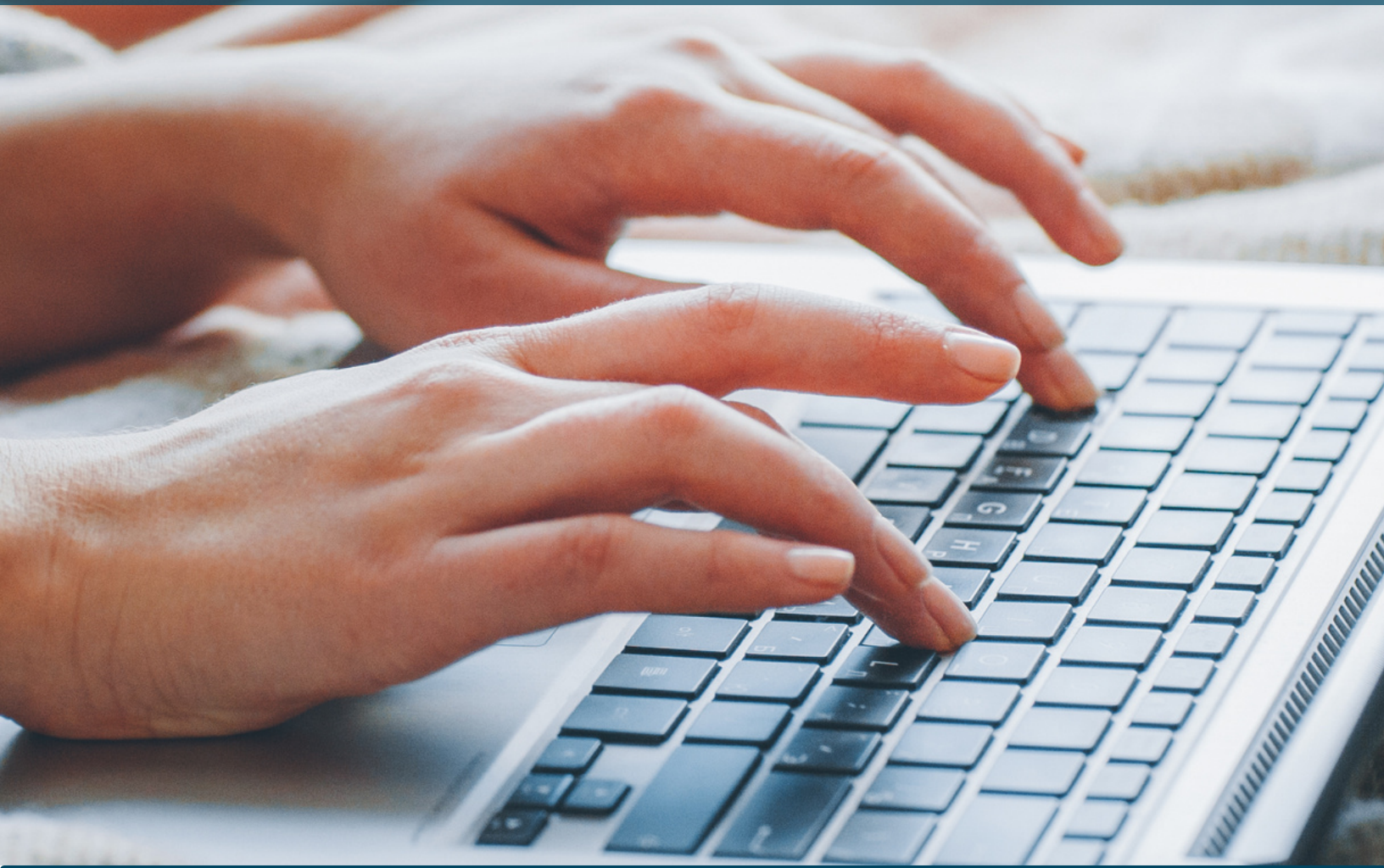


Protecting Children Online: The Importance of Content Moderation

PART 1 OF 2



A Joint Publication of:



International Centre™
FOR MISSING & EXPLOITED CHILDREN

The Issue

The digital landscape has changed dramatically in recent years, and the COVID-19 pandemic has significantly changed the way children use the internet. With many schools closed or operating remotely, children have become more reliant on digital technology for learning and socializing with friends.

The pandemic has also accelerated the shift toward increasing children's dependence on digital technology. With the widespread use of smartphones and social media platforms, children are now able to access information and connect with others from around the world at their fingertips. This has led to increased screen time and online activity, as well as greater exposure to the risks associated with online behavior.

Though the internet offers many benefits and opportunities for children, there are also many risks to navigate. Online risks for children include exposure to inappropriate content, cyberbullying, online grooming by predators, and addiction to technology.

According to a 2021 study by UNICEF, approximately 1 in 3 internet users worldwide are children, and 1 in 5 children have experienced cyberbullying. The same study found that 71% of children aged 3-4 years globally use the internet, while 92% of children and youth aged 15-24 years are online.

In 2021, the United States National Center for Missing & Exploited Children (NCMEC) received over 29 million reports of suspected Child Sexual Abuse Material (CSAM) online, a 35% increase from the previous year. According to a 2020 report by the Internet Watch Foundation (IWF), 68% of the online sexual abuse material they identified featured children under the age of 10 years, with a 360% increase in this type of imagery of 7- to 10-year olds in the first half of 2022 alone. These statistics show that online child safety is a significant issue globally, and efforts are needed to protect children online.

Parents concur. A 2021 report by the Pew Research Center found that 81% of parents in the United States are concerned about the potential risks their child may face online, including exposure to inappropriate content and online predators. While a poll of global internet users found that 91% of respondents felt strongly pessimistic about the current situation of child safety and protection of rights in the United States.

