CSAM Hosting Monitor

Report March 2023

Qasim Lone, Carlos H. Gañán, Michel van Eeten

TU Delft

Contents

Executive Summary

Since October 2018, TU Delft has been developing and operating a monitor to analyze where Child Sexual Abuse Material (CSAM) is being hosted in the Netherlands. This was done in close collaboration with EOKM. Looking back at 2022, the monitor's main findings are:

1. In the public-private roundtable to fight CSAM hosting, started in 2018, a norm was agreed upon: the hosting industry should remove CSAM within 24 hours after the NTD. In November 2022 – December 2022, EOKM conducted manual checks how fast providers and domain owners removed CSAM material after receiving an NTD, based on a sample of URLs prepared by TU Delft. We checked 2971 URLs located at 39 hosting providers and 188 domains. We found that just 40% of all CSAM was removed in 24 hours. Compared to 2021, the removal speed got significantly worse. In 2021, 87% of all CSAM was removed in 24 hours. In 2022, 56% remains online for 48 hours or longer, whereas in 2021, this was just 10%.
2. We do not know why the overall performance got worse, but we do note that many of these providers were not really in the picture before. They might have less experience with removing CSAM and they also were not activated by government pressure. The providers that were in the picture earlier, achieved much higher removal rates. The exception is KnowSRV, who is a steady presence in the top 5, yet its removal speed deteriorated a lot in the 2022 measurement compared to 2021.
3. There is huge variance among providers in terms of performance. In fact, the population of providers appears to be polarized: most providers either remove almost everything in 24 hours or they remove almost nothing. There are only a handful providers in between. More precisely: 17 of 39 providers in the uptime measurement manage to get more than 80% of all URLs in our measurement removed within 24 hours. In fact, 16 of them remove more than 90% in 24 hours. Then there are number of providers who have bad performance: 16 of 39 providers remove less than 20% of all CSAM in the first 24 hours. Only 6 providers (10%) remove between 20-80% in 24 hours. So it seems that once providers become active, they can manage to get the bulk of the material removed swiftly.
4. In 2022, we saw an overall lower volume of CSAM being hosted in the Netherlands, compared to 2021 and 2020. It is not clear whether this is a structural change or a temporary phenomenon, as illustrated by the uptick in volume in December 2022. The process of discovering CSAM URLs is mostly a black box, though we suspect the detection effectiveness of automated systems, like those run by the U.K.'s IWF, might go down when the landscape changes, as it did in 2022.
5. The top 5 of hosting providers with the most CSAM changed in a significant way. The number 1 from 2020 and 2021, NFOrce, dropped to place 5, with just a fraction of the volume of previous years. Instead, a completely different provider entered the picture: Amarutu Technology. Amarutu only had a negligble amount of CSAM in the years before, but entered the picture in early 2022 and then grew rapidly to become the 2022 number 1 CSAM hoster. KnowSRV, Telegram Messenger (no relation to the instant messaging service), and Hostslim make up the rest of the top 5, just like in 2021.
6. At the level of domains, we also saw a trend break in 2022. Of the top 10 domains with most CSAM, none of them were in the top 10 in 2021 or 2020. In fact, most of these domains were not seen at all in those years. Along the same lines, we observe that none of the domains in the top 10 from 2021 are still present in the 2022 top 10. In fact, 9 of the 10 disappeared completely.
7. Taken together, these two observations combined suggest that the landscape has changed. Were there was some continuity among the set of dominant providers and domains between 2020 and

2021, this continuity was disrupted. This might be the consequence of providers that were dominant in those years taking action, most notably NFOrce. NFOrce has severed ties with the clients who were running services, like image hosting, that were abused by the criminals sharing CSAM. Once NFOrce dropped from the picture, the overall volume of CSAM went down and other providers and domains came into view.

8. One way to summarize the results of the past three years of public-private partnership (PPP) against CSAM hosting in the Netherlands is that the providers that have been the focus of the PPP have been responsive. They already took action, but many increased their efforts. Takedown speeds were (very) high and in some cases providers have parted ways with legitimate customers who were being abused by people sharing CSAM. The domains that left with these clients, migrated to outside the Netherlands, They might be less reachable for NTDs and less active in removal than during their presence in the Netherlands. With their improved performance, they became less important in CSAM hosting. Providers that have been outside the engagement with the PPP are now dominating the Dutch landscape.

1

Introduction

For years, the Netherlands has been identified as a prime location for the hosting of child sexual abuse material (CSAM). Data on the hosting of CSAM is collected via INHOPE, the global network of national hotlines to combat CSAM. In their annual report over 2020, INHOPE listed the Netherlands as the second country worldwide and the leading hosting location in Europe, harboring 76% of all CSAM in Europe.[1]

As we noted in our previous report, these problems are not caused directly by the hosting companies. Their customers might operating legitimate image-hosting websites where unknown users upload and share CSAM. It is increasingly recognized, however, that providers do have to take part of the responsibility to combat the hosting of CSAM and make their networks more safe and secure. Providers are at least expected to respond swiftly to Notice & Takedown (NTD) requests, which are issued by INHOPE hotlines and which ask them to remove the CSAM material from their network. The response from providers varies widely, ranging from vigilant to slow to negligent. In rare and extreme cases, providers even offer 'bulletproof' services for criminals, knowingly facilitating the hosting of CSAM.

In response to the prominent presence of CSAM in Dutch hosting networks, a 2018 roundtable (March 27, 2018) brought together a broad coalition of representatives from the Dutch government, industry and academia.[2] The goal of this public-private initiative was to identify more effective ways to combat the hosting of CSAM in the Netherlands. This led to the articulation of several shared "ambitions". A key ambition with wide industry support is to create more transparency regarding which industry providers are involved in hosting CSAM and how swiftly they remove CSAM from their network once they are notified of its presence via an NTD request. As a result of the roundtable process and the government initiatives, the first public TU Delft report that monitored the hosting of CSAM in the Netherlands was sent to parliament in October 2020 by the minister of Justice and Security.[3] A second report was released in March 2022.[4] The current report is a follow-up of those two reports. Some of the text on the underlying methodology is included again for the reader's convenience. All three reports are based on data from Expertisebureau Online Kindermisbruik (EOKM), the Dutch national hotline and member of INHOPE. We report on the patterns in volume, location, providers and domains. Overall, we will answer the following research questions:

[1] INHOPE Annual Report 2021, available online at: https://inhope.org/media/pages/articles/annual-reports/8fd77f3014-1652348841/inhope-annual-report-2021.pdf

[2] https://zoek.officielebekendmakingen.nl/kst-31015-150.pdf

[3] https://https://www.rijksoverheid.nl/documenten/rapporten/2020/10/08/csam-hosting-monitor---rapport-september-2020

[4] https://www.rijksoverheid.nl/documenten/rapporten/2022/06/29/tk-bijlage-1-csam-monitor-tu-delft

1. Which Dutch providers host CSAM materials in their network? The first objective is to identify where CSAM materials are hosted and by whom. We will combine the data from EOKM with IP addresses, AS numbers and WHOIS registration data to identify the hosting providers operating the networks where the CSAM is located.
2. How does the location of CSAM material in the Dutch market change over time? We survey the trends over the past two years (2020-2022).
3. How can the distribution of CSAM material over providers and domains be understood? We will take the output of questions 1 and 2 and interpret these findings in the context of the overall Dutch hosting landscape.
4. How fast do the hosting providers and their customers respond to an NTD request to remove CSAM? We report on what portion of CSAM is removed within 24 hours of the NTD request and how different providers perform in terms of meeting the 24-hour removal norm. For this, manual checks were conducted by EOKM in November-December 2022, in collaboration with TU Delft.

The remainder of this report consists of the following parts. We first explain the methodology of the CSAM hosting monitor (Chapter 2). We then turn to the results on the measurement of the NTD takedown speed in August 2020 (Chapter 3). Next, we present the results on the location of CSAM in the Dutch hosting landscape and identify how this pattern has changed over time (Chapter 4). We end with a few concluding notes on the next steps for the monitor.

2

Methodology

The main goal of this project is to monitor the location of CSAM material across the Dutch hosting market and to track the speed of responding to notice and takedown requests (NTD). We first outline what data we use on the detection of CSAM and how to attribute CSAM material to the relevant hosting provider and domain.

2.1. CSAM Data

CSAM material is located at URLs (Uniform Resource Locators) – simply put, a link to a webpage or picture. The URL can point to a single picture or it can link to a page that contains more CSAM. These URLs are submitted to EOKM. The process that follows is visualized in Figure 2.1. EOKM processes and checks the URLs to confirm whether they actually contain CSAM. Once checked, an NTD is sent to the hosting provider and domain owner asking them to remove the material at the specified URL. The TU Delft monitor tracks the number of NTDs sent to each provider and domain owner, as well as the speed with which the CSAM material was taken offline. The latter is based on manual checks done on a sample of NTDs/URLs – as will be discussed in Chapter 3.

The core data on the online location of CSAM is provided by the EOKM. EOKM receives reports of suspected CSAM from volunteers via its “meldpunt” (Dutch hotline) and from other INHOPE member hotlines through a system called ICCAM. A few of the hotlines contribute most of the reports, because they have the mandate and tools to conduct automated scans and web crawls, thus discovering more CSAM than other hotlines. The Canadian hotline has been one of the main contributors.

In March 2020, the Canadian hotline left the INHOPE network. This means that EOKM is no longer receiving those CSAM reports and thus sends no notifications for them. The Canadian hotline sends its own NTDs to providers. This development has two implications: first, Dutch providers now get some NTDs from EOKM and some from the Canadian hotline; second, our report is undercounting known CSAM for Dutch providers, since we only have access to the EOKM data.

