Exploring the removal of online child sexual abuse material in the United Kingdom: Processes and practice

Elena Martellozzo[1]
Middlesex University, UK

Jeffrey DeMarco[2]
Middlesex University, UK

Abstract

This paper explores the processes involved in the removal of online child sexual abuse material. It specifically focuses on the work of the Internet Watch Foundation (IWF) in the UK who are partially responsible for the removal of this content. The empirical work was carried out between May 2017 and September 2017 and explored whether the IWFs processes at removing online child sexual abuse material were both useful and effective to the police and the wider communities. The research applied a mixed methods approach: Semi-structured interviews with employees (N=10) and key stakeholders (N=9), seeking to explore the strengths and challenges of both the task and the IWFs remit. Both employees and stakeholders saw the value in the original and innovative 'space' and 'approach' the IWF took in removing child sexual abuse material. This included the evolving nature of their tools, from the original URL list filter, to the more adaptable image hashing process. However, challenges around transparency, visibility and partnership were also raised. With online child sexual abuse ever evolving, it is important to consider novel ways in which intervention and prevention of victimization and offending can take place. Where a multi-disciplinary approach is needed in supporting victims, this research provides and insight into how one such organisation uses tools and techniques, different to traditional statutory services or law enforcement responses.

Keywords: Child sexual abuse material online; Internet Watch Foundation; Child Protection; Cyberspace; Child Abuse

[1] Associate Professor in Criminology, Middlesex University, The Burroughs, London NW4 4BT, United Kingdom. E-mail: e.martellozzo@mdx.ac.uk
[2] Lecturer in Psychology, Middlesex University, The Burroughs, London NW4 4BT, United Kingdom. E-mail: j.demarco@mdx.ac.uk

Introduction

The sexual abuse of children and the production of child sexual abuse material (CSAM) existed before the advent of the internet (Martellozzo 2013). However, for some individuals, the internet acts as a catalyst for the emergence of dormant sexual interests (Taylor, M and Quayle, E 2003). For others, with a well-established sexual interest in children, it represents an interactive medium that easily supplies child abuse material (2003). The online circulation of this material, independent of motivations behind its production and distribution, has serious and long-term consequences on children and their families. It needs to be taken very seriously. CSAM first surfaced as a societal problem in the late 1970s and this has been reflected in the increase in legislation worldwide (Adler 2007). However, despite increased legislation against CSAM, its production and distribution across the globe has not ceased.

The primary aim of this article is to present research findings on the utility and effectiveness of the processes involved in removing illegal sexual content by the Internet Watch Foundation (IWF) in the United Kingdom. It is structured as follows: firstly it provides an overview of how CSAM has been classified over the years; it then focuses on the links between the IWFs primary provisions, including Notice and Takedown requests (NTD); the management and evolution of their URL list; the 'hashing' of content; and IWFs effectiveness in and the provision of support to multi-stakeholders in securing and implementing a safer cyberspace. In the past years, several debates have been put forward when discussing the most appropriate definition or terminology to adopt for child pornography without minimising its seriousness and criminal nature. Most organisations and scholars across Europe now utilise the term CSAM, as it ensures that this crime is not confused with other forms of legal adult pornography. This is also the term the authors use throughout this article.

Understanding CSAM

To distinguish between abusive content and to facilitate the sentencing process, Combating Paedophile Information in Europe's (COPINE) typologies have been integrated into English and Welsh law (Quayle, Vaughan et al. 2006; Quayle, Taylor 2003). These are based upon ten categories of images that may be sexualised by an adult perceived to have a sexual interest in children. Figure 1 below illustrates the differences between these ten categories.

Table 1 Original COPINE Scale

LEVEL	DESCRIPTION
INDICATIVE	Non-erotic and non-sexual images of children—can come from commercial use or family photo albums. Possession of collection highlights inappropriateness
NUDIST	Pictures of children semi-nude or nude but in nudist settings
ERO	Photos or images of children in normal environments where the pictures of their underwear are generated
POSING	Pictures of the child posing in various states of dress however the collection of the images together suggests intent
EROTIC POSING	Pictures where the child is deliberately posing in various states of dress ina sexual manner
EXPLICIT EROTIC POSING	Focus of images are on the genitals of the child, regardless of state of dress
EXPLICIT SEXUAL ACTIVITY	Pictures showing mutual and self-masturbation, oral sex and intercourse by a child—will not involve an adult in the images
ASSAULT	Children being touched sexually by an adult
GROSS ASSAULT	Depiction of sexual assault including penetrative sex, oral sex or masturbation, involving and adult
SADISTIC/ BESTIALITY	Either images of children being tortured (e.g. whipped, bound) or images depicting sexual activity between children and animals

The degree of sexual assault becomes more severe as one climbs the scale. In 2002, the scale was converted to become one that could be graded by the Sentencing Advisory Panel (SAP). They range from level one, where children pose nude or semi-nude, through to level four, where penetrative sexual assault takes place. Level five, where acts of sadism and bestiality are exercised, is considered the ultimate level (Quayle 2008). Although this categorisation may contribute to the sanitation of indecent images, it represents an important step forward in recognising the seriousness of what images depict (Martellozzo, Taylor 2009).

