OSINT by Crowdsourcing: A Theoretical Model for Online Child Abuse Investigations

Kemal Veli Açar[1]
Turkish National Police, Turkey

Abstract

For various reasons, the backlog of seized devices has increased to unparalleled levels, sometimes leading to years of delays in the trial process, even for online child abuse investigations. In addition to child abuse materials, the digital belongings of an abuser might contain small yet important pieces of information, such as nicknames, e-mails and place names. A thorough digital examination and an appropriate analysis of this set of information might reveal the exact locations or the real identities of criminal associates or victimized children. All these digital clues should be investigated properly through open sources, to identify the real identities of possible owners and discover whether their relationships with the suspect are crime-related. Considering the fact that thousands of such digital traces are present in all seized material for each separate investigation, a thorough examination of every piece of information related to a criminal case becomes a tremendously challenging issue for law enforcement agencies (LEAs). While the resources of LEAs are clearly insufficient, automated methods alone are not adequate to respond fully to current needs. In this regard, this article proposes a new and unorthodox way of handling the ever-growing workload of online child abuse investigators more effectively, by tapping into the energy of a carefully selected crowd of volunteers. After a brief literature review on related subjects, such as open source intelligence (OSINT) and crowdsourcing, in terms of their technical, legal and organizational aspects, the proposed theoretical model will be elaborated. Likely concerns and probable bottlenecks relating to the same respective aspects regarding successful actualization of the model will be identified and thoroughly discussed.

Keywords: OSINT, Crowdsourcing, Digital Forensics, Online Child Sexual Abuse, Criminal Investigations.

Introduction

Since the advent of the Internet, online child sexual abuse has become a global concern, continually growing and diversifying. In addition to the dissemination of child abuse materials throughout cyberspace (Taylor & Quayle, 2003), new forms such as online grooming (Whittle, Hamilton-Giachritsis, Beech, & Collings, 2013), sexual extortion (Kopecký, 2017) and webcam child prostitution (Açar, 2017a) have come to the attention of the public as online offences against children. While some aspects of the subject matter,

[1] Unit Manager, Department of Cyber Crime, Turkish National Police, Eymir Mah., 49, Sokak, 06834 Gölbaşı/Ankara, Turkey. Email: kemalveli.acar@egm.gov.tr

such as the psychosocial and legal dimensions, have been studied extensively in the literature, by comparison, the development of practical methods to advance the global fight against online child sexual abuse seems to be neglected (Açar, 2017b).

High capacity but low-cost data storage options have resulted in the common use of terabyte-sized hard disks and cloud services. It is now unexceptional for a person to have a large amount of digital information spread throughout separate personal devices and cloud-based personal accounts. In addition to the unprecedented increase in average data per person, awareness of digital evidence has grown substantially among law enforcement agencies (LEAs). Currently, even in relation to petty crimes such as shoplifting, investigators are fully aware of the fact that the digital belongings of the suspect and victim may be highly beneficial for increasing the rates of apprehension or conviction. For these reasons, the backlog of seized devices has increased to unparalleled levels, leading to delays in the trial process of up to years, even for online child abuse investigations (Goldberg, 2015; Netclean, 2017). In order to solve this problem to some extent, a range of technological methods has been proposed, such as data mining, data reduction and triage (Quick & Choo, 2014; Lillis, Becker, O'Sullivan, & Scanlon, 2016). However, unlike in the case of less serious criminal acts, for online child abuse investigations reliance on triage and completely automated processes may have grave consequences.

In addition to child abuse materials, the digital belongings of an abuser might contain more minor yet important pieces of information such as nicknames, e-mails and names of places. A thorough digital examination and appropriate analysis of this set of information might reveal the exact locations or real identities of criminal associates or victimized children. While highly convenient for investigators, triage is essentially a quick, on-the-spot digital forensic evaluation, for time-sensitive cases in particular (Rogers, Goldman, Mislan, Wedge, & Debrota, 2006). In most of these cases, investigators do not even know whether they will be able to extract incriminating evidence from seized digital materials before the operation takes place. However, for online child abuse investigations, investigators generally have a solid idea of what and where the suspect might possess the illicit materials beforehand. Even in the absence of such firm knowledge, the possibility of failing to notice incriminating evidence of significant importance makes LEAs extra cautious. Consequently, due to these complications, triage is an extremely risky method of digital forensics examination and rarely preferred in relation to online child abuse investigations.

On the other hand, completely automated methods of extracting digital evidence are highly efficient for revealing some types of information, such as e-mail addresses and Internet browser history. However, finding such information through the examination of digital evidence is only the beginning of the judicial process. E-mail addresses, nicknames and other relevant details should be properly investigated through open sources, to identify the real identities of possible owners and discover whether their relationships with a suspect are crime-related (Gibson, 2016). For each clue, on average, an hour-long open source intelligence investigation (OSINT) should be conducted, including the time-consuming report writing stage. Taking into account the fact that thousands of such digital traces are present in every item of seized material for each separate investigation, a thorough examination of every piece of information regarding a criminal case becomes a tremendously challenging issue for LEAs.

Offenders do not generally operate on separate devices for legal and illegal activities. Therefore, concerning a connection with a specific criminal event, a huge accumulation

of small clues needs to be investigated. However, while the scale of data for examination has been growing continuously, the resources of LEAs have not kept up with this unprecedented rise. Thus, as an unfortunate yet unavoidable option, LEAs prioritize cases in which victimized children are involved. In such serious cases, they examine every single item of digital evidence to reveal all the associations of suspects, to the greatest extent possible. On the other hand, for "regular" offenders, who form the majority, a digital forensic examination essentially means a search through their digital belongings for child abuse materials. Understandably, first and foremost, the allocation of limited resources should be deployed for the identification of victims, not for low-risk offenders. Some police forces have used evidence-based assessment tools instead of relying on their own experience and judgment. For example, the Kent Internet Risk Assessment Tool (KIRAT) in the UK aims to identify likely contact offenders among online abusers (Long, Alison, Tejeiro, Hendricks, & Giles, 2016). However, even though they might seem less important at first glance, a thorough digital examination of such "ordinary" abusers' seized materials might lead to the identification of more dangerous offenders or victimized/vulnerable children. Since online child abusers consider themselves more a community of outcasts than a gang of outlaws (Durkin, Forsyth, & Quinn, 2006), and they have relatively tight online relationships with each other, the digital materials of a seemingly unimportant suspect might contain more valuable information than first evaluated by LEAs.

Due to the aforementioned reasons, in terms of conducting a more extensive inspection for each case in a timely manner, it is unreasonable and quite impractical to reinforce LEAs with huge additional human and financial resources. However, there might be feasible ways to employ more human-centric solutions to this problem without either skyrocketing costs or disregarding the advantageous opportunities presented by cutting-edge technological innovations. In both online and offline environments, ordinary people have persistently demonstrated admirable commitment to the cause of protecting the wellbeing of children and an inexhaustible willingness to participate in related activities, without demanding any pecuniary benefits, since the emergence of the Internet. However, except for highly controversial undercover operations by concerned civilians (Adler, 2011) and the individual reporting of online child abuse incidents to hotlines (International Association of Internet Hotlines, 2016), the large-scale exploitation of the potential benefits of a more controlled and systematic method seems to be ignored.

In this regard, this article proposes a new and unorthodox way of handling the ever-growing workload of online child abuse investigators effectively by tapping into the energy of a carefully selected group of volunteers. The main aim outlined by the author is essentially to influence further research into such practical measures of digital forensic examination and to have an impact on policymakers about unconventional methods of crime prevention in general, and online child abuse investigations in particular. After a brief literature review on related subjects, such as OSINT and crowdsourcing, in terms of technical, legal and organizational aspects, the proposed theoretical model will be elaborated. Then, likely concerns and potential bottlenecks on the same respective aspects regarding the successful actualization of the model will be identified and thoroughly discussed. Since it is neither reasonable nor feasible to foresee all the ramifications of the actualization of such a theoretical model in complete and precise detail, the most obvious and important aspects will be highlighted.