Approximately 1 in 3 internet users worldwide are children.



The Need for a Trust & Safety Function

Now more than ever, Trust & Safety (T&S) functions within technology companies are crucial to ensure the protection and well-being of children on digital platforms. With the growing use of social media and other online platforms, children are increasingly relying on these platforms to connect with others, share their opinions, and engage with content.

This increased reliance on digital platforms also means that there is a greater risk of encountering harmful or abusive content.

T&S teams play a critical role in mitigating these risks by developing and enforcing policies and guidelines to keep children safe. They also work to identify and remove harmful content, prevent abusive behaviors, and provide support to children who have been affected by harmful content or behavior. This work is often carried out in partnership with third-party companies specialized in hiring and deploying appropriately skilled people, or developing specific technology, for policy enforcement.

In doing so, T&S teams help to build trust among users, particularly children, who are more likely to continue using a platform if they feel safe and supported. That is why the recent downsizing and elimination of T&S teams across the technology industry is so alarming.

Without these specialized T&S teams, how can the tech industry create and maintain a safe online environment for children?

Aware that the pandemic has created new opportunities for bad actors online to exploit children's vulnerabilities, responsible companies in the tech industry have taken steps to protect children online, including introducing more effective age verification measures, adding content moderation coverage to deal with the increase in the volume of violations, and optimizing reporting systems.

Many social media platforms have age restrictions and require users to verify their age, and some platforms have implemented features to help prevent cyberbullying, such as comment filters and reporting tools. Screen time limits for child users of social media are also being implemented.

Additionally, some companies have developed parental control tools to help parents monitor and restrict their children's online activity. However, there is still room for improvement, with a multi-faced response to these complex challenges.



Platform Policies and Content Moderation

– An Important Tool

Digital platforms often have specific community guidelines to cover what type of content, conduct, and contact is allowed on their sites, in order to protect users and foster a safe online community.

Child safety policies cover topics such as prohibiting the creation and distribution of CSAM. Online grooming, defined by [Thorn](#) as “tactics abusers deploy through the internet to sexually exploit children,” is often specifically called out as a prohibited behavior in major platform policies. Other policies, such as those covering harassment and bullying, are applicable to all users, including children.

Content moderation – the review of user-generated content against platform policies, conducted through a combination of human and technology review – is essential for effective policy enforcement. Platforms often differ in their specificity around child safety policies; this requires a balanced approach. Specificity is good so that there is clarity around what is and isn’t allowed. However, too much specificity can be disempowering for content moderators in making accurate judgements based on full context and following the spirit of the policy.

Content moderation helps protect children online by removing or blocking harmful content, such as cyberbullying, hate speech, and inappropriate or sexually explicit material, from online platforms.

By doing so, content moderation creates a safer online environment for children by reducing children’s exposure to harmful content and behavior. This can also help prevent re-victimization: by quickly removing known violating content, platforms prevent continued harm to children that is present when content continues to circulate across multiple platforms.

Platforms that proactively detect violations and work with partners to enforce policies in a timely and effective manner, can significantly improve trust on their platform. This can help promote positive online behavior and encourage users to engage in respectful and appropriate communication. Thought leadership produced by [MIT Technology Review Insights](#), “Humans at the center of effective digital defense,” in association with Teleperformance highlights the essential work of content moderators and the broader content moderation landscape in more detail.

References:

- UNICEF, Children in a Digital World 2021, at <https://www.unicef.org/media/113106/file/ChildreninaDigitalWorld2021StatisticalBrief.pdf>
- Patricia Davis, NCMEC, Borderless: A Series on the Global Battle to Protect Children Online, UNITED STATES NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN (NCMEC), Feb. 2, 2023, at <https://www.missingkids.org/blog/2023/ncmec-leads-global-fight-to-protect-kids>
- IWF, 20,000 reports of coerced ‘self-generated’ sexual abuse imagery seen in the first half of 2022 show 7- to 10-year-olds, at <https://www.iwf.org.uk/news-media/news/20-000-reports-of-coerced-self-generated-sexual-abuse-imagery-seen-in-first-half-of-2022-show-7-to-10-year-olds/>
- Risa Gelles-Watnick, Explicit content, time-wasting are key social media worries for parents of U.S. teens, PEW RESEARCH CENTER, Dec. 15, 2022 at <https://www.pewresearch.org/fact-tank/2022/12/15/explicit-content-time-wasting-are-key-social-media-worries-for-parents-of-u-s-teens/>
- CGTN, Over 90% feel pessimistic about current state of child protection in the U.S., Mar. 4, 2023, at <https://news.cgtn.com/news/2023-03-04/CGTN-poll-Over-90-pessimistic-about-child-protection-in-the-U-S--1hU1sOCMScw/index.html>
- The Verge, Elon Musk’s Twitter layoffs leave whole teams gutted, at <https://www.theverge.com/2022/11/4/23439790/elon-musk-twitter-layoffs-trust-and-safety-teams-severance>
- Cormac Keenan, New features for teens and families on TikTok, TIKTOK NEWSROOM, Mar. 1, 2023, at <https://newsroom.tiktok.com/en-us/new-features-for-teens-and-families-on-tiktok-us>
- eSafety Commissioner, Safety by Design, at <https://www.esafety.gov.au/industry/safety-by-design>
- Thorn, What is Online Grooming, at <https://www.thorn.org/blog/online-grooming-what-it-is-how-it-happens-and-how-to-defend-children/#:~:text=Online%20grooming%20is%20a%20term,to%20keep%20a%20child%20silent.>
- MIT Technology Review Insights, https://www.teleperformance.com/media/p5unc5w/gmt_mit_wp_humans-at-center-of-effective-digital-defense_december_2022.pdf

A Joint Publication of:

