

CANADIAN CENTRE for CHILD PROTECTION®

Helping families. Protecting children.

PROJECT ARACHNID: ONLINE AVAILABILITY OF CHILD SEXUAL ABUSE MATERIAL



An analysis of CSAM and harmful-abusive content linked to certain electronic service providers

SUMMARY DOCUMENT



Note to reader:

This document is a condensed summary of a comprehensive report titled, *Project Arachnid*: Online Availability of Child Sexual Abuse Material

Full details related to the results, methodology, limitations, and recommendations can be found in the complete report at protectchildren.ca/PAReport

The Canadian Centre for Child Protection would like to thank the international child protection entities working collaboratively within Project Arachnid to scale up its capacity to globally reduce the availability of online child sexual abuse material.

For a full list of those classifying material in Project Arachnid, visit projectarachnid.ca

© June 8, 2021, Canadian Centre for Child Protection Inc. (C3P). All rights reserved. Data relied upon to produce this report is held by C3P, and all analysis was conducted internally by C3P staff. Reasonable efforts have been made to ensure the accuracy of all information herein. The names of the companies reflected in the data are either the names as represented by the host on the site or in terms of service posted on the site, or the name of the URL if no other name could be determined. "CANADIAN CENTRE for CHILD PROTECTION" and "Cybertip!ca" is registered in Canada as a trademark of, and Project Arachnid is used as a trademark of, C3P. All third-party trademarks included within the report are the property of their respective owners, and their inclusion is not meant to imply endorsement or affiliation of any kind. Trademark symbols, if applicable, are not included in any data tables.

ABOUT THE CANADIAN CENTRE FOR CHILD PROTECTION

The Canadian Centre for Child Protection Inc. (C3P) is a national charity dedicated to the personal safety of all children. C3P operates Cybertip.ca, Canada's tipline to report child sexual abuse and exploitation online, as well as provide other intervention, prevention, and education services.

In January 2017, C3P established Project Arachnid, a web platform designed to detect known images of child sexual abuse material (CSAM) and issue removal notices to electronic service providers (ESPs) where possible.

C3P also supports survivors whose child sexual abuse was recorded and distributed online. Through our work with survivors, crucial contextual information about the nature of child sexual abuse is collected and shared with stakeholders vested in the safety and protection of children.

WHY WE WROTE THIS REPORT

C3P is issuing this report to highlight how systemic failures of the technology industry and inaction by governments have severely hindered the fight against the proliferation of CSAM on the internet. CSAM perpetuates a cycle of harm to children globally by stripping them of their personal safety and right to privacy, while inflicting great and lasting harm.

Reducing the availability of this material must be a core pillar of any child protection framework. Key to achieving this goal is a deep understanding of the role that internet-based companies, especially those that accept usergenerated content, play in facilitating access to and the dissemination of CSAM and harmful-abusive content.¹

Primary sources of data are mostly held by ESPs who are privately run, and do not tend to proactively release meaningful information about the distribution, moderation and removal of the content hosted on their platforms. This lack of transparency prevents a true understanding of the scale of the threat, and impedes the development of legislative and regulatory responses, as well as remedies for victims of these crimes or abusive behaviours.

Under these circumstances, developing sound evidence-based policies or regulation poses a real challenge. This report fills in some of the gaps by using company-specific data on the accessibility of CSAM and harmful-abusive content on certain platforms, all of which is independently collected by Project Arachnid. This report also offers a road map for governments seeking accountability on behalf of children through the responsible regulation of ESPs.

¹ The term harmful-abusive images of children encompasses all images or videos associated with the abusive incident, nude or partially nude images or videos of children that have become publicly available and is used in a sexualized context or connected to sexual commentary. It also includes publicly available images or videos of children being physically abused, tortured or restrained.

WHAT THIS REPORT IS ABOUT

This report is an analysis of three complete years (2018 to 2020) of data collected by Project Arachnid on the availability of publicly accessible CSAM and harmful-abusive content on the clear and dark web.

Project Arachnid, operated by C3P, is a victim-centred tool that crawls the open web² in search of images of CSAM. When CSAM or harmful-abusive content is detected, a removal request is sent to the ESP most likely to have the most immediate control or custody of the media. This automated process is triggered thousands of times per day.

As a result of these activities, C3P possesses records related to more than 760 ESPs across the globe. This report provides a detailed analysis of these records, as well as details on the following key transparency metrics:

- **Removal times:** Elapsed time for an image or video at a specific URL to become inaccessible on a web page once a removal request has been issued;
- Image recidivism: The rate at which images that were previously the subject of a removal notice for a respective ESP re-emerges on their service and is re-detected by Project Arachnid.