To minimise the challenges presented by the classifications of images and potential subjective assessment, the Sentencing Council proposed a simplified version of the five levels comprising A, B and C (Kloess, Woodhams et al. 2017). Category A images tend to involve sadism or bestiality and would involve penetrative sexual activity with a child. Category B images would depict non-penetrative sexual activity. Category C relates to indecent images not falling within categories A or B. The Sentencing Advisory Panel adopted by the Court of Appeal, advised that all child abuse images should be allocated to one of these three categories as of 1st April 2014 (Sentencing Guidelines Council 2012). The classification of CSAM set out by the Sentencing Council Guidelines assist the IWF in assessing this material. However, understanding these categorisations and is only one element of the safeguarding processes for young people in cyberspace. Other tools and processes allow the IWF to provide a comprehensive and technological partial solution to the spread of CSAM.

The role of the Internet Watch Foundation

The arrival of the Internet has facilitated the production and distribution of indecent images of children on an enormous scale, in a short space of time, and at almost no cost. The repercussions of this phenomenon on the children depicted on those images are serious and long term. In the UK, the IWF is the leading organisation responsible for removing tens of thousands of webpages, images and videos of child sexual abuse content from the web each year. By removing such content, it helps victims depicted in those images and it contributes to make the internet a safer place for anyone to use.

The IWF was founded in the UK in 1996 to operate a 'hotline' for reports of CSAM from the public. It employs trained analysts whose role, amongst others, is to identify where CSAM is hidden through the analysis of these reports and to, consequently, remove any illegal images. Furthermore, in the past ten years, the IWF has started to identify the language sex offenders use and developed a vast database of around 450 keywords and phrases use by sex offenders to refer to illegal content. The keyword list is shared with the IWF's members, such as Google, Microsoft, Facebook etc. who can help limit the searches on their services and identify where the illegal content is being shared (Burgess 2020). This is in addition to a database of hashed images that companies use to stop existing abuse content being uploaded, also discussed later in this article. Information regarding the child abuse images hosted in the UK is communicated to the UK police.

If the sites are in the UK, then the police may act upon them directly, whereas if they are hosted elsewhere in the world, a report will be passed to the relevant authorities in that country or to the International Association of Internet Hotlines (INHOPE) representative in the suspect country. Within the UK, the IWF will also pass reports directly to internet service providers (ISP). The role of the IWF has expanded over time, to encompass its own proactive investigative work. The importance and effectiveness of this work will be presented, reported upon and discussed.

Notice and Take Down

Notice and Take Down (NTD) is the procedure followed by many ISPs in responding to court orders and other allegations of the presence of illegal content on their website or platform. This includes content that is sexual and abusive in nature, but also commercial, such as copyright infringement. Illegal content is removed by the host, following receipt of a notice (content may not be removed if notice is deemed vexatious or inaccurate). In relation to the IWF, the presentation of an NTD request to any ISP leads to the removal of CASM in the UK. The inner-workings, effectiveness and efficiency of this process have not been researched to date. Reviewing current evidence base around the removal of CSAM, can assist in providing transparency on the form and function of the key processes used.

Moore, Clayton and colleagues (2009) discussed the impact of incentivising NTD requests. This refers to the removing parties 'gain' in the removal of the content. For example, a bank would be highly motivated to remove a scam that was using their branding, but not that of competitors. With CSAM, it was perceived that ISPs or social media organisations would be highly motivated to remove this content due to the reputational risk hosting that content may have. However, the authors argued that there was a lack of standardised practice. These practices differed on three levels: the incentives, or motivation for removal; the legal framework for action; and the speed at which material can be removed. The authors suggested that the speed at which various types of content was removed highlighted that the requester's incentives outweigh other factors, from the potential penalties, to the methods used to obstruct take-down. The effectiveness of removal seemingly depends more on the incentives for this to happen, than on issues such as the legal basis or the type of material involved.

It is impractical for ISP's to police the entirety of the content that their users place on the Internet, so it is generally perceived by their organisation as unjust to bear the

entirety of liability. However, the ISPs are in an unrivalled position to suppress content held on their systems by removing access to resources, such as access to the platform. Therefore, many content removal regimes make ISPs liable for content once they have been informed of its existence. If they fail to 'take-down' the material, sanctions may be brought against them.

While the police are vital to the process of dealing with child sexual abuse images, there are jurisdictional issues around working across international borders. It is not within the IWFs remit to issue takedown notices to anyone outside the UK. However, according to INHOPE best practice, the IWF do follow up and send an email for advocating the removal of overseas content.

Hash Value Systems

Hash value systems are broadly used by bodies such as the IWF and differentiate from URL list blocking in that the illegal file itself is targeted rather than where it is located (McIntyre 2013) . Hash values are used not to block access to images but to ensure they can be located and removed at source. This approach relies on the use of hash values which can uniquely identify a particular file or photograph (Martin 2018). Bodies such as the NCMEC in the US and the IWF have compiled databases of hash values known to match to CSAM files and can compare the hash values of files stored or spread by users. If there is a match, they will be able to identify the file as CSAM. The results can then be presented in textual format to an analyst, and if based on this information it is decided to analyse the items further, they will be downloaded for review. In the USA, AOL pioneered the use of this strategy through its Image Detection and Filtering Process (IDFP) which it has run since 2004. The IDFP scans all emails sent by AOL members, generating hash values for any images being transmitted. Those hash values are then compared with an internal database containing the hash values of child abuse images previously dealt with by AOL and will remove any content that has a match. In the USA, the Cyber Tip Line at NCMEC will then be notified by receiving a report containing the image, username, email address and zip-code of the user. The next stage of the process requires the NCMEC to notify law enforcement who can then request the user's full details from the ISP. This system has resulted in numerous convictions within the US. In 2008, the government passed the PROTECT Our Children Act which authorises the NCMEC to provide hash values to Internet providers to spot and remove child abuse material.