EOKM staff checks the reported URLs and verifies that it does in fact contain CSAM material. Next, it checks that it is hosted in the Netherlands. If not, the URL is sent via ICCAM to the relevant hotline in another country. If the reported material is confirmed as being CSAM and as being hosted in the Netherlands, then EOKM then decides on whether to send the link to the National Police (law enforcement agencies, LEA) for investigative purposes or to issue an NTD request to the hosting provider, domain owner and, in some cases, the registrar for the domain name. When EOKM sends a link to the police, it will eventually also issue an NTD but with a bit of a delay, typically of one day. There are several reasons to send a link to the police. It could be previously unknown material (rather than known CSAM), it could be material with a possible Dutch connection or the link might otherwise be relevant to ongoing

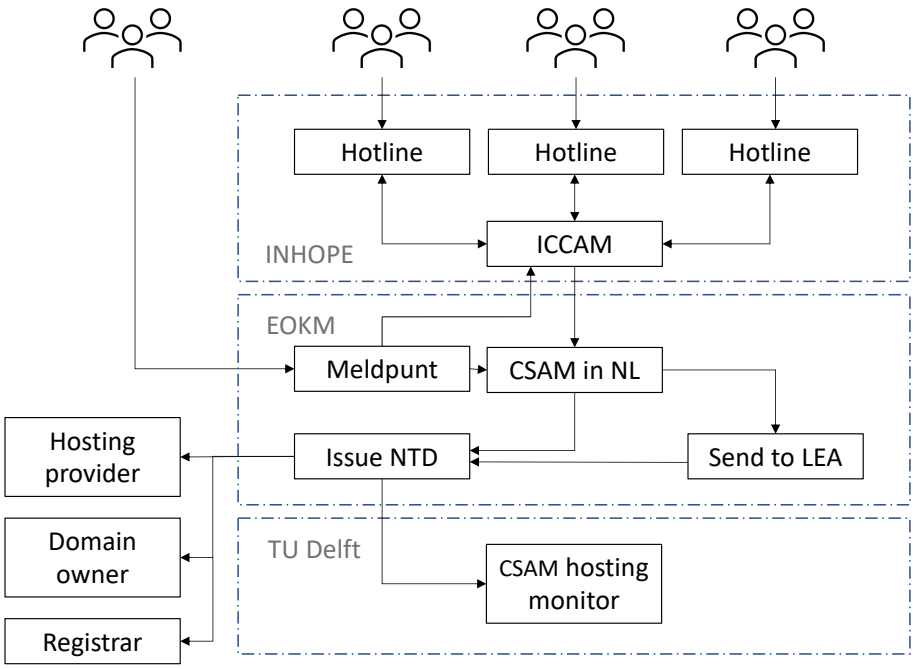

Figure 2.1: Overview of CSAM NTD process and TUD Monitor

investigations.

From the data at EOKM, TU Delft then extracts the relevant properties to feed into the CSAM hosting monitor. Until January 2020, the processes at EOKM contained many manual actions, like manually crafting the NTD emails every day. In January 2020, an automated system called SCART (based on AbuseIO[1]) has been taken into operation. It automates certain steps after the manual verification of the suspected CSAM.

For the monitor, we had to develop different techniques and tools to process the data from EOKM, depending on whether the NTD process was conducted manually or via the new system, SCART. With support and feedback from EOKM and the SCART developers, we wrote software scripts to consistently enrich the reported CSAM URLs and count the number of URLs per hosting provider and per domain. This software runs on EOKM infrastructure, so that no CSAM URLs or other CSAM-exposing data needs to leave the network. Only the final aggregated outputs per hosting provider and per anonymized domain are stored at TU Delft.

2.2. CSAM Hosting Monitor

The data going into the scripts for the monitor contains the IP addresses where the CSAM was located, the domain, the time stamp, the approximate country location for that IP address (based on GeoIP), the Autonomous System (AS) in which the IP address resides, and the organization to which the corresponding IP address belongs, according to WHOIS data.

WHOIS data is derived from databases that contain information about the registered users or assignees of an Internet resource, such as domain names, IP addresses, and Autonomous System Numbers (ASN), which are used for routing traffic across networks. WHOIS for IP addresses and Autonomous System Numbers are operated by Regional Internet Registries (RIR). In the case of Europe, the RIR is RIPE NCC (Réseaux IP Européens Network Coordination Centre). WHOIS data for domain names is maintained by the registry for the top-level domain under which the domain name is located. For .nl do-

[1] https://abuse.io/

mains, for example, SIDN is the registry and its WHOIS system provides information about the domain owner.

WHOIS data on IP addresses identifies what entity administers the network in which the IP address is located and specifically to which hosting provider the IP address has been assigned by the RIR. The hosting provider may further assign the IP addresses to one of its customers, which might itself be a company. In terms of responsibility, there are two scenarios. If the hosting provider agrees with the customer that the customer takes responsibility for handling potential abuse issues, then the provider updates the WHOIS record at the RIR to reflect this. The WHOIS record will then show the customer name and contact information, including an email address where abuse can be reported. This is called a 'sub-allocation'. If no such sub-allocation is made in WHOIS, the provider retains the responsibility for dealing with abuse of resources at those IP addresses.

In Figure 2.2, we summarize our process to map CSAM URLs to hosting providers. We begin by selecting only the URLs that are verified to contain CSAM and for which NTDs were sent to the hosting providers and/or domain owners. (In a fraction of cases, no NTD was sent to the provider, because the CSAM had already been taken down before the NTD could be sent. Also, before the automation via the SCART system, during peak time some URLs did not lead to an NTD because the limited availability of EOKM staff to manually send the NTD.)

Next, we perform IP WHOIS queries at the RIPE NCC database.[2] This provides us with the organization name of the entity to which the IP address has been assigned. It also provides us with the abuse contact address. We compare the abuse contact addresses from WHOIS with the contact address used for the NTD request. If these are not the same, then we do not include the URLs in our count of URLs for that provider, because we cannot verify that the provider actually received the NTD at the address that the provider has specified in WHOIS.

In a very small fraction of cases, we did not find an organization record in RIPE NCC WHOIS. We then mapped the missing organization using the MaxMind IP2ORG database[3] and manually validated this mapping by comparing the organization name to the domain in the abuse contract email address and the Autonomous System owner name. If these were consistent, we attributed the URL to that provider.

In addition to identifying the relevant hosting provider for the URLs, we also identify which domains contain CSAM. We extract the Fully Qualified Domain (FQDN) from the URL as a domain name. A URL is composed of sub-domain and the parent domain. For instance, a sub-domain `abc` and parent domain `example.com` has the fully qualified domain name `abc.example.com`. We used FQDNs to identify domains since, in some cases, sub-domains are hosted in different hosting providers. We assign a unique randomly-generated number to each of the FQDN to obfuscate the domain names.

2.3. Tracking CSAM Over Time

Ideally, the monitor would track CSAM hosting in the exact same way over time. This would then give us a continuous timeline of CSAM hosting in the Netherlands. This is now supported by the SCART system at EOKM. That being said, the data going into SCART from the ICCAM system is not consistent over time. Since most of the EOKM CSAM data comes from ICCAM, changes in how INHOPE data sharing works have a serious effect on what is observed in the Netherlands. The data in ICCAM comes from other INHOPE hotlines. The hotlines in UK and Canada proactively search for CSAM with webcrawlers. When their crawlers visit more domains hosted in the Netherlands, then more CSAM would be found and thus more URLs would flow from ICCAM to EOKM and our monitor. The opposite change also occurs. As we mentioned before, in March of 2020, the Canadian hotline left INHOPE en thus also stopped sharing the CSAM it discovered in the Netherlands with EOKM. Instead, the Canadian hotline decided to send

[2] `https://apps.db.ripe.net/db-web-ui/query`

[3] `https://dev.maxmind.com/geoip/legacy/csv/#GeoIP_Organization_Edition_CSV_Database_Fields`

