

CSAM Data

Factcheck of recent European
Commission statements

Maaïke van Kesteren
Michel van Eeten
Rolf van Wegberg

Contents

1 Introduction	3
2 Statistics Cited Around EC Proposal	5
2.1 Victims of sexual abuse in all aspects	5
2.2 Victims of distribution of online sexual abuse material	5
2.3 The role of Europe and the Netherlands	6
2.4 Scale of CSAM and trends	6
3 Reflection	9

1

Introduction

New EU legislation to stop and prevent child sex exploitation online was suggested by the European Commission on May 11, 2022.¹ With 2022 being the European Year of the Youth, the European Commission (EC) has stated a dual goal: Providing better internet for children and preventing and combating child sexual abuse and exploitation.² The aim is to ensure that children are protected, empowered and respected online.

Part of controversy around the proposed legislation are the proposed measures concerning encryption. The EC stated that most of the Child Sexual Abuse Material (CSAM) is shared in interpersonal communication and therefore should be detected by service providers and hosting companies. In an interview with the Dutch newspaper Trouw, European Commissioner Ylva Johansson argued that the spread of CSAM is migrating from websites to closed, and often encrypted, private message platforms such as on Whatsapp or Telegram.³

In light of this statement of migration of distribution techniques, the Expertisebureau Online Kindermisbruik (EOKM) has asked Delft University of Technology to create a brief overview of the different statistics and statements used in the debate, as well their origin and factual accuracy.

Section 2 will provide an overview of some of the reported statistics and state their source and representation in media. In addition, the statistics and methods reported by the most important institutions within the online CSAM landscape will be presented and compared, in which their assumptions and limitations will be highlighted. Section 3 presents some brief conclusions.

¹See <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022PC0209>

²See https://ec.europa.eu/commission/presscorner/detail/en/qanda_22_2977 and <https://audiovisual.ec.europa.eu/en/video/I-224942>

³<https://www.trouw.nl/nieuws/pakt-nieuwe-eu-wet-kindermisbruik-aan-of-zorgt-deze-voor-massasurveillance~b4fe2f3c/>

2

Statistics Cited Around EC Proposal

In the debates that take place amongst the many parties united in the fight against CSAM, plenty of statistics and statements are presented. It is often unclear from which sources these originate. In addition, some statistics require nuance, whereas others are simply misrepresented in media. Therefore, this section will take several statements and attempt to identify their source. If we can identify the original source, we will compare the statement to that source and check how the statistic is represented in the media. These statements originate from the press conference on a package on children rights by Dubravka Šuica, Vice-President of the European Commission, and Ylva Johansson, European Commissioner, as well as from the above-mentioned interview with Commissioner Ylva Johansson in newspaper Trouw.¹

2.1. Victims of sexual abuse in all aspects

According to the Council of Europe, 1 in 5 children is victim of sexual abuse.² This includes all forms of sexual abuse, such as unwanted touching, rape, sexual assault, grooming, exhibitionism, exploitation in prostitution and pornography and online sexual extortion and coercion. In 70-85% of the cases, the child is abused by someone they know, like a family member or sports coach. The press conference represents this statement correctly. However, in the interview it is stated that 1 in 5 children is victim of *online* sexual abuse. This is a misrepresentation of the source.

Statement: 1 in 5 children is victim of sexual abuse online – False

2.2. Victims of distribution of online sexual abuse material

In both the interview and the press conference it is stated that 85 million photos and videos of CSAM have been intercepted in 2021.³ This is indeed the number reported by the National Center for Missing and Exploited Children from the U.S. The total amount of intercepted content worldwide is evidently larger. It is important to note that in the US 29.1 million of the 29.3 million reports came from Electronic Service Providers (ESPs). ESPs form a large set, but they include amongst others Facebook, Google, Reddit and Discord. In other countries in the world, ESPs are not obliged to detect and report CSAM. Of the reports in the U.S., 0.5 million came from Snapchat, 1.3 million from Whatsapp, 3.4 million from Instagram and 22.1 million reports, by far the most, came from Facebook.