The analysis of more than 5.4 million records presented in this report is based on a sub-section of the clear web. Project Arachnid's reach does not extend to peer-to-peer sharing networks, semi-closed platforms (e.g., Facebook*, Twitter*) or membership-based sites. As a result, figures provided in this report are certain to be a gross underestimation of the true extent of availability of this material on the internet.

This report provides key insights into the central role played by lesser-known ESPs, whose lack of action or delayed action contribute to making CSAM and harmful-abusive content available online.

OVERALL FINDINGS

This report's overall findings demonstrate that expecting ESPs to voluntarily invest the resources needed to reduce the availability of CSAM is simply not working.

With unacceptably long delays for removing flagged content and previously flagged content re-emerging on websites, it is clear ESPs are collectively failing to prioritize the safety and privacy of children online.

Overwhelmingly, the results point to a need for government to enact meaningful laws, regulation and policies that impose legal requirements on ESPs and promote greater accountability, especially for ESPs that accept user-generated content.

This report and its recommendations are intended to provide governments and policymakers with certain key insights required to make informed decisions most likely to be effective in reducing the availability and distribution of CSAM on the internet.

² The open web refers to the publicly accessible areas of the clear and dark web.

KEY FINDINGS AND WHY THEY MATTER

Finding: Significant image recidivism with ESPs

Nearly half of all media (48%) that triggered the issuance of a removal notification to an ESP, had previously been flagged on that ESP's service by Project Arachnid. Certain ESPs have image recidivism rates in excess of 80 percent, meaning that in some cases offending images are repeatedly resurfacing on their systems.



WHY IT MATTERS

This finding suggests many ESPs are either not using or are making poor use of proactive media detection technology designed to block or remove user-generated CSAM content. It also suggests ESPs may not be adding flagged images to their internal banned media lists to prevent the future re-uploading of that specific content.

The finding also suggests hosting providers are not requiring, as part of their contractual agreements with customers, the use of basic proactive detection tools.



2

Finding: Lengthy removal times for many images

Over the period studied in this report, the median removal time for content targeted by Project Arachnid was 24 hours. Alarmingly, however, 10 percent of actioned media — representing thousands of victims — took seven weeks (42 days) or longer before becoming inaccessible.

It is important to recognize that calculated removal times presented in this report are based on when a notification was issued to an ESP. In reality, the media being targeted for removal were visible on the internet for an unknown amount of time prior to detection by Project Arachnid. So while the removal time upon notification is known to Project Arachnid, only the ESP knows how long the media were accessible on the internet.



WHY IT MATTERS

The finding that half of targeted media were removed in 24 hours must be considered in the broader context of the problem. In isolation, this statistic is encouraging as it suggests Project Arachnid is an effective tool for achieving relatively prompt image removals for a significant portion of targeted media. However, it masks the broader issue: Many ESPs remove media within a day of notification, but in the absence of any regulatory requirements, they have no commercial or legal interest in investing in measures to prevent the images from surfacing or re-surfacing in the first place. This is laid bare by the correspondingly high image recidivism rates.

From a victim harm-reduction standpoint, once an intimate or abusive image is published on the internet, expeditious removal times are needed to cut off further distribution or access. The victims depicted in media with longer removal times experience greater levels of harm.

This finding suggests many ESPs may not devote adequate resources to handle the volume of complaints or removal requests from the public.



3

Finding: Adolescent victims being left behind

Images of post-pubescent victims were found to have significantly longer removal times (90th percentile: 56 days for post-pubescent CSAM vs. 40 days for pre-pubescent CSAM) and higher rates of image recidivism (73% for post-pubescent CSAM vs. 46% for pre-pubescent CSAM) when compared to images of pre-pubescent victims.

Project Arachnid also detects significantly less post-pubescent CSAM (n=120,173) compared to prepubescent CSAM (n=3,403,748) and harmful-abusive content. This finding however is not likely to be representative of the true volume of adolescent material on the internet.



WHY IT MATTERS

Adolescents experience longer and repeated cycles of victimization.

Longer delays for older children suggests ESPs may view the removal of media depicting adolescents (post-pubescent CSAM) as potentially less urgent as opposed to images that visually may be unambiguously illegal (pre-pubescent CSAM).

The inherent challenge in image categorization with unidentified post-pubescent victims, the patchwork of legal standards related to CSAM and certain law enforcement practices have caused image categorization efforts to ultimately skew toward younger victims.

Since most image detection technologies, including Project Arachnid, rely on the digital fingerprints of previously verified images to uncover suspect media on the internet, the nature of what is detected through automation reflects this bias toward content depicting younger victims. The result — many adolescent victims are being left behind.

In addition to the technological barriers that exist, C3P analysts also report significant resistance from certain ESPs. In some cases, an ESP will challenge the validity of a removal notices, believing the victims depicted in the images are not minors even when the victim is known to C3P.





Finding: Single ESP linked to disproportionate amount of CSAM

Nearly half (48%) of all media detections by Project Arachnid are linked to a file-hosting service operated by French telecommunications giant Free, owned by the Paris-based parent company Iliad Group. More than 18,000 archive files, collectively containing nearly 1.1 million media files of apparent CSAM or harmful-abusive content were flagged to the company between 2018 and 2020.

In many cases, Project Arachnid's web crawler has detected links to these archived files across many areas of both the clear web and dark web. Given these many access points to the media archives, the total known availability of CSAM and harmful-abusive images on Free's hosting service surpasses 2.7 million media detections.

Beginning in 2018, C3P began corresponding directly with company officials, providing them with lists of direct links to the file archives containing CSAM being hosted on their system.

Project Arachnid has continued to detect and issue notices on newly uncovered CSAM and harmful-abusive media to the company. As of May 18, 2021, nearly 3,000 archives for which removal notices were issued between 2018 and 2020 (inclusively) were still publicly accessible, according to Project Arachnid records.



WHY IT MATTERS

The file hosting service provided by Free — offered at no cost, requiring no account and with generous storage capacity — is a popular tool discussed on both clear and dark web forums dedicated to the sexualization and abuse of children.

Rather than viewing content temporarily embedded on a web page, users must download the media, generating new copies of the files on their local computers. In this context, even if the media hosted at the source is eventually removed, several other privately held copies are likely to exist and may very well re-emerge at a later date on the internet.

In addition, the volume of images linked to Free suggests a significant portion of CSAM and harmful-abusive content is made accessible by a relatively small portion of ESPs whose service design and offerings are viewed favourably by individuals acting in an illegal or exploitative manner. This means that certain key strategic actions can cause a signification disruption in the availability of this material on the internet.

5

Finding: Dark web is main conduit for CSAM, but not preferred hosting location

The vast majority of media detected by Project Arachnid (97%) is physically hosted on the clear web. However, the dark web, specifically the Tor network (the most popular subset of the dark web), appears to act as the main conduit for directing individuals to areas on the clear web where CSAM is found.

Typically, distributors of this material will upload encrypted, password-protected archive files containing hundreds of images or videos onto a free file-hosting service. Once uploaded, the distributor will then turn to forums on the dark web and provide members with access to the direct download link and password for the archive file.

Tor, accessible only through specialized browsers, anonymizes the web traffic between a user and the website they are visiting. The process through which traffic is anonymized and encrypted, however, comes at a cost — substantially slower page loading and media download speeds. This explains why those interested in distributing large multimedia collections often choose to upload their content on archive or image-hosting services on the clear web, where download speeds are much faster.

Certain ESPs — including some specifically highlighted in this report — allow anonymous users to upload content directly to their platforms from the Tor network.



WHY IT MATTERS

The relationship between the clear and dark web is important to understand when crafting regulation designed to combat the spread of CSAM.

Failing to adopt network security measures designed to block suspect traffic, or users attempting to mask their true IP address often invites unwanted activity from individuals intent on exploiting an ESP's platform.

This finding also presents an opportunity. Since the majority of images detected are on the clear web, they can be linked to an ESP and can therefore be targeted for removal. This also means that imposing effective government regulation on the technology industry at large is likely to have a significant impact on the reduction of the availability of CSAM on the internet.

DATA ANALYSIS AT A GLANCE

Between 2018 and 2020, Project Arachnid's crawling activities have detected more than 5.4 million images or videos of verified CSAM or harmful-abusive content (**Figure 1.0**) on the services of more than 760 ESPs operating across the globe.

As of the writing of this report, C3P is facing a backlog of more than 32.8 million suspect media that have yet to be assessed. This is because the rate at which Project Arachnid detects suspect media far outpaces the human resources available to assess the content.