The author fully acknowledges that the main structure introduced in this article would yield horrendous outcomes in the form of violations against basic human rights in the wrong hands. An oppressive or anti-democratic state could deploy this idea to detect, track and prosecute the political dissenters or unfavored minorities such as homosexuals. By the help of an "online troll army" disguised as sensible citizens, bits of information provided by digital forensics tools or mass data acquisition through open sources could be used against the mentioned social groups in different ways such as preventing them exercising their freedom of speech in online environments anonymously or revealing entire nodes of a social network to hunt them in real life. Throughout the history, almost all technological innovations have been employed by malevolent authorities for nefarious goals and there is no plausible excuse to believe that the introduced model will stand as an exception. Even with the right intentions in place, the implementation of this idea might lead to unforeseen yet disastrous results, particularly in the absence of appropriate safeguards. Nevertheless, as is the case with the wiretapping, this theoretical model offers a double-edged sword. If it is wielded in a right way with the strict guidelines and unwavering commitment to the basic human rights, it might produce remarkable effects on the fight against crime. But, if it is wielded wrongly, it will definitely turn people's life into a hell in horrible ways. Here, the author elaborates a potentially unsafe solution to an ever-growing problem with a very optimistic focus but without ignoring the grave consequences of this solution in case of misuses and abuses by the state actors, which has been a default danger for almost every technological novelty. Thus, the commitment of the author to solve the issue at hand shouldn't be misinterpreted as the complete disregard for human rights or uninformed naivety of an enthusiastic practitioner.

Literature Review

Since Jeff Howe coined the term in 2006 (Howe, 2006), crowdsourcing has gained significant attention from both scholars and business people. Initially, the term simply referred to a type of outsourcing through which some business-related activities of a company are carried out by an anonymous crowd, in response to an open call in online environments (Howe, 2008). Still relevant and popular for the time being, Amazon Mechanical Turk, iStockphoto and Threadless are probably the best-known examples of commercial crowdsourcing. However, as online communities also began to solve problems regarding scientific and humanitarian issues such as classifying galaxies (Raddick et al., 2009) and gathering accurate information during a social crisis (Okolloh, 2009), the definition of the term has evolved into a wider meaning than just an alternative way of business-making. According to Brabham (2008), crowdsourcing is an online, distributed problem-solving and production model that exploits the collective contributions of online crowds to serve the organizational aims of the requester. In this respect, regardless of the presence of an open call and the prerequisite anonymity of the crowd, if an online group carries out any problem-solving activity desired by system owners, this could be considered a form of crowdsourcing, as Doan, Ramakrishnan, & Halevy (2011) have claimed. Although Estelles & González (2012) have tried to confine the term within a narrower meaning by establishing the basic characteristics of crowdsourcing, as the empirical implementations of the phenomenon diversify with the emergence of new forms, a broader definition has become more relevant and suitable in practice.

Unlike crowdsourcing, OSINT has a long history dating back to World War II (Burke, 2007). In its broadest sense, OSINT refers to the concept of the acquisition and analysis of

unclassified information from a variety of publicly available sources, such as traditional media outlets and online environments (Mercado, 2009). Owing to the unprecedented growth of user-generated content, following the rise of Web 2.0 in particular, a huge cache of open source data has been created by ordinary users. It is estimated that approximately 1.7 megabytes of new information will be created every second for every human being on the planet by the year 2020 (Keijzer & Klingebiel, 2017). While at first glance this outlook presents a considerable opportunity for intelligence agencies to perform their duties more efficiently and at a lower cost, the complete acquisition and meaningful interpretation of such enormous amounts of data has become an extremely challenging task. For that reason, governments have built huge data centres to store a constantly growing amount of open source data (Bamford, 2012) and have purchased or developed state-of-the-art software to analyse and evaluate the data appropriately (Risen & Lichtblau, 2013). However, since the exponential increase of user-generated content in particular is likely to continue in the future, this exhaustive approach to the issue requires the constant allocation of immense resources. Unfortunately, due to the clouds of secrecy surrounding the intelligence community, the accurate evaluation of the effectiveness of such comprehensive and expensive management and analysis of open source data remains unexplored.

Inevitably, the idea of employing OSINT as a complementary tool for the process of criminal investigations and measures for crime prevention has recently gained popularity among LEAs. In addition to manual forms of utilizing open source data for solving individual cases (Ramwell, Day, & Gibson, 2016), technological solutions relying on automated big data analysis have been developed, such as CAPER and ePOOLICE, both EU-funded initiatives. Among other aspects, in essence these projects aim to detect emerging trends pertaining to organized crime (Pastor & Larsen, 2017) and to demonstrate relationships between criminal entities in a visually comprehensible way (Brewster, Andrews, Polovina, Hirsch, & Akhgar, 2014) through the acquisition of different formats of open source data by web crawling. Additionally, in combating online child sexual exploitation, Charalambous et al. (2016) rehashed the popular idea of web crawling through cyberspace for a pre-determined set of keywords such as '12yob' or '9yog' in order to identify malicious web sites and their users (Westlake, Bouchard, & Frank, 2012). To sum up, even if their success has been limited due to heavy dependence on web crawling, current automated solutions fundamentally deal with the detection of unknown cases, rather than alleviating the enormous backlog of old cases.

On the other hand, crowdsourcing criminal investigations is rarely proposed as a concept (Huey, Nhan, & Broll, 2013; Brabham, 2013), let alone put into practice. Applied instances of this crossover have remained exceptional and had a negligible impact on the general trend so far. For example, Internet Eyes and FaceWatch are online platforms on which users watch CCTV footage belonging to private enterprises and identify criminal activities such as shoplifting and damage to property, in return for a small fee (Trottier, 2014). Thus, the responsibility for watching hours of footage has shifted from business owners and LEAs to crowds. As a very recent example of online child abuse investigations, Europol initiated a campaign entitled "Stop Child Abuse – Trace An Object" at the beginning of June 2017 (Muraszkiewicz, 2018). The unidentified objects in child abuse materials such as t-shirts, toys and bags were announced to an anonymous crowd in the hope that crowd members give useful information on the origin country of

such objects and thus the origin country of victims and abusers. At the moment, through the International Child Sexual Exploitation Database (ICSE DB) managed by the Interpol (Calcara, 2013), a limited numbers of law enforcement agents examine the child abuse materials to detect victims, abusers and crime scenes. While it's too early to see the outcomes of Europol's campaign, more eyes on the issue might enhance the current efforts of victim identification by LEAs.

However, using a much wider interpretation, scholars have incorrectly defined some sensational incidents as a form of crowdsourcing criminal investigations. For instance, after the Boston Marathon bombings in 2013, the FBI requested visual materials obtained by bystanders during that day with an open online call. With the indisputable assistance of an anonymous online crowd, the perpetrators were identified within three days (Tapia, LaLone, & Kim, 2014). Despite major similarities to the basic characteristics of crowdsourcing, the FBI solved the case by analysing all images and identifying the suspects themselves. The crowd was not directly involved in or even invited to the problem-solving phase of the process. In essence, instead of locating every bystander and doing large numbers of interviews with them, the FBI used an unconventional yet highly efficient method of gathering information about a particular incident from eyewitnesses. It could be said that they merely did traditional police work in an uncommon way. However, even though it was ultimately an enormous failure, Anonymous and Reddit also conducted their own parallel investigations to attempt to solve the case, and theoretically, their efforts were closer to the concept of crowdsourcing than that of the FBI. In any case, the Boston Marathon bombings revealed a highly enthusiastic and very capable online crowd that strives to deliver justice (Nhan, Huey, & Broll, 2017).

The Theoretical Model

As interviews conducted by Huey et al. (2013) with police officers about "cyber-vigilantes" have clearly demonstrated, LEAs have coherent legal and practical objections to the idea of active civilian participation in criminal investigations. Despite all the good intentions a crowd bring to the scene, they are not fully aware of some crucial concepts surrounding the judicial process, such as protecting the confidentiality of information against third parties, the admissibility of evidence to a court and other legal liabilities. On a similar note, most members of a crowd also lack the necessary educational background and sufficient experience of technical issues required to prepare a complete OSINT report. While a crowd might have a ceaseless desire and ample time to spend on delivering justice, LEAs have both the legal authority and practical training to be involved in criminal investigations. Undeniably, an open call to anonymous crowds is not a viable option for combining the energy of these two groups appropriately, since it might easily disrupt the confidentiality of a criminal investigation and cause irreversible damage. However, a carefully selected, properly trained and tightly controlled crowd of volunteers might complement the efforts of LEAs in online child abuse cases by revealing small clues from digital forensic examination without compromising the entire judicial process.

The following sections elaborate on such a theoretical model from different angles, such as the organizational, technical and legal aspects. The author entirely acknowledges that there are many potential combinations to accomplish this utopian idea within different organizational settings, dissimilar types of legal rules and various technological solutions. On the other hand, while the author completely places his faith in the ethical uses of such model in accordance with the universal standards brought by international human rights

law, the implementation of this idea could also easily turn into a tasteless dystopia quickly. Since it is neither possible nor feasible to predict all the alternative options and to elaborate necessary safeguards to prevent the violations of basic human rights by state actors for the model in just one article, the essential features and challenges will be described in the hope that further research will explore other possibilities and preventive measures in detail.