Following this template, the New York Attorney General's office established its own hash value database, which is now being used by Facebook amongst others to detect and remove child abuse images (Attorney General Cuomo 2010).
Hash value systems concentrate on the location of images rather than the images themselves. Therefore, a review of each web address is required and unlike hash value systems, will fail to detect the same image if it is moved to a new location on the web. Hash value does not rely on the image location and will correctly identify and remove files even though they are being transmitted from a new location, whereas DNS or URL based blocking will fail.

URL List blocking

The URL list is a list produced by analysts working at the IWF who have located an identified webpage that are hosting CSAM in image of video form through active searching for illegal content. The list is updated twice daily, where new sites are added and those where content has been taken down are removed, so the list is always current for ISPs to apply to the filtering processes. As the IWF analysis proactively search and identify the sites, every page is assessed by an individual who is trained on identifying the content, and grading it using the SAP three categories, as previously discussed. Every IWF member has access to the list, and therefore the ability to block access to the webpages hosting this content-with an additional option of preventing the sites returning in searches run by internet users.

Methodology

The research applied a mixed methods approach of three stages, including the application of 'triangulation' to ensure the richness, reliability and validity of the findings. This included:

Stage 1 involved a targeted literature review. Documents were selected through discussions with the IWF and the research team and the synthesis of the literature intended to highlight emerging concepts in policy relevant to the prevention of online sexual exploitation and abuse of young people. The literature also provided current insights on child protection legislation and the NTD processes. The document review also provided the research team with key themes and ideas for inclusion in subsequent stages of the study.

Stage 2 included ten semi-structured interviews with employees working in various roles across the IWF and nine semi-structured interviews with key stakeholders. This data collection process took place between May 2017 and September 2017, within the premises of the IWF and remotely, with the key stakeholders. Employees were questioned on a variety of topics about the IWF and their role in the removal of CSAM, its visibility and transparency, and a primary focus on the main processes and functions and their effectiveness. These included, but were not limited to, the maintenance and distribution of their URL list; the issuing of NTD requests to website hosts and ISPs; and the continued development of the Hash List tool. Strengths and limitations of these processes were identified; with recommendations on areas for improvement. Key stakeholders were questioned on a range of topics and issues pertaining to their understanding of the effectiveness of the IWF across its key objectives. This included a focus on NTD practices; the construction, implementation and evolution of the URL list; and the hash list. Finally, the perceived transparency of practice and visibility to partners, stakeholders and the public was discussed.

Stage 3 involved a quantitative stage of the study included two surveys split between stakeholders (N=36) and IWF employees (N=31) to explore the processes involved in their roles and effectiveness of the organisation in general. This stage was implemented after the qualitative stage was completed and is not reported on here.

All participants (both stakeholders and employees) were recruited through the IWF technical research lead and through a combination of snowball and opportunity sampling. Both employees and stakeholders were invited to volunteer for the interviews and to complete the questionnaire.

Analysis

The qualitative data was transcribed and input into NVivo 11 for organisation and analysis. Using ‘thematic analysis’ a process of familiarization with the data, leading to the generation of initial codes and the subsequent searching for themes along these codes was followed. Once the research team felt confident with the emerging data, the themes were reviewed and refined, themes were validated, and analysis conducted for the saturation of data. Transcribed interview data was managed and analysed using a structured approach, where key topics emerging from the data were identified through

familiarisation with the transcripts. An analytical framework was drawn up in the form of a matrix where columns represented the key sub-themes or topics and the rows represent individual focus groups. Data were summarized in the appropriate matrix cell for analysis. For both the employees and stakeholders' data, key topics from the interview schedules were used for the first level coding, dividing the analysis into categories across 3 key areas: *organisational roles, views and experiences of the organisation, key processes and activities*. It should be noted that many pieces of discourse/text from participants were multi-labelled to provide more complex analysis and a glimpse into the relationships between themes.

Ethics

All materials were designed by the academic research team and conformed to ethical guidance of the British Society of Criminology (2017) and the British Psychological Society (2014) . Each stage of the research was approved by the Ethics Committee of the School of Law at Middlesex University and by an independent advisory board, convened for this research project. In accordance with ethics guidelines, participants were given a briefing introducing the researcher and the background of what was involved. Information sheets were also shared in both hard copy during the interview, as well as an electronic version. Each participant was explained the voluntary nature of the study prior to commencing the interview. Informed consent was taken verbally on site with both sets of participants. It was explained to all participants that they could cease engagement in the interview at any time. They were informed that the discussions would be audio-recorded for analysis, and that their identities and responses would remain anonymous. Topic guides were created to direct the flow of discussion and prompt new leads.

Following delivery and collection of information, all participants were debriefed and provided with contact details for the researcher for follow-up questions.

Findings

Employees expressed their views on the IWF's processes as outlined in the previous section, geared towards the removal of CSAM. These revolved around the URL list, NTD and the organisations unique hash list.

