects known child abuse imagery using a triple verified database. The Instant Image Identifier is associated with the three pairs of eyes (3Is) principle, meaning that CSAM images are triple verified in the classification process. Three people (three pairs of eyes) have classified each image. We have implemented this tool across all our platforms as another layer of protection for our sites, allowing us to check whether uploaded files occur in datasets of known online child sexual abuse material.

**Safeguard** – Safeguard is Aylo's proprietary image recognition technology designed with the purpose of combatting both child sexual abuse imagery and non-consensual content, by preventing the re-uploading of previously fingerprinted content to Aylo platforms.

**Age Estimation** – Aylo utilizes age estimation capabilities to analyze content uploaded to Aylo platforms using a combination of internal proprietary software and external AI solutions, provided by AWS and Private ID, in an effort to strengthen the varying methods used to prevent the upload and publication of potential or actual CSAM.

**Yoti** – Aylo uses **Yoti**, a leading digital identity verification solution trusted by governments, regulators and industries around the world, to verify the identity of content creators. Yoti verifies identity by validating government issued photo ID and comparing photo identification and a live face scan, of those who apply to become verified uploaders before they can be approved for eligibility to upload any content to our platforms.

**Spectrum Labs AI** – **Spectrum Labs AI** uses natural language processing techniques to make sense of text-based data, including video titles, subtitles, tags, user comments and descriptions. We use it to detect, among other things, mentions or references of anything related to underage, non-consent, revenge porn, doxing, solicitation of sex, CSAM and hate speech in comments entered on our platforms.

**ActiveFence** – Aylo partners with **ActiveFence** to support in identifying and responding to safety risks on and off their platforms. **ActiveFence** is used to monitor the threat landscape for human exploitation and CSAM. **ActiveFence** uses intelligence techniques to map networks and monitor deep and dark web forums for threat actors and translates those insights into actionable intelligence for our platforms.

## CSAM REPORTING

The National Center for Missing and Exploited Children (NCMEC), the independent authority on combating CSAM on the internet, annually lists the number of reports of potential CSAM it receives per platform (major social and content platforms, including Pornhub) and the time it takes the platform to remove the potentially violative material after it is notified. Source: **NCMEC 2023 Report**

Platform	Number of Notices Sent From NCMEC	Average Days to Respond / Take Down	Number of CyberTipline Reports Submitted
Pornhub	2	.9	2,487
Facebook	115	6.08	17,838,422
Instagram	30	4.21	11,430,007
WhatsApp	-	-	1,389,618
Twitter / X	631	1.29	597,087
Google	274	4.21	1,470,958