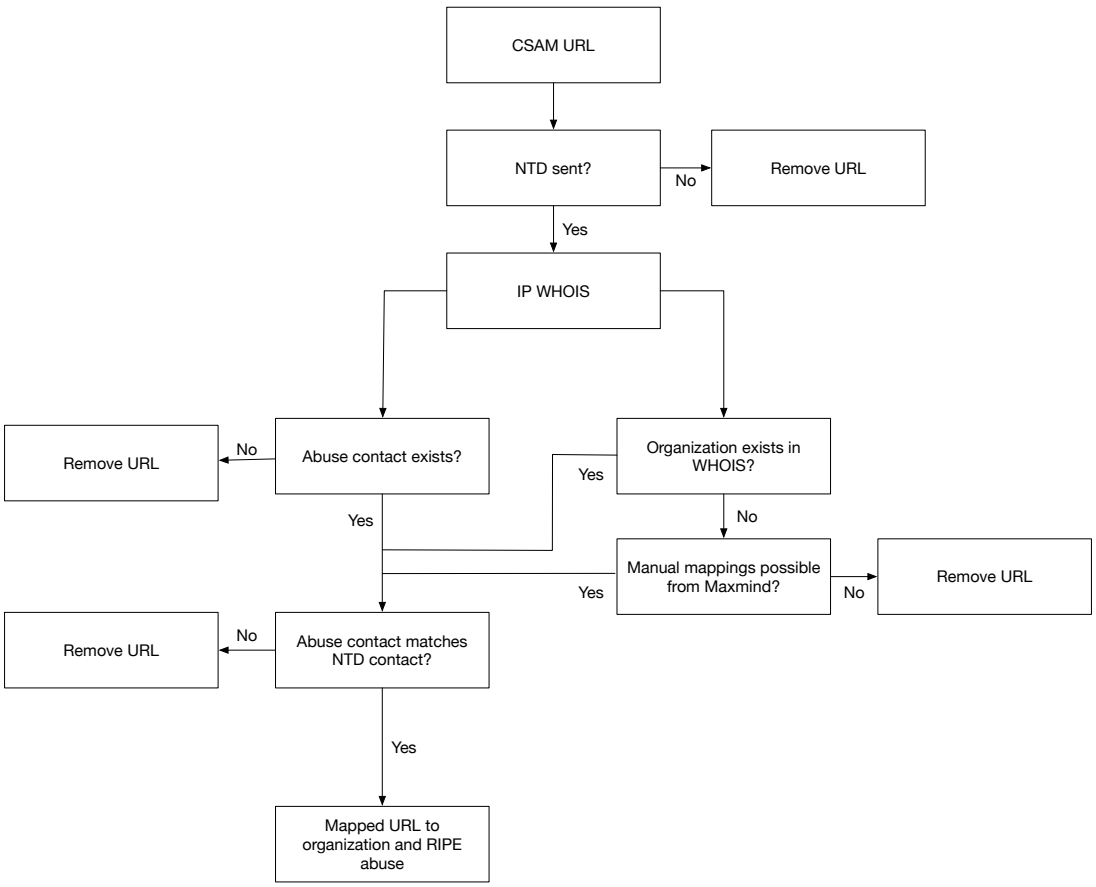

Figure 2.2: Mapping of CSAM URLs to hosting providers

out its own NTDs. This reduced the visibility of the monitor on CSAM in the Netherlands.

2.4. NTD Takedown Speed

Together with EOKM, we have also conducted a measurement of NTD responsiveness. The industry norm agreed upon in the public-private working group was that CSAM should be removed by providers within 24 hours after the NTD. We present the measurement and underlying methodology in more detail in Chapter 3.

Because the uptime measurements are critical to evaluating the responsiveness of the hosting providers to CSAM NTDs, this process relies completely on manual checks by EOKM staff. We do not use automated measurements. While automated checks are partially implemented in SCART, they are not yet reliable enough to use as evidence on provider compliance with the industry norm.

2.5. Corroborating NTD data with providers

Over the course of 2020 and 2021, the monitor had been working with data that was exported from SCART and that was assumed to contain all NTDs to providers. In the Fall of 2021, IP Volume, one of the providers that was included in the monitor, asked us to receive the NTD data that the monitor had used for its network.[4] That data was shared. IP Volume then ran a comparison of the shared data

[4]We should note that there are two providers called IP Volume: IP Volume Inc and IP Volume LTD. The former is registered in the Seychelles, while the latter is registered in the United Kingdom. When we refer to IP Volume in this report, we mean IP Volume Inc.

against its own internal logs of the NTDs it had received. The comparison found that most of the data was consistent, but it also unearthed three sources of discrepancies.

First, for some URLs, the provider logged different time stamps for receiving the NTD than the SCART export data logged for sending the NTD. We investigated this and there was indeed a discrepancy. However, the exact moment of the NTD is only important for the uptime measurement, not for the count of how many URLs were hosted by the provider. The discrepancy was corrected and we received time stamps that were consistent with those reported by the provider.

Second, some NTDs were in our SCART export data, but not received by the provider. This turned out to be caused by the following scenario: the URL was first sent to the police. Normally, the URL would then be sent in an NTD the next day. Before this NTD was sent, EOKM staff did another check to see whether the material was still online. They found it had already been taken down, so there was no need anymore to send the NTD. These URLs did, however, appear on the provider network and thus should be included in the count of URLs for the provider, even though the provider never received an NTD for them.[5] This issue did not impact the 2021 uptime measurements. We double checked and found that for all URLs included in the uptime measurement, NTDs had been sent. We used the corrected NTD data and time to calculate the removal speed.

Third, the provider found NTDs that it had received, but that were not in our SCART export data. Further investigation made clear that these NTDs were incorrectly missing from the data that was exported from SCART to the TU Delft monitor. This problem affected all providers. It means we undercounted the number of URLs for the providers. This omission has been corrected. The current report on 2022 is based on a more complete NTD export data from SCART. . We write “more complete”, because a small discrepancy still exists after the corrections. For a very small number of URLs, the SCART data does not have a record of an NTD directed at the abuse email of the provider, but the provider did state that they received an NTD. The volume of these exceptions is too small to impact the findings in this report.

We re-calculated the numbers for 2020-2021 from the corrected data and found that the overall findings from our 2020 CSAM hosting monitor report were unchanged. The top 5 remained the same in each year and the percentage of URLs hosted by each provider only changed marginally and did not impact the conclusions.

Because this check with IP Volume was very valuable at improving the quality of the monitor, we repeated this process with two other providers. These checks confirmed that the corrected export data was consistent with the provider findings. For NForce, we also corrected five URLs in the uptime measurement, since it was demonstrated that the domain left the NForce network during the measurement. Thus, the uptime that was observed was no longer associated with the NForce network.

In sum, regarding the discrepancies, we found that (1) there were inconsistencies in the time stamps of the NTDs for the uptime measurement, which we were able to correct; (2) providers sometimes do not receive an NTD, yet we are correctly counting them in the monitor, since the URL was hosted by them; and (3) there were NTDs missing in the monitor data, thus undercounting the total number of URLs hosted by the provider. This undercounting was fixed in subsequent data exports and does not affect the current report.

In 2022, we did encounter a different problem. Because of a system malfunction during the Jan-May 2022 period, we received the SCART data without the IP address, so only containing the URLs, time stamp and NTD recipient. This hampers our independent verification of the location of the CSAM. As explained in Section 2.2, we independently verify the hosting location for all URLs, double checking the attribution done by SCART. For this purpose, we run our own scripts to enrich the data on the same day as the NTDs are sent. This enrichment process does an independent lookup as to which provider is

[5] Before this discrepancy was discovered, we assumed that for all URLs we included an NTD was sent. Indeed, we stated this assumption in our 2020 report. We now know that this was incorrect. We should have stated that for a fraction of URLs no NTD was sent.

hosting the URL. We cannot do this backward in time when the IP address is missing. Since the URL might be hosted at a different IP address at a later date, we need to look up WHOIS and BGP data at the moment when the NTD was sent.

To handle the missing IP address for some URLs, we checked whether we did a lookup for a different URL on the same domain in a period of 10 days around the NTD date. For some portion of the missing data, we found such a URL. This allowed us to reliably map these missed URLs to the respective provider. For 6% of all URLs in 2022, we could not reliably map the missing IP address in the dataset. This means we had to drop them from our analysis.

For the analysis of the 2022 data, we shared a draft report with NForce and LeaseWeb, since they had earlier indicated they were interested in comparing the monitor results with their internal data. Leaseweb's data was fully consistent with the SCART data. In the case of NForce, the SCART data on the reported URLs was also consistent with NForce's internal data. For the uptime measurement, however, a slightly more complex picture emerged.

NForce informed us that they have a system to check the status of reported URLs. This system includes automated and human verification. According to their logs, all URLs in our sample were removed well within 24 hours. We observed 19 URLs that EOKM observed as being online between 24-28 hours were all on a single webdomain. NForce thinks that these URLs might have appeared to be online during the EOKM check because the webdomain sets long cache times. If EOKM did not forcibly cleared the browser cache before checking the status of the URLs, the browser might have shown content that was actually no longer online. Similarly, they stated that the one URL that was online for more than 48 hours was also a false positive. Here, their speculation is also that the human check resulted in an error. For this URL, this webpage replaces the removed video with other random video content from elsewhere on the site. So it might appear at first glance that the same video is still online. Only some tiny fine print at the bottom of the page informs the user that the original content is unavailable.

It is important to discover and fix data quality issues and we are grateful for the efforts made by the providers. While CSAM data faces grave restrictions in terms of sharing, we advise to conduct these cross checks with providers for subsequent analyses of the SCART data, e.g., by the new public authority that is currently being instituted.

2.6. Limitations

In this section, we discuss some of the measurement challenges that the monitor encounters. We take these into account when interpreting the results. First, the monitor data is based on the data received by EOKM. The core of the CSAM reports are based on ICCAM data. Two INHOPE contributors, Canada and U.K., proactively search for CSAM. This has several implications. First, it is unclear how their crawlers and manual searches detect the new content. There is a possibility the crawlers focus their search for content on hosting providers and domains where material had been detected before, rather than elsewhere. This might result in a biased picture of the problem, since material at those hosting providers is more likely to be discovered than elsewhere.

Second, there are fluctuations driven by increases and decreases in intake of data from the INHOPE network. Since March 2020, the Canadian hotline is no longer a member of INHOPE. This means that IWF, the U.K. INHOPE member, is now the major source of data. It is not transparent how this data is collected. It is thus also not clear what is causing the changes in CSAM volume to occur. We cannot tell whether this reflects changes in the presence of CSAM in Dutch networks or changes in the detection of CSAM. In all likelihood, it is a combination of these effects.

Another issue that we observed is that some domains respond with multiple IP addresses when they are queried via DNS. In other words, a domain name has identical copies of the same website residing on multiple servers, which in some cases maps to different hosting providers. These domain owners might be using a well-known methodology, "round-robin DNS," in which a different IP responds to every new query. It provides domain owners with the option of load balancing the across different providers

and improves the fault tolerance of their website in case of downtime at one provider. This setup means the same domain, and thus the same material, is mirrored across different providers. For the monitor, we have attributed the URL based on the DNS lookup conducted at the time of the URL being prepared for NTD.