IWF employee perspectives

URL List

The efficiency and the continual maintenance of the URL list was crucial in ensuring the list was current, and avoided accidental blocking of websites for unnecessarily long periods of time:

> *"…it is really important for us to make sure it doesn't stay on our list any longer than it needs to…"* **(E2)**

The membership base is required to download the latest list daily which has implications for the valid and reliable administration of the list, if only simply in order to keep up-to-date:

> *"…we need them to be downloading the latest list because the list can fluctuate so much, and we can have hundreds of URLs go on in one day..."* **(E3)**

Many members do not adhere to their obligations to download the list regularly, which can cause significant difficulties in its application and utility.

The lists function in protecting the public from CSAM was critical, and apart from protecting the public, supported the IWF team and law enforcement in reducing their workload:

> *"Where people are not stumbling across this content means that law enforcement resources are hopefully not going to be wasted on people who are not necessarily seeking this out…"* **(E9)**

Updating the URL list regularly ensured that sites on these lists were active, and that any URLs that have been blocked are removed from the list, avoiding the list becoming too long:

> *"It is checked by the analysts every day that the content is current and live."* **(E4)**

However, a weakness of the URL list was that perpetrators can navigate around its protective features, especially among more 'tech-savvy' users:

> *"…it is not a fool proof solution and if people are determined to get past it, they will…"* **(E9)**

Recent research from the Independent Inquiry into Child Sexual Abuse (IICSA) demonstrated that most perpetrators engaging in online childhood sexual abuse activities are manifesting increased technological skill (DeMarco, Sharrock et al. 2017). There is an argument that while the URL list is no longer appropriate in its function, perpetrators may no longer behave in a manner that would be successfully mitigated by this process.

When signing up to be a member of the IWF, there is a contractual agreement that the URL list is downloaded at least once per day. However, there are circumstances in which members do not adhere to their contractual agreements, and the URL list is not downloaded as promptly as required:

> *"The issues come more in terms of downloading the URL list, for example, because they have contractual agreements to download at least once a day and many of our members download twice or three times a day, but we can see from our stats that some of them are not downloading once a day and therefore we know they are not protecting their networks as well as they could in terms of having the latest URLs on the list."* **(E3)**

This latter point was raised as a point of conern by employees. Not only does the lack of uptake mean that some content remains active and available, but it also led to frustration for the analysts working through cumbersome and time-consuming process. Failure to download lists in a timely fashion thus wastes human resources.

NOTICE AND TAKEDOWN (NTD)

A key aspect of the NTD process was to follow up all requests swiftly. This ensured that the host received the notice, and encouraged them to act promptly:

> *"…as far as I am concerned, it just has so many strengths. We look at our take down times and we can effectively take down in 2 minutes if the person is at the other end of the phone…"* **(E2)**

Members were quick to act on these requests. Many employees appreciated this and attributed it to the fact that organisations who have signed a legally binding contract to access IWF services, were motivated at working towards reducing online CSAM:

> *"I think most members are members because they want to get rid of this, it is bad for business, and no one wants this on their servers unless your business is this. For the members, if we are notifying them direct, we will get emails*

> *straight back saying they've got that and they are pretty keen to get rid of this."* **(E5)**

A primary challenge identified in relation to the NTD process was the IWF's lack of legal power in enforcement. Having to go through the police meant that good relationships must be maintained with these agencies, in order to streamline the process:

> *"…the IWF has no legal power to enforce the take down so if someone refused to comply with our request, we would have to go back to [the local police force] and say we have been in constant contact with this person. [If] they are ignoring our requests or saying they won't take it down until they have spoken to their customer and then we would rely on [local police force] to make contact directly…"* **(E1)**

Having to rely on law enforcement agencies to make contact with the organisation issued with a notice also placed limitations on the speed with which the report could be completed.

> *"I suppose a weakness is, well it is not really a weakness but a problem with it is that we have to spend a lot of time checking whether it is still up or whether it has been removed and seeking people to encourage to remove the content so it is quite labour intensive, is my impression but to me is seems necessary…"* **(E7)**

The reasons for why sites may be reluctant to take down content included the financial remuneration involved in hosting commercial CSAM sites on their servers. Servers hosted in jurisdictions with a lack of awareness or processes to take down of CSAM can prove difficult in complying with requests—especially if the content is depicting offending/victimization from elsewhere. For example, there are several low-income countries where there is no national policing infrastructure to deal with the takedown or investigation of online CSAM and as such, the crimes are reported centrally through Interpol. In this case, it is difficult for the IWF to work at the local or national level.

HASH LIST

The main strength of the hash list identified by participants was its ability to prevent a high volume of content being widely distributed, improving the level of protection of repeat victimisation for those whose images are online:

> *"Its potential could be incredible. So, if a child or adult victim (adult now but was a child victim) of sexual abuse whose images are on the internet, it offers real hope that their images can be as close as being wiped as possible…"* **(E2)**

Hashing not only benefited the victims in the images but also decreased the need for analysts to view the same images repeatedly:

> *"…we can prevent our analysts from seeing the same image over and over again, so we are looking after them better…* **(E3)**

Employees also reported that the grading/categorization process of images via hashing was simple and rapid, observing that content which appeared repeatedly or is very similar in nature, can be quickly and effectively categorised:

> *"…you know what you are going to categorise for some of them just as soon as they come up because you know that image and you know what is going to come up, so you can categorize quite a lot of it quite quickly. I wouldn't have thought there was any way of doing it any quicker really."* **(E5)**

