GDPR Compliance in Cybersecurity Software: A Case Study of DPIA in Information Sharing Platform

Martin Horák
Institute of Computer Science
Masaryk University
Brno, Czech Republic
horak@ics.muni.cz

Václav Stupka
Institute of Computer Science
Masaryk University
Brno, Czech Republic
stupka@ics.muni.cz

Martin Husák
Institute of Computer Science
Masaryk University
Brno, Czech Republic
husakm@ics.muni.cz

ABSTRACT

In this article, we discuss the issues of GDPR's impact on cybersecurity software and operations, namely automated information sharing. We illustrate the topic on an example of an intrusion detection alert sharing platform. First, we had to investigate the risks to privacy in the alert sharing platform and ensure its compliance with the GDPR's obligations. Second, fears and uncertainties emerged in the alert sharing community regarding the GDPR and its obligations and, thus, willingness to share the information was negatively impacted. We conducted DPIA to investigate risks related to information sharing in cyber security and dismiss doubts within the community. Although our results suggest that the risks are not high, we point out that the hype around GDPR caused substantial development of the sharing platform. The DPIA helped in a deeper understanding of risks and their management and is a solid argument for information sharing in cyber security under GDPR.

CCS CONCEPTS

• **Security and privacy** → *Intrusion/anomaly detection and malware mitigation*; • **Social and professional topics** → *Privacy policies.*

KEYWORDS

Intrusion detection, Information sharing, GDPR, Personal data, Privacy, CSIRT

ACM Reference Format:
Martin Horák, Václav Stupka, and Martin Husák. 2019. GDPR Compliance in Cybersecurity Software: A Case Study of DPIA in Information Sharing Platform. In *Proceedings of the 14th International Conference on Availability, Reliability and Security (ARES 2019) (ARES '19), August 26–29, 2019, Canterbury, United Kingdom.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3339252.3340516

1 INTRODUCTION

The impact of the General Data Protection Regulation (GDPR) on digital society is unquestionable, and the cybersecurity community is no exception. Nevertheless, European cybersecurity community ended up in an interesting situation, exploring ways to comply with GDPR's obligations in its constituency, while also complying itself. The cybersecurity procedures and tools were questioned for compliance with GDPR, and many potentially problematic issues emerged. Cybersecurity community, most apparently in the field of network security and operations of Computer Security Incident Response Teams (CSIRT), had to accept the widened definition of personal data that, under GDPR, includes IP addresses and other identifiers used in network communication. Such data identify regular users, whose privacy should be preserved, but also malicious actors, whose activities should be investigated and mitigated.

The issues revolving around GDPR are most apparent in collaboration of cybersecurity teams. CSIRTs often share information on current threats, events, and countermeasures to warn others, effectively mitigate the threats, and contribute to improving the global security situation. There is a plethora of procedures, tools, platforms, and communities, where almost any imaginable piece of information or knowledge may be shared. Prime examples are collaborative intrusion detection systems [20] or intrusion detection alert sharing platforms [10], where information on security events are shared and analyzed, often by multiple peers. However, some of the data shared and analyzed by these platforms are personal data. Thus, legal aspects of such tools and operations should be taken seriously and, if possible, performed *by design*, i.e., during the development of the tools.

In addition, the sharing community became aware of GDPR's obligations, and voices of fear and uncertainty appeared. Different legal frameworks protect certain categories of data at the EU level, such as intellectual property, trade secrets, and classified information, but the personal data protection is referenced as the most potential hurdle to information sharing [16]. Therefore, it is important to moderate these obstacles so that the security community would not avoid sharing information.

To reflect the needs of the cybersecurity community concerning the novel legal framework, we conducted the Data Privacy Impact Assessment (DPIA) procedure to investigate the risks of information sharing in cyber security. Our findings regarding the investigated system provided valuable feedback to the developers and helped in further improving the platform and disrupting any fears and doubts potential users may have.

This paper is organized into five sections. A brief overview of related work can be found in Section 2. Subsequently, we illustrate information sharing in cyber security on the example of intrusion detection alert sharing platform in Section 3. In Section 4, we describe the concepts of Data Privacy Impact Assessment (DPIA) and apply them in the assessment of the sharing platform. Finally, Section 5 concludes the paper and summarizes the lessons learned from our case study.

ARES '19, August 26–29, 2019, Canterbury, United Kingdom

2 RELATED WORK

Legal aspects of alert sharing platforms belong to a broader discussion about cyber security, privacy, and law. Some of the basic concepts of this discussion can be found in the paper of Sokol et al. [17] In this paper, we can also found argumentation about IP addresses which authors regard as a personal data even in the meaning of the current Data Protection Directive. The questions of how data protection limits the operation of the cybersecurity alert sharing platform was partially answered in our previous work [18]. In this paper, we identified the purpose and legal ground for data processing. Since the selected legal ground is a legitimate interest, they also provide the proportionality test and related argumentation for necessity, legitimacy, and balance of the data processing. Serrano et al. [15] focus on the legal aspects of the cyber security data sharing between organizations subject to different legal frameworks. They also propose a solution in the form of Information Exchange Policy (IEP), which is set up by the organizations themselves. Bhatia et al [4] discussed the privacy risks in sharing the data in cyber security from the US perspective. Their work is based on surveying security professionals and illustrates the trade-offs between data usage and user's willingness to share their data. Finally, the European Network and Information Security Agency (ENISA) published a study which deals with legal and regulatory factors and performs an assessment of what effects these factors have on cross-border data sharing between CERTs [8]. One of the factors the study focuses on is the European legal framework governing data protection and privacy. The study provides an example of a legal checklist for privacy and data protection. The concrete solution was recently proposed by Hellwig et al. [11], who presented a GDPR-compliant module for supporting the exchange of information between Computer Emergency Response Teams (CERT). Legal implications of information sharing, along with implementation issues and lessons learned from the ECOSSIAN project, were recently summarized in the book by Florian Skopik [16].

Nevertheless, the issue of profiling was not deeply examined yet. In some cases, the analytics operations of the platforms may fall within the definition of profiling in the General Data Protection Regulation (GDPR). The Article 29 Working Party in its guides [2] states that profiling indicates that planned processing may pose a high risk for data subjects' rights and therefore a Data Protection Impact Assessment (DPIA) should be conducted. We could argue that in case of a security alert sharing platform, this is not necessary, but that was not our goal. We conducted DPIA of our platform to distract the concerns arising from the possible violation of GDPR. As from the very essence of security alert sharing platform follows that for its proper function, it is necessary to encourage subjects to join the platform and share their data. In this article, we describe our process of DPIA of security alert sharing platform.

3 INFORMATION SHARING IN CYBER SECURITY

Collaboration and exchange of information and knowledge are vital for enforcing cyber security. Cybersecurity communities share threat intelligence, vulnerability reports, and forensic data, as well as incident response guidelines, reports, and best practices. One of the possible scenarios of collaboration in cyber security is the exchange of information on observed security events, typically in the form of alerts raised by intrusion detection systems. Sharing such information allows for understanding the situation on a larger scale, e.g., for advanced intrusion detection, security situation assessment, and early warning. Alerts often overlap with Indicators of Compromise (IoC), alerts often include IoCs, and IoCs are often based on observed events described in alerts. The main difference is that IoC is an artifact (IP address, malware signature, botnet C&C, etc.), while alert refers to an event in time, including events associated with IoC (malware signature detected, communication with a blacklisted IP address or botnet C&C observed, etc.). In this paper, we discuss primarily alert sharing platforms, where alerts from intrusion detection systems and similar data are shared. There is a plethora of existing tools, platforms, and sharing communities. Readers interesting in extensive comparison of standards, tools, platforms, and formats are kindly referred to the white papers by ENISA [9, 10]. Selected aspects of information sharing in large scale and with focus on threat intelligence were recently discussed in a book by Florian Skopik [16]. From research papers, we would recommend a survey of collaborative intrusion detection systems by Vasilomanolakis et al. [20].

3.1 The Assessed Alert Sharing Platform

To illustrate the sharing of security alerts, we briefly describe the assessed sharing platform. SABU[1] is an intrusion detection alert sharing platform developed by CESNET and Masaryk University. The intended scope of its usage is the network of CESNET, Czech national research and education network (NREN), and its partners. The platform is based around a central hub for sharing alerts in the platform, where the alerts are exchanged. Further, data enrichment, analysis, and reporting facilities are also part of the platform. A schema of the sharing platform alert sharing platform presented in Figure 1.

A central component of the SABU alert sharing platform is Warden[2], a central sharing hub that receives the alerts from sending connectors and distributes them to receiving connectors. The sending connectors are typically intrusion detection systems, honeypots, and other tools that raise cybersecurity alerts. Moreover, sending connectors may also transfer data from other alert sharing platforms. The receiving connectors are either reporters or active network defense interfaces. Reporters, such as Mentat[3], create periodical reports or send alerts to registered users who process them manually. In contrary, interfaces to active network defense tool receive the data and convert them into blacklists, firewall and spam filter rules, etc. An analytical component iABU is both receiving and sending connector. It is an instance of the AIDA framework[4] that performs syntactic and semantic checks, alert aggregation and correlation, visualization, and advanced data analytics [12]. The processes of alert analysis loosely follow fundamental concepts proposed by Valeur et al. [19], who proposed basic procedures and algorithms for alert correlation in general. Finally, a reputation database NERD[5] keeps reputation scores of IP addresses found in

[1] https://sabu.cesnet.cz/en/start
[2] https://warden.cesnet.cz/en/index
[3] https://mentat.cesnet.cz/en/index
[4] https://github.com/CSIRT-MU/AIDA-Framework
[5] https://nerd.cesnet.cz/

