Trust

Transparency Reporting Implementation Guide

TECH COALITION

TRUST
Voluntary Framework for
Industry Transparency

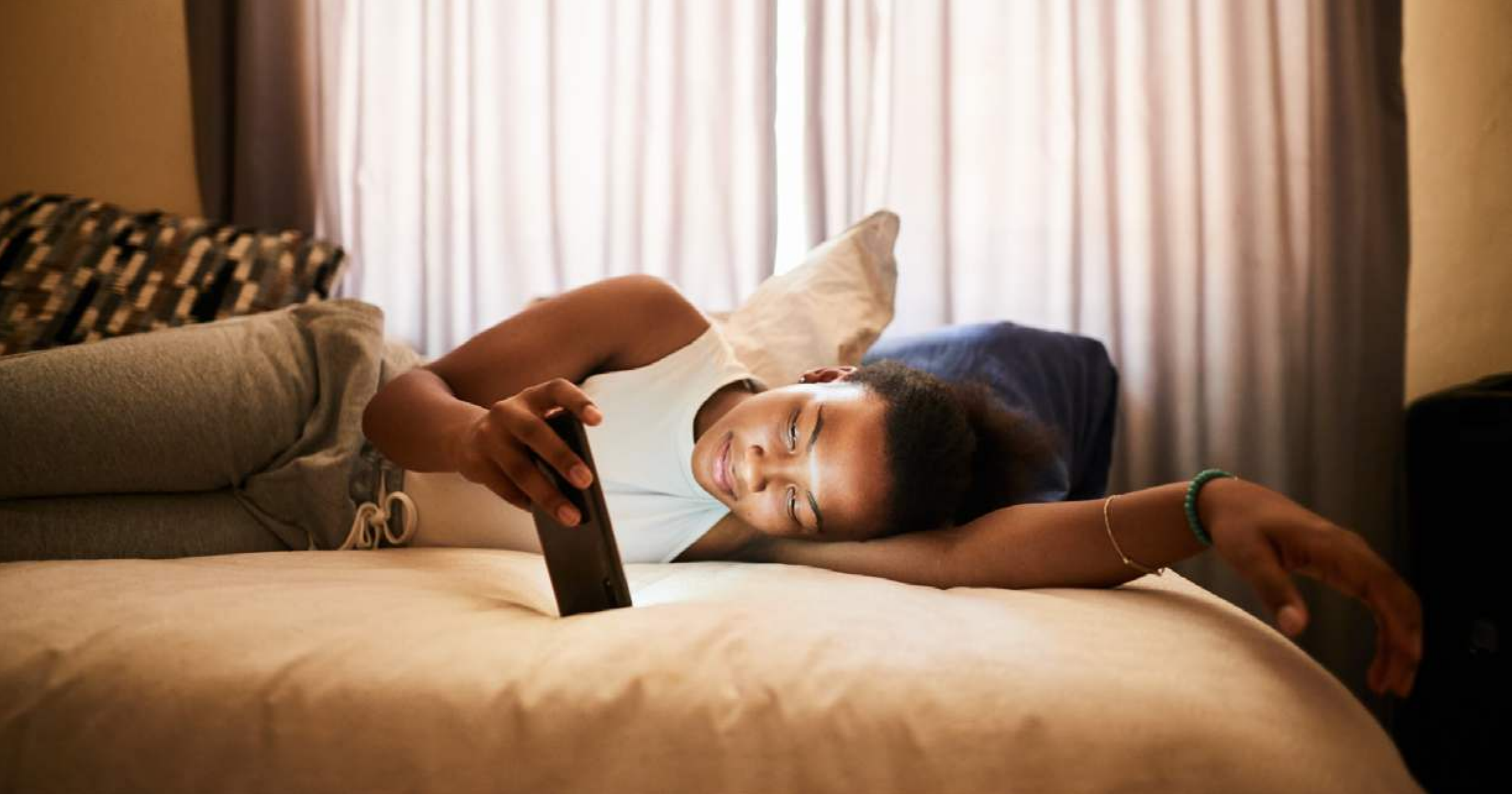
Table of contents

Introduction

The Tech Coalition's Trust: Voluntary Framework for Industry Transparency (the Framework) provides principles-based guidance to tech companies seeking to build trust around their efforts to address online child sexual exploitation and abuse (CSEA) risks on their services.