There is a lack of standard practice across agencies with the categorisation of images, and participants stated that it was important to consider the context of the picture when categorising:

> *"[Child Exploitation and Online Protection] unit for example will say a topless picture of a child is an indecent image but we would say not necessarily as there are so many contextual issues around that."* **(E9)**

Feelings of frustration were expressed when content continued to reappear online and could not be permanently removed from the internet regardless of the effort made by the analysts:

> *"It has been a difficult task for our analysts because hashing is pretty heavy going and from my point of view, I did have real concerns about the welfare of the analyst team and how much we push them in terms of hashing which is why we developed the welfare programme"* **(E3)**

This was perceived to be due to the lack if implementation by these large member companies, and an ignorance of the agreed protocol in that they were not being appropriately subscribed to.

> *"…it is a huge frustration when we know they aren't doing what they are supposed to be [doing], even though they are saying they are doing it, but it is manifestly obvious that they are not because otherwise we wouldn't be seeing this particular content on these sites."* **(E9)**

There was also a concern with the fact that only members who subscribed to the service through the IWF were eligible to use the hash list, suggesting that this extensively limits its level of its application.

> *"The cons are that companies have to be a member of ours in order to use it, but once they are a member, they have access to every service, we don't sell the hash list as an individual thing…"* **(E2)**

The positive impact that the maintenance and distribution of the URL list and the hash list may have on the ability for the IWF to fulfil its remit was discussed by employees. These tools help in combating the distribution of CSAM for the general internet user. But a second factor is that this limits the amount of damage on victims of repeated viewing of CSAM content:

> *"…we know it really impacts victims to think that people might be seeing that content over and over again, so I think that it is still a powerful tool…"* ***(E9)***

One of the main strengths identified, was the technical capabilities of the hash list; that when applied as specifically and widely as intended, will aid with law enforcement and the policing of content that is being uploaded:

> *"…provides opportunities for law enforcement because if you are an ISP or a member that has the ability to track your users, are uploading images that match to what is on the hash list then there is an immediate ability to identify the distributor of that content because in theory you have their IP address…"* **(E1)**

Proactive investigative work

Most staff commented positively on the expansion of the proactive work undertaken by IWF analysts. Other hotlines in Europe currently work in a reactive way in the sense that they take action only when they have received a report by an internet user, but they do not search actively for child sexual abuse content on the internet. Therefore, the proactive work that the IWF does is unique; carried out competently and successfully and is valued positively by its staff members:

> *"…the expansion and of course the proactive work has been hugely successful because 100% of it is actionable material which we get removed. We still do*

the public reporting but the amount that is actionable is still about 10%". ***(E7)***

Key challenges

Two primary challenges were identified by employees. Firstly, that there was often a difficulty in maintaining charitable status whilst operating within a business-oriented framework; thus, ensuring that the tools are dispatched as intended, and that the finances cover all the running costs:

> *"It is always a difficult balance to run as a charity and if you look at our charitable objectives, we are trying to get our URL and hash list out as far wide as possible in order to protect the networks and take down this horrendous content, but we are still a business."* **(E3)**
>
> *"…we have to be able to sustain the number of analysts we have got and the only way we can do that is to raise income and that is membership fees…"* ***(E3)***

It also became evident that a computer software tool had been developed specifically for the analysts to use in their day-to-day job when working on the hash list. However not everyone utilises it because they did not like using the Internet Browser that hosts the tool:

> *"The analysts won't use it because they don't like chrome and it is a really frustrating situation. Some of them will use chrome and some of them won't so it is not being used. I personally think the answer to that should be it doesn't matter if you like it or not, all respect to you but this is the tool you have been provided with and you will use this to do your job"* **(E9)**

Whether this is a widespread issue among all employees or not, it seems an important issue to address. If digital tools are being specifically developed to enhance and improve productivity, but some staff are non-compliant with their use, further training or discussion about alternative browsers are required for optimal efficiency. In addition, it was evident that the hash list was not the sole list in operation for identifying and removing CSAM. However, many existing lists were not linked, suggesting that much of the work done across organisations was occurring in silos, and duplicating efforts.

Stakeholder Perspectives

This section presents findings about the URL list, NTD process and the Hash list as discussed by stakeholders working in some capacity with the IWF. The stakeholders ranged from having no formal links with the IWF to long standing interactions in delivering their remit. These included representatives from government, technology industry, criminal justice, advocacy for children's rights/third sector and mental health practitioners.

URL List

A good understanding of how the URL list was constructed daily by the IWF and subsequently implemented by its members was shared by the stakeholders:

> *"…each URL leads to the resource that has been detected and allows for its particular resource to be blocked. It is then made available to the IWF partners and ISP's".* **(S5)**

The daily process of building and adhering to the URL list from member's perspective was described with confidence in the credibility and accuracy of the IWF's processes and trust in their expertise at the blocking, filtering and removal of content:

> *"…the list is applied to all internet traffic emanating from* [OMITTED] *so it's as close as it can be to 100%. Operationally, we apply the list once a day and I think it is offered twice a day. We offer it once a day and whilst over the last 13 years there have been one or two incidents, we basically accept the list as dead, we don't question it, we don't interrogate it unless there is an issue raised…"* **(S9)**

Stakeholders believed that the URL list was diligently updated and that this meant it was a very reliable resource in both its application and utility:

> *"The way that it works is that the list is used to access… sorry for blocking materials on websites and the key point here is that the list is permanently updated and that is what the IWF is doing a very good job at".* **(S3)**

The data available for extrapolation from the list was highly commended by a range of stakeholders, along with the enthusiasm from within the IWF to share relevant data such as current trends in the area; website function and type (e.g. social media, image boards, news groups etc.); countries hosting the most CSAM; victim profile and demographic information (e.g. age, sex and race); and location of abuse (e.g. bedroom, outdoor). This information may allow law enforcement to remain current with changes and evolutions in both perpetrator and victim behaviour.