EOKM maintains a 'green list' of cooperative domains. They are quick in removing the content, and most of the reported content for these domains is usually illegal. EOKM sends them an NTD right away, without performing additional checks. For the other domains, there is more investigation for each URL. In peak times, this might mean that it might take a while before EOKM has time to conduct this investigation. In that period, the URL might have gone offline. Once the investigation starts, the URL is no longer valid and thus no NTD is sent. This would mean that this domain owner and provider get slightly fewer NTDs than the 'green list' domains and providers get, even if they originally had the same amount of material. Given that we count on the basis of NTDs, it means that the 'green list' domains and providers get a slightly higher count, compared to those not on the green list.

A final and important challenge is to conduct automated uptime measurements of the URLs received by EOKM. While automated measurements would be ideal, as it would give us fine-grained data on how long a URL stays online after an NTD, a variety of technical challenges prevent this. One key challenge is the use of CAPTCHAs by the domain owners. CAPTCHAs are tests embedded on a web page where a visitor has to conduct a simple task, to determine whether or not the user is human. The current tracking scripts cannot automatically solve the CAPTCHAs and without this, the page content cannot be retrieved. For this reason, our report includes a manually conducted uptime measurement, conducted by EOKM in close collaboration with TU Delft. We describe this in more detail in Chapter 3. On one side, the manual process helped us by-pass CAPTCHAs. On the other side, the timing is less precise. The check after 24 hour actually takes place a bit later, say after 26 hours. It was never conducted earlier than 24 hours. So the error in the measurement is always to the benefit of the provider. In other words, we overestimate takedown speed. A related issue is the NTDs that were sent out on Friday. Since EOKM did not conduct checks during the weekend, these URLs were revisited on Monday. If they were removed at that time, then we counted them as removed within 24 hours. Again, we erred on the side of caution.

In summary, the monitor encounter certain challenges in measuring and monitoring the domains with CSAM data. While these issues have an effect on the specific amounts of CSAM that are calculated, these issues are not large enough to change the longitudinal picture and the concentrations in CSAM material that the monitor reports.

3

CSAM Takedown

In the public-private roundtable to fight CSAM hosting, started in 2018, a norm was agreed upon: the hosting industry should remove CSAM within 24 hours after an NTD request was received. As in our previous reports, we include a manual check on the uptime of CSAM material based on a sample of URLs prepared by TU Delft. EOKM staff manually verified whether the CSAM at a URL was still online 24 and 48 hours after the NTD was sent, to measure how fast material was removed. In this chapter, we describe the results of the most recent manual checking of takedown speed, conducted in November – December 2022. We analyze how many providers and domain owners adhere to a 24h window for removal of CSAM.

3.1. Measuring Takedown Speed

How long does it take before a domain owner or hosting provider has taken down the CSAM material? Between between November 04, 2022 and December 22, 2022, we daily selected a sample of URLs to be checked manually by EOKM staff. These URLs were included in NTDs sent the day before (or the Friday before, when sampling on a Monday).

Before selecting a URL for the sample, we first checked whether it was still open in the system. In some cases, we found that EOKM staff had already done a manual check that same day and found that the material had been removed. In those cases, we logged that these URLs were removed within 24 hours.

On some days, the volume of URLs coming in through ICCAM was relatively small. In that case, all URLs could be checked. On other days, we had to sample the incoming new URLs so as to not overburden EOKM staff. We did stratified sampling per provider so that even providers with a small number of URLs would be included in the sample. A non-stratified random sample would have basically meant the sample would be consist mostly of the top 5 providers in that period and we would not get any reasonable measurement for the other providers. After selecting the sample and sending it to EOKM, their staff would visit all URLs in the sample after 14:00h, which was more than 24 hours after the NTD had been sent. If the material was still online, it was not compliant with the industry norm. Then a second visit would be scheduled for the next day, at least 48 hours after the NTD.

To measure takedown speed, we combine two datasets: (1) the URLs that were already taken down before we could select it in our sample (these URLS are all removed within 24 hours); and (2) the URLs from our sample, where EOKM conducted additional manual checks to see if the material remained online. We represent the takedown time in three categories: (i) content removed within 24h, so within the industry norm; (ii) content removed within 48h; and, finally, content that remained online for 48 hours and more. Some URLs in the last category were revisited several times more. The pattern was highly

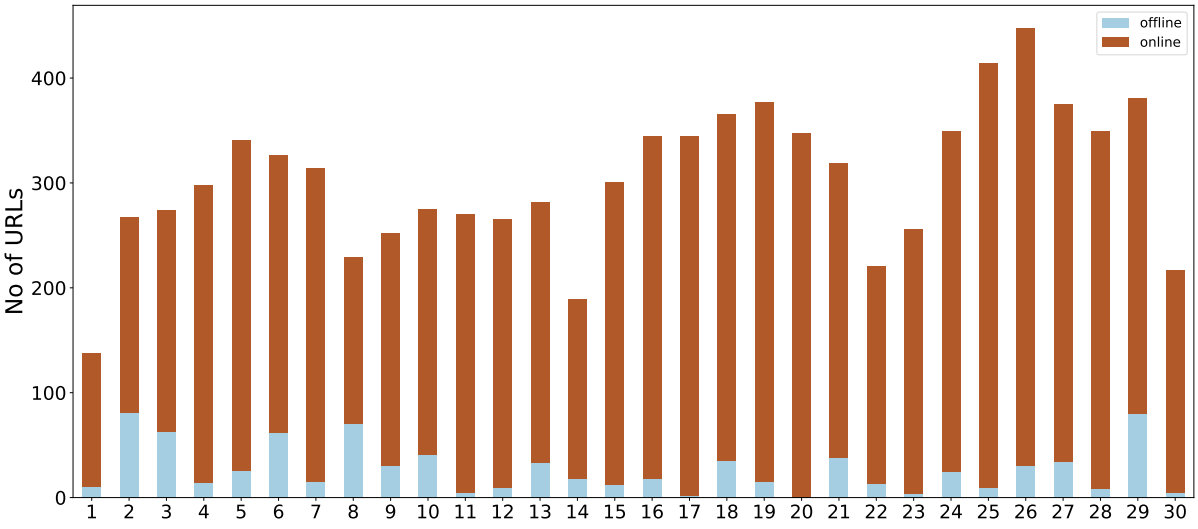

Figure 3.1: Number of URLs visited per day between Nov 2022 – Dec 2022 for checking whether the CSAM content was still online

varied. Some went offline one day later, after 72 hours, some stayed online for the whole period where EOKM did checks, in an extreme case going up to 21 days.

Figure 3.1 shows the number of URLs that were selected by TU Delft and manually checked by EOKM every day during the measurement in November 2022 – December 2022. The number varies per day because, for certain days, there were more URLs that were verified as CSAM. We also monitored previously online URLs for at least one more check before dropping it from further checks. The color indicates how many URLs were seen on that day as online versus offline. Most URLs were found to be online. This is because most URLs in that period were reported from two hosting providers: KnownSRV and Amarutu Technology. As we will see below, they leave a significant portion of the content online for longer than 48 hours.

We cannot determine which actor actually took down the CSAM. The NTD is sent to the hosting provider and, if EOKM also has an abuse contact address for the domain, also to the domain owner. In rare cases, it is send to the domain registrar. While we cannot observe from the outside who actually took the action to remove the CSAM material, the standard practice is that the hosting provider forwards the NTD to its customer, the domain owner. If EOKM has an abuse contact address for the domain, then the domain owner itself will also receive the NTD directly, at the same time as the hosting provider. The domain owner is typically the entity that actually takes the CSAM offline. In some cases, the hosting provider might intervene itself, if it feels the domain owner is not acting responsibly. Under these industry practices, the takedown speed of a hosting provider is mostly dependent on the takedown speed of its clients, the domain owners, though it can also intervene directly or put pressure on the clients. To take this joint responsibility into consideration, we will also show how many domains were associated with each provider for the URLs in the uptime measurement, in the next section. Different domains roughly correspond to different customers, who might have different response times.

3.2. Takedown Speed per Hosting Provider

In Table 3.1, we present the results of the manual uptime measurement, combining both datasets (as explained above). In total, we have data for 2971 URLs located at 39 hosting providers and 188 domains. It is important to note that because of the sampling approach discussed in the previous section, the number of URLs in the sample should *not* be read as indicating the total volume of URLs for those providers, nor of the relative amount of URLs for each provider. We have over-sampled smaller providers and under-sampled larger ones.

In terms of the results, we found that 40% of all CSAM is removed in 24 hours. A further 4% is

removed between 24-48 hours, and 56% remains online for 48 hours or longer – in some extreme case it was still online when we stopped the measurements.

As in our previous reports, we find a high variance in performance. We discuss the performance of the top 5 hosting providers with the most CSAM URLs. Chapter 4 will present those numbers in more detail, but for now we only need to know that the top 5 is: Amarutu Technologies (1), KnownSRV (2), Telegram Messenger (3), HostSlim (4) and NForce (5).

Amarutu, the number 1 hoster of CSAM in 2022, performs atrociously bad. Just 4% of all URLs are removed in 24 hours and over 95% remains online for more than 48 hours. This indicates they basically do nothing to get this material removed swiftly. This performance was just as bad a year ago, in the 2021 measurement.