Statement: 85 million photo and videos of CSAM have been intercepted in 2021 – True

¹See <https://audiovisual.ec.europa.eu/en/video/I-224942> and <https://www.trouw.nl/nieuws/pakt-nieuwe-eu-wet-kindermisbruik-aan-of-zorgt-deze-voor-massasurveillance~b4fe2f3c/>

²See <https://human-rights-channel.coe.int/stop-child-sexual-abuse-in-sport-en.html>

³See <https://www.missingkids.org/gethelpnow/cybertipline/cybertiplinedata>

2.3. The role of Europe and the Netherlands

Regarding where this material is hosted, several statements are presented. First, it is stated that 90% of CSAM material is hosted in the E.U. A fact sheet by the European Commission attached to the press conference in which the package on children's rights was presented states, however, that 60% of child sexual abuse material worldwide is hosted by E.U. servers⁴. So the quoted number of 90% is not supported by the source.

Statement: 90% of CSAM material is hosted in the European Union – False

In addition, the Netherlands is reported to host 45% of CSAM material worldwide, 65% of the material shared in the U.S., and 66% of material hosted in Europe. Meaning this material is supposed to be hosted on a Dutch server. The INHOPE network reported that in 2021 76.01% of CSAM material in Europe was hosted in the Netherlands.⁵ This was a reduction relative to 2020, when they reported 90% of the content in Europe to be hosted in the Netherlands. While the number fluctuate across sources, the claim that a large portion is hosted in the Netherlands is consistent with the various sources.

Statement: 45% of CSAM material worldwide is hosted in the Netherlands – Unclear, but it is substantial

2.4. Scale of CSAM and trends

Moreover, it is stated that the E.U. reports of CSAM of the member states have increased with 6,000% the last ten years. The fact sheet of the European Parliament presents 23,000 reports in the European Union in 2010, against more than 1 million reports in 2020.⁶ This is an increase of 4200%.

Statement: Reports of CSAM in the E.U. member states have increased with 6,000% the last ten years – False

It is important to note that there is a distinction between a report and content, such as a photo. In the CSAM landscape, three major parties report yearly on their findings. These are the National Centre of Missing and Exploited Children (NCMEC) in the U.S., the Internet Watch Foundation (IWF) in the U.K. and the INHOPE network, consisting of hot lines around the world. The Dutch EOKM and the IWF are part of this network. Data gathering methods differ per organisation. The INHOPE network, like the EOKM, depends on reports from volunteers and from other hotlines.⁷ NCMEC receives mandatory reports from ESPs, as stated previously. These reports make up 99% of their found content. Remarkably, Whatsapp and SnapChat are also amongst the reporting organisations despite their end-to-end encryption of their messages. This is because a user of these platforms can report a message to the service provider. The IWF is the only organisation of the CSAM detection organisations that actively searches for the material using a crawler. They do so since 2014, after having been asked by the UK government.⁸ How this search is executed exactly is unknown, but 56.62% of their reports in 2019 were actively searched for.⁹

Moreover, detection of long-term trends is difficult, due to the different types of reporting of data by different hot lines and the availability of year reports. Only the year reports of the IWF are available from 2010 to 2021. NCMEC has reports between 2018-2021 available, whereas the INHOPE network has reports dating back to 2014. However, the way they report their findings changes over the years. All hotlines distinguish between reports (a person or entity reporting CSAM to the hotline) and files (i.e. the actual reported material like photos). One report can contain a single file or thousands of files. The hotlines do not always distinguish between files and illegal files. Several conclusions can be drawn however, when this data is considered.

As shown in Figure 2.1, reports do not equal material. While the reports NCMEC received between 2018 and 2021 indeed slightly increase, the number of files found display no clear trend. 85 million files found in 2021 is equal to the amount found in 2018. The dip in 2019 and 2020 could perhaps be due to the corona pandemic, but the numbers in 2021 are no increase compared to those of 2018.