> *"…I've gone to them and asked what is trending at the moment, can you let me know and they'll send me exactly that so again from my perspective, they haven't been shy about sharing with me".* **(S2)**

There were a number of limitations identified that impacted upon stakeholders depending on their area of expertise and affiliation. For example, when asked what they thought worked well or less well with regards to the URL list, stakeholders did not feel like they were able to comment on the list for reasons such as lack of transparency:

> *"I have no idea because there is no transparency. The IWF says it works well but I have no idea if that is true or not. I am prepared to believe it works well but I have no idea".* **(S1)**

Stakeholders suggested that of those organisations that avail of the URL list, there was a perception that the list lacks in depth and breadth and suggested that the IWF should increase the number of links on the list:

> *"To be honest, it is not that big, and I was surprised by how small it was…having seen the list that I was presented with, it didn't teach me anything new to be honest but that is not to say that the work they are doing isn't good – it is".* **(S2)**

Technical limitations included the increasing trend of web sites to employ SSL certificates, thus encrypting information on the web page and making it impossible to add an offending HTTPS URL to the IWFs URL list. At the time of writing, only

HTTP URLs could be added to the list. At the advent of the Internet, network administrators used HTTP domains to share information they put out on the Internet however as the world wide web grew over subsequent years and became ubiquitous, these addresses have increasingly been changing to more secure HTTPS domains.

> *"Well the big-ticket item is clearly the change of traffic from HTTP to HTTPS so the list as we see it today has obviously got a limited shelf life as we go forward…"* **(S9)**

There are also concerns that ultimately, the list is easily compromised by motivated users, and therefore potentially lacks integrity in preventing content from being accessed by offenders:

> *"…when you look at objectives of the charity then you have a tension, don't you because it is all very well saying we block stuff but there is a sort of pretence here; everybody knows that the IWF list is limited and it is easy to get around, but industry supports it…"* **(S9)**

This concern coupled with the questionable relevance of URL lists and the advance of encryption does further support the discussion surrounding changing perpetrator patterns and how much prevention/blocking the URL list is reliably able to cover.

Notice and Takedown

The IWF was applauded for the swiftness of NTD and it was highlighted as one of the organisation's greatest strengths. The contact and links to individuals in law enforcement and industry ensured that they could remove content expeditiously when needed:

> *"I think what is important is the speed because behind any image or video there is a child who is subject to a crime; they are a victim. What is important is that this material remains online as less as possible".* **(S3)**

The IWF's NTD was discussed as a key process about the relationship between hotlines, industry and law enforcement. It is central to the relationship with hosting providers in the UK, the working relationship with law enforcement and vitally, to the facilitation of seizing material before it is promptly taken down. Therefore, it is seen as a larger element than a process offered by the IWF, but a binding action plan that

brings diverse organisations together in working towards the ultimate goal of removing illegal content:

> *"So, hotlines are partners for the police in ensuring that they receive the images or the reports that are confirmed to be child sexual abuse material and the hotlines are also instrumental in the process of NTD because the do request the hosting internet service provider to take down the material expeditiously…".* **(S3)**

While an additional concern related to the lack of uniformity of the process and clarity on what is occurring internationally with regard to legislation reinforcing a NTD in different jurisdictions. In several countries, conforming to a NTD request is done voluntarily and is not a legal requirement was a complaint which was raised:

> *"The biggest problem for NTDs is the voluntariness of the procedure where there is no legal model reinforcing it. Self-regulation by the industry, in respecting NTDs, only works as long as they wish to be compliant. There is a definite requirement for legal process to be possible where such compliance is not guaranteed. Providing for such legal process through legislation or case law is necessary where it is not available".* **(S5)**

This again implies some criticism of the NTD processes, especially if the effectiveness of its implementation in the UK is compared to other jurisdictions. It is clear that the IWF benefits from the strength of its relationships and reliance on strong legislation, but industry compliance is also a big factor.
In order for NTDs to be continually effective, relationships with law enforcement agencies across the globe would need to be established:

> *"The necessity to liaise with LE, at a national level, and to build partnerships with them from within the countries implementing NTDs cannot be underestimated. Only through building national partnerships and networks within the INHOPE model can it be expected that NTDs will continue to be honoured and effective".* **(S5)**

Stakeholders suggested that there has been a shift away from NTD processes operated by the IWF to focusing on filtering and blocking through their URL list. Caution is needed when judging the list, to whether imagery is being blocked and no longer being hosted or displayed through UK servers or if the content has simply gone elsewhere on the internet.