Over the three year period, media targeted for removal by Project Arachnid reached 626,110 (**Figure 1.0**). The significant discrepancy between media that was ultimately targeted for removal and total media detections on the internet is due to three factors:

- 1 Archive files that may contain collections of hundreds of images are often treated by Project Arachnid as a bulk removal initiative. This means a removal request may relate to several images, but represent only a single record.
- On many occasions issuing a removal notice was no longer required, since the offending media was removed or was not accessible by the time it was reviewed. This is a consequence of the assessment backlog.
- 3 Some media were found on the dark web and therefore the identity of the ESP is unknown. No action beyond data collection can occur in these situations.

Figure 1.0 Project Arachnid results summary: 2018 to 2020



5,416,713 Total verified media

Verified media targeted for removal

Ö

Removal time (median / 50th percentile)

42 days Removal time (90th percentile)



Recidivism rate

The contents of archive files (which may contain several thousand images) are not reflected in the "Verified media targeted for removal" figure since removal notices are issued for the archive file itself.

For the purposes of this report, ESP-specific information presented in **Table 1.0** is provided for those with 5,000 or more media or files that have triggered the issuance of one or more removal notice. However, for technical reasons related to the handling of media contained inside of archive-type files, records related to one ESP in particular — French telecommunications company *Free* — is tracked differently by Project Arachnid and is therefore not reflected in **Table 1.0**.

Table 1.0 shows key metrics related to the volume of media detected, removal times and recidivism rates for select ESPs. Note that the volume of media detected for a given ESP is driven by a multitude of factors, such as tips from the public, the nature of the website and the nature of the content. For these reasons, exercise caution when comparing ESP-specific figures.

³ Suspect media is derived only from websites that host known CSAM, and the term refers to any media that is reasonably suspected to be CSAM but which has not been through the assessment process.

Considering the chain of ESPs linked to CSAM

While this report highlights the intersections key ESPs have had with CSAM and harmful-abusive content, it must be noted that relying on the presented data alone does not paint a complete picture of the role each ESP's activities play in enabling access to this material.

Removal notices issued by Project Arachnid are generally issued to ESPs based on a combination of factors. Evaluations based on which ESP possesses the most immediate control over the targeted media, responsiveness to removal notices and the availability of contact information help quide where notices are ultimately sent.

The company-specific figures in this report reflect those companies to which notices are sent which is not reflective of the broader chain of ESPs associated with facilitating the public display of each detected image or video.

Figure 2.0 illustrates a crucial point – the existence of a single image or video on the internet, ultimately requires a coordinated series of services by a number of companies, all of which generally have some ability to mitigate or stop the proliferation of CSAM or harmful-abusive content on specific services.

C3P expects to adjust its data collection practices to better capture and report on the associations between ESPs and the availability of CSAM on the internet in future reporting.

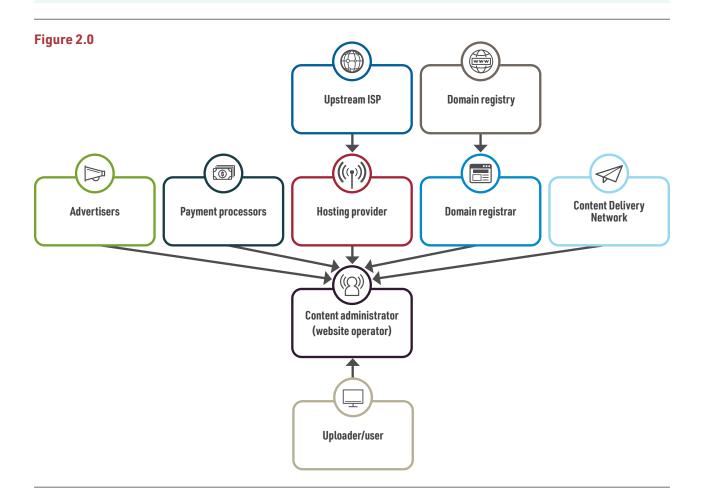


Table 1.0 shows key metrics related to the volume of media detected, removal times, and recidivism rates for each ESP.