The Framework draws on the experience of Tech Coalition members, multi-stakeholder conversations, and extant practices in transparency reporting in relation to online harms.

This guide serves to aid industry in implementing the Framework and moving toward better alignment in transparency reporting. A reporting template is also available to help companies get started on their first child safety transparency report.



Overview

Transparency reporting in this context refers to reports that explain a company's approach to addressing online child sexual exploitation and abuse (CSEA), which should highlight the company's policies, explain its processes, and document the outcomes of its efforts.

Transparency is an essential component of industry efforts to combat online CSEA. It drives accountability and plays a critical role in building trust with users, regulators, and the general public. The importance of transparency reporting in helping advance this fight is recognized in the **Voluntary Principles to Counter Online Child Sexual Exploitation and Abuse:** Principle 11 of that document reflects the need for companies to regularly publish data or insights on their efforts.

This Transparency Guide includes sections on:

1. Why Transparency Matters
2. Audiences for transparency reporting
3. Principles to guide transparency reporting
4. Step by step transparency report process

1. Why Transparency Matters

Transparency reporting creates an opportunity for companies to provide information to the public about corporate governance, values, and priorities, which can build consumer trust. Publishing this information can then support communications and policy teams in effectively engaging with regulators, policy makers, and other external stakeholders. A transparency report can provide statistics that can be critical in helping tell a company's story on how it approaches and addresses a particular public policy issue or topic, like child safety. The process of developing a transparency report also can serve as an internal assessment tool for a company to take inventory of practices that work and where improvements can be made.

Company transparency reporting on online CSEA serves three main purposes:

1. Explains a company's policies and actions to address the risk of harm resulting from online CSEA;
2. Demonstrates accountability for a company and develops trust by showing how it has implemented those policies and actions, and driven any improvements; and
3. Shares knowledge and meaningful data to support the global fight against CSEA.

We explore the value of transparency reporting in detail in a Tech Coalition Member Resource: Why Transparency Matters. Join the Tech Coalition to access this and other member resources.

Reports initially were undertaken by the tech industry to provide more information to users on government surveillance requests and other government requests for user data.

Transparency reporting has expanded for many companies to include information related to terms of service enforcement (which can include reporting on child safety efforts) and other areas. This guide will focus primarily on the issues related to transparency reporting on a company's child safety efforts; however, many of these logistical issues may also be relevant for other areas of transparency reporting.

In the area of child safety, regulatory agencies around the world are considering transparency reporting as a key component of best practice measures and/or regulatory compliance for online platforms. For example, the Five Country Ministerial Voluntary Principles to Counter Online Child Sexual Exploitation and Abuse contains a specific principle on transparency reporting (see Principle 11: Companies seek to regularly publish or share meaningful data and insights on their efforts to combat child sexual exploitation and abuse), and in 2020, NCMEC published its first ever transparency report documenting how many CyberTips were filed by various providers.

Examples of existing transparency reports specific to child safety or with child safety components include:

- Adobe
- amazon
- CLOUDFLARE
- Discord
- Dropbox
- GIPHY
- Google
- MEGA
- Meta
- Microsoft
- Pinterest
- Snap Inc.
- TikTok
- twitch
- X
- verizon
- yahoo!
- yubo
- zoom
- TECH COALITION

2. Audiences for Transparency Reporting

While a transparency report may be reviewed for a wide range of purposes, the primary audiences include:

- **The general public and users (including parents)** – who are able to better understand the policies and actions of the services they or their children use;
- **Victims and survivors** – who can see what is being done to prevent and limit potential harm;
- **Government representatives, ranging from policy-makers to law enforcement** – who are able to see what actions companies are taking to address a societal harm;
- **The child safety community, including child safety non-government organizations and other civil society groups** – who are able to review actions being taken and identify effective techniques, as well as opportunities to better fight online CSEA; and
- **Other companies** – who can learn from others, share best practices, and help normalize dialogue about this challenge.