⁴See https://ec.europa.eu/commission/presscorner/detail/en/fs_22_2978

⁵See <https://inhope.org/media/pages/articles/annual-reports/8fd77f3014-1652348841/inhope-annual-report-2021.pdf>

⁶https://ec.europa.eu/commission/presscorner/detail/en/fs_22_2978

⁷Annual reports retrieved from <https://inhope.org/EN/articles/annual-reports>

⁸See <https://www.iwf.org.uk/media/y3rldq1m/2013-annual-report.pdf>

⁹See <https://www.yumpu.com/en/document/read/63240486/iwf-annual-report-2019>

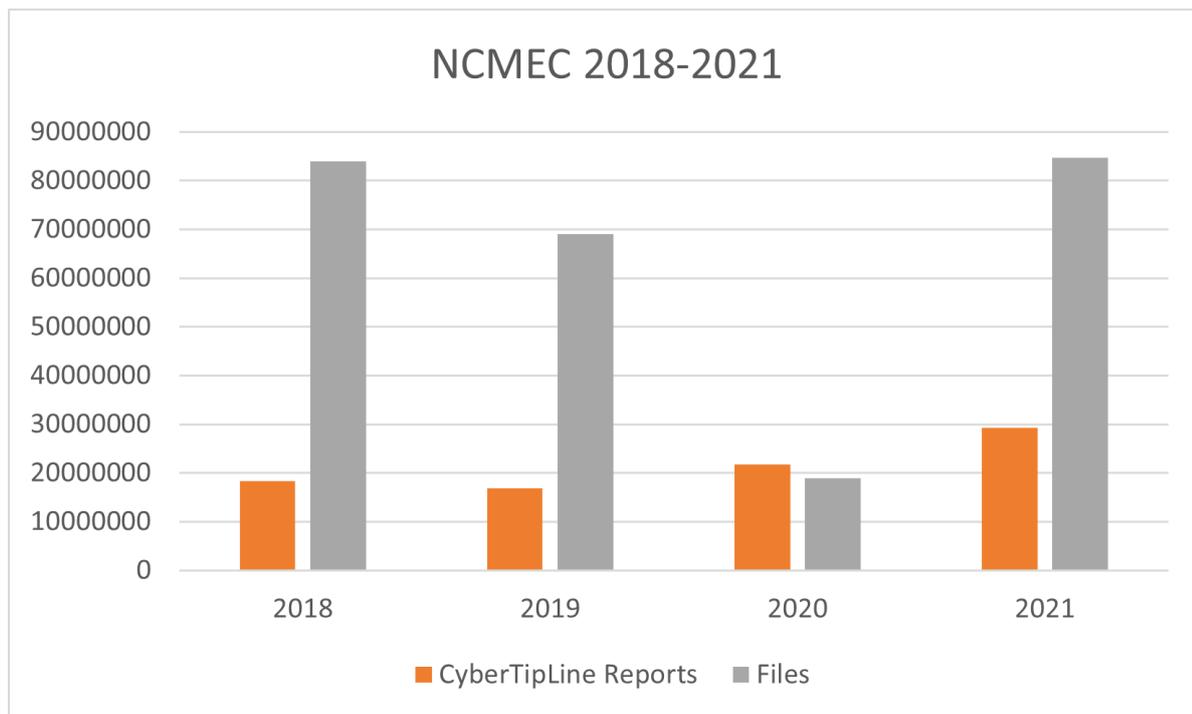


Figure 2.1: CSAM Reports and Files from 2018-2021 as reported in NCMEC Annual Reports. Annual reports are retrieved from <https://www.missingkids.org/footer/about/annual-report.pdf> The annual report of 2021 does not specify files, these are mentioned here <https://www.missingkids.org/gethelpnow/cybertipline/cybertiplinedata>

The IWF, as shown in Figure 2.2, does not report not on the number of files, but on criminal content URLs. In their year reports, they do not identify specific chat services, but they do specify newsgroups (internet discussion groups). Peer-to-peer networks are also mentioned in some reports. It is unclear if this refers to bittorrent and other file sharing protocols or to encrypted services. In any case, they are only specified in the reports of 2014 and 2015. The numbers are negligible. Only 1 (2014) and 8 (2015) reports are attributed to these networks, compared to a total of roughly 75,000. As stated previously, IWF actively searches since 2014, when the graph clearly displays an increase in reports and criminal content URLs. Due to their active search, it is not possible to state that the amount of content found is increasing because there is more material on the Internet, or because their search is becoming more effective at finding the existing material.