Table 1.0

Summary of ESP-specific results: 2018 to 2020					
ESP name	Service type	Verified media detected	Removal time (median / 50 th percentile)	Removal time (90th percentile)	Recidivism rate
Imagevenue	Content administrator	143,892	1 day	1 day	87.5%
Serverel	Hosting service	72,412	6 days	60 days	93.5%
CloudFlare	Content delivery network	49,183	1 day	27 days	48.6%
Incrediserve LTD	Hosting service	39,400	3 days	53 days	34.0%
Trichan	Content administrator	34,157	1 day	138 days	26.2%
NFOrce Entertainment B.V.	Hosting service	23,211	8 days	70 days	5.9%
ImgOutlet.com	Content administrator	18,582	2 days	4 days	4.9%
ImgView.net	Content administrator	10,640	2 days	6 days	5.3%
FranTech Solutions	Hosting service	9,729	13 days	40 days	65.3%
ImgDew.com	Content administrator	9,192	2 days	6 days	5.8%
Host Sailor	Hosting service	8,740	1 day	15 days	68.6%
ColoCrossing	Hosting service	7,809	27 days	127 days	35.5%
ALFA TELECOM s.r.o.	Hosting service	7,472	1 day	4 days	93.6%
DataWeb Global Group B.V.	Hosting service	7,103	1 day	2 days	11.4%
ImgMaze.com	Content administrator	6,841	2 days	6 days	4.5%
Liteserver Holding B.V.	Hosting service	6,766	1 day	43 days	86.4%
ImageBam	Content administrator	6,339	1 day	1 day	3.2%
OVHcloud	Hosting service	6,281	3 days	23 days	11.4%

Table includes only ESPs with 5000 or more media subject to a removal notice.

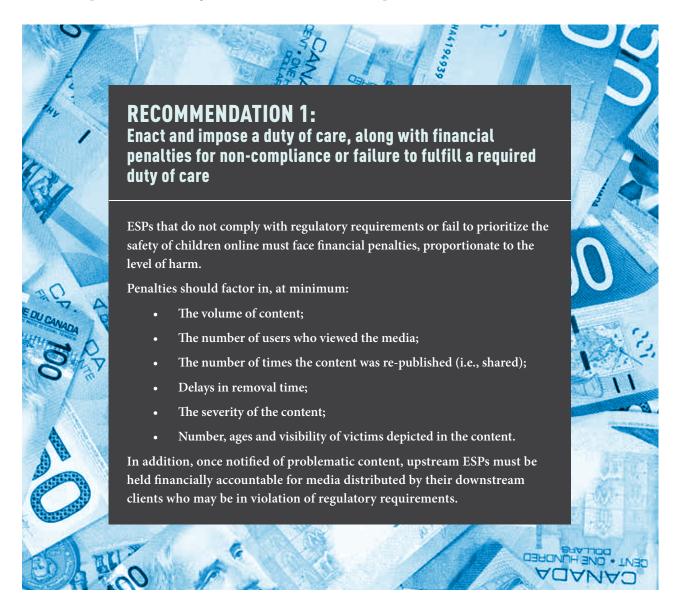
The ESP Free (dl.free.fr), which has among the greatest volume of detected media is not included in this table due to a different data management process for this specific ESP.

RECOMMENDATIONS

The borderless nature of the internet means that ESPs intent on avoiding regulation will inevitably seek a path of least resistance by structuring or restructuring their operations, or locating/relocating their content to jurisdictions willing to tolerate their activities.

For these reasons, the fight against online child exploitation requires a coordinated and international response from governments willing to adopt global standards for the distribution of content on the internet.

The following list of recommendations are based on C3Ps extensive experience in reducing the availability of CSAM and harmful-abusive content on the internet. Governments and policymakers should view these as critical components in the development of effective regulation of ESPs as it relates to the protection of children:



RECOMMENDATION 2:

Impose certain legal duties on upstream electronic service providers and their downstream customers

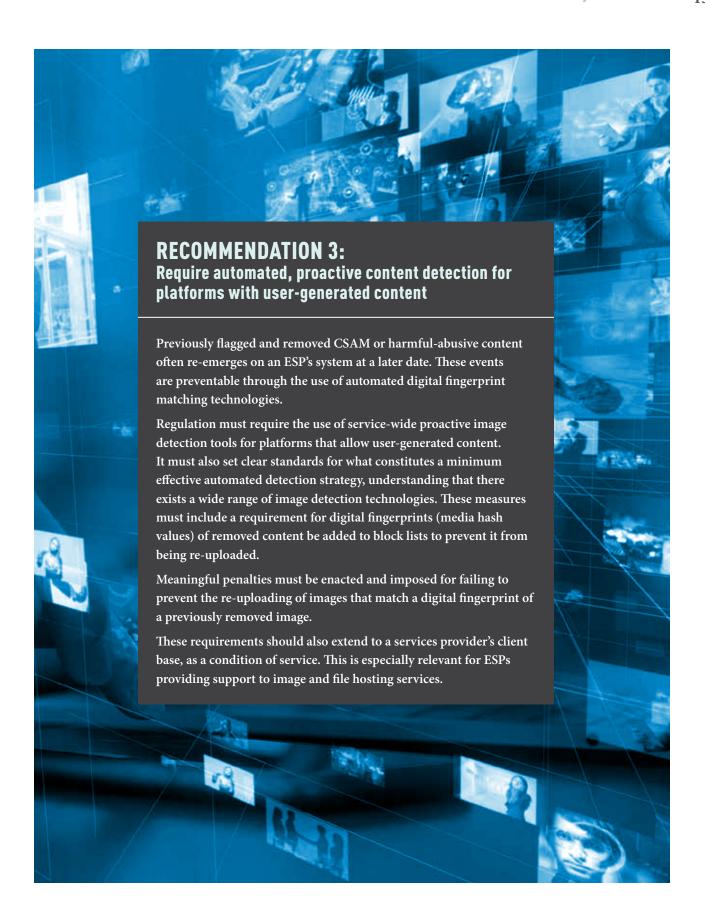
The operations of the internet traverse multiple jurisdictions and there are little to no coordinated regulatory or legislative requirements for internet based companies. Instead, the digital ecosystem is primarily structured through a myriad of complex and interrelated contracts made between various entities, each of which may be in different legal jurisdictions and have different tolerance levels for illegal content.