While transparency reports are not generally directed to children, children are at the center of why companies should provide transparency reports. Companies should separately provide appropriate resources to help children, their families and/or caregivers understand and mitigate the risks of CSEA on their service.

3. Principles to Guide Transparency Reporting

Trust: Voluntary Framework for Industry Transparency (the Framework) has been developed by the Tech Coalition to provide principles-based guidance to tech companies seeking to build trust around their efforts to address online child sexual exploitation and abuse (CSEA) risks on their services. It is a voluntary framework, drawing on the experience of Tech Coalition members, multi-stakeholder conversations, and extant practices in transparency reporting in relation to online harms. In joining the Tech Coalition, member companies have demonstrated their commitment to combating CSEA, and to their accountability for those efforts.

The Framework aims to:

- Encourage companies to provide online CSEA transparency reporting;
- Support the development and improvement of transparency reports by providing a variety of options; and
- Increase consistency across reporting, to better enable information-sharing and accountability.

The Framework sets out five principles as a general basis for considering how to approach transparency reporting. Given the diversity of online services available, a principles-based approach provides companies with the ability to tailor an individual strategy while still achieving the overall goal of providing greater accountability and information about their efforts to address CSEA. Each service may face unique risks, depending on its purpose and features, user base, business model, and a range of other factors. Effective practices for one service will not necessarily suit another, and highly prescriptive approaches to trust and safety practices may be too narrow, or have unintended consequences.

This Framework aims to recognize the diverse service landscape and provide flexibility for different companies to adapt their transparency reporting practices, while also defining a common framework. This commonality is important, because without some consistency in measures and approach, it becomes challenging to maximize the value of transparency reporting and to get a complete picture of efforts across companies and begin to understand prevalence and other key trend data.

1. Reporting should support trust and accountability
Transparency builds trust and demonstrates a willingness to be held accountable for decisions and actions. Transparency reports should therefore be designed to build trust with users, suppliers, employees, investors and government authorities by demonstrating that the company has appropriate policies and procedures, and is applying them consistently and fairly. The data-gathering process should also provide insights that allow a company to continuously improve their policies and practices and systems and processes. And, as outlined earlier, transparency reporting on CSEA supports efforts to address an important, whole-of-society issue.

2. Reporting should reflect the unique nature of each company's service(s)
There is a diverse range of digital products and services that may be impacted by CSEA, ranging from social media, to communications, gaming, productivity, cloud storage, infrastructure provision, to newer concepts like the "metaverse," and beyond. Each service differs with regard to how risks manifest on its platform, the specific steps it can take to address the risk, and the results of those protective measures. Each service also has its own unique technology infrastructure. Thus, each service's transparency reporting efforts should be proportionate and tailored to its specific business case, risk profile, practices, and technology.

3. Reporting will depend on service maturity
Every company is at a different stage of maturity (broadly defined) and capacity. Maturity may also differ among differing services offered by a single company. Transparency reporting follows the development and implementation of risk mitigation, content policy, tools and process. It requires building out data collection, reporting and analysis capacity to produce the relevant information. New and smaller companies or services may prioritize building their risk mitigation capacity (including child-safe design and other elements of safety by design) and should be given adequate time to develop the capacity to produce and publish transparency reports. As companies and services mature, they should equally seek to mature and grow their transparency reporting capability in a proportionate manner. Companies may also seek to go beyond the Framework as their business and approach matures.