1. Organizational Aspect

According to Doan et al. (2011), there are four challenges to be solved for the smooth implementation and successful progression of any crowdsourcing system: recruitment and retention of the user base during the life cycle of the system, clarification of the responsibilities and duties of users, combining their inputs efficiently in a meaningful way, and lastly, evaluating their contributions through manual and/or automated methods. Depending on the different approaches to any of these four separate challenges, crowdsourcing systems might take diverse forms. For example, prominent crowdsourcing system Amazon Mechanical Turk solves the recruitment and retention issue by paying a token amount of money to users for every task they handle (Buhrmester, Kwang, & Gosling, 2011). A divisible problem, such as translating a book into another language, is split up into many small tasks and distributed to Turkers through the system. To summarise, the contribution of Turkers is solving these micro-tasks, which generally require little manual effort or expertise. The same micro-tasks are assigned to multiple users, to increase the accuracy of the solution required. Since each task is assigned to several people, from a technical point of view it becomes easier to identify apathetic users who do not do the task at all, or malicious users who intentionally perform the task in the wrong way.

As noted earlier, a highly enthusiastic crowd of volunteers has been impatiently waiting to contribute more to the fight against online child sexual abuse, by offering their time and expertise. Unlike the emerging commercial crowdsourcing platforms which face a situation of a 'tragedy of the commons' due to underqualified participants or inefficient contributions (Simula, 2013) from this passionate crowd, finding the required number of qualified people with the necessary diverse skill-sets would probably not be a challenge for LEAs. Albeit exceptional, when compared with the anonymous communities of well-known online environments, specially tailored and vetted crowds for particular purposes could also be assembled. For instance, the US Army invited only active soldiers in the field to contribute their ideas to a project entitled ArmyCoCreate (Kietzmann, 2017). In a similar way, with this project, the inclusion of many senior members of long-standing initiatives such as non-governmental organizations, coupled with an extensive background check on relatively newer potential contributors by LEAs would probably prevent the infiltration of the crowd by malicious and/or incapable people.

However, in terms of effective crowd management, affiliation with and withdrawal from such a community still require the close inspection and ceaseless attention of related government authorities. Primarily to enhance the operational security of the group, as a side benefit, one of the most suitable options for the administrative aspect of the model is to keep a registry of the volunteers with a government authority other than LEAs, such as police organizations or the ministry of justice. Therefore, the LEAs would know a user only by the name automatically assigned by the system, and the management of the crowd would fall under the responsibility of an entirely different command. Thus, in the case of

unauthorised disclosures from LEAs or volunteers, only the nicknames of the crowd would be exposed, as long as the registrant office kept their real identities secret.

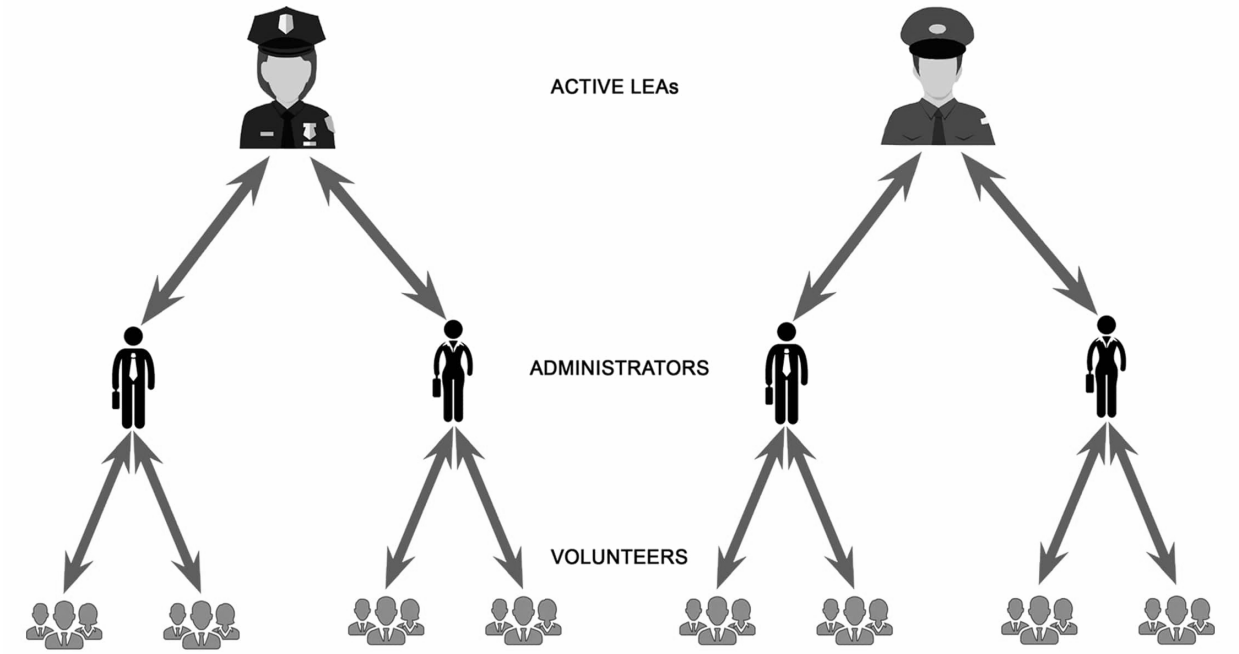
In accordance with the findings of Doan et al. (2011), the next steps for building a successful crowdsourcing system are specifying the responsibilities of the group and ensuring the optimal combination and accurate evaluation of their inputs. As mentioned before, the duty of the crowd is essentially to conduct open source intelligence about the pieces of information they receive. Since this relies heavily on technological and legal aspects such as the interface of the pertinent software and the admissibility of the OSINT reports, the details regarding the responsibilities of the crowd will be elaborated in the following sections. While combining and evaluating the inputs are also relevant for these aspects, drawing up a framework for the organizational structure in some detail will facilitate full comprehension of the proposed model.

Depending on several attributes, such as the capabilities of the crowd, the outcomes desired by the system owners, and the complexity of the assigned tasks, quality control approaches to crowdsourcing systems and task allocation among the crowd can be variable (Allahbakhsh et al., 2013). For example, Threadless assesses a contribution's quality by relying on the majority consensus about it. If a particular design receives overwhelmingly positive feedback from the crowd, it is judged a work of fine quality. On the other hand, while the crowd on Threadless deals with only one design at a time as a task, Amazon Mechanical Turk distributes the same micro-tasks to multiple users and combines their contributions to ensure the quality of each particular outcome. Then, these separately solved micro-tasks are combined to form a solution to the original problem for the requester. However, due to its exceptional nature, the proposed model differs from conventional methods of crowdsourcing systems by offering two layers of expert reviews of the contributions of bottom-level users. In essence, ordinary users report to administrators, and administrators report to active LEAs. Lastly, since it is based on a volunteer model and thus has no budget constraints (Karger & Shah, 2014), the same task can be assigned to as many people as the system owners want, to improve the quality of combined OSINT reports.

As shown in Fig. no. 1, at the bottom of the organizational structure there are ordinary users who conduct OSINT on the information they received through the interface. Subsequently, the OSINT reports go to the first layer of experts, consisting of senior members of related initiatives, retired LEAs and other more reliable volunteers. In addition to the basic evaluation of the quality of contributions, this group of administrators combines the reports pertaining to the same micro-tasks to form a complete OSINT report for a particular issue and reassigns some tasks to other users for deeper examination, to ensure the overall quality, if necessary. Likewise, in terms of functionality in particular, an evaluation of the combined reports is done by active LEAs. They can also assign specific information to another administrator with a different set of ordinary users for further consideration. On the other hand, the legal evaluation of the combined reports is done only by experts from active LEAs, since they have the legal liability, proper training and adequate work experience for ensuring the admissibility of digital evidence to a court. Since an OSINT report might contain falsified information by third parties, such as fake social media accounts or digitally manipulated images, the content of each report is not as legally reliable as physical equivalents such as DNA and fingerprints, thus not directly acceptable as evidence by itself (Sampson, 2016). Therefore, if LEAs consider all or some parts of a report useless or questionable in terms of detecting the perpetrators or securing a

conviction against an arrestee, they can easily dismiss it. Lastly, between the administrators of volunteers and the experts from active LEAs, big data analysis on all reports and crosschecking them through non-OSINT government databases can be integrated into the system as supportive instruments. Such innovative reinforcements to this human-centric model might facilitate the extraction of previously unknown links within a case or between seemingly unrelated cases. It could also speed up the evaluation process of combined reports by complementing their accuracy. These technological options will be elaborated in the following section.

Figure 1. Outline of the organizational structure of the proposed model



Lastly, for the sake of the security of the crowd, some restrictions on communications should be introduced. Bottom-level volunteers must not make contact with each other and must not know each other's real identities, since one malicious user could compromise the confidentiality of the whole community. In a similar way, communication between administrators should be banned, and contact with active LEAs must be made only by administrators. Furthermore, an administrator should always deal with the same set of users and similarly, an active LEA should communicate only with the same administrators, to prevent the compromising of the whole system in the case of a cyber-attack by malicious third parties, or the unauthorised disclosure of classified information by internal actors. These communication rules can be easily and effectively put in place and controlled, through relatively simple technological measures.

2. Technological Aspect

Brabham (2013) has examined the necessary conditions and best practices of how crowdsourcing systems should be established for government institutions in three phases, entitled planning, implementation and post-implementation. While this research places a

strong emphasis on the democratic essence of crowdsourcing as a way of more active citizen participation in political issues, the best practice number four of the implementation phase in particular is precisely applicable to the proposed model: "Invest in usable, stimulating, well-designed tools". Currently, there are many OSINT tools available for volunteers, ranging from a simple Google search to expensive sets of software. However, the appropriateness and effectiveness of any given toolkit might change over time since they depend on the current legal and technical circumstances (Revell, Smith, & Stacey, 2016). For example, a popular OSINT application can become rapidly outdated due to new legislation that restricts the acquisition of online personal data or the emergence of advanced techniques for capturing open sources. Therefore, in addition to the possibility of legal complications, leaving the decision about compiling the best OSINT toolkit to the users would definitely hinder the overall quality of their contributions. While some volunteers have the prerequisite knowledge of OSINT, the financial resources and previous technical experience for assignments, the majority might not even possess more than good intentions and spare time. This broad gap between members should be closed as much as possible by the government authorities. For that reason, LEAs must develop a specifically tailored OSINT toolkit for volunteers and also update it to new versions as finer tools emerge, or related legal amendments occur. Owing to such a usable, well-designed toolkit, differences in performance within the community would be marginally low, but in addition, users could draft more OSINT reports at any point without wasting time for the selection of the necessary tools.

Some essential features of the crowdsourcing system, such as the submission of OSINT reports, distribution of task assignments and training modules for members could be integrated into this user-friendly application. Additionally, as a convenient method of communication, volunteers could make contact with administrators, representatives of the registry office and legal advisors through the same interface, which could also serve as an entry point for the OSINT toolkit. Similarly, administrators could also keep in contact with active LEAs and legal advisors. Thus, in addition to the expeditious and reliable dissemination of information within the organization, secure communication channels between the layers of the system would be enabled and smoothly controlled by government authorities. However, as with every work environment, there would be some vulnerabilities to be taken into consideration before the development and implementation phases. For example, albeit unlikely, there is a genuine risk that online child abusers might infiltrate this community with the aim of irreversibly damaging operational security and the positive public image of the system. Therefore, to minimize the occurrence of such a threat, the interface might regularly check each user's computer for child abuse materials or any similar illicit activity. Having already been deployed by prominent figures within LEAs and the private sector for a long time (Stock, 2014; Carr, 2017), fuzzy hash matching technologies such as visual fingerprints from Videntifier Technologies (Lejsek et al., 2009) and PhotoDNA from Microsoft Inc. (Ith, 2015) might be incorporated into the interface. However, extra care should be taken when such measures are introduced. Since it is common to be exposed to child abuse materials while conducting OSINT on possible abusers, specially appointed inspectors should thoroughly examine the content of users` daily activities as to whether the presence of such materials is work-related or not. Furthermore, another probable risk for operational security is that some volunteers might encounter information relating to their acquaintances, such as work associates or family members and inform/extort such people. To prevent similar misuses of classified

information from a technical viewpoint, the interface could periodically scan each user's computer before assigning a particular task. Thus, if an e-mail or a name is linked to a user, the system would automatically assign that particular task to another user, where no visible connection to the related information has been identified as a result of the scan. Undoubtedly, the scope of scanning should be predefined by the system owners and should be done only after the written consent of the volunteers is properly obtained. If negative, the result of every session of scanning might be sent to the users in the name of establishing a more trustful and healthier relationship with the crowd. Such transparency on this issue also would definitely facilitate the creation of a feeling of security among users.

Shown in Figure 2, as a fine example of the fusion of OSINT and non-OSINT data, compiled OSINT reports can be crosschecked through classified government databases in order to complement, confirm or negate the contributions of the crowd (Day, Gibson, & Ramwell, 2016). In an automated fashion, this fusion not only strengthens the admissibility of OSINT reports but also saves time for LEAs, by shortening the duration of the total inquiry period needed for all reports. Furthermore, as noted earlier, after the administrators have compiled the OSINT reports, by combining the outputs of volunteers, big data analysis might also take place in the workflow.

Figure 2. Big data analysis and crosschecking through databases between administrators and LEAs

There are two major advantages of big data analysis for the active LEAs: fast and efficient analysis of related information within a case and exploration of previously unknown connections to seemingly unrelated cases. Firstly, since thousands, at least, of e-mails and nicknames can be associated with a particular case, manual examination of OSINT reports by active LEAs would be a time-consuming task, even within the proposed model. Additionally, due to the overabundance of digital information, a crucial link for the conviction of a suspect within a particular case might be overlooked easily during a manual examination. Therefore, such an outcome contradicts the main objectives of building this theoretical model, such as saving time on criminal investigations and

ensuring a thorough inspection of digital evidence. Among other things, big data analysis of OSINT reports of a particular case visualises the connections between pieces of related information, in order to accelerate the evaluation process carried out by active LEAs. Thus, such an analysis can effortlessly give a complete account of the communications of a suspect both in chronological and geographical order by just a click. Last but not least, a whitelist and blacklist of e-mails and nicknames such as the consumer support e-mails of prominent online environments and known aliases of wanted criminals can be defined beforehand to increase the speed and maximize the benefits of big data analysis. By doing so, while a blacklist could be highlighted by the system, a whitelist could be ignored and not sent to the volunteers for subsequent assignments.

Furthermore, big data analysis might reveal previously unknown connections between different cases by tying seemingly unrelated incidents together. As aforementioned, since abusers relatively have tighter relationships with each other, one nickname or e-mail can be extracted from multiple digital belongings. For example, combining unconnected chat records with the same person from three different seized computers might provide a better account of the real identity of an unknown serious offender, and where he/she might be located. Thus, by gathering and analysing these separate clues, LEAs can deepen investigations at the right points to reach the more serious or previously uncovered abusers. Moreover, serving as an intelligence gathering purpose for entire child abuse investigations, big data analysis might show trends in illicit activity in various ways, such as the websites an abuser has preferred, the extent of open source data they have exposed and the average number of criminal associates they have. In terms of developing more effective measures of crime prevention, such information can help LEAs to make both strategic and tactical decisions in a swifter and more efficient manner, in the short and the long term. Transforming the findings of digital forensics examination into an intelligence asset has been proposed before by a few scholars (Ribaux, Walsh, & Margot, 2006; Quick & Choo, 2014; 2017). However, due to major technological and legal obstacles surrounding the issue, these ideas have not yet been sufficiently actualized. This theoretical model can offer an indirect way to transform the information of digital evidence into an intelligence asset, and could be a good starting point for more direct ways in the future.

3. Legal Aspect

Mostly from a business-making perspective, (Wolfson & Lease, 2011) examined the legal pitfalls of crowdsourcing in terms of employment law, patent law, data security, copyright and crowd-funding. In this study, the first two recommendations to system owners are more relevant for the proposed model than the other legal areas mentioned: "Be Mindful of the Law" and "Use Contracts to Clearly Define Your Relationships". These two separate pieces of advice overlap for the theoretical model, since the related legal background intrinsically forms an indispensable part of the contracts regarding relationships between authorities and volunteers. For example, after a law provides the foundation for the proposed model by authorization, largely based on this essential legal source, a secondary set of legal rules should be produced by legislators to define the details of other important aspects such as the daily workflow of the system, the selection and withdrawal of volunteers, and legal exemptions.

As is the case with any government-led activity, appropriate legal authorisation must be given to both volunteers and active LEAs so that they can conduct their duties safely, according to a pre-determined set of rules. If the current legal background of a country

does not grant the prerequisite authorisation, the drafting of new laws from scratch, or the making of amendments to existing ones must be carried out. In addition to the essential legal background, many other issues should be clarified in a detailed manner, as far as possible, to minimize legal hesitations to an acceptable level. Nonetheless, covering all these legal aspects of the theoretical model sufficiently not only exceeds the size and scope of this article but also seems impractical because the materialization of such an unorthodox measure might cause previously unforeseen legal ramifications. However, some undeniably visible matters such as limited exemption from criminal charges, misuse of authorisation and classified information, and general management of the crowd can be discussed to a certain extent.