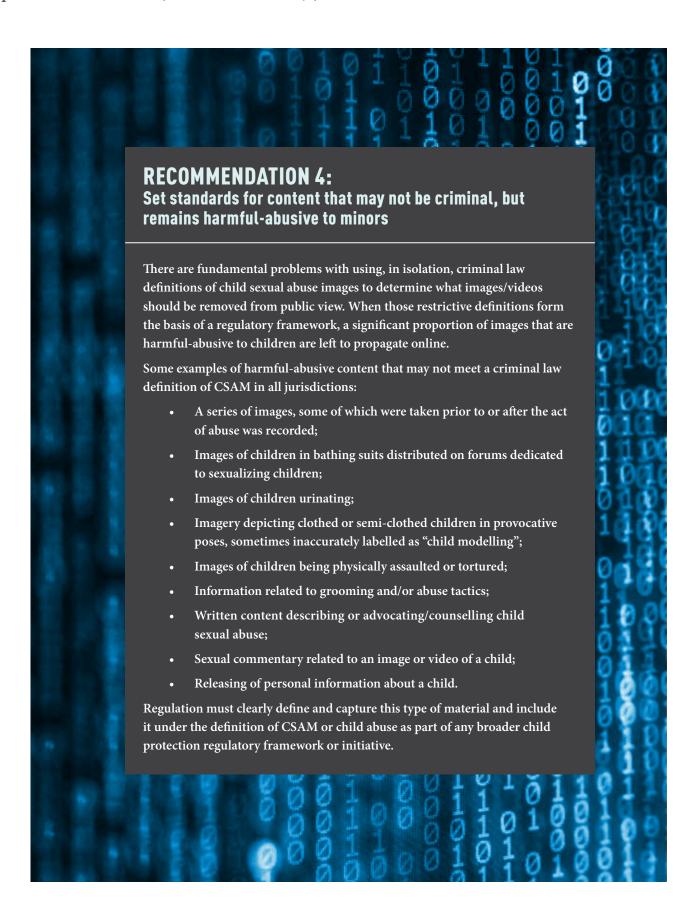
All of the companies bound by these contractual arrangements are necessary to make content ultimately accessible to an end user. As a result, to address a particular problem, every entity within the system must be bound by enforceable contractual terms that address the problem and also be required to impose and enforce similar contractual terms against its own customers. If any entity in the chain is not bound by such terms, or is not willing or able to enforce its own terms against its customers, that gap can be exploited thereby enabling the problem to flourish.