KnownSRV, the number 2 hoster, performed better than Amarutu, but still poorly: 37% of the URLs still online after 48 hours. Just 58% was removed in 24 hours. This is worse than in 2021, when 83% was removed within 24 hours. So no improvement here.

Telegram, the number 3, removed 87% in 24 hours. This was only 6% in 2021, so we see a dramatic improvement. HostSlim, the number 4 in 2022 and number 3 in 2021, shows the even better performance: 96% of all URLs are removed in 24 hours. Only 1% remains online longer than 48 hours. This is a big improvement compared to the 2021 measurement: at that time, only 38% was removed within 24 hours. NForce, in 2022 the number 5 provider, but in 2021 the number 1, managed to remove 78% in 24 hours. When we checked our findings with them, they stated that their internal data showed that 100% was removed (see Section 2.5 for more details). This performance would be consistent with 2021, when it managed to 100% of the URLs removed within 24 hours. Stark Industries, ServerStack, Leaseweb and 13 other providers managed to get to a perfect record of 100%, which demonstrates the effectiveness of their efforts to remove CSAM.

Interestingly, the population of providers appears to be polarized: most providers either remove almost everything in 24 hours or they remove almost nothing. There are only a handful providers in between. More precisely: 17 of 39 providers in the uptime measurement manage to get more than 80% of all URLs in our measurement removed within 24 hours. In fact, 16 of them remove more than 90% in 24 hours. Then there are number of providers who have bad performance: 16 providers remove less than 20% of all CSAM in the first 24 hours. Only 6 providers (10%) remove between 20-80% in 24 hours. So it seems that once providers become active, they can manage to get the bulk of the material removed swiftly. Inactive providers leave most of the material online.

Within our uptime measurement, the number of domains with CSAM per provider varies substantially. In some providers, like HostSlim, all URLs are located on just a single domain. At the other extreme, we find Aeza Group, which is hosting in total 31 domains with just 44 CSAM URLs in total. More domains means the provider is dependent on more clients (domain owners) to get the material removed, unless the provider is willing to take unilateral action.

3.3. Takedown Speed per Domain

Since hosting providers are typically dependent on domain owners for the removal of the CSAM, we also analyze the data at the level of domains. For this purpose, we only use the data from our manual sample. In our sample, we have 2971 URLs across 188 domain owners. In Table 3.2, we present the results for the 20 domains with the most URLS. All other domains are combined in the last row.

We see a pattern where three distinct groups are visible. Some domain owners are highly responsive, such as the owner of domain 2105, which removed a large set of URLs all within 24 hours. Interestingly, this domain is hosted by KnownSRV, so its poor performance (58% removed in 24 hours) is because its customers (domain owners) are very different in their responsiveness to NTDs. Then there are a few domains which are responsive, but slow. Domain 1981, for example, removed everything within 48 hours, but only 34% within 24 hours.

Yet the largest group of domain owners is very unresponsive domain owners: 15 out of 20 providers left more than 50% online for more than 24 hours and 11 providers even left more than 90% online beyond 48 hours. For 5 domains, 100% is left online untouched for more than 48 hours.

In short: our results show that just 2 domains in the top 20 remove more than 90% within 24 hours. A further 3 domains remove more than 90% in 48 hours. And 15 domains are very unresponsive. Most of their material remains online beyond 48 hours. A single unresponsive domain owner with a substantial amount of CSAM can significantly impact the performance of a hosting provider. The provider then has to pressure the domain owner into responsiveness or drop them as a client. Some providers took the latter action.

3.4. In Sum

In summary, we observe that there is a lot of variance in how fast providers and domain owners get CSAM material removed. Out of the 39 providers in our uptime measurement, 16/39 (41%) manage to get more than 90% of all URLs in our measurement removed within 24 hours. Another group of 16 providers shows the opposite pattern: they remove less than 20% of all CSAM in the first 24 hours. Overall, about 56% of CSAM remain online after 48 hours and this is located on specific domains that are unresponsive and where the hosting provider is also not intervening.

Hosting Provider	Total Urls	Domains	Offline within 24h	Offline between 24-48h	Online for 48h or more
KnownSRV Ltd.	1129	26	655 (58.02%)	55 (4.87%)	419 (37.11%)
Amarutu Technology	1019	7	41 (4.02%)	1 (0.1%)	977 (95.88%)
HostSlim B.V.	183	2	175 (95.63%)	6 (3.28%)	2 (1.09%)
Serverel Inc.	139	21	10 (7.19%)	2 (1.44%)	127 (91.37%)
NFOrce	92	2	72 (78.26%)	19 (20.65%)	1 (1.09%)
Stark Industries	61	1	61 (100.0%)	0 (0.0%)	0 (0.0%)
Telegram Messenger	55	2	48 (87.27%)	5 (9.09%)	2 (3.64%)
VDSina	48	12	3 (6.25%)	2 (4.17%)	43 (89.58%)
IROKO Networks	43	2	26 (60.47%)	1 (2.33%)	16 (37.21%)
AEZA GROUP LLC	41	16	9 (21.95%)	15 (36.59%)	17 (41.46%)
ServerStack, Inc.	26	1	26 (100.0%)	0 (0.0%)	0 (0.0%)
LeaseWeb Netherlands	14	1	14 (100.0%)	0 (0.0%)	0 (0.0%)
EUROHOSTER Ltd.	13	1	13 (100.0%)	0 (0.0%)	0 (0.0%)
IWS NETWORKS LLC	12	2	2 (16.67%)	0 (0.0%)	10 (83.33%)
SpectraIP B.V.	10	2	0 (0.0%)	0 (0.0%)	10 (100.0%)
DataWeb Global Group	9	3	4 (44.44%)	3 (33.33%)	2 (22.22%)
HZ Hosting Ltd	9	2	0 (0.0%)	0 (0.0%)	9 (100.0%)
IT-DELUX ltd.	9	3	1 (11.11%)	5 (55.56%)	3 (33.33%)
Des Capital B.V.	9	7	1 (11.11%)	1 (11.11%)	7 (77.78%)
WIBO Baltic UAB	8	1	8 (100.0%)	0 (0.0%)	0 (0.0%)
UA-HOSTING SIA	6	1	6 (100.0%)	0 (0.0%)	0 (0.0%)
The Infrastructure Group	6	5	1 (16.67%)	1 (16.67%)	4 (66.67%)
Zenex 5ive Limited	4	1	4 (100.0%)	0 (0.0%)	0 (0.0%)
G-Core Labs S.A.	3	1	2 (66.67%)	1 (33.33%)	0 (0.0%)
SOLLUTIUM EU	3	1	3 (100.0%)	0 (0.0%)	0 (0.0%)
Abelohost BV	2	2	0 (0.0%)	1 (50.0%)	1 (50.0%)
LLHOST INC. SRL	2	2	0 (0.0%)	0 (0.0%)	2 (100.0%)
ONLINE SAS	2	2	0 (0.0%)	0 (0.0%)	2 (100.0%)
Global Layer B.V.	2	2	0 (0.0%)	1 (50.0%)	1 (50.0%)
ColocationX Ltd.	2	2	0 (0.0%)	1 (50.0%)	1 (50.0%)
Totaaldomein BV	2	1	2 (100.0%)	0 (0.0%)	0 (0.0%)
Depositphotos inc.	1	2	0 (0.0%)	0 (0.0%)	1 (100.0%)
Host Sailor Ltd	1	2	0 (0.0%)	0 (0.0%)	1 (100.0%)
INXY LTD.	1	1	1 (100.0%)	0 (0.0%)	0 (0.0%)
Scaleway	1	1	1 (100.0%)	0 (0.0%)	0 (0.0%)
Hostiserver	1	1	1 (100.0%)	0 (0.0%)	0 (0.0%)
Baykov Ilya Sergeevich	1	1	1 (100.0%)	0 (0.0%)	0 (0.0%)
Alex Group LLC	1	1	1 (100.0%)	0 (0.0%)	0 (0.0%)
ABC Consultancy	1	1	1 (100.0%)	0 (0.0%)	0 (0.0%)

Table 3.1: Uptime measurement for hosting providers (November 2022 – December 2022)

Domain Identifier	Total Urls	Offline within 24h	Offline between 24-48h	Online for 48h or more
2149	841	3 (0.36%)	0 (0.0%)	838 (99.64%)
224	217	207 (95.39%)	1 (0.46%)	9 (4.15%)
2293	183	175 (95.63%)	6 (3.28%)	2 (1.09%)
963	157	153 (97.45%)	2 (1.27%)	2 (1.27%)
1864	108	0 (0.0%)	9 (8.33%)	99 (91.67%)
2105	106	106 (100.0%)	0 (0.0%)	0 (0.0%)
307	86	40 (46.51%)	8 (9.3%)	38 (44.19%)
2097	83	83 (100.0%)	0 (0.0%)	0 (0.0%)
2146	82	11 (13.41%)	0 (0.0%)	71 (86.59%)
2251	61	61 (100.0%)	0 (0.0%)	0 (0.0%)
1800	59	6 (10.17%)	21 (35.59%)	32 (54.24%)
1906	45	0 (0.0%)	0 (0.0%)	45 (100.0%)
321	43	38 (88.37%)	5 (11.63%)	0 (0.0%)
1767	42	0 (0.0%)	0 (0.0%)	42 (100.0%)
1330	40	0 (0.0%)	0 (0.0%)	40 (100.0%)
2046	37	26 (70.27%)	0 (0.0%)	11 (29.73%)
1981	37	18 (48.65%)	19 (51.35%)	0 (0.0%)
2156	34	0 (0.0%)	1 (2.94%)	33 (97.06%)
1803	27	0 (0.0%)	1 (3.7%)	26 (96.3%)
2163	25	11 (44.0%)	1 (4.0%)	13 (52.0%)
Others	658	255 (38.75%)	46 (6.99%)	357 (54.26%)

Table 3.2: Uptime measurement for domains (Nov 2022 – Dec 2022)

4

CSAM Landscape in the Netherlands

In this chapter, we explore which providers in the Netherlands host CSAM. The main goal of this chapter is to understand the current state of CSAM and how it has evolved. Because data from before 2020 was based on a manual NTD process, it cannot directly be compared to 2021 and 2022, since in those years an automated process was in place with the SCART system. So in this chapter, we focus on the data from January 2020 to December 2022 as provided by SCART. The aim of this chapter is the following questions:

- How has the volume of CSAM changed over time?
- How is the CSAM content distributed over hosting providers? How has this changed?
- How is the CSAM content distributed over domains? How has this changed?