In the UK, child abuse investigators and other professionals dealing with the issue are exempt from charges under the Sexual Offences Act 2003 if the exposure to child abuse materials is work-related (Lyle, 2016). Similar legal safeguards should be introduced for the volunteers of the proposed model, but with relatively limited protection. These legal exemptions should be meticulously formulated, in a nuanced yet balanced way, in order both to deter malicious actions by members and to provide a legally stable work environment for the crowd. If the safeguards were too lax, it would not be possible to punish users who disguised themselves as vigilantes to view child abuse materials. If they are too strict, then the unavoidable exposure to such materials during OSINT research might become a criminal act in itself. These two extreme situations are equally dangerous for the sustainability of the crowdsourcing system in the long term. Government authorities neither want their platform transformed into a safe haven for online abusers nor want it to become crippled due to the legal hesitations of the users or unnecessary persecution of them.

Furthermore, exploiting the OSINT interface or the crowdsourcing system by the volunteers for other purposes than assisting LEAs should be an important concern of legislators during the implementation phase of the model. First and foremost, in some situations, users can obstruct justice by misusing the information they have obtained through the system. For example, users could protect people they know from criminal investigation and make them wipe any incriminating digital evidence of a malicious deed. As mentioned before, technical solutions can minimize such incidents to a certain extent. In a similar way, a malicious volunteer might try to extort money from a possible abuser by intimidation. Since the same task is assigned to multiple users simultaneously and the relationship of particular information to a suspect is not known by volunteers beforehand, such extortion attempts would probably have a slight effect on operational security and are likely to be revealed by LEAs in a short time. Still, in both cases, the perpetrators should receive a heavy punishment due to the substantial damage they created to the trustworthiness of the whole system. Secondly, albeit not as crucial as criminal misuses, another legal issue that needs to be clarified in detail is whether the volunteers could take advantage of the OSINT interface for personal purposes. If personal use is allowed, the conditions and limitations of such permission should be described clearly. Possibly, unless the interface is employed in the commission of a crime such as cyber-stalking or sexual extortion, volunteers should be able to conduct OSINT freely for private reasons.

Moreover, probably in the secondary set of legislations, administrative issues such as the submission to and withdrawal from the crowd, standards of the award system if applicable and other codes of conduct should be elucidated in a detailed way as much as possible.

Lastly, and maybe most importantly, since both administrative rules and criminal laws are subject to many changes throughout the life cycle of the proposed model, the authorities should keep volunteers continually informed about the legal aspects, through remote trainings over the interface. Similarly, periodic exams might be conducted to measure the level of understanding of the crucial legal concepts surrounding the model. Then, according to the results of such exams, additional special trainings on complex matters might be prepared to keep volunteers more aware of the legal aspects. Thus, both unintentional and intentional violations can be minimized to the extent that this unusual system sustains its continuation and eventually becomes a widely accepted instrument within the judicial process.

Discussion

Understandably, it is almost impossible to envision all likely problems during the implementation of the proposed model, let alone the negative outcomes of materialization. However, even at first glance, some possible issues surrounding the theoretical model are more visible and obvious than others. In accordance with the previous elaboration of the structure, the most likely objection points from various pressure groups and probable bottlenecks in the system will be identified and examined in three stages, as organizational, technological and legal aspects.

a. Organizational Aspect

From the outset, while the recruitment of volunteers for the proposed crowdsourcing system seems less problematic than the other organizational complications, retention of them within the system over many years or keeping their performance above a desired level after the overall enthusiasm of the first months begins to fade might pose a significant challenge, leading to an unmanageable lack of efficiency in the short term. In that case, the slower turnover rate of daily access and/or insufficient contributions from the community might hinder the sustainability of this unorthodox measure in the longer term. In the eyes of policymakers, particularly in the short term, such a solution should produce unprecedented benefits steadily, to appease the anticipated worries of some pressure groups, such as taxpayers, Internet privacy advocates and members of the opposition party. Undoubtedly, the crowd has the prerequisite motivation, spare hours and other resources from the beginning. However, due to the fragile nature of the system at its initial stages, volunteers should be kept sufficiently motivated by appropriate measures taken by the authorities to ensure survival in the short term and support the firm continuation of the proposed model in the long term.

What might motivate participants in a crowdsourcing initiative to contribute to a particular project, and general methods of maintaining or increasing their efforts have been an important area of concern for scholars. Besides monetary benefits, Braham (2010) has stated that the love of community at Threadless and the addictive nature of the activity on the site have an impact on the motivation of users. Similarly, according to Chandler and Kapelner (2013), the amount of the perceived meaning of a task positively affects the participation of users and both the quality and quantity of their outputs. In accordance with these findings, Rogstadius et al. (2011) have also confirmed that in a non-profit setting, people produce outputs with more accuracy. Lastly, among the aforementioned best practices of successful government-led crowdsourcing activity, Brabham (2013) advises the acknowledgement of the wishes of users to follow through with obligations in

the post-implementation phase. Thus, the needs of participants are met and the same crowd is then encouraged to contribute to similar government-led activities in the future. Even simple forms of acknowledgement such as a virtual rank among members or a basic certificate of achievement might serve as a badge of honour and encourage members to continue.

As noted earlier, some restrictions on interactions between members of a crowd should be introduced, to ensure the safety of users and the security of the crowdsourcing system. Even casual online chats between parties might lead to unwanted consequences, let alone offline encounters. While these limitations minimize the unwanted effects of possible external or internal compromises, they also weaken the bonding of each user to the crowd. In a normal setting, members should frequently communicate with each other, both through online and offline channels, preferably revealing their real-life identities. Compared with anonymous workers, Huang and Fu (2013) found that increasing social transparency between workers improved the quality of their contributions. On a similar note, the most successful users should be rewarded, both to maintain the motivation level of the most enthusiastic participants and to encourage the less motivated to contribute more to the system. The acknowledgement of such members' efforts should be visible to all users and sometimes even to the public, via social media or traditional media outlets. When gaining monetary benefits is not a particularly prime objective for a crowd, sources of high intrinsic motivation such as reputation and social relations come under the spotlight (Borst, 2010). However, due to aforementioned security concerns, it could become extremely challenging for official authorities to provide appropriate methods for maintaining the motivation level of a crowd. Still, some solutions such as creating virtual ranks among members depending on their contributions and privately rewarding the best users on a weekly or monthly basis could be trialled, in the absence of widely accepted motivational measures.

2. Technological Aspect

Regardless of the specific features of a selected or developed software set for the user interface and system structure, all probable technological combinations must overcome a core challenge: providing a user-friendly yet highly efficient platform without compromising the overall security of the model. As long as the system owners conform to this crucial rule in the process of building the base of the model, deciding other technical details becomes a secondary concern of less importance. In that regard, to find a balance between security and simplicity has been a long-standing objective and challenge for software developers since the dawn of the digital age. However, for the proposed model, security is not an optional feature to be partially ignored in the name of marketing it to more consumers at the start and patched afterwards, in the event that any vulnerability emerges. For that reason, unlike the commercial equivalents, this universally accepted approach to design should be complied with more strictly, since monetary compensation would not be enough to cover the potential damage following any exploitation of vulnerability. Nonetheless, the triumph of the theoretical model also relies heavily on sustained active user participation, so they should be able to contribute effortlessly and constantly through an easy-to-use interface.

As mentioned earlier, the compartmentalization of volunteers into separate and unconnected divisions might minimize the outcomes of both internal and external cyber-

attacks to some extent and serve as a structural defence line for the model. In addition, periodic scanning of volunteers' computers to detect child abuse materials or other related information such as the e-mails and nicknames of suspects could be an appropriate way to establish control over users. However, the frequency, scope and duration of this scanning should be set so that both members' daily and crowdsourcing activities are not affected negatively. For example, as a general convenience for users, automated scans might only start or continue when the computer screen is off. However, since the storage capacity, software configuration and other technical capabilities of each user's computer are wildly different, it is practically impossible to guarantee a smooth scanning process and personal cyber security for each volunteer every time. To solve this problem, the government could assign a pre-configured high capacity workstation to every member of the crowdsourcing system, pre-installed with antivirus software in particular. Thus, any unwanted possible results of performance and security issues regarding the differences between the computer systems of the user base might be prevented, or at least minimized, beforehand. Undoubtedly, volunteers should also use these workstations for their daily activities. A personal computer would probably make them contribute more, compared with a stand-alone computer, since even a five-minute coffee break could turn into a session of OSINT gathering. Lastly, since the phase in which big data analysis and the fusion of OSINT and non-OSINT data has already occurred securely within the cyber-domain of government facilities, this article will not deliberate on those parts of the model.