Hash list

The IWF's hash list is understood to be one of the latest developments within the company and involves 'scraping' the Internet for hash numbers that match known CSAM, consequently allowing individual companies to filter out inappropriate content before it is uploaded to their servers:

> *"So, if you have a database of verified child sexual abuse material on which you take the hash list, if I compared the hash of a new image with the database and the check is positive then I know for sure that that image is known child sexual abuse material…it again can be used by the law enforcement authorities but also by the social networking site and those that host files because it will allow them to understand if they are dealing with material which they shouldn't be dealing with".* **(S3)**

However, there has not been universal implementation by those stakeholders queried:

> *"…we don't use it and we have no use for it at the moment. I believe there is possibility for us to use it, but I have not got widespread acceptance of that yet".* **(S9)**

The hash lists utility in preventing a proportion of CSAM from being uploaded online was seen as its ultimate benefit, while also preventing re-victimization as the tool prevents images from being uploaded onto their services:

> *"…you don't see the image and that is important so when you pass hashes on, you are not passing child sexual abuse images on, you are passing on a hash which is a mathematical formula from the image, so you don't spread child sexual abuse material…"* **(S3)**

This process is important in enabling analysts to surmise if they are dealing with historical images that have been on the Internet for an extended period of time, and are known to law enforcement; or if they are dealing with new CSAM that has never been processed.

"... if an operator is able to say yes this is an image that is portraying an abuse and they extract the hash list and they know that that hash is not contained in the database then they will inform the law enforcement and authorities and the law enforcement will take it up by including and pursuing the crime so…" **(S3)**

The hash list also provides social networking sites with the awareness of CSAM on their platforms, without the necessity for viewing images thus helping preserve the welfare of IWF employees:

"...this hash list is actually very important, and it again can be used by the law enforcement authorities but also by the social networking site and those that host files because it will allow them to understand if they are dealing with material which they shouldn't be dealing with". **(S3)**

While there was a view that the hash list had the most potential of the three processes, stakeholders felt there was a lack of guidance or mandate from governmental agencies about how they should implement the list, as it is a significantly different process to the used URL list:

"I believe that perhaps the hash data set represents the future and I have said that and discussed that with the IWF and I think these are the things we need to have a go at but it is questionable about how a company like [OMITTED] *might use it at the moment and there isn't a clear steer from government saying to ISPs that we want you to apply the list and this is what we want you to do with it because it is very different to the URL list".* ***(S9)***

It was explicitly stated that the need to integrate new technology did outweigh the technical challenges in implementing such changes. However, as well as the investment in new technology, specialist training and additional human resources were required to understand new technology and their integration into established systems.

"I think when you make those decisions to download hashes, it does require a substantial amount of investment in terms of being technically capable of doing it, committed to doing it long term, cross checking hashes, understanding the overlap with the existing hashes that we may have and then kind of a quality check and this is an ongoing resource that needs to be dedicated to it so I think from the [OMITTED] *side, we want to be mindful that when we engage in additional commitments that we really understand*

beforehand that what is going to be the impact on our technical and ops team". **(S8)**

It was suggested that the IWF should be mindful of technical changes they make as an organisation, as those decisions can impact its members in a number of ways, and potentially impede meeting their obligations in dealing with CSAM:

"…I think changes were made in the IWF side to allow for automatic download but the cost of making these changes requires again further investment from technical teams…it may be on the IWF's side but it has a ripple effect for people who are committed to doing this work so I think considering the benefits of making those changes and how it will impact existing processes that are in place across the board…" **(S8)**

There is a need for the IWF's hash list and the Home Office's Child Abuse Image Database (CAID), to be linked to ensure that images are shared appropriately between agencies and across sectors, nationally and internationally. Hash lists also need to be expanded and shared between agencies

"…the Child Protection System works off a database with over a million images (there are a lot more than that) but what it is, is a starting point… I mean at one point we had over 5,000 IP addresses linked to indecent images of children from that database and we can't deal with that so do we need a bigger hash set to be run across that system? Probably not, because a million is quite a lot but listen, they've got a database to work from, it can't be a bad thing". **(S2)**

Stakeholder expressed caution when discussing the advantages of hashing. It is important not to be over-reliant on hash lists as perpetrators may learn to circumnavigate them to avoid detection. A cautious approach to detecting and reporting offenders or criminal images was therefore expressed in order not to alert the offender community of special measures being used across organisations to reduce the amount of content available and circulating online.

"…we are going to have a big group of offenders in there who are probably finding computers interesting and hearing about hashes etc. so they are going to educate themselves about the hash, how do I get hidden, if I send something I will just change a pixel, so we have to be aware of not becoming over reliant in simple terms is what I am saying". **(S3)**

There is recognition that the hashing of videos has emerged as an added complication, which can keep CSAM offline and will be an added challenge for the IWF to tackle going forward:

> *"This is an interesting thing because you now get a hash of a video because someone out there will see a part of a film and that specific bit has turned them on and, so they edit the video".* **(S2)**

Discussion

What becomes clear from the evidence presented is the importance of using technology as an asset in combatting online CSA. The IWF has demonstrated that they are able to contribute innovatively and work with key agencies, such as the police, in tackling online CSA. They have, through their cross-disciplinary partnerships, demonstrated a willingness to innovate and collaborate with the technology industry and their membership base about how to improve their services and, ultimately, be more effective. This can be achieved through joint-partnership with equivalent organisations in the third sector, and the creation of formal alliances. This could also allow shared learning and knowledge transfer across processes and procedures so that other organisations can assist with this worth cause. Specifically, it may be helpful for the IWF to highlight where they can provide better information to relevant authorities, but also where they themselves could benefit from additional support and assistance.