4. Reporting should be regular and evolve over time
Companies should aim to provide reporting on a regular cadence: whether annually, biannually, or quarterly. More frequent reports provide an opportunity to share more up-to-date information, so companies should aim to report at least annually. Companies should also seek to provide comparative information and metrics across periods to enable readers to track trends and to draw comparisons over time. Having said that, transparency reporting is a dynamic and iterative process. Each company begins transparency reporting with the resources at hand and builds from there, gaining insights from both internal and external feedback. Companies should seek views and feedback from a range of stakeholders.

5. Reporting should not compromise privacy or safety
Companies should strive to be as open as feasible, while not compromising other important interests, including privacy and safety. When data is appropriately aggregated and anonymized, the provision of transparency reporting provides important visibility into the way a company is protecting the rights of its users, including the rights of children. However, transparency reporting should not in any way infringe on individual privacy. Great care should be taken to ensure that any data provided cannot be tied to specific individuals. Transparency reporting should also be carefully calibrated to ensure that the information released does not compromise a company's safety efforts and inadvertently enable bad actors to subvert safety measures.

4. Step by Step

Step 1: Tracking Metrics in a Case Management System
The first step in transparency reporting is to build or leverage a case management system that allows your team to log and pull data in a structured manner. Companies in the earlier stages of their maturity model may opt to log their cases manually in a spreadsheet or in other web-based platforms. Companies further along in their maturity model may have a way to automatically pull these metrics based on how an agent resolves a case. It is important to consider the operational impact (e.g., impact on case resolution or indexing times) that the data collection process can have on your team.

1. EARLY STAGE
Spreadsheet solutions with minimal automated workflows and only manual mechanisms to pull data
2. EMERGING STAGE
Third-party solution or in-house solution with some customization available; semi-automated reporting (manual pulls still required)
3. MATURE STAGE
Fully customized, in-house solution with automated metric retrieval methods

See Appendix A: Considerations for your Case Management System for example database fields and how they may correlate to metrics that are included in your transparency report.

Step 2: Planning your transparency report
The next step in this process is determining what information and data should be included in your transparency report in order to provide a fulsome picture of the child safety detection and reporting mechanisms in place at your company.

The Tech Coalition's Trust: Voluntary Framework for Industry Transparency is designed to help companies organize reporting in a way that will allow different audiences to understand how companies combat CSEA on their service(s).

The Framework recommends a reporting structure that:

1. covers policies and practices, allowing a company to describe their approach to CSEA and what they prohibit on their services;
2. provides a descriptive summary of the processes and systems the company uses in combating CSEA;
3. includes numerical reporting on the outcomes of the company's overall approach.

The Tech Coalition has created the *Trust: Transparency Reporting Template* to help companies get started on their first child safety transparency report

Download and complete a *Trust: Transparency Reporting Template* here.

Something to consider as your company embarks on the path of creating or updating its child safety transparency report is whether to publish company-wide data or data broken down by product or area. In making this decision, one thing to keep in mind is balancing the goal of offering greater transparency without providing detail in such a way that it may have the unintended consequence of allowing offenders to elude detection.

As noted above, one of the key purposes of transparency reporting is to create an opportunity for companies to provide information to stakeholders about corporate governance and values. Presentation and visualization of both a company's core policies and principles, and statistics documenting outcomes, are critical in this effort.

The next step in this process is for a company to consider what data to publish in a transparency report. In making this decision it is important to remember that once you publish a transparency report, the expectation will be that you will continue to publish transparency reports in regular intervals (e.g. semi-annually or annually as described above).

This information will need to be updated regularly, so it may be prudent to keep at least the first iteration of a transparency report streamlined so as to assure repeated, updated publication of the statistics is possible in the future.