Despite all the security measures mentioned, the technological aspect remains the Achilles' heel for the model. Particularly at the initial stages, when doubts about the system's usefulness are significantly high, even a minor exposure through cyber-attacks can irreversibly damage the reputation of the system to the point of complete eradication. To minimize the possibility of such an event, the materialization of a crowdsourcing system might start with a limited number of volunteers. Developers and administrators could assess potential organizational bottlenecks and technical vulnerabilities more accurately and more securely through such a test sample. As the system evolves into a more efficient and safe environment at every step, batches of newcomers might join the crowd. Therefore, both public perception of the system and the operational security of the model would remain undisturbed during the most critical initial phase. All being well, this strengthened state at the beginning stages would provide the requisite stamina for the system to survive any external attacks and unintentional disclosures in the long term.

3. Legal Aspect

At first sight, overcoming legal challenges seems to be the easiest obstacle to overcome on the path to the sound implementation of such an unorthodox idea. If the current legal background of a country does not grant the prerequisite authorisation, relevant bodies such as legislators and policymakers could swiftly draft laws or legal documents at a similar level, to allow the materialization of the model. If the legal background already exists in any form, secondary legal regulations such as directives and written briefings could easily be prepared or amended, to ensure the operational security and overall functionality of the proposed model. However, creating new legislation or adapting older laws to new phenomena is not as undemanding as it might appear on paper.

Firstly, public sentiment about such an idea should be extremely visible and undeniably strong to make it a high priority for the representatives of voters. However, apart from in a few countries, deploying innovative methods for the fight against online child sexual abuse is not a pressing concern or a popular issue at present (Açar, 2017b). It would be a huge challenge to persuade the policymakers of a specific country that they need such a drastic measure, unless a global or local wave of awareness raising about the issue erupts, moving society's opinion in a different direction. Methods once considered controversial might then begin to seem more reasonable and viable for today, in the eyes of the public.

Secondly, even if the legal process of implementation begins promisingly, an ensuing heated public debate about such an unusual model might easily arouse the sentiment of the people to the extent that policymakers avoid or delay any meaningful legal changes indefinitely. Without a doubt, the privacy rights of suspects will be the foremost point of opposition. Even though the system divides the digital evidence into millions of pieces, from a legal viewpoint, the human rights attached to a single piece are no less important than the whole of the information. Furthermore, some pressure groups such as privacy advocates might interpret the limited exemption of volunteers from the crime of accessing child abuse materials differently. By exaggerating the low possibility of a malicious infiltrator among the crowd, they might represent the exemption issue to society as a definitive way of becoming viewed as a government-certified paedophile. Moreover, from a different perspective, in relation to conspiracy theories some dissidents might claim that this idea is just a starting point for establishing a cyber-army of volunteers in the future, for other types of cyber crime, or even for intelligence-gathering purposes (Aschmann, Van Vuuren, & Leenen, 2015). While such expansion of operational aims is not completely impossible, assuming that society would readily accept wildly different forms of OSINT by crowdsourcing, just because the first one was successful, is an oversimplification.

Despite all the extensive remote trainings on legal aspects and the presence of a user interface that minimizes legal complications by technical design, some volunteers would still delve into the "grey zones" of OSINT (Hribar, Podbregar, & Ivanuša, 2014) or would execute self-assigned undercover operations in the name of protecting children more actively. When the irresistible urge to create a meaningful contribution combines with impatience about delivering justice, a common trait among volunteers involved with the global fight against online child sexual abuse (Açar, 2017b), questionable actions or even criminal deeds might easily occur along the way. Firstly, the grey areas of OSINT include a variety of non-public but open source data accessible only through special channels, such as pre-prints, dissertations and registries. While this is legal by definition and consists of only nine percent of entire online open source data, the inclusion of such data in OSINT reports might give the wrong signal to a crowd that they can freely seek other ways to gather information in addition to the special interface. Secondly, some users might create fake social media profiles and befriend possible users, with the motive of obtaining more accurate information about them. In terms of intelligence gathering, such actions create incredibly valuable information. However, due to the entrapment defence (Roiphe, 2003), it also can render the information useless in the end, by harming admissibility. As public support for a crowdsourcing system grows and volunteers become experienced to the extent of professionalism in the long term, grey zones and undercover operations might be cautiously included into the model. However, in the short term, such actions should be punished appropriately, to protect the whole system. While entering into the

grey zones of OSINT might merely result in an unofficial warning by administrators or temporary expulsion from the system, unquestionable activities such as undercover operations should lead to permanent expulsion. Except in serious cases, criminal charges should not be filed, because this type of incident occurs as a consequence of the good intentions of volunteers, not through malicious intent.

Last but not least, the author has put so much confidence in the governmental bodies so far because it is implicitly assumed that highly ethical authorities and utterly competent individuals would execute this idea with no misuse or systemic abuse. However, the reality couldn`t be more different than the immensely optimistic picture which the article has drawn. Somewhere in the world, there will always be rogue ruling elites, unethical government institutions and dishonorable individuals who abuse well-intentioned innovative designs for putting their twisted vision into effect. Averting widespread misuses of state-backed organizations heavily relies on a concerted global action mainly comprises restrictive regulations such as international sanctions. On the other hand, for the countries which have shown firm commitment to the defence of human rights, possible systemic abuse and individual misuses could be minimized with some legal and technical measures. For example, frequent inspections by an independent auditing authority, which preferably includes civilian experts from well-established privacy advocates, might ensure the overall quality of operations. As long as the benevolent motives stay intact, many similar measures could be chosen among a plethora of options for exercising strict control over the system.

Other Implicatıons for Practice

In addition to digital forensic examination, the scope of the model might be expanded to solve other bottlenecks in information-gathering regarding online child sexual abuse investigations. For example, in 2016 alone the National Center for Missing and Exploited Children (NCMEC) received 8.2 million reports through CyberTipline (National Center for Missing and Exploited Children, undated). When a user uploads child abuse materials to the online platforms of US-based prominent Internet Service Providers (ISPs) such as Facebook, Twitter and Google, the subscriber information for that particular user is sent to NCMEC within a short time. While individual reporting about certain incidents such as sexual extortion and online grooming by anonymous users is possible, the majority of the current workload of NCMEC comprises reports automatically detected by ISPs through previously mentioned hash matching technologies. As an unfortunate yet inevitable option, NCMEC and DHS pick and thoroughly investigate the serious cases in which new child abuse materials are uploaded or real child victims are involved. Then NCMEC send the related information either directly to the relevant country or via the Department of Homeland Security (DHS) (Açar, 2017b).

Since the same restrictions on resources also apply at the receiving end of NCMEC reports, there is little left to be done by other countries. Through national databases, a quick inquiry about subscriber information such as IP addresses, e-mail addresses and phone numbers is generally the first and only thing that foreign LEAs do. If an abuser is smart enough to hide himself/herself behind de-anonymizing technology such as The Onion Router (TOR) network or any Virtual Private Network (VPN), he/she can avoid detection by LEAs for a long time. Even if a particular social media account belonging to an abuser is shut down due to the uploading of known child abuse materials, he/she can effortlessly proceed to spread malicious content by setting up new fake accounts with new e-mail addresses each time. Owing to the de-anonymizing technologies, the subscriber

information of every new account will likely end up on the desk of a different country's LEAs, who do not know anything about the hidden history of the abuser. Therefore, unfortunately, in a case where an abuser mistakenly leaves an online trace of crucial information about his/her real identity or exact location connected with one of his/her initial accounts, it might easily be overlooked by LEAs since they cannot conduct OSINT for every report.

Unfortunately, "ordinary" abusers can avoid punishment until they evolve into "serious" abusers by sharing new child abuse materials or victimizing real children. In addition to changing this regrettable progression, CyberTipline and similar kinds of reports might be included in the workflow of the theoretical model, because it would not be surprising to see user names or e-mail addresses of such abusers in the digital belongings of a suspect as well. Besides the exploration of previously unknown ties between seemingly unrelated CyberTipline reports, as is the case with digital forensic examination, valuable intelligence on the behaviours of "ordinary" abusers can be collected in this way. Lastly, since it literally belongs to the same domain, persuading policymakers and obtaining public support would not bring any additional challenge to the implementation phase.

As mentioned earlier, it is possible to launch similar crowdsourcing systems for assisting criminal investigations after the proposed model has gained wide acceptance and strong trust from both the public and policymakers. While the technical and administrative challenges and the possible solutions likely remain the same, depending on the particular type of crime, essential changes must still be made. For example, since anti-terrorism and property crimes are fundamentally different from each other, their crowdsourcing systems would probably also be different in many crucial aspects. Furthermore, such theoretical solutions could even be introduced for the purposes of intelligence gathering in the future. Certainly, the possibility of such immense expansion would face insurmountable opposition from many different pressure groups from various angles.