As highlighted in this article, where their processes are praised, it is in their originality and how they contribute to on-going issues facing government and law enforcement in dealing with the volume of online CSA. However, their staff and stakeholders also discussed that as good as their processes and procedures are, there is area for improvement and transparency. There is a perception that some of these need streamlining, so policy colleagues and the wider technology industry can see how these are being delivered. This is important when considering a wider 'public health' approach in that the more the IWFs practice is endorsed by partnerships, the more likely the public will be to recognise and engage with them as a suitable and legitimate organisation. Although not presented here, the findings suggested that the IWF could do a great deal more in communicating their remit and solidifying their position in the public sphere.

Also, as with all technologies, things change. As evidenced in the findings, there is a belief internally and externally that, although the URL list has been commended, it should no longer play a priority within the activities of the IWF; this is due to changing victim behaviours, offender abilities, and general online behaviours. The hash list is perceived to be the way forward, however this as well requires substantial more attention and linking up then in its current form. However, as recently reported by Weird (Burgess, 2020), the IWF has invested time and resources to develop a database of keywords that have been created and utilised by sex offenders to refer to CSAM, whilst avoiding detection. Indeed, a greater understanding of sex offenders' behaviour, including the terminology they use, can help find websites and locate images that had not been seen before.

Conclusion

In a society where the boundaries between crime, abnormal behaviour and technology are increasingly being blurred, the IWF has a strong foundation in serving as a go-between for criminal justice agencies, policy and legislation development, and statutory services safeguarding children and young people. This paper aimed to showcase the important and novel work they are offering, while also suggesting areas where changes can be made, and partnerships strengthened in order to offer as holistic and useful service as possible. Moving forward, the IWF will continue to have a vital role to play in supporting communities in navigating the risk on online CSA. As they continue to evolve, so will the online risks for children and opportunities for offenders and as such, their continued development and innovation will be critical within the wider preventative models.

Bibliography

ADLER, A., 2007. All porn all the time. *New York University Review of Law & Social Change,* **31**, pp. 695-911.

ATTORNEY GENERAL CUOMO, 2010-last update, Attorney General Cuomo announces expansion of ground-breaking initiative to eliminate sharing of thousands of images of child pornography on social networking websites. Available: https://ag.ny.gov/press-release/2010/attorney-general-cuomo-announces-expansion-groundbreaking-initiative-eliminat [Nov, 2019].

BRITISH PSYCHOLOGICAL SOCIETY, 2014-last update, BPS Code of Human Research Ethics.

BRITISH SOCIETY OF CRIMINOLOGY, 2017-last update, British Society of Criminology Statement of Ethics
. Available: British Society of Criminology. 2017. British Society Of Criminology Statement Of Ethics. [online] Available at: <http://www.britsoccrim.org/documents/BSCEthics2015.pdf> [Accessed 12 April 2017]. [12 January, 2020].

BURGESS MATT, 2020, April 23. Researchers have finally cracked the secret paedophile code. *Wired*.

DEMARCO, J., SHARROCK, S., CROWTHER, T. and BARNARD, M., 2017. Behaviours and characteristics of perpetrators of online-facilitated child sexual abuse and exploitation. *A Rapid Evidence Assessment, NatCen.Prepared for Independent Inquiry into Child Sexual Abuse (IICSA)*.

KLOESS, J.A., WOODHAMS, J., WHITTLE, H., GRANT, T. and HAMILTON-GIACHRITSIS, C.E., 2017. The challenges of identifying and classifying child sexual abuse material. *Sexual Abuse*, pp. 1079063217724768.

MARTELLOZZO, E., 2013. *Online child sexual abuse: Grooming, policing and child protection in a multi-media world.* London: Routledge.

MARTELLOZZO, E. and TAYLOR, H., 2009. Cycle of abuse. *Index on Censorship,* **38**(1), pp. 117-122.

MARTIN, D., 2018. Demystifying Hash Searches. *Stanford Law Review,* **70**(2), pp. 691-733.

MCINTYRE, T.J., 2013. 12. Child abuse images and cleanfeeds: assessing internet blocking systems. *Research handbook on governance of the Internet,* , pp. 277.

MOORE, T., CLAYTON, R. and ANDERSON, R., 2009. The economics of online crime. *Journal of Economic Perspectives,* **23**(3), pp. 3-20.

QUAYLE, E., 2008. The COPINE project. *Irish Probation Journal,* **5**(9), pp. 65-83.

QUAYLE, E. and TAYLOR, M., 2003. Model of problematic Internet use in people with a sexual interest in children. *CyberPsychology & Behavior,* **6**(1), pp. 93-106.

QUAYLE, E., VAUGHAN, M. and TAYLOR, M., 2006. Sex offenders, Internet child abuse images and emotional avoidance: The importance of values. *Aggression and violent Behavior,* **11**(1), pp. 1-11.

SENTENCING GUIDELINES COUNCIL, 2012-last update, Sexual Offences Guideline Consultation. Available: https://consult.justice.gov.uk/sentencing-council/indecent-images-children/supporting...

TAYLOR, M AND QUAYLE, E, 2003. Child Pornography: An Internet Crime,

British Journal of Social Work, **33**(8), pp. 1129-1130.