We first discuss patterns in the volume of CSAM in the Netherlands and how this material is distributed over the hosting landscape. Next, we look at the domains where this material has been hosted.

4.1. CSAM Volume

In Figure 4.1, we show the number of URLs reported to hosting providers per month. The amount of CSAM detected and notified about in the Netherlands has fluctuated a lot over the past two years. In 2020, we saw many months with more than 20,000 URLs. This was also the case in early 2021. After that, the volume diminishes. It is tempting to interpret this as a declining trend. Yet, in December 2022, we suddenly see an uptick again, back to 20,000 URLs. That being said, in the last 1,5 years, the overall volume was lower than in the first 1,5 years. It is unclear if this is a persistent trend or a temporary situation.

The fluctuations are the result of several interacting factors that cannot be disentangled. There might be more material available online at certain times, but there is also better detection going on, so more material is discovered. There are also fluctuations in the intake of data from the INHOPE network. Since Canadian hotline left in 2020, fewer URLs are reported to ICCAM. IWF, the U.K. INHOPE member, is now a major source of data for ICCAM. The detection process at IWF is not public and we cannot evaluate how it impacts the discovery of CSAM in the Netherlands.

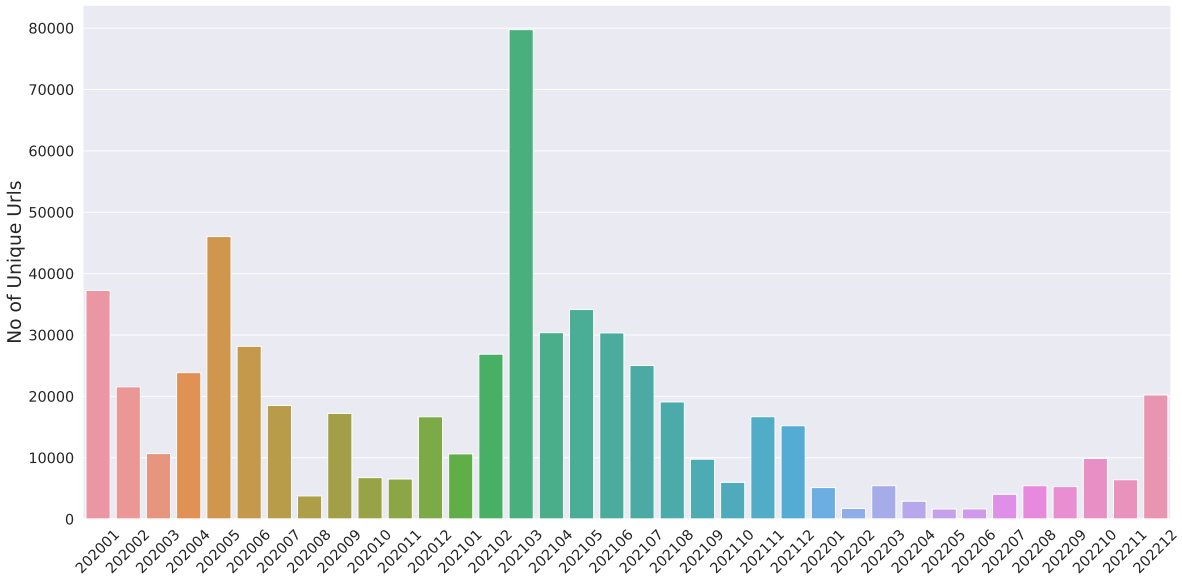

Figure 4.1: Number of URLs notified per month after SCART is deployed

Jan(2020) - Dec(2020)		Jan(2021) - Dec(2021)		Jan(2022) - Dec(2022)	
Provider	**No of URLs**	**Provider**	**No of URLs**	**Provider**	**No of URLs**
NFOrce	216,183 (92.76%)	NFOrce	269,423 (91.33%)	Amarutu Technology	30,895 (51.13%)
KnownSRV	5938 (2.54%)	KnownSRV	8570 (2.90%)	KnownSRV	10,850 (17.95%)
IP Volume	5558 (2.38%)	HostSlim	2602 (0.88%)	Telegram Messenger	7112 (11.77%)
Leaseweb NL	1907 (0.81%)	Telegram Messenger	2552 (0.86%)	HostSlim	4976 (8.23%)
HOSTKEY	1255 (0.53%)	MSK.HOST Hosting	1629 (0.55%)	NFOrce	1559 (2.58%)
Others	2200 (0.94%)	Others	10,199 (3.45%)	Others	5031 (8.32%)

Table 4.1: Number of CSAM URLs hosted in Hosting Providers

4.2. Distribution of CSAM Across Hosting Providers

In Table 4.1, we show distribution of CSAM across hosting providers in the period January 2020 – December 2022. Notwithstanding the fluctuations in the data, the distribution of URLs across providers reflects some stable patterns. The material is concentrated in a small number of providers. Four providers have the overwhelming majority of all URLs: 98.49% in 2020, 95.97% in 2021 and 89.08% in 2022. In the first two years, the bulk was hosted by NForce and KnownSRV. Yet, in 2022, the fraction hosted by NForce drops dramatically, from 91.33% to a mere 8.23% of all content. We already signalled this shift in our previous report. At the end of 2021, NForce asked some legitimate domain owners who were being abused by people uploading CSAM to leave NForce. KnownSRV remained in the number two position, with a similar number of URLs (ranging from 5,000-10,000). This now represents a higher share of the total amount of content, because the total number of URLs is much lower in 2022 than it was in the years before. Amarutu, a new provider in the top 5, moves into position one, seemingly out of nowhere.

If we look at the composition of the top 5, we can see some other shifts (see Table 4.2. While NForce and KnownSRV are in there consistently, other companies shift positions over time. In 2021, three providers move into the top 5: HostSlim, Telegram Messenger (no relation to the encrypted chat service), MSK.HOST Hosting. These providers were not completely new. They had been hosting small quantities of CSAM in earlier periods. Though they have a slightly larger share of material than before, it still represents only a small fraction of the overall volume. HostSlim and Telegram Messenger remained in the top 5 in 2022, with higher absolute numbers of URLs.

Leaseweb NL and IP Volume dropped out of the top 5 in 2021 and did not return in 2022. Beyond the top 5 providers, there is a long and fluctuating list of providers that have a small number of reported URLs. In total, 52 hosting providers received at least one NTD in 2020, 87 did so in 2021 and 2022.

Hosting Provider	2022	2021	2020
Amarutu Technology	1		
KnownSRV	2	2	2
Telegram Messenger	3	4	
HostSlim	4	3	
NFOrce	5	1	1
MSK.HOST Hosting		5	
IP Volume			3
Leaseweb NL			4
HOSTKEY B.V.			5

Table 4.2: Top 5 hosting providers across three time periods

Figure 4.2 shows that the picture over the past two years has been mostly stable month to month, until December 2021. Until that point, NForce had the bulk of all material. In many months, their share consisted of over 90% of the URLs. The remaining fraction was being hosted by some recurring providers, like KnownSRV, and a long list of fluctuating providers.

This pattern changed dramatically during the last months of 2021. The portion of CSAM hosted by NForce dramatically decreased. They ended up with less than 20% in January 2022 and nearly disappeared during the rest of the year, except for some small fractions in later months. Over the same period, Amarutu emerged as the new number 1, while being more or less absent in prior years. Knownserve remained a stable presence, but as we discussed earlier, its relative share increased, as it visible in the figure.

All in all, we see that over the course of three years, the picture of which providers host the bulk of the material has completely changed. Comparing 2020 to the second half of 2022, the whole top 5 has changed, except for KnowSRV. We know that this is partially in response to the government initiative to engage the private sector. NForce said that they consciously decided to sever ties with some customers who were operating legitimate domains, like image hosting domains, that were being abused by miscreants sharing CSAM. Other providers also indicated they increased their efforts to prevent the presence of CSAM and, if detected, to get it removed swiftly.

4.3. Distribution of CSAM Across Domains

To better understand the concentration of CSAM at hosting providers, we are taking a closer look at the domains that host most of the domains. We know from the annual report of INHOPE and EOKM that websites, including image-hosting and file-hosting websites, make up the bulk of the discovered CSAM material.[1] This pattern holds true globally, as well as in the Netherlands.

The variance across these domains is enormous. Some of domains contain thousands of URLs with CSAM, while others contain only a handful. The networks of the top 5 hosting providers have a disproportionately large portion of all these domains. In Figure 4.3, we plotted the portion of domains located with the top 5 providers from 2022, going back to 2021 and 2020. The portion hosted by the top 5 roughly fluctuates around 40%. While that signals some concentration, it also shows that the number of domains outside the top 5 is larger. As said, most of these domains contain only small amounts of CSAM.

Looking at the domains, we gain some more insight into how the landscape has changed in 2022. Figure 4.4 plots what portion of all URLs were located on the top 15 domains from 2022. As we can see, these domains all emerged as CSAM locations quite recently. Before December 2021, they were

[1] See p. 36 of the annual report of INHOPE for 2021 (`https://inhope.org/media/pages/articles/annual-reports/8fd77f3014-1652348841/inhope-annual-report-2021.pdf`)

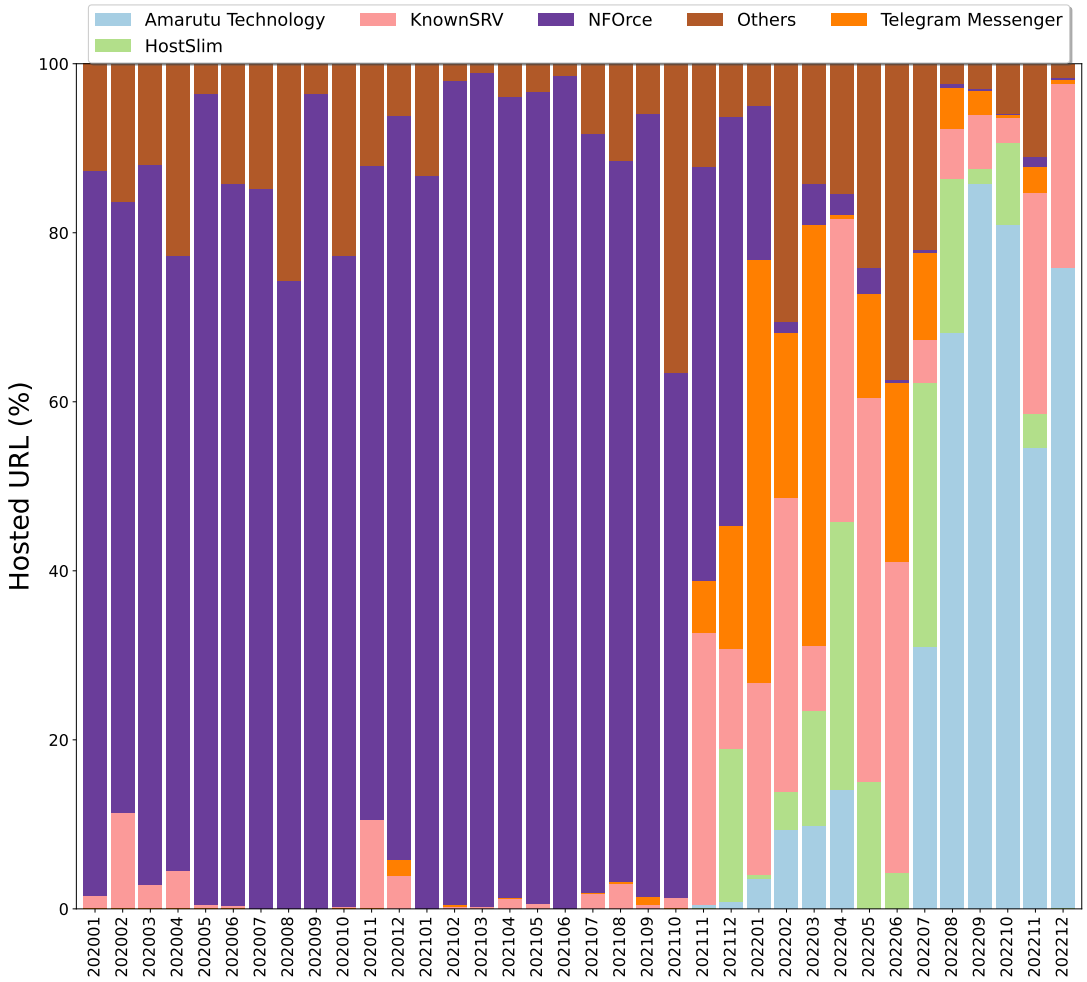

Figure 4.2: Number of URLs notified per hosting provider for each month, for the top 5 of 2022

mostly absent from the CSAM landscape. Even month to month, we see large fluctuations. Domain contains a lot of CSAM in December 2021 and January 2022 and then all but disappears. We see similar patterns for most other domains. This means it is hard to pinpoint where the problem is occurring. This makes it harder for hosting providers to engage with their customers (domain owners) and hold them accountable, as the CSAM content shift quickly to other domains. Along the same lines, it makes it harder for the government to hold hosting providers accountable, since the problem is shifting even within its customer base.

Even though the CSAM is concentrated in a small number of providers, the pattern at the level of domains is much more dynamic. In Table 4.3, we take the top 10 domains in 2022 (left columns) and then look at their position for 2021 and 2020 (middle and right columns). As can be seen, none of the domains from the top 10 of 2022 were in the top 10 in 2021 or 2020. If we look in 2019, none of these domains were anywhere near the top 10. The domains that were seen before were ranked much lower and had only a tiny fraction of all CSAM in those earlier years (1.20% in 2021 and 0.99% in 2020).

In Table 4.4, we make the same comparison for the top 10 domains from 2021. Only one of them is seen again in 2022, ranked 11 with only 1.55% of all URLs. The rest is absent. In 2020, so the year before the 2021 top 10 domains were established, four of those domains were already in the top 10 and four others also had a small fraction. So there was a bit more continuity in CSAM hosting from 2020 to 2021 than from 2021 to 2022. For completeness sake, we also present the comparison for the top 10 from 2020 (Table 4.5). It confirms the same observation: four domains remain in the top 10 in the next year (2021), and the six other domains still host at least some content in 2021. Yet they all dissappear going into 2022.

In sum, at the level of domains, 2022 brought about a major shift in the CSAM landscape, even more

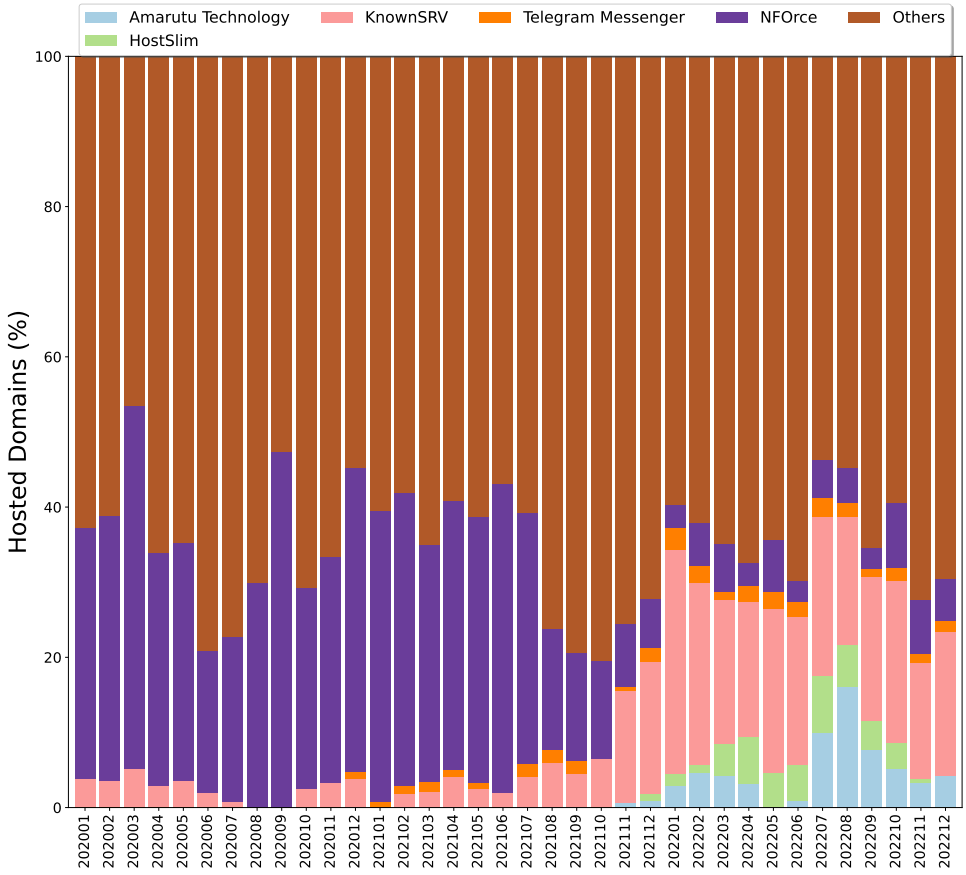

Figure 4.3: Proportion of all domains with CSAM per provider (Top 5 Hosting provider 2022)

than at the level of hosting providers. This does pose a bit of a puzzle, as we observed in earlier years: if CSAM-hosting domains are so dynamic, then why is the pattern at the level of hosting providers a bit more stable? This would suggest that the new CSAM-hosting domains end up with the same providers, to some extent. We have no good explanation for this. We know that it is not the effect of their size. If the hosting providers are simply the largest providers, they would attract a disproportionate number of domains to their network. But they are not the largest provider, so that explanation does not hold. The only other potential factor we observed is that some new domains are owned by the same people as the old domains, according to WHOIS – e.g., some owner of an image-hosting domain sets up another image-hosting or file-hosting domain. Yet this is too anecdotal to really explain the pattern.