Conclusion

Following the public outcry after the infamous Marc Dutroux case (Esposito, 1998), Belgium and The Netherlands drafted new legislations to remove the inefficiencies of their judicial process. To this day, these countries have remained at the front line of the global fight against online child sexual abuse by employing unorthodox methods of crime fighting (Vendius, 2015). In a similar way, the abduction and murder of Amber Hagerman and Megan Kanka triggered many legal changes in the US (Zgoba, 2004). Unfortunately, the pattern of creating new laws and methods for ensuring the wellbeing of children both in online and offline environments is heavily dependent on the occurrence of some sensational incidents. Policymakers should not wait for the horrific victimization of some children before initiating unusual policy changes to tackle online child sexual abuse.

This article has proposed a theoretical model to solve one of the current "time bombs" of digital forensics: the enormous backlog of online child sexual abuse cases. While the resources of LEAs are insufficient, automated methods alone are not adequate to respond fully to needs. Instead of waiting for this issue to be revealed to the public in the form of a series of tragic events, as a bare minimum policymakers should discuss innovative methods that combine the irreplaceable efforts of the human mind with the computational energy of automation. In this respect, this article has elaborated on a crowdsourcing system composed of volunteers and active LEAs. Through a user interface, volunteers would

conduct OSINT on pieces of information such as e-mails and nicknames extracted through a digital forensic examination process. To increase the speed and efficiency of the model, big data analysis and the fusion of OSINT and non-OSINT data would also be integrated within the workflow.

In the age of mass surveillance by the private sector and intelligence agencies, the author is fully aware of the possible misuses and systemic abuse of such system by state actors and ill-intentioned individuals. As is the case with any unorthodox measure, particularly in relation to the legal aspects, this model is bound to create heated debate. In fact, many reasonable arguments could be advanced concerning the potential dangers of such a model. Still, no matter how low the odds are, there might be a chance to apply such unorthodox ideas into real life without severely compromising the basic human rights. Otherwise, if this "time bomb" explodes without any solution in place beforehand, the possible dangers might pale into insignificance when compared with the resultant irreversible damage to children and to society. For this reason, instead of merely responding to the current problems with insubstantial answers, policymakers and scholars should be actively discussing theoretical solutions such as these, regardless of their feasibility and simplicity.

References

AÇar, K. V. (2017a). Organizational Aspect of the Global Fight against Online Child Sexual Abuse. *Global Policy*, *8*(2), 259-262.

AÇar, K. V. (2017b). Webcam Child Prostitution: An Exploration of Current and Futuristic Methods of Detection. *International Journal of Cyber Criminology*, *11*(1), 98-109.

Adler, A. (2011). To Catch a Predator. *Columbia Journal of Gender and Law*, *21*(2), 130-151.

Allahbakhsh, M., Benatallah, B., Ignjatovic, A., Motahari-Nezhad, H. R., Bertino, E., & Dustdar, S. (2013). Quality control in crowdsourcing systems: Issues and directions. *IEEE Internet Computing*, *17*(2), 76-81.

Aschmann, M., Van Vuuren, J. J., & Leenen, L. (2015). Cyber armies: the unseen military in the grid. In J. Zaaiman & L. Leenan (Eds.). *ICCWS 2015 - The Proceedings of the 10th International Conference on Cyber Warfare and Security* (pp. 20-29). UK: Academic Conferences Limited.

Bamford, J. (2012, March 15). The NSA is Building the Country's Biggest Spy Center (Watch What You Say). *Wired*. Retrieved from http://www.wired.com/threatlevel/2012/03/ff_nsadatacenter.

Borst, I. (2010). Understanding Crowdsourcing: Effects of motivation and rewards on participation and performance in voluntary online activities, ERIM PhD Series in Research in Management, 221, Rotterdam.

Brabham, D.C. (2008). Crowdsourcing as a Model for Problem Solving: An Introduction and Cases. *Convergence: The International Journal of Research into New Media Technologies* *14*(1), 75–90.

Brabham, D. C. (2010). Moving the crowd at Threadless: Motivations for participation in a crowdsourcing application. *Information, Communication & Society*, *13*(8), 1122-1145.

Brabham, D. C. (2013). *Using Crowdsourcing In Government*. IBM Center for the Business of Government. Retrieved from

http://www.businessofgovernment.org/sites/default/files/Using%20Crowdsourcing%20In%20Government.pdf.

Brewster, B, Andrews, S., Polovina, S., Hirsch, L., & Akhgar, B. (2014). Environmental scanning and knowledge representation for the detection of organised crime threats. In N. Hernandez, R. Jaschke & M. Croitoru (Eds). *21 st International Conference on Conceptual Structures, ICCS 2014: Graph-Based Representation and Reasoning* (pp. 275-280). New York, NY: Springer.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data?. *Perspectives on psychological science*, *6*(1), 3-5.

Burke, C. (2007). *Freeing knowledge, telling secrets: Open source intelligence and development*. CEWCES Research Papers. Paper no. 11. Retrieved from: http://epublications.bond.edu.au/cewces_papers/112007.

Calcara, G. (2013). Role of Interol and Europol in the Fight against Cybercrime, with Particular Reference to the Sexual Exploitation of Children Online and Child Pornography. *Masaryk University Journal of Law and Technology*. 7(1), 19–33.

Carr, J. (2017). A Brief History of Child Safety Online: Child Abuse Images on the Internet. In J. Brown (Ed.). *Online Risk to Children: Impact, Protection and Prevention*. (pp. 5-21). Chichester: John Wiley & Sons.

Chandler, D., & Kapelner, A. (2013). Breaking monotony with meaning: Motivation in crowdsourcing markets. *Journal of Economic Behavior & Organization*, *90*, 123-133.

Charalambous, E., Kavallieros, D., Brewster, B., Leventakis, G., Koutras, N., & Papalexandratos, G. (2016). Combatting cybercrime and sexual exploitation of children: an open source toolkit. In B. Akhgar, P. S. Bayerl & F. Sampson (Eds). *Open Source Intelligence Investigation* (pp. 233-249). New York: Springer.

Day, T., Gibson, H., & Ramwell, S. (2016). Fusion of OSINT and non-OSINT data. In B. Akhgar, P.S. Bayerl & F. Sampson (Eds). *Open Source Intelligence Investigation* (pp. 133-152). New York: Springer.

Doan, A., Ramakrishnan, R., & Halevy, A. Y. (2011). Crowdsourcing systems on the world-wide web. *Communications of the ACM*, *54*(4), 86-96.

Durkin, K., Forsyth, C. J., & Quinn, J. F. (2006). Pathological internet communities: A new direction for sexual deviance research in a post modern era. *Sociological Spectrum*, *26*(6), 595-606.

Esposito, L. C. (1998). Regulating the Internet: The new battle against child pornography. *Case Western Reserve Journal of International Law*, *30*(2), 541-565.

Estellés-Arolas, E., & González-Ladrón-de-Guevara, F. (2012). Towards an integrated crowdsourcing definition. *Journal of Information science*, *38*(2), 189-200.

Gibson, H. (2016). Acquisition and preparation of data for OSINT investigations. In B. Akhgar, P. S. Bayerl & F. Sampson (Eds). *Open Source Intelligence Investigation* (pp. 69-93). New York: Springer.

Goldberg, A. (2015, November 8). Child abuse cases delayed by police backlog. *BBC*. Retrieved from: http://www.bbc.com/news/uk-34713745.

Howe, J. (2006, January 6). The rise of crowdsourcing. *Wired*. Retrieved from: https://www.wired.com/2006/06/crowds.

Howe, J. (2008). *Crowdsourcing: How the power of the crowd is driving the future of business*. New York: Crown Publishing Group.

Hribar, G., Podbregar, I., & Ivanuša, T. (2014). OSINT: a "grey zone"?. *International Journal of Intelligence and CounterIntelligence*, *27*(3), 529-549.

Huang, S. W., & Fu, W. T. (2013). Don't hide in the crowd!: increasing social transparency between peer workers improves crowdsourcing outcomes. In: W.E. Mackay, S. Brewster & S. Bødker (Eds.). *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 621-630). New York: ACM.

Huey, L., Nhan, J., & Broll, R. (2013). 'Uppity civilians' and 'cyber-vigilantes': The role of the general public in policing cyber-crime. *Criminology & Criminal Justice*, *13*(1), 81-97.

International Association of Internet Hotlines. (2016). *Annual Report 2015*. Retrieved from: http://www.safeline.gr/sites/default/files/INHOPEAnnualReport2015.pdf.