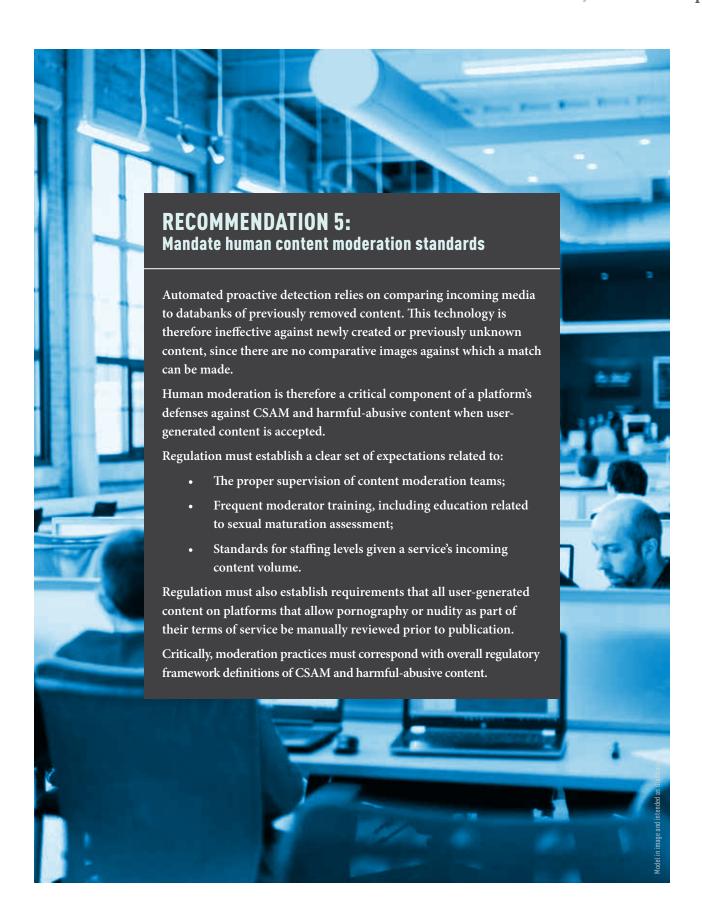
Similar to the way in which many nations have adopted legislative and regulatory control to ensure consumer protection in the areas of insurance, sale of goods and privacy issues, so too must they establish a framework to manage what the internet has become. Legislative and regulatory requirements that set out minimum base standards that are non-negotiable must be enacted. Each and every provider in the chain must be legally obligated to adhere to those base standards within their own operations, and to impose those same standards on their customers. Each ESP must be capable of being linked to at least one real person and nations must put an end to the endless legal loopholes that have enabled companies to evade legal liability for the harm they are facilitating by coordinating rules across jurisdictions.

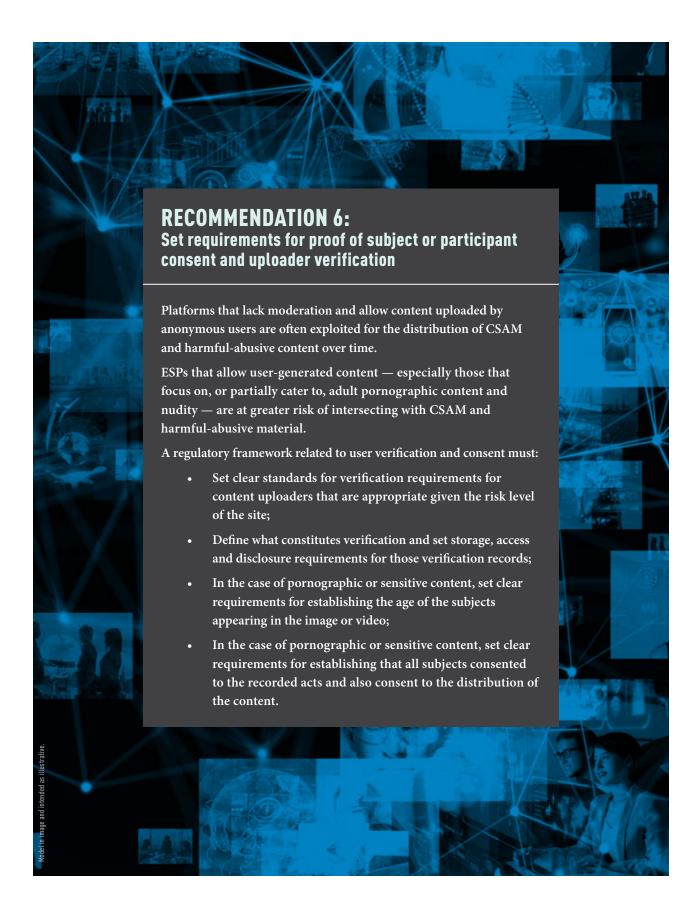
The rules must apply, at a minimum, to those that provide image or file-hosting services and include at least the following elements:

- Prescribed definitions and removal requirements for CSAM and harmfulabusive content:
- Required accountability measures to be taken by the provider in the event of illegal or harmful-abusive content being hosted by the customer of the provider;
- Significant and meaningful liability/penalties for any provider that fails to take certain actions when its customer violates the removal requirements.









