Jan 17th, 12:00 AM

# Help Wanted - Challenges in Data Collection, Analysis and Communication of Cyber Threats in Security Operation Centers

Ali Sercan Basyurt
*University of Duisburg-Essen*, ali-sercan.basyurt@uni-due.de

Jennifer Fromm
*University of Duisburg-Essen*, jennifer.fromm@uni-due.de

Philipp Kuehn
*TU Darmstadt*, kuehn@peasec.tu-darmstadt.de

Marc-André Kaufhold
*TU Darmstadt*, kaufhold@peasec.tu-darmstadt.de

Milad Mirbabaie
*Paderborn University*, milad.mirbabaie@uni-paderborn.de

# Help Wanted - Challenges in Data Collection, Analysis and Communication of Cyber Threats in Security Operation Centers

Ali Sercan Basyurt[1], Jennifer Fromm[1], Philipp Kuehn[2], Marc-André Kaufhold[2], and Milad Mirbabaie[3]

[1] University of Duisburg-Essen, Department of Computer Science and Applied Cognitive Science, Duisburg, Germany
{basyurt,fromm}@uni-due.de
[2] Technical University of Darmstadt, Department of Computer Science, Darmstadt, Germany
{kaufhold,kuehn}@peasec.tu-darmstadt.de
[3] Paderborn University, Department of Information Systems, Paderborn, Germany
{milad.mirbabaie}@uni-paderborn.de

**Abstract.** Security Operation Centers are tasked with collecting and analyzing cyber threat data from multiple sources to communicate warning messages and solutions. These tasks are extensive and resource consuming, which makes supporting approaches valuable to experts. However, to implement such approaches, information about the challenges these experts face while performing these tasks is necessary. We therefore conducted semi-structured expert interviews to identify these challenges. By doing so, valuable insights into these challenges based on expert knowledge is acquired, which in return could be leveraged to develop automated approaches to support experts and address these challenges.

**Keywords:** Cyber security · Cyber threat communication · Security Operations Center · Data collection · Data analysis

## 1 Introduction

Through the increasing digitalization across all aspects of life the threat posed by cyberattacks is also increasing. The Federal Criminal Police Office in Germany (BKA) noted in their annually published Cybercrime Federal Situation Report 2020 that 108,474 cases of cybercrime were registered while only 32.6 % of these cases were solved. Compared to previous years, the number of criminal activities is constantly increasing at a higher rate than cases being solved which highlights the severity of this type of crisis [1]. Thus, Security Operation Centers (SOCs) have been established "as an organizational unit operating at the heart of all security operations" with the objective to "detect, analyze, and respond to cybersecurity threats and incidents employing people, processes, and technology"[2]. Depending on their overall scope, SOCs offer both preventive, reactive and security quality services for specified stakeholders such as authorities, citizens, or enterprises [3].

To achieve their objectives, SOCs can utilize different types of information and communication technologies (ICT), including Security Information and Event Management (SIEM) technology to "analyze security event data in real time; and to collect, store, analyze and report on log data for regulatory compliance and forensics"[4]. However, as SIEMs focus on data from host systems and applications, as well as network and security devices, they are not suitable to monitor the threat and vulnerability landscape that is publicly discussed among security experts. In complement, Open Source Intelligence (OSINT) solutions allow to collect data from various open, public and social sources, including blogs, feeds, newsletters, social media, and websites [5]. However, by including more data, more resources and potentially tool support become necessary to handle the influx of data which might not be available to SOCs [6].

There are different information systems (IS) facilitating the collection of open [7] or social data [6], but they rely on information exclusively provided by accounts that were identified as experts in the field of cybersecurity. Furthermore, some commercial tools such as "Buffer" and "Hootsuite" allow to automatically post messages across different platforms, but do not provide technical support for the automated creation of effective warning messages. To the best of our knowledge, there is a lack of integrated tools tailored to the needs of SOCs, which facilitate data collection, analysis and communication across different channels. Our work is intended to work towards the creation of such a integrated approach by uncovering challenges throughout the entire process, from data collection to analysis and communication of the cyber situation. This leads to the following research question: **Which challenges do SOCs face when collecting and analyzing data as well as communicating cyber situations?**

In order to contribute to research in this field, we first conducted a literature review to review related work on data collection, analysis, and the communication of cyber threat information (Section 2). We then performed semi-structured interviews (Section 3) to identify areas of improvement in the collection and analysis of data and communication of cyber situations based on information acquired from cybersecurity experts (Section 4). We then discuss our findings (Section 5) and conclude our paper by depicting the novelty of our findings as well as outlining avenues for future research (Section 6).

## 2 Related Work

This section presents related work on the tasks of cyber threat collection, analysis and communication conducted by SOCs to enhance situational awareness (SA) [8] amongst involved stakeholders, i.e., the SOC ifself (by threat collection and analysis) and its clients (by threat communication). The term of cyber situational awareness was established as a subset of SA and is often related to knowing what is happening in the own network [9], but also "requires to gain a common operational picture of the threat environment in which the constituency is operating" [10]. In this work, we do not analyze the subset of network awareness, but focus on threat awareness enabled by open, public and social data.

### 2.1 Collection and Analysis of Cyber Threat Information

The cybersecurity domain spreads its information on various closed and open sources. While closed sources are protected by some kind of barrier, *e.g.*, a paywall, open sources are open to any access. Although closed sources offer a niche in the world of threat intelligence, since they offer distinct information, their volume is much smaller than the available information in open sources [11]. Social media sources like Twitter do have a very active cybersecurity community, which have been the focus of different studies in the past [12–19]. Especially for the use cases of threat event detection [17, 19], exploit prediction [14, 15] or hacker demasking [20] Twitter has been shown useful. Other sources in the domain of cybersecurity used for information extraction are blogs [21], bug reports, and security advisories [22], forums [7], the dark web [7, 23], or official security information sources like vulnerability databases [24].

When utilizing open sources to achieve cyber situational awareness, SOCs face challenges due to the high volume, velocity, and variety of open and social big data [25]. During daily monitoring tasks and large-scale incidents, the issue of information overload becomes apparent, defined as "information presented at a rate too fast for a person to process" [26, p. 823]. To mitigate the potential negative impact on decision making, several technical studies suggest to "transform the high volume of noisy data into a low volume of rich content" [27, p. 1] which is useful for response organizations. For instance, Mittal *et al.* [6] implemented "CyberTwitter" to generate alerts for cybersecurity threats and vulnerabilities extracted from Twitter. Further studies highlight the relevance of a thought-out usability, configurable filtering mechanisms, duplicate detection, grouping of similar messages, or information summaries as technological measures to overcome information overload and prioritize relevant information [28–30].

Still, open sources are prone to information disorder [31], including misleading, fabricated or manipulated content. In social media, information quality has been operationalised as a multidimensional construct of credibility, relevance, completeness, comprehensibility, and timeliness of information [32]. While the relevance and timeliness is addressed by data collection and filtering, existing research shows that machine learning algorithms, indicator-based assistance systems, or the improvement of media literacy can contribute to the assessment of credible information [33–35]. To ensure the completeness and comprehensibility of information for situational awareness and decision-making [8], SOCs require useful (visual) information overviews. For instance, Onorati *et al.* [36] provide a semantic visualization tool for Twitter, which utilizes the principles of visual analytics by combining "automatic analysis techniques with interactive visualisations for an effective understanding, reasoning and decision making on the basis of very large and complex data sets" [37, p. 157].

### 2.2 Effective Cyber Threat Communication

Once SOC staff have gained an understanding of the current cyber situation, the effective design and dissemination of cyber warning messages is of great

importance to reach as many affected stakeholders as possible and trigger security protection behaviors. Many IS studies on the effective design of cyber warning messages are based on protection motivation theory [38–42]. According to this theory [43], effective cyber threat communication should convey a high threat severity and actionable coping responses, so that fear is triggered in those affected, but also self-efficacy that required actions can be carried out. Johnston *et al.* [42] extended protection motivation theory and found that the communication of informal sanctions imposed by peers increases the intention to comply with security guidelines in organizations more than formal sanctions enforced by the organization. Further studies showed that concrete fear appeals with strong arguments increase protection motivation more than abstract fear appeals with weak arguments [39–41]. Meanwhile, studies suggested guilt and motivational appeals as more effective emotional appeals [38, 44].

Recent studies also emphasized the importance of actor-specific cyber threat communication, as certain message characteristics were found to be more effective when they aligned with individual characteristics of affected stakeholders. For example, Johnston *et al.* [45] found that an alignment of a fear appeal's rhetorical style (we vs. you) and organizational identification (low vs. high) increased security compliance behavior. Likewise, Plachkinova *et al.* [46] showed that loss-framed messaging was more effective for end-users with low initial security concerns, while messages focusing on desirable outcomes were more effective for end-users with high initial security concerns. With regard to the effective dissemination of cyber warning messages, neuro information systems research showed that a habituation effect already occurs after the second exposure, resulting in less attention being paid to the message [47]. In line with this, one study presented polymorphic cyber warning messages as an effective measure to counteract the habituation effect [48].

Previous research shows that effective cyber threat communication requires a lot of effort from SOCs, as messages should ideally be tailored to different audiences and designed in different ways to counteract the habituation effect. This highlights the importance of technological support for this work process, but no solutions specifically for automated cyber threat communication exist at this time. There are already commercial products that allow the automated distribution of messages on different social media channels (*e.g.*, "Buffer", "Hootsuite"), but they do not provide support for actor-specific content adaptation.

## 3 Method

To identify requirements and areas that can be improved in SOCs, we conducted semi-structured interviews with SOC members in Germany. This ensured that valuable insights into cybersecurity as well as into the inner workings, needs, and processes of SOCs could be gained. The interviews were conducted, recorded, and transcribed in German. Our nine participants came from different SOC types, i.e., authority-integrated units (P4, P5, P6, P9), a government-owned enterprise (P2), and private service providers (P1, P3, P7, P8). Regarding their roles,

five of them acted as cyber incident responders (P1, P3, P6, P7, P8), three as team leaders (P2, P4, P9), and one as public safety answering point for cyber incidents (P5). We chose semi-structured interviews, because specific questions for key areas where the experts' contribution is important can be determined in advance. Still, these also allowed the interviewer to ask follow-up questions if the participant provided unforeseen insights.

The interview guide included 29 questions which can be divided into six sections (see Table 1). The first section was dedicated to acquiring general descriptive information about the participant and their role in their organization, while the second section aimed at gaining insights into the process of how organizational cyber incidents are reported. The third and fourth section dealt with the collection of data on cybersecurity and the analysis of the cyber threat situation. Afterward, information about the communication of cyber threats was acquired. Lastly, ethical, organizational, and legal aspects were addressed.

**Table 1.** Interview guide: sections and exemplary questions

| **Interview sections** | **Exemplary questions** |
|---|---|
| Descriptive information | What does your normal day-to-day work usually look like? |
| Reporting of organizational cyber incidents | What data do you need or transmit in order to deal with cyber threats? |
| Collection of data on cybersecurity | What data sources are you analyzing to identify these threats and security vulnerabilities? |
| Situation picture for the cyber threat situation | To what extent is a situation picture of a cyber threat and security gaps created? |
| Communication of cyber threats | To what extent do you adapt cyber alerts to different communication channels and affected actors? |
| Ethical, organizational, and legal aspects | What are organizational and legal requirements for data processing or stakeholder communication? |

The interviews were conducted between February and April 2021 using Skype, with a length between 59 minutes to 109 minutes. Prior to the interview the questions were sent to the participants and they were given the chance to ask questions via E-Mail to ensure that they comprehend them. Using MAXQDA, the interviews were analyzed by means of a qualitative content analysis [49]. As a first step in the coding process, the transcripts of three interviews were read by four researchers and based on text fragments, each researcher inductively derived their own coding guide. These guides were discussed among the researchers and combined into one coding guide, which then was used by one researcher to label sections from all nine interview transcripts with matching codes from the code guide. Afterward, a second researcher coded the transcripts with the same coding guide. When differences in coding were found these differences were discussed until a uniform code was assigned. If a text section contained relevant information that could be considered separately, multiple codes were assigned for

this section. In cases where multiple codes were assigned, a comparison between these codes was made and when it was plausible they were merged to a superior code category.

## 4 Findings

An overview of the ICT use of the German state SOCs we analyzed in our study is depicted in Figure 1. The process can be divided into the steps of acquisition, analysis, and response. First, incidents are either reported by customers (via mail or telephone) or detected by software (such as intrusion detection). After initial information about the incident is gathered, SOCs use a ticketing and reporting system to collect their evidence for incident response. Second, this evidence is collected and analyzed using awareness-focused (*e.g.*, manufacturer websites, security advisory feeds, and social media channels such as Twitter where data is collected manually) and collaboration-oriented (*e.g.*, malware information sharing platforms, the collaborative chat of the German administrative CERT network) channels. Third, the collected evidence is then used to inform a certain stakeholder with specific recommendations, to provide (daily) reports for selected stakeholders, or to issue a general warning for multiple stakeholders in case larger-scaled ICT infrastructures are threatened.

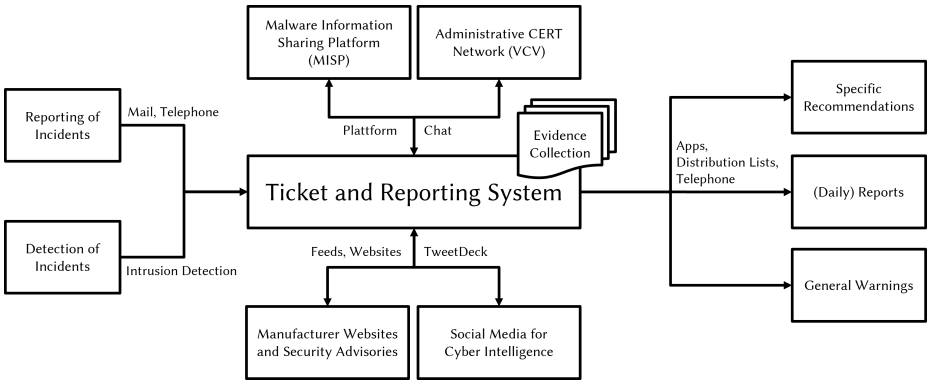


**Fig. 1.** ICT infrastructure of German state-level SOCs

### 4.1 Collection of Open-Source Data

For the collection of open-source data **legal restrictions** such as for the automated collection and long-term storage of data were named as challenging by six participants. One participant described the legal restrictions exemplary for Twitter data as follows: "We are not allowed to read your Twitter account automatically.The Office for the Protection of the Constitution says no. I can

however monitor your Twitter account manually but not automatically. There are just too strong legal concerns." (P3)

The **lack of application programming interfaces (APIs)** to acquire data from different sources but also the heterogeneity available APIs and their restrictions was described as another challenge in the data collection process by four participants. One individual explained, "the data basis on which we carry out our analyses is based on different data and our impression is all the interesting data, such as Instagram data, but also Facebook data are only available to a limited extent, and it is a huge effort to get the relevant information" (P1). Moreover, three participants expressed the **diversity of data formats** and the lack of standardized incident reports as another significant challenge. This increases the difficulty to collect and process data because "manufacturers are constantly changing their formats, or they adapt them, i.e., you can't even build a tool that works, because then you have the problem that Microsoft suddenly changes the format of their security information from one day to the next" (P2). Lastly, the sheer **number of available data sources** was named as a challenge by three experts. One of them elaborated that he has to visit " 56 websites in the morning, from software manufacturers, for example, to just look, are there reports that have not yet been recorded" (P5)

With regard to potential for improvement, seven of nine experts highlighted the **automation of data collection and deletion of data that is not needed anymore** and explained that "if something existed, where the information provided by all kinds of people, for example from the top 100 security experts worldwide, would be combined from social media. I think that would be very, very helpful" (P7). In direct relation with the sentiment of combining data four participants described the desire of **unifying data streams by summarizing different data sources** as a potential improvement area by further conveying that "an analysis tool where various information channels converge would be of course really attractive" (P1). Moreover, three experts expressed the desire for a **standardized format for data and reports** because data from different sources is also structured differently which makes it hard to automatically process. One participant further explained that the available information is "extremely difficult to process automatically, that is the problem because they are not beautiful Extensible Markup Language (XML) data, they are free texts that are structured in completely different ways" (P2).

### 4.2 Analysis of the Cyber Situation

Five of the nine SOCs experts identified the **high effort to assess data collected from different sources manually** as a significant challenge for the analysis of the cyber situation. One participant elaborated that "[...] the processing of the filtered results is then a manual process [...] we currently have tweets from over 500 Twitter accounts, and of course they have also followers or other sources so that an unbelievable amount of data is present" (P2). Related to this, interviewees raised concerns about how to assess the **trustworthiness**

**of sources and information** since "it is, of course, possible that you have information from a trustworthy partner, that is still false" (P2). In addition, three experts named **irrelevant and redundant information** to be an issue for the second category by exemplarily expressing that they "[. . . ] have to filter out what is in the area of interest, what is not in the area of interest" (P6) and "data that is collected twice or three times or comes in waves" (P9).

Moreover, four participants identified **legal restrictions with regard to personal data** as a critical challenge for the analysis of the cyber situation because these limit their possibilities for data analysis. It was stated that "the e-government law has been in place since the beginning of the year. That puts us quite tightly in terms of handling personal data. [. . . ] This means that if we have to process such data, then we need the approval of a person who is qualified to act as a judge" (P2).

Out of nine experts, four described the **automatization of processes** such as the filtering of redundant data as an important improvement that they desired for their current work. An expert elaborated: "It would be helpful if, at some point, we had such a tool at our disposal, where incoming information could be prepared and cataloged automatically. [...] But any help or support or automation would, of course, be very welcome at this point" (P4). In accordance, P2 valued the idea of a **visual dashboard** which monitors different open data (e.g., Pastebin, RSS feeds) and social media (e.g., Facebook, Twitter) sources, if it supports the assessment of the criticality and urgency of information.

### 4.3 Communication of Cyber Threats

For the communication of cyber threats, six participants described **reaching specific target groups with warning messages** as challenging because currently not all SOCs use multiple channels to reach specific groups. An expert further described that depending on their IT knowledge and skills, users often have additional questions when receiving a general warning message: "There are often inquiries from individuals because the people out there are not all technically savvy, and then we have to give advice and provide support over the phone" (P2). Furthermore, the **effectiveness of a warning message and how to assess it** were described as a challenge by five experts since message receivers do not always follow the given advice. An expert explained: "the implementation of given advice [...] is more likely to be the problem. The users on the other side of the system might not see the urgency" (P5). Six experts found that **their own lack of knowledge about the software used by target groups** and their technical skills makes it challenging to identify relevant recipients for warning messages: "So it's just too much to know all the software we use here. I know maybe 80 percent, maybe 90, but I can't get to a hundred" (P3).

Moreover, three participants described the **lengthy approval process of warning messages** as challenging due to having to wait for approval from multiple individuals. "What is always very time-consuming from my point of view is the internal approval of warning messages, because there are also other

things that the approving parties need to take care of" (P5). Lastly, in regard to the communication of warning messages, three participants described **the manual effort needed to publish security warnings** as a challenge. One of them explained: "It is the case that an HyperText Markup Language (HTML) code can be generated from the ticket, which then has to be imported relatively easily, but manually, into the TYPO3 system. But this is important because this TYPO3 system does not always do what it is supposed to do, and therefore you sometimes have to readjust it manually" (P2).

Two experts described the desire for **publishing security warnings automatically** as an area of improvement, although they were not sure about the technical implementation: "that [publishing warning messages] requires a lot of manual effort, but there is probably no other way?" (P2). Furthermore, three experts desired **customizable warning messages for target groups** in the form of selecting keywords to generate a message for a group instead of writing individual messages. One expert described a potential form of automation as follows: "The message is automatically created, through the case, category and target being automatically identified [...]" (P4). Finally, two SOCs members wished for an **approval platform for warning messages** that speeds up the approval process of a message, for example, by creating "[. . . ] a platform on which to revise this document without constantly sending it back and forth. Maybe that would make it a little easier" (P5).

Figure 2 shows which challenges could be potentially addressed by the suggested improvements.

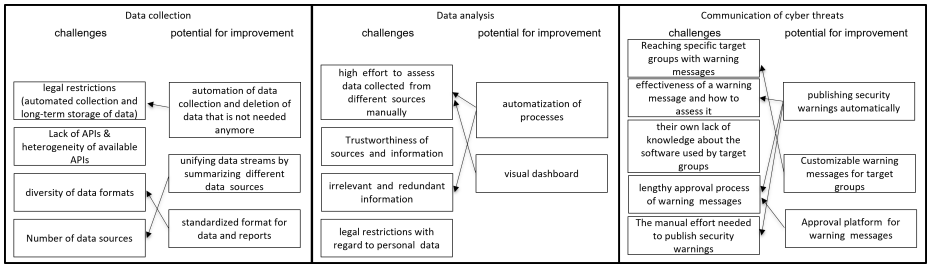


**Fig. 2.** Mapping of challenges to improvement suggestions

## 5 Discussion

The interviews provide findings regarding challenges and areas for improvement identified by SOCs experts for the collection of open-source data, analysis of the cyber situation, and communication of cyber threats. For the communication of cyber threats five challenges were identified wheras four challenges were named for each remaining task indicating that currently each of these tasks is associated

by SOCs with hurdles. It can be deduced that the experts wish for improvements for the collection of data and communication of cyber threats the most based on the three improvements suggested for each category, compared to only one improvement suggestion for the analysis of the cyber situation. These findings underline that there is still a need for a tool or other measures to fully support SOCs members in their daily work, starting from the collection and analysis of data up to the communication of cyber threat information. An example for such a tools architecture based on our findings is conceptually visualized in Figure 3.

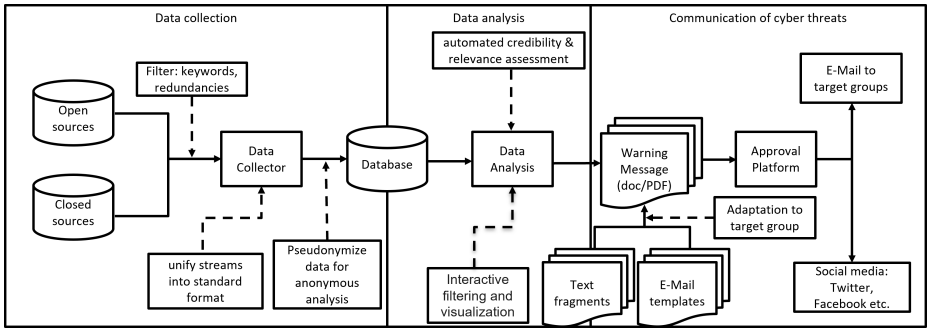


**Fig. 3.** Exemplary tool architecture based on expert interview results

To address the challenge of various input formats for the collection of open-source data, a standardized form of incident reporting in a standardized data format could be developed. However, this leads to the issue of how to motivate the different data sources to use this form and format. A solution for this could be to retrieve the data from different sources in their original form and to automatically transfer the information into the standardized format. Furthermore, the experts named the sheer volume of data sources to be challenging in the data collection process in alignment with Olshannikova *et al.* [25]. This issue could be addressed by an application that as previously suggested collects data automatically from predefined sources in real-time with the addition of filtering and deleting data that is not needed anymore as suggested by experts as a possibility for improvement. This solution is in alignment with findings in previous studies where similar suggestions to overcome information overload were made, such as filtering and removing duplicate information [28–30]. The issues mentioned by experts regarding the APIs, their variety, their scarcity, and the restrictions imposed on their use must be addressed with their corporate publishers because they have the rights to address them. The legal restrictions for the automated collection of data are determined by politicians and institutions, therefore, they must be abided by while working with this type of data. However, a precautionary step to ensure the anonymity of users whose data is collected could be by pseudonymizing the personal information of users before they are stored in

a database. By doing so, data that is not usable due to infringing on privacy regulations could be made available for further use.These aspects are visualized in the data collection section of Figure 3.

Regarding the analysis of the cyber situation, the experts named several challenges that could be addressed by the desired automatization of processes during the analysis section of their work. This includes challenges such as assessing the credibility of information as well as assessing its relevance. These issues could be addressed by an application that automatically retrieves data based on keywords and other parameters to obtain only relevant information from predefined sources that the SOCs currently review manually. In order to reduce false information and assess the trustworthiness of sources, the data from one source could be cross-checked automatically with other sources and if the information is reported by several sources, it can be considered credible together with the source. However, the legal restrictions for the analysis of data are in the hands of politicians and institutions, and are therefore outside the scope of our abilities. But as previously suggested the pseudonymization of user information before storing and analyzing it could be an approach to address this issue. These automatizations to address the challenges during the analysis process are included in the data analysis section of Figure 3.

One major problem for the communication of cyber threats was described to be reaching specific target groups with effective warning messages. Different groups requiring different formulations and information as part of the message which recognize their IT knowledge and skills. This issue can be addressed by the suggested improvement of customizable warning messages for target groups by creating an application with prewritten text fragments and related keywords for each fragment so that a warning message can be generated for a threat and group, by selecting relevant keywords. By creating messages tailored to target groups the effectiveness of a warning message could be improved due to providing them with information that is relevant and understandable for them. Different emotional appeals could be considered while creating these messages for different target groups to increase the effectiveness of the threat message, as was described in existing informations systems literature [38, 44–46]. Furthermore, the suggested solution to generate warning messages would also address the challenge presented by the high manual effort necessary to create warnings by reducing the process of creating a message to selecting keywords and proofreading the message before publishing it. To acquire knowledge about the technologies used by the target groups and the effectiveness of a message, the target groups must be further consulted. Lastly, to shorten the lengthy approval process for warning messages that has been described as a challenge, a platform could be created where parties approving an alert are able to make corrections to the alert and approve it by signing off on the alert on the platform, rather than sending alerts back and forth with suggestions and an approval alert. Fully automating the communication of a warning message seems currently not feasible due to it requiring the approval from multiple individuals and multiple iterations. Additionally, patches for threats might not be available immediately and

the exploitation of a threat could possibly have dire consequences which makes a careful assessment before communicating a warning message important. After the approval of a message, it could be automatically published to predefined channels to address the desire for automation for the publication of messages that were expressed by participants. This would also further reduce the manual effort. These components of the communication process are display in the communication of cyber threats section of Figure 3.

By conducting these interviews with SOCs experts an empirical research contribution according to Wobbrock and Kientz (2016) is made through new insights into challenges and areas where improvements are desired based on the opinions of experts and by developing a conceptual tool architecture that addresses these challenges and areas for improvements shown in Figure 3 [50]. They are focused on the collection and analysis of data and communication of cyber situations. The practical implication of our research is that the additional insights into these areas can be leveraged to implement new approaches as well as to further improve existing ones such as "Discover" [7] and "CyberTwitter" [6] for the execution of the above-mentioned tasks.

## 6 Conclusion

We identified multiple challenges and areas for improvement for SOCs when collecting and analyzing data as well as communicating the cyber situation through expert interviews with SOCs members. We found multiple challenges and possibilities for improvements for each of these tasks and suggested possibilities for addressing these challenges by bridging the areas for improvement with the challenges such as the development of a tool that automatically cross-checks retrieved information about a threat with different sources to assess the trustworthiness of the information and their source.

This study contributes to research by identifying these challenges and potential areas for improvement based on information provided by experts and SOCs members. Furthermore, these areas for improvement and the suggestions made for how to address them represent the practical implications of our work which can be used to develop a tool to provide targeted support to SOCs employees for example by implementing the conceptual tool described in Figure 3.

Nevertheless, the explanatory power of our results is limited, due to the number of expert interviews conducted. The sample of interviews was purposefully chosen, however, they only included SOCs from Germany which also presents a limitation of our research since our results may not cover challenges and areas of improvement perceived by SOCs members across the globe. The tool in Figure 3 has not been validated with experts yet. However, this will be done in future work. Based on these limitations, we suggest that additional interviews should be conducted with SOCs from different cultural backgrounds to strengthen the findings of our work. Additionally, we suggest developing prototypes that address our findings and presenting them in evaluation studies with SOCs.

## References


1. Bundeskriminalamt: Cybercrime bundeslagebild. Tech. rep., Bundeskriminalamt (2020)
2. Vielberth, M., Böhm, F., Fichtinger, I., Pernul, G.: Security operations center: A systematic study and open challenges. IEEE Access 8, 227756–227779 (2020)
3. Riebe, T., Kaufhold, M.A., Reuter, C.: The Impact of Organizational Structure and Technology Use on Collaborative Practices in Computer Emergency Response Teams: An Empirical Study. Proceedings of the ACM: Human Computer Interaction (PACM): Computer-Supported Cooperative Work and Social Computing 5(CSCW2) (2021)
4. Montesino, R., Fenz, S., Baluja, W.: Siem-based framework for security controls automation. Emerald Group Publishing Limited (2012)
5. Alves, F., Bettini, A., Ferreira, P.M., Bessani, A.: Processing tweets for cybersecurity threat awareness. vol. 95, p. 16 (2021), https://www.sciencedirect.com/science/article/pii/S0306437920300727
6. Mittal, S., Das, P.K., Mulwad, V., Joshi, A., Finin, T.: CyberTwitter: Using Twitter to generate alerts for cybersecurity threats and vulnerabilities. In: Proceedings of the 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2016. pp. 860–867 (2016)
7. Sapienza, A., Ernala, S.K., Bessi, A., Lerman, K., Ferrara, E.: DISCOVER: Mining Online Chatter for Emerging Cyber Threats. In: Companion of the The Web Conference 2018 on The Web Conference 2018 - WWW '18. pp. 983–990. ACM Press, Lyon, France (2018), http://dl.acm.org/citation.cfm?doid=3184558.3191528
8. Endsley, M.R.: Toward a theory of situation awareness in dynamic systems. vol. 37, pp. 32–64. SAGE Publications Sage CA: Los Angeles, CA (1995)
9. Franke, U., Brynielsson, J.: Cyber situational awareness - A systematic review of the literature. Computers and Security 46, 18–31 (2014), http://dx.doi.org/10.1016/j.cose.2014.06.008
10. Ruefle, R., Dorofee, A., Mundie, D., Householder, A.D., Murray, M., Perl, S.J.: Computer security incident response team development and evolution. IEEE Security & Privacy 12(5), 16–26 (2014)
11. Bouwman, X., Griffioen, H., Egbers, J., Doerr, C., Klievink, B., van Eeten, M.: A different cup of TI? The added value of commercial threat intelligence. pp. 433–450 (2020), https://www.usenix.org/conference/usenixsecurity20/presentation/bouwman
12. Alves, F., Ferreira, P.M., Bessani, A.: Design of a Classification Model for a Twitter-Based Streaming Threat Monitor. In: 2019 49th Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshops (DSN-W). pp. 9–14 (Jun 2019), iSSN: 2325-6664
13. Tundis, A., Ruppert, S., Mühlhäuser, M.: On the Automated Assessment of Open-Source Cyber Threat Intelligence Sources. In: Krzhizhanovskaya, V.V., Závodszky, G., Lees, M.H., Dongarra, J.J., Sloot, P.M.A., Brissos, S., Teixeira, J. (eds.) Computational Science – ICCS 2020. pp. 453–467. Lecture Notes in Computer Science, Springer International Publishing, Cham (2020)
14. Chen, H., Liu, R., Park, N., Subrahmanian, V.S.: Using twitter to predict when vulnerabilities will be exploited. In: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 3143–3152. ACM Press, New York, New York, USA (2019), http://dl.acm.org/citation.cfm?doid=3292500.3330742

15. Sabottke, C., Suciu, O., Dumitras, T.: Vulnerability disclosure in the age of social media: Exploiting twitter for predicting real-world exploits. pp. 1041–1056 (2015)
16. Ritter, A., Wright, E., Casey, W., Mitchell, T.: Weakly supervised extraction of computer security events from twitter. In: WWW 2015 - Proceedings of the 24th International Conference on World Wide Web. vol. i, pp. 896–905 (2015)
17. Le Sceller, Q., Karbab, E.B., Debbabi, M., Iqbal, F.: SONAR: Automatic Detection of Cyber Security Events over the Twitter Stream. In: Proceedings of the 12th International Conference on Availability, Reliability and Security. pp. 1–11. ACM, Reggio Calabria Italy (Aug 2017), https://dl.acm.org/doi/10.1145/3098954.3098992
18. Yagcioglu, S., Seyfioglu, M.S., Citamak, B., Bardak, B., Guldamlasioglu, S., Yuksel, A., Tatli, E.I.: Detecting Cybersecurity Events from Noisy Short Text. In: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers). pp. 1366–1372. Association for Computational Linguistics, Minneapolis, Minnesota, USA (2019), http://aclweb.org/anthology/N19-1138
19. Riebe, T., Wirth, T., Bayer, M., Kuehn, P., Kaufhold, M.A., Knauthe, V., Guthe, S., Reuter, C.: CySecAlert: An Alert Generation System for Cyber Security Events Using Open Source Intelligence Data. In: International Conference on Information and Communications Security (ICICS) (2021)
20. Jones, K., Nurse, J.R.C., Li, S.: Behind the Mask: A Computational Study of Anonymous' Presence on Twitter. vol. 14, pp. 327–338 (May 2020), https://ojs.aaai.org/index.php/ICWSM/article/view/7303
21. Liao, X., Yuan, K., Wang, X., Li, Z., Xing, L., Beyah, R.: Acing the IOC game: Toward automatic discovery and analysis of open-source cyber threat intelligence. In: Proceedings of the ACM Conference on Computer and Communications Security. vol. 24-28-Octo, pp. 755–766. ACM Press, New York, New York, USA (2016), http://dl.acm.org/citation.cfm?doid=2976749.2978315
22. Feng, X., Liao, X., Wang, X.F., Wang, H., Li, Q., Yang, K., Zhu, H., Sun, L.: Understanding and securing device vulnerabilities through automated bug report analysis. pp. 887–903 (2019)
23. Robertson, J., Diab, A., Marin, E., Nunes, E., Paliath, V., Shakarian, J., Shakarian, P.: Darkweb Cyber Threat Intelligence Mining. Cambridge University Press, Cambridge (2017), http://ebooks.cambridge.org/ref/id/CBO9781316888513
24. Kuehn, P., Bayer, M., Wendelborn, M., Reuter, C.: OVANA: An Approach to Analyze and Improve the Information Quality of Vulnerability Databases. In: Proceedings of the 16th International Conference on Availability, Reliability and Security. p. 11. ACM (2021), https://doi.org/10.1145/3465481.3465744
25. Olshannikova, E., Olsson, T., Huhtamäki, J., Kärkkäinen, H.: Conceptualizing Big Social Data. vol. 4, pp. 1–19. Springer International Publishing (2017)
26. Hiltz, S.R., Plotnick, L.: Dealing with Information Overload When Using Social Media for Emergency Management: Emerging Solutions. In: Comes, T., Fiedrich, F., Fortier, S., Geldermann, J., Müller, T. (eds.) Proceedings of the International Conference on Information Systems for Crisis Response and Management (ISCRAM). pp. 823–827. ISCRAM Digital Library, Baden-Baden (2013)
27. Moi, M., Friberg, T., Marterer, R., Reuter, C., Ludwig, T., Markham, D., Hewlett, M., Muddiman, A.: Strategy for processing and analyzing social media data streams in emergencies. In: 2015 2nd International Conference on Information and Communication Technologies for Disaster Management (ICT-DM). pp. 42–48 (2015)

28. Plotnick, L., Hiltz, S.R.: Software Innovations to Support the Use of Social Media by Emergency Managers. vol. 34, pp. 367–381. Taylor & Francis (2018), https://doi.org/10.1080/10447318.2018.1427825
29. Alam, F., Ofli, F., Imran, M.: Descriptive and visual summaries of disaster events using artificial intelligence techniques: case studies of Hurricanes Harvey, Irma, and Maria. pp. 1–31. Taylor & Francis (2019), https://doi.org/10.1080/0144929X.2019.1610908
30. Kaufhold, M.A., Rupp, N., Reuter, C., Habdank, M.: Mitigating Information Overload in Social Media during Conflicts and Crises: Design and Evaluation of a Cross-Platform Alerting System. vol. 39, pp. 319–342. Taylor & Francis (2020)
31. Wardle, C., Derakhshan, H.: Information disorder: Toward an interdisciplinary framework for research and policy making. vol. 27 (2017)
32. Moi, M., Habig, T., Schubert, A., Brune, M., Witter, F., Kiel, M.: EmerGent Deliverable 4.5: Information Quality Criteria and Indicators. Tech. rep., University of Paderborn, Paderborn (2017), http://www.fp7-emergent.eu/d4-5-information-quality-criteria-and-indicators/
33. Viviani, M., Pasi, G.: Credibility in social media: opinions, news, and health information—a survey. vol. 7, pp. 1 – 25. Wiley Periodicals, Inc (2017), http://dx.doi.org/10.1002/widm.1209
34. Hartwig, K., Reuter, C.: TrustyTweet: An Indicator-based Browser-Plugin to Assist Users in Dealing with Fake News on Twitter. In: Proceedings of the International Conference on Wirtschaftsinformatik (WI). Siegen (2019)
35. Mihailidis, P., Viotty, S.: Spreadable Spectacle in Digital Culture: Civic Expression, Fake News, and the Role of Media Literacies in "Post-Fact" Society. vol. 61, pp. 441–454 (2017)
36. Onorati, T., Díaz, P., Carrion, B.: From social networks to emergency operation centers: A semantic visualization approach. vol. 95, pp. 829–840. Elsevier (2019)
37. Keim, D., Andrienko, G., Fekete, J.d., Carsten, G., Melan, G.: Visual Analytics: Definition, Process and Challenges. pp. 154–175 (2008), http://hal-lirmm.ccsd.cnrs.fr/lirmm-00272779
38. Menard, P., Bott, G.J., Crossler, R.E.: User motivations in protecting information security: Protection motivation theory versus self-determination theory. vol. 34, pp. 1203–1230. Taylor & Francis (2017)
39. Orazi, D.C., Warkentin, M., Johnston, A.C.: Integrating construal-level theory in designing fear appeals in is security research. vol. 45, p. 22 (2019)
40. Schuetz, S.W., Benjamin Lowry, P., Pienta, D.A., Bennett Thatcher, J.: The effectiveness of abstract versus concrete fear appeals in information security. vol. 37, pp. 723–757. Taylor & Francis (2020)
41. Jansen, J., van Schaik, P.: The design and evaluation of a theory-based intervention to promote security behaviour against phishing. vol. 123, pp. 40–55. Elsevier (2019)
42. Johnston, A.C., Warkentin, M., Siponen, M.: An enhanced fear appeal rhetorical framework. vol. 39, pp. 113–134. JSTOR (2015)
43. Rogers, R.W.: A protection motivation theory of fear appeals and attitude change. vol. 91, pp. 93–114. Taylor & Francis (1975)
44. Park, E., Kim, J., Shan, E.: Guilt appeals and information security policy compliance. In: Americas Conference on Information Systems (2018)
45. Johnston, A.C., Warkentin, M., Dennis, A.R., Siponen, M.: Speak their language: Designing effective messages to improve employees' information security decision making. vol. 50, pp. 245–284. Wiley Online Library (2019)
46. Plachkinova, M., Menard, P.: An examination of gain-and loss-framed messaging on smart home security training programs. pp. 1–22. Springer (2019)

47. Anderson, B., Vance, T., Kirwan, B., Eargle, D., Howard, S.: Users aren't (necessarily) lazy: Using neurois to explain habituation to security warnings. In: International Conference on Information Systems (2014)
48. Anderson, B.B., Vance, A., Kirwan, C.B., Eargle, D., Jenkins, J.L.: How users perceive and respond to security messages: a neurois research agenda and empirical study. vol. 25, pp. 364–390. Taylor & Francis (2016)
49. Mayring, P., Fenzl, T.: Qualitative inhaltsanalyse, handbuch methoden der empirischen sozialforschung. In: :, pp. 534–547. Springer-Verlag (2014)
50. Wobbrock, J.O., Kientz, J.A.: Research contributions in human-computer interaction. vol. 23, pp. 38–44. ACM New York, NY, USA (2016)