Ith, T. (2015, July 15). Microsoft's PhotoDNA: Protecting children and businesses in the cloud. *Microsoft*. Retrieved from: https://news.microsoft.com/features/microsofts-photodna-protecting-children-and-businesses-in-the-cloud/#b9EvHwoGUbEgSvr8.97.

Karger, D. R., Oh, S., & Shah, D. (2014). Budget-optimal task allocation for reliable crowdsourcing systems. *Operations Research*, *62*(1), 1-24.

Keijzer, N., & Klingebiel, S. (2017, March 19). *Realising the Data Revolution for Sustainable Development: Towards Capacity Development 4.0*. Partnership in Statistics for Development in the 21st Century Discussion Paper No. 9.

Kietzmann, J. H. (2017). Crowdsourcing: A revised definition and an introduction to new research. *Business Horizons*, *60*(2), 151-153.

Kopecký, K. (2017). Online blackmail of Czech children focused on so-called "sextortion"(analysis of culprit and victim behaviors). *Telematics and Informatics*, *34*(1), 11-19.

Lejsek, H., Jóhannsson, Á. Þ., Ásmundsson, F. H., Jónsson, B. Þ., Daðason, K., & Amsaleg, L. (2009, October). Videntifier™ forensic: a new law enforcement service for automatic identification of illegal video material. In *Proceedings of the First ACM workshop on Multimedia in forensics* (pp. 19-24). New York: ACM.

Lillis, D., Becker, B., O'Sullivan, T., & Scanlon, M. (2016). Current Challenges and Future Research Areas for Digital Forensic Investigation. *arXiv preprint arXiv:1604.03850*.

Long, M., Alison, L., Tejeiro, R., Hendricks, E., & Giles, S. (2016). KIRAT: Law enforcement's prioritization tool for investigating indecent image offenders. *Psychology, Public Policy, and Law*, *22*(1), 12-21.

Lyle, A. (2016). Legal considerations for using open source intelligence in the context of cybercrime and cyberterrorism. In B. Akhgar, P.S. Bayerl & F. Sampson (Eds). *Open Source Intelligence Investigation* (pp. 277-294). New York: Springer.

Mercado, S. (2009). A Venerable Source in a New Era: Sailing the sea of OSINT in the Information Age. In C. Andrew, R. J. Aldrich & W. K. Wark (Eds.). *Secret Intelligence: A Reader* (pp. 78-89). London: Routledge.

Muraszkiewicz, J. (2018). Crowd Knowledge Sourcing–A Potential Methodology to Uncover Victims of Human Trafficking. In G. Leventakis & M. R. Haberfeld (Eds.). *Societal Implications of Community-Oriented Policing and Technology* (pp. 23-30). New York: Springer.

National Center for Missing and Exploited Children. (n.d.) Key Facts. *NCMEC*. Retrieved from: http://www.missingkids.com/KeyFacts.

Netclean. (2017). The Netclean Report 2016. *Netclean*. Retrieved from: https://www.netclean.com/wp-content/uploads/2016/12/NetClean_Report_2016_English_print.pdf.

Nhan, J., Huey, L., & Broll, R. (2017). Digilantism: An analysis of crowdsourcing and the Boston marathon bombings. *The British Journal of Criminology*, *57*(2), 341-361.

Okolloh, O. (2009). Ushahidi, or 'testimony': Web 2.0 tools for crowdsourcing crisis information. *Participatory learning and action*, *59*(1), 65-70.

Pastor, R. P., & Larsen, H. L. (2017). Scanning of Open Data for Detection of Emerging Organized Crime Threats - The ePOOLICE Project. In H. Larsen, J. Blanco, P. R. Pastor & R. R. Yager, (Eds.). Using Open Data to Detect Organized Crime Threats: Factors Driving Future Crime (pp. 47-71). Springer International Publishing.

Yager (Eds.). *Using Open Data to Detect Organized Crime Threats* (pp. 47-71). New York: Springer.

Quick, D., & Choo, K. K. R. (2014). Impacts of increasing volume of digital forensic data: A survey and future research challenges. *Digital Investigation*, *11*(4), 273-294.

Quick, D., & Choo, K. K. R. (2017). Pervasive social networking forensics: intelligence and evidence from mobile device extracts. *Journal of Network and Computer Applications*, *86*, 24-33.

Raddick, M. J., Bracey, G., Gay, P. L., Lintott, C. J., Murray, P., Schawinski, K., & Vandenberg, J. (2009). Galaxy zoo: Exploring the motivations of citizen science volunteers. *arXiv preprint arXiv:0909.2925*.

Ramwell, S., Day, T., & Gibson, H. (2016). Use cases and best practices for LEAs. In B. Akhgar, P. S. Bayerl & F. Sampson (Eds). *Open Source Intelligence Investigation* (pp. 197-211). New York: Springer.

Revell, Q., Smith, T., & Stacey, R. (2016). Tools for OSINT-Based Investigations. In B. Akhgar, P. S. Bayerl & F. Sampson (Eds). *Open Source Intelligence Investigation* (pp. 153-165). New York: Springer.

Ribaux, O., Walsh, S. J., & Margot, P. (2006). The contribution of forensic science to crime analysis and investigation: forensic intelligence. *Forensic Science International*, *156*(2), 171-181.

Risen, J., & Lichtblau, E. (2013, June 8). How the US uses technology to mine more data more quickly. *The New York Times*. Retrieved from: https://www.nytimes.com/2013/06/09/us/revelations-give-look-at-spy-agencys-wider-reach.html.

Rogers, M. K., Goldman, J., Mislan, R., Wedge, T., & Debrota, S. (2006). Computer forensics field triage process model. *Journal of Digital Forensics, Security and Law. 1*(2), 19-37.

Rogstadius, J., Kostakos, V., Kittur, A., Smus, B., Laredo, J., & Vukovic, M. (2011). An assessment of intrinsic and extrinsic motivation on task performance in crowdsourcing markets. In *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media* (pp. 321-328).California: The AAAI Press.

Roiphe, R. (2003). The Serpent Beguiled Me: A History of the Entrapment Defense. *Seton Hall Law Review*, *33*(2), 257-296.

Sampson, F. (2016). Following the Breadcrumbs: Using Open Source Intelligence as Evidence in Criminal Proceedings. In B. Akhgar, P.S. Bayerl & F. Sampson (Eds). *Open Source Intelligence Investigation* (pp. 295-304). New York: Springer.

Simula, H. (2013). The rise and fall of crowdsourcing?. In: R. H. Sprague (Ed.). *Proceedings of the 46th Hawaii International Conference on System Sciences* (pp. 2783-2791). Washington: IEEE Computer Society.

Stock, J. (2014, December 10). *A Way Forward for Stronger International Cooperation*. Speech presented at Global Summit to Tackle Online Child Sexual Exploitation in the UK, London. Retrieved from: https://www.interpol.int/en/content/download/27460/368410/version/1/file/Global%20Summit%20to%20Tackle%20Online%20Child%20Sexual%20Exploitation%20-%20Keynote%20address%20J%C3%BCrgen%20Stock%20INTERPOL%20Secretary%20General-1.pdf.

Tapia, A. H., LaLone, N. J., & Kim, H. W. (2014). Run amok: Group crowd participation in identifying the bomb and bomber from the Boston marathon bombing. In S.R. Hiltz, M.S. Pfaff, L. Plotnick & P. Shih (Eds.). *Proceedings of the 11th International Conference on Information Systems for Crisis Response and Management* (pp. 265-274). Pennsylvania: The Pennsylvania State University.

Taylor, M., & Quayle, E. (2003). *Child pornography: An internet crime*. New York: Brunner-Routledge.

Trottier, D. (2014). Crowdsourcing CCTV surveillance on the Internet. *Information, Communication & Society*, *17*(5), 609-626.

Vendius, T. T. (2015). Proactive Undercover Policing and Sexual Crimes against Children on the Internet. *European Review of Organised Crime*, *2*(2), 6-24.

Westlake, B., Bouchard, M., & Frank, R. (2012). Comparing methods for detecting child exploitation content online. In: M. Memon & D. Zeng (Eds.). *2012 European Intelligence and Security Informatics Conference* (pp. 156-163). Washington: IEEE Computer Society.

Whittle, H., Hamilton-Giachritsis, C., Beech, A., & Collings, G. (2013). A review of online grooming: Characteristics and concerns. *Aggression and violent behavior*, *18*(1), 62-70.

Wolfson, S. M., & Lease, M. (2011). Look before you leap: Legal pitfalls of crowdsourcing. *Proceedings of the Association for Information Science and Technology*, *48*(1), 1-10.

Zgoba, K. M. (2004). Spin doctors and moral crusaders: The moral panic behind child safety legislation. *Criminal justice studies*, *17*(4), 385-404.