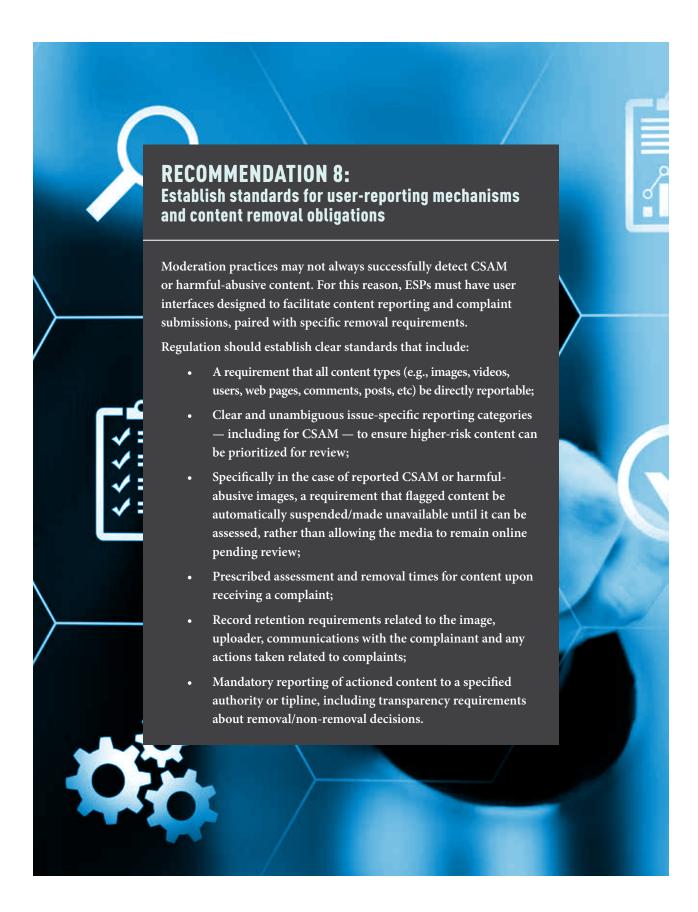
RECOMMENDATION 7:

Establish platform design standards that reduce risk and promote safety

In addition to proactive and reactive moderation measures, platforms must further reduce the prevalence of CSAM or harmful-abusive content by cultivating an environment that discourages users from exploiting their service for this purpose.

Regulation should establish requirements for:

- Prohibiting user-generated content where the uploader originates from an IP address associated with a Tor exit node, VPN service or other IP concealment techniques;
- Blocking search terms and forum/chat names that are associated with CSAM or harmful-abusive content;
- Removing or suspending accounts that distribute or access CSAM or harmful-abusive content;
- Segregating children and adults in the digital space by design. When not feasible, additional rules and protections must be implemented;
- Requiring platforms to provide an easily accessible and responsive mechanism for users to contact content administrators for lodging complaints;
- Measures, such as user age verification, for preventing children from accessing adult or mature content.



CONCLUSION

Many internet companies are failing to prioritize the safety and privacy of children online. A digital ecosystem enabled by jurisdictional uncertainty, along with a lack of clear regulation or transparency, has significantly contributed to the proliferation of CSAM and harmful-abusive content on the internet.

The findings contained in this report, which is based on three years of data collected by Project Arachnid, analyzed details on 5.4 million images or videos of CSAM and harmful-abusive content related to more than 760 ESPs.

The report established there exist high levels of image recidivism and often long delays in removal times for many internet companies. This suggests many ESPs are not deploying sufficient resources to ensure their platforms are free of, or dramatically limit, the presence of CSAM and harmful-abusive content on their services.

Other key insights discussed in this report include:

- The role the dark web plays in facilitating access to CSAM on the clear web;
- How a relatively few ESPs can have a significant impact on the availability of CSAM on the internet;
- Why statistics related to adolescent victims dramatically underrepresent the true scale of harm they experience;
- The central role lesser-known ESPs play in making CSAM and harmful-abusive content available on the internet;
- The importance of considering the broader chain of ESPs that facilitate the availability of CSAM on the internet.

The report strongly suggests expecting industry to voluntarily invest in resources to prevent the spread of CSAM and harmful-abusive content has been a failure. It points to a pressing need for consistent, enforceable and global standards that impose accountability requirements on ESPs.

Flowing from the findings, a set of eight key evidence-based recommendations are presented for governments seeking to reduce the availability and distribution of CSAM on the internet, and to adopt measures that prioritize the safety of children.

This report is both a road map and an opportunity to properly extend the duty of care we owe to children in the online world.

fodel in image and intended as illustrative.