Facts and statistics to include in a streamlined model could be the following:

- Number of CyberTip reports sent to NCMEC or relevant authorities
- Number of accounts disabled for CSAM and other child exploitative conduct
- Number of pieces of content removed or sites deindexed, if applicable

Step 3: Ensuring Executive Buy-In

Obtaining internal executive buy-in for the launch of a transparency report will require identifying key stakeholders and ultimate decision-makers (e.g. GC, CEO, etc.) in your company. In order to ensure leadership is aligned with the transparency strategy, it is important to highlight the transparency landscape for these key individuals and outline the pros and cons, including:

Pros

- Facilitate better understanding of the nature of problem
- Better control your company's narrative and get ahead of questions from regulators/media
- Maintain parity with (or surpass) peer companies in the same industry
- Promote public trust and be part of the solution to a difficult public safety challenge

Cons

- May lead to unwanted PR/attention for a difficult subject
- May invite more detailed questions and scrutiny
- Will require continued investment to maintain transparency reporting

Step 4: Launching the report with a communications strategy

It is important to develop a proactive and reactive strategy with your communications team and also local heads of offices, so they are prepared to answer follow-up questions from the media and regulators.

A blog post along with the release of a transparency report is a common method to express trends and accomplishment of goals. There are many other formats in which company policies can also be expressed such as Community Guidelines, FAQs on a Transparency Report, internet safety resources website, a general policies page, privacy policy, or terms of service.

Additionally, your company may want to consider including an accompanying FAQ section with the transparency report. The FAQ section could proactively answer common questions that may come up due to the publication of this additional child safety related data. This will also likely help your communications team who may otherwise face common and repeating questions pertaining to understanding the transparency report.

The FAQ section could answer questions such as:

- What is CSAM?
- What does your company do when it detects CSAM (or other online child exploitation) on its platform?
- Where does your company report CSAM? Here would be a good place to explain reporting statute requirements governing providers and what NCMEC is.
- What is a CyberTip report?
- What are ways to report illegal content or takedown requests to your company?

The FAQ section can also link to any blog posts, acceptable use policies, or community guidelines generated by your company, describing policy positions on child safety related matters. Because transparency reports need to be regularly updated and are heavily data driven, having a streamlined approach to what is included in those reports — and augmenting those with relevant blog posts that expand more on child safety policies your company puts forward — is something to consider. A blog post may also be a good place to consider highlighting examples of success stories, for example, incidents where children were rescued or an offender network disrupted. For reference, look at Tech Coalition member transparency reports for sample FAQ sections.

If it is necessary to provide a reactive statement to the launch of a transparency report, consider the following example:

> *[Company X] seeks to eradicate CSAM content from our platform(s). We leverage a combination of automated scanning and expert human review to detect and remove this abhorrent material from our platforms. We also support cross-industry efforts to combat CSAM, including the Tech Coalition, of which [Company X] is a member. Ensuring a safe user experience for our customers, including children, has always been one of our core values, and one that we will continue to pursue with rigor.*

Step 5: Iterating on your Transparency Report
Iteration is a key part of any transparency report program. Set yourself up to develop successful iterations of your report by considering the following practices.

Cadence: Many companies currently publish transparency reports in other areas (i.e. government data requests) every six months, and as noted above, NCMEC has begun publishing CyberTips reporting statistics by company on a yearly basis.

Website Analytics: Consider collecting analytics on your Transparency Report pages to understand how readers are entering and navigating through your Transparency Report. Consult with Legal or your Privacy team to understand the types of information that you can collect.

Strategic Planning: To ensure that you continue to reduce the manual hours required to launch a report, consider building your transparency report resource requests (engineering, site development, ops, data analyst) into your strategic planning discussions within your company.

Feedback: Engage with your communications team to understand how civil society groups, customers, regulators, and the press are reacting to your published report. Use the feedback as an opportunity to think about new types of metrics to potentially include in the next version of your report.

Appendix A: Considerations for your Case Management System

Please note that the information provided here is only a suggestion and that different companies will have to tailor their case management system based on their products/ services, their existing staffing, and the types of metrics that they want to collect internally (for operational reasons) and publish externally. It is also worth noting that some of these fields are suggested for internal tracking purposes only, particularly to help identify trends and product and/or detection enhancements that may be appropriate. Tips for effective data collection:

- Use Clear Naming Conventions for your database fields.
- Avoid Storing Multiple Values in a Single Field.
- Document internal policies for collecting data.

TYPE	FIELDS IN CASE MANAGEMENT SYSTEM	METRICS MAINTAINED
Unique case ID	• Unique case ID	N/A
How the content was discovered	• Detection Source (xDNA, internal classifier, manual report) • Match source (NCMEC NGO list, NCMEC industry list, keywords list) • Service/product where the match was detected	• CSAM reactively surfaced (user or third party report) versus proactively surfaced (e.g. hash matching or machine learning classifier) • Number of CyberTip reports calculated on a company-wide level versus a product or service level
Matched asset information	• Asset Type (text, image, or video) • Match classification (A1, A2, B1, B2) or any other classification your company chooses (memes, involving infants, beastiality, etc.) • Whether the asset is new CSAM identified (content not in a hash list) • Classification of the offense for each asset (e.g., CSAM, grooming, sextortion, trafficking) • Types of files/content that were preserved	• Number of CSAM detected broken down by file type — such as by video, images, text, audio, gifs etc. • Number of reports broken down by CSAM file classifications, such as A1, A2, B1, B2, or any other classification your company chooses to provide to NCMEC (e.g. meme imagery, CSAM involving infants, beastiality, CSAM that has gone viral) • Number of reports categorized by type of offending content (e.g. CSAM, grooming, sextortion, trafficking) • Number of new CSAM content identified, which can lead to the identification of new victims and hashing of new CSAM
Moderation actions taken	• Moderation actions taken on the reported content (including under what policy) - (e.g., removal, deletion, deindexing from search) • Moderation action taken on the account that posted the content (including under what policy) include account warning, temporary or permanent suspensions	• Number of accounts disabled or deactivated for CSAM • Number of removals or account disables for material other than CSAM but that may still be harmful to children and violative of your terms of service or acceptable use policy. • Amount of content removed at the domain level or deindexed from surfacing on your search engine (if applicable)
CyberTips filed	• Field indicating whether a CyberTip was reported to NCMEC or international equivalent • Field indicating whether a supplemental report was filed	• Number of CyberTips sent to NCMEC (this could be broken down by number of CSAM imagery files, number of reports, or overall number of accounts/URLs reported) • Number of supplemental CyberTips filed — e.g. reports that are generated after further investigative work is done on an initial CyberTip report
Outcome	• CyberTip outcome data (active investigation, arrest data, children saved) • Follow-up legal process per CyberTip and the legal process type • Country where the CyberTip was submitted to **Note:** For important cases and outcomes, we suggest also having your team save illustrative examples that you may want to highlight to your department leadership, executive team, or externally.	• Number of child rescues generated from your CyberTip reports (if given feedback from NCMEC and/or law enforcement agencies) • Amount of legal process received associated with your CyberTip reports (e.g. search warrants, subpoenas, court orders, MLAT requests etc...) • Country where each CyberTIp was submitted)
Appeals	• Appeals metric if available (volumes of appeals received, number of moderation decisions overturned)	• Number of appeals and whether the appeals were approved or rejected

Appendix B: External Resources

Voluntary Principles to Counter Online Child Sexual Exploitation & Abuse, Five Country Ministerial

Santa Clara Principles on Transparency and Accountability in Content Moderation 2.0

The Transparency Reporting Toolkit, Harvard Berkman Center

Voluntary Transparency Reporting Framework pilot, OECD

TECH COALITION

About Tech Coalition

The Tech Coalition facilitates the global tech industry's fight against the online sexual abuse and exploitation of children. We are an alliance of technology companies of varying sizes and sectors that work together to drive critical advances in technology and adoption of best practices for keeping children safe online. The Tech Coalition convenes and aligns the global tech industry, pooling their knowledge and expertise, to help all our members better prevent, detect, report, and remove online child sexual abuse content. This coalition represents a powerful core of expertise that is moving the tech industry towards a digital world where children are free to play, learn, and explore without fear of harm.

To learn more visit www.technologycoalition.org