Finally, it is important to make a distinction between a file and a victim. Not every detected file implies a new victim. A lot of the material that is being reported is not new, but has been seen before. INHOPE reported in 2021 that of the 443,705 illegal files that were reported in their network, 82% contained unknown content. Unknown content refers to any files that have not been previously seen. NCMEC reports that of the previously stated 85 million files, a total of 22 million was unique.¹⁰ In other words, material that is made is shared more than once.

Alongside the interview with commissioner Johansson in Trouw, the EOKM was cited as saying that in the Netherlands detecting CSAM is becoming increasingly difficult due to a migration of the material to chat services like Telegram. This is difficult to factcheck, because the involved parties do not specifically report on this and not much research can be found. The NCMEC and the EOKM are the only hotlines that specifically addresses the chat services in their year report.¹¹ As stated previously, NCMEC receives mandatory reports from all ESP's, including chatservices like Whatsapp. The EOKM reports that next to 413,848 reports containing 822,970 URLs, there have been a mere 234 reports regarding distribution of CSAM in the forms of stickers, images and video's in WhatsApp and Telegram

¹⁰<https://www.missingkids.org/gethelpnow/cybertipline/cybertiplinedata>

¹¹See <https://jaarverslag2021.eokm.nl/indexEOKM-EN.html>

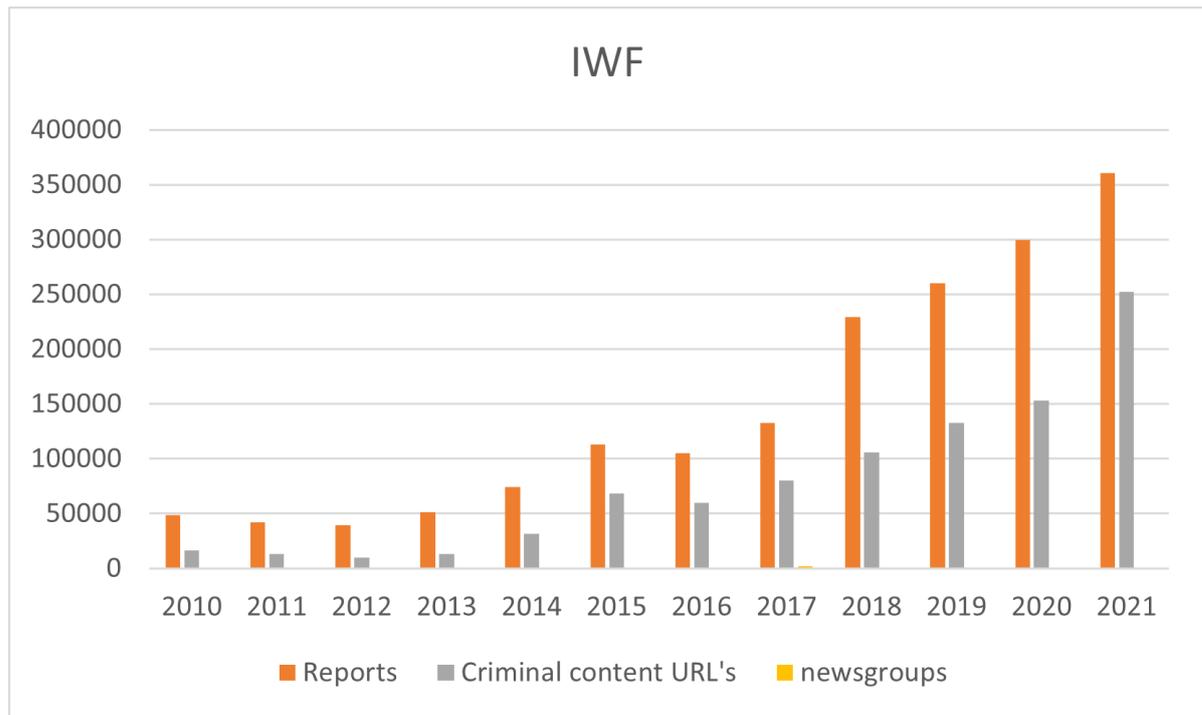


Figure 2.2: CSAM Reports and files from 2010-2021. Annual Reports retrieved from <https://inhope.org/EN/articles/annual-reports>

groups. INHOPE¹² and IWF¹³ do not report anything specific on chat services. INHOPE does state that the amount of reports related to Social Networking was less than 1% in 2021. NCMEC did receive most of their reports from Facebook, which at the moment is not end-to-end encrypted. They do however express their concern regarding the possibility of end-to-end encryption of Facebook, stating they expect half of their reports will disappear if Facebook becomes end-to-end encrypted.¹⁴

Further research on this topic in the Netherlands uncovered one study by Pointer, the research team of the Dutch broadcasting company KRO-NCRV. They found 2,463 videos and 3,718 images were sent in a so-called “expose” chats on Telegram¹⁵. In most of these cases, the content depicted a minor. These images and videos were often made by the victim, but then shared by the recipient without the consent of the victim, thus breaking the trust and confidentiality in which these were made, not mention breaking the law.

All in all, there is no clear evidence in the reports or research that supports the claim that there is a shift of distribution of CSAM from websites on the internet to encrypted chatservices.

Statement: Distribution of CSAM is shifting from imagehosters to encrypted chatservices – Not supported by the evidence

¹²See <https://inhope.org/media/pages/articles/annual-reports/8fd77f3014-1652348841/inhope-annual-report-2021.pdf>

¹³See <https://www.iwf.org.uk/about-us/who-we-are/annual-report-2021/>

¹⁴See <https://www.missingkids.org/theissues/end-to-end-encryption>

¹⁵See <https://pointer.kro-ncrv.nl/naaktbeelden-honderden-slachtoffers-telegram-groepen-gedeeld>

3

Reflection

This report was written in light of the proposed legislation by the European Commission to combat and prevent Child Sexual Abuse Material. Some of the statements from the press conference and the following interview in Trouw were fact checked. Most emphasis was given to the final statement that distribution of CSAM is shifting from websites to interpersonal communication in encrypted Telegram and WhatsApp groups. Finally, the reported numbers and methods of the most important hot lines have been compared in order to see if claimed trends are supported by the evidence.

Of the 6 statements from the Commission's supporting arguments for new measures against encrypted communication services, 3 were incorrect, 1 lacked supporting evidence and 2 were likely correct. An overview of the fact-checked statements and the conclusion drawn is presented in Table 3.1.

None of these considerations take away from the observation that many children fall victim to sexual abuse, and some of this abuse is posted online. The exact extent of the online share of all forms of sexual abuse is unclear. Moreover, there is no clear evidence to support the statement that distribution of CSAM is increasingly taking place through encrypted messaging services. The volume of reports on encrypted services received by the hot lines are very small in compared to the reports they receive on websites. Furthermore, there is no evidence of a clear increasing trend in those numbers.

Overall, it has become clear that the methods of data gathering differ to such an extent that the data cannot be directly compared across the various sources or even within the same source over various years. Varying from anonymous reports and collaboration with the police in the Netherlands, to mandatory reports by ESPs in the US, to actively searching with a crawler in the UK. In short, one should therefore tread very carefully when using these statistics for arguing trends in the online hosting of CSAM.

Statement	Verdict
1 in 5 children is victim of sexual abuse online	<i>False</i>
85 million photo and videos of CSAM have been intercepted in 2021	<i>True</i>
90% of CSAM material is hosted in the E.U.	<i>False</i>
45% of CSAM material worldwide is hosted in the Netherlands	<i>Unclear, but it is likely substantial</i>
Report of CSAM in the E.U. member states have increased with 6,000% the last ten years	<i>False</i>
Distribution of CSAM is shifting from imagehosters to encrypted chatservices	<i>Not supported by the evidence</i>

Table 3.1: Assessment of statements and verdicts