The key takeway for policy and the PPP combating CSAM hosting is that the landscape has changed a lot over the last year and this is partially due to its own efforts to encourage the leading hosting providers to take more action. They did and as a result, the problem is now at other providers. We also saw a much lower volume in 2022, but it is too early to tell whether this is a persistent change, given the uptick in volume in December 2022.

Domain	2022			2021			2020		
	Rank	Total Urls	%	Rank	Total Urls	%	Rank	Total Urls	%
2074	1	9729	16.10						
2149	2	9169	15.17						
321	3	7004	11.59	14	2536	0.85	31	318	0.14
2146	4	5113	8.46						
2046	5	4005	6.63						
1864	6	2815	4.66						
963	7	1414	2.34	37	529	0.18	64	56	0.02
2061	8	1246	2.06						
307	9	1205	1.99	40	502	0.17	18	1942	0.83
1844	10	950	1.57						
Total		**42,650**	**70.57**		**3567**	**1.2**		**2316**	**0.99**

Table 4.3: Comparison of top 10 domains relative to 2022

Domain	2022			2021			2020		
	Rank	Total Urls	%	Rank	Total Urls	%	Rank	Total Urls	%
2				1	165,300	56.26	1	49613	21.28
82				2	15642	5.32	2	25559	10.96
7				3	9021	3.07	25	1022	0.44
820				4	8168	2.78			
656				5	7928	2.70	20	1721	0.74
247				6	6796	2.31	5	15681	6.73
1681	11	935	1.55	7	6003	2.04			
364				8	5095	1.73	9	7734	3.32
413				9	4765	1.62	11	7609	3.26
970				10	4633	1.58	79	35	0.02
Total		**935**	**1.55**		**233,351**	**79.41**		**108,974**	**46.75**

Table 4.4: Comparison of top 10 domains relative to 2021

Domain	2022			2021			2020		
	Rank	Total Urls	%	Rank	Total Urls	%	Rank	Total Urls	%
2				1	165,300	56.26	1	49613	21.28
82				2	15642	5.32	2	25559	10.96
601				25	9021	0.46	3	24940	10.70
357				41	8168	0.17	4	16617	7.13
247				6	7928	2.31	5	15681	6.73
789				12	6796	1.20	6	12592	5.40
556				11	6003	1.22	7	11341	4.86
671				21	5095	0.72	8	11152	4.78
364				8	4765	1.73	9	7734	3.32
242				29	4633	0.24	10	7675	3.29
Total					**204,628**	**69.63**		**182,904**	**78.45**

Table 4.5: Comparison of top 10 domains relative to 2020

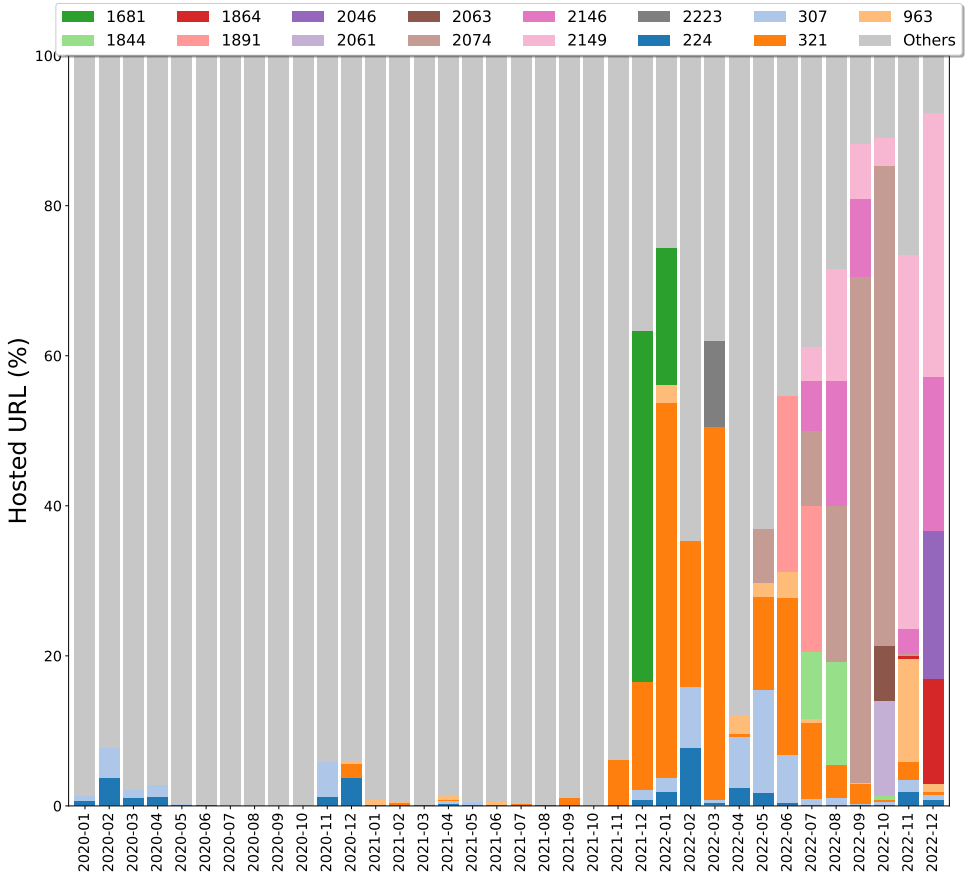

Figure 4.4: Percentage of URLs hosted per domain (Top 15 domains from Jan 2022-Dec 2022)

5

Conclusion

The CSAM Hosting Monitor is one instrument in a broader set of public-private actions to combat the hosting of CSAM in the Netherlands. It helps stakeholders by making transparent which hosting providers have most of the material in their networks, how fast they are in removing the material once they receive an NTD, and how the location of the material changes over time.

In principle, the monitor would also track the progress of the overall approach to combat CSAM hosting. Simply put: is the amount of CSAM hosted in the Netherlands going down? And if so, which providers and domains are improving the most and which ones are not improving? These are critical questions to guide the future actions of the public-private partnership to combat CSAM hosting. In reality, however, the data on these questions is surrounded with uncertainty and noise. The volume of CSAM hosting in the Netherlands is highly volatile. There are a lot of interacting factors that together drive thay volume. Better detection leads to more known CSAM. This is driven, among other things, by the crawlers operated by IWF. These crawlers are a black box. One month they detect a lot, the next month very little. Also, when the location of the CSAM changes, the crawler might look for it in the wrong place, at least temporarily. A second factor is the behavior of the criminals – that is, the people uploading the CSAM. They choose certain services at one point and they might decide to move their material somewhere else, when needed. Thus, providers might suddenly be struck by a spike in CSAM, even though they did not change anything on their end. A third factor is the actions of the providers and the domain owners, like the adoption of the hash-check service or severing ties with legitimate customers because their services are being abused by criminals.

These factors cannot be disentangled in the data. So yes, sometimes there appears to a clear downward trend in the volume of CSAM. For example, when then-minister Grapperhaus announced in June 2020 that he had “turned over the hour glass” and that the clock was ticking for providers and their customer domains to act, it did seem like the volume of CSAM subsequently went down.[1] A major part of the downward trend, however, was driven by changes in the data supplied to EOKM. It is unclear that it has anything to do with the public-private actions. The reduction was also short-lived. In the course of 2021, the volume went up again, even though providers were doing more than before, rather than less.

An important countermeasure has been the launch by EOKM of the so-called hash-check service. Some hosting providers have been urging their customers who struggle with CSAM on their domain to adopt this hash-check service.[2]. The list of domains that are using the hash-check service is confidential,

[1] See `https://www.rijksoverheid.nl/documenten/kamerstukken/2020/07/07/tk-voortgangsbrief-aanpak-online-seksueel-kindermisbruik-en-kindersekstoerisme`.

[2] See for more information the letter to parliament by the minister of Justice and Security on the progress made in the fight against CSAM, available at: `https://www.rijksoverheid.nl/documenten/kamerstukken/2020/07/07/tk-`

as is the volume of queries that domains run against the service. So it is not possible for us to connect the trends in volume for certain domains against the list of users of the service.

The top 5 providers with most CSAM changed substantially in 2022. NForce, the former number 1, has moved to place 5, after asking certain customer domains to leave — domains that were struggling with preventing CSAM uploads. A completely new hoster took the number 1 position: Amarutu Technologies.

This shift had several effects. First of all, some of the domains that left NForce moved to other countries. In some cases, these countries have no CSAM hotline, so the material is less rapidly removed – or not at all. Amarutu Technologies has a very poor record of taking down CSAM: only 4% of the URLs are removed in 24 hours. Perhaps this reflects their lack of experience in dealing with this material or the fact that this provider has not (yet) been the focus of government engagement. This stands in contrast to NForce, which in 2021 got 100% of all URLs taken down within 24 hours. The number 2, KnownSRV, has been consistently in the top 5 over the last two years. Its volume is not going down. And now its performance in removing CSAM also got significantly worse compared to 2021: only 58% is now removed within 24 hours.

One way of summarizing these results is to say that certain providers were the focus of the public-private initiative to combat CSAM hosting. They have been responsive to the policy. They were already actively removing the material, but many of them increased their efforts. In response, the problem has shifted to different providers, providers who have not been effectively engaged, either because they ignored earlier signals that they need to act or because they just emerged as important providers of CSAM hosting.

For the monitor, the next steps are to transfer the developed tools and lessons to the new public authority that will oversee the removal of CSAM. If and how the monitor will be continued by the authority, is still being explored.

voortgangsbrief-aanpak-online-seksueel-kindermisbruik-en-kindersekstoerisme. See also: https://www.meldpunt-kinderporno.nl/over-ons/hash-database/