Adult Content Standards

IWF Internet Watch Foundation

EFFECTIVELY TACKLING CHILD SEXUAL ABUSE ONLINE:

v2 June 2024

A Standard of **Good Practice** for Adult Services

Supporting Document

Working together to STOP child sexual abuse on adult content websites

SUPPORTING DOCUMENT

Adult Content Standards

Overview

This document provides supporting information for the Internet Watch Foundation's (IWF) Standard of Good Practice for Adult Services.

The IWF intends to conduct an annual review of the supporting document to ensure it is relevant and effective in addressing evolving challenges and changes to the online child sexual exploitation and abuse (CSEA) landscape.

A. Principles

The Advisory Board has set out six principles for adult services for tackling child sexual abuse (CSA), to apply in all countries where services are provided, across all sites:

1. **Adopt a zero-tolerance approach to child sexual abuse.**
2. **Ensure transparency, with enforceable terms of service.**
3. **Operate with accountability, with clear reporting mechanisms.**
4. **Embrace technological tools and solutions.**
5. **Collaborate with specialists.**
6. **Embrace regulatory and safety initiatives, including voluntary principles.**

B. Specific steps to achieve good practice

Baseline Standard

a. Adult services must comply with legislation and regulation[1]. This includes requirements to assess the level of risk on their services and embrace safety by design as an approach to mitigate harms in the configuration of their platform or service.

- Adult services must comply with Digital Millennium Copyright Act (DMCA) takedown notices by removing any copyright-infringing material and must have counter-notification and repeat infringer processes in place.
- Adult services must ensure that users can report claims of copyright infringement.
- Adult services must display Terms of Service and Community Guidelines which list prohibited activities and must operate consistently with them. These must be easily located on the platform.
- Adult services must publish child sexual abuse (CSA) and Non-Consensual Intimate Imagery ("NCII") policies and must make them easily accessible.

b. Adult services must adopt mechanisms to detect and prevent known child sexual abuse material (CSAM), as deemed appropriate by the IWF.

- Adult services must adopt the Association of Sites Advocating Child Protection's (ASACP) free service which allows adult services to add Restricted to Adults (RTA) labels to their site pages, or a similar service which allows parents to enable parental control filtering of the platform.
- Adult services must subject all uploaded content to assessment by AI tools which specialise in age estimation before it is published.
- Adult services must implement a solution allowing uploaded content to be assessed as being or containing artificially created or manipulated content or media and must either prevent the publishing of such content entirely or have a policy in place to deal with it.
- Adult services must implement a system or support one or more third party systems which allow users to hash their own content to prevent it from being uploaded, without sharing the content with the platform.
- Adult services must implement an upload filtering and content fingerprinting mechanism, allowing to prevent re-upload of previously fingerprinted content. Using this mechanism, adult services must fingerprint all content removed for containing CSA or potential CSA.
- Adult services must scan uploaded content against databases of previously identified and reported content hashes, to prevent the upload of CSA material.
- Adult services must maintain and regularly update terms on banned word and URL lists and must actively prevent or block the use of such terms on its platform.
- Adult services should moderate or deploy tools to assess user-submitted text fields for multiple bad behaviours.

1. Legal requirement: Implicit

c. **Adult services must ensure that all materials published comply with the following standards (taken from the British Board of Film Classification's (BBFC) standards for an R18 certification) and therefore exclude:**
   - Material which is in breach of the criminal law;
   - Material (including dialogue) likely to encourage an interest in sexually abusive activity, which may include adults role-playing as non-adults;
   - The portrayal of sexual activity which involves real or apparent lack of consent; any form of physical restraint which prevents participants from indicating a withdrawal of consent;
   - The infliction of pain or acts which are likely to cause serious physical harm, whether real or (in a sexual context) simulated. Some allowance may be made for non-abusive, consensual activity;
   - Penetration by any object likely to cause physical harm;
   - And sexual threats, humiliation or abuse which do not form part of a clearly consenting role-playing game.

d. **Adult services must publish transparency reports every six months[2].**
   - The transparency report must include:
     - A record of all content removed since the end of the last reporting period and the reason it was removed.
     - The percentage of content that was removed before publishing.
     - The sources of the removed content, including the percentage of each source.
     - The volume of URLs de-indexed from search engines for containing CSA.
     - The average time taken to respond to users who reported an issue relating to CSA.
     - Subject to confidentiality agreements, adult services should be transparent about legal requests to which they respond.

e. **Adult services must establish and maintain a dedicated portal to facilitate accessible and secure communication between law enforcement and the platform regarding specific investigations, ensuring a channel for victim redress is readily available.**

f. **Adult services must have a clear reporting function for users to flag harmful content[3].**
   - Adult services must make visitors to the platform aware of how to report material for removal if it violates the Terms of Service.
   - Content removal forms must be easy to find and complete and can be used to report any content or media, including videos and images.
   - Adult services must include a flagging feature to allow account holders to flag content, users, and comments.
   - Adult services should review all reported or flagged material within 24 hours.
   - Adult services must maintain a feedback loop with reporters of harmful content by informing them of the ultimate result of the request.

g. **Adult services must subject anyone visiting, publishing, or appearing in material on the platform to age verification measures to confirm they are over 18 years old at the time of production before content is permitted to be published[4].**
   - Adult services must display clear messaging on their services or sites

2. Legal Requirement: Digital Services Act. Article 24 (2)
3. Legal Requirement: Online Safety Act, Part 3, Section 20, Subsection 3
4. Legal Requirement: Online Safety Act, Part 5, Section 81, Subsection 2

to declare that content only features performers aged over 18 at the time of production.

- When required by law, adult services must display an age disclaimer asking the visitor to confirm that they are 18 years or older, on the landing page or any direct URL link from a search engine before revealing any adult material.
- Adult services should validate payment beneficiaries against the account holder.
- Adult services should include information on how parents can enable controls in popular browsers and operating systems to prevent children from accessing the platform.

h. Adult services must ensure that consent is secured for all individuals appearing in content. All individuals must be allowed to withdraw consent at any time. Cases which involve professional contracts should be assessed on a case-by-case basis.

- Adult services must require an explicit statement from content publishers that they have ID and consent documents on file for all individuals appearing in the content they are uploading, at the time of registration and again at the time of each individual upload.
- A process should be in place to audit and review ID and consent. Failure by an uploader to provide evidence of ID and record of consent for a participant should result in the removal of all content featuring that participant.
- Adult services should provide downloadable forms and support an electronic signature process to help uploaders comply with this requirement.

i. Adult services must not support technologies that obscure the content of messaging and other communications that would inhibit moderation.

j. Adult services must not adopt, or encourage the adoption of, technologies that can be used to bypass content filtering and content blocking mechanisms, whether for accessing their services or hosting them.

- Parental controls are an example of a blocking mechanism that must not be bypassed.
- A non-exhaustive list of examples of technologies that could be used to evade content filtering or blocking from the user device includes:
  - Virtual private networks;
  - Encrypted DNS protocols such as DNS-over-HTTPS and DNS-over-QUIC;
  - Oblivious DNS and Oblivious HTTPS;
  - Apple Private Relay.
- An example of a technology that could be used to evade content filtering or blocking at the content host is the Encrypted Client Hello (ECH) extension to the TLS 1.3 protocol.

Higher Standard

a. Human moderators must review all content before it can be published on the platform.

- Human moderators should not be held to minimum review quotas and should be trained to escalate any content which may present a CSA risk.
- Human moderators should be subjected to quality assurance standards.
- Live streams should be subject to real-time human moderation.
- Direct or private messaging should be subject to banned word detection and human review.

b. Human moderators must be well-supported.

- Human moderators should be carefully inducted into the role and trained to make accurate assessments for content that shows CSA, including from external expert organisations where appropriate.

c. Adult services must deploy tools and changes on all historical content and accounts, even if this requires removing old content.

- Adult services must implement standards equally to every owned site.

d. Platforms should have clear and effective deterrence messaging display if a user attempts to search for words and terms associated with underage content.

- Deterrence messaging should signpost to free, confidential, and anonymous support services, such as the Lucy Faithfull Foundation's Stop It Now helpline.

e. Adult services should seek to engage with other relevant organisations to use their tools, participate in initiatives, and seek their expertise.

- Adult services must register with the IWF and National Center for Exploited and Missing Children (NCMEC) and report all potential instances of CSA to the National Crime Agency (NCA).
- Adult services should signpost to services such as Take it Down and Stop NCII.

f. Adult services should share intelligence with other platforms about threats to children and emerging issues on their sites or services.

Organisations and suggested partners

Internet Watch Foundation (IWF): A UK-based non-profit which works globally to eliminate child sexual abuse images and videos from the internet.

BBFC: The UK's independent regulator of film and video content, and leading authority in the regulation of pornography.

Childnet: A non-profit organisation working to help make the internet a safe place for children.

Lucy Faithfull Foundation: UK-wide child protection charity dedicated solely to preventing child sexual abuse.

Marie Collins Foundation: A charity committed to harnessing the voices of victims and survivors of Technology-Assisted Child Sexual Abuse and exploitation.

SWGfL: A charity dedicated to empowering the safe and secure use of technology globally.

National Center for Missing & Exploited Children: US-based private, non-profit corporation whose mission is to help find missing children, reduce child sexual exploitation, and prevent child victimisation.

Thorn: A non-profit organisation that builds technology to defend children from sexual abuse.

ECPAT International (End Child Prostitution and Trafficking): A global network dedicated to ending the sexual exploitation of children.

Yoti: A digital identity company that makes it safer for people to prove who they are.

Onfido: Provides an artificial intelligence-based technology with facial biometrics.

IDnow: Provides an identity verification platform.

Tools

Hash Lists:
When an image containing CSA is identified, it can be "hashed" (turned into a unique string of numbers or "digital fingerprint"). This can then be deployed to find duplicates of the same image and ensure it is not uploaded again.

By using hash lists, adult services can stop users from uploading, downloading, viewing, sharing, or hosting known images and videos showing CSA.

URL Lists:
A URL list contains a list of webpages showing confirmed CSA. The IWF URL List, for example, allows companies to block access to these criminal webpages, whilst IWF analysts work to have the actual image or video removed from the internet.

For adult content providers, URL lists can be deployed to ensure that users are not sharing links to criminal webpages in comments on videos.

Suggested tools:

- IWF's Member Services
- Google's Content Safety API/CSAI Match
- Microsoft's PhotoDNA
- Thorn's Safer
- Microsoft's Azure Facial Age Estimation
- Adobe's Spectrum Text
- South West Grid for Learning's STOPNCII
- Association of Sites Advocating Child Protection's (ASACP) Restricted to Adults (RTA) label tool

Register your interest

If you are interested in becoming an IWF member and/or would like to find out more information about our standard of good practice, please email **members@iwf.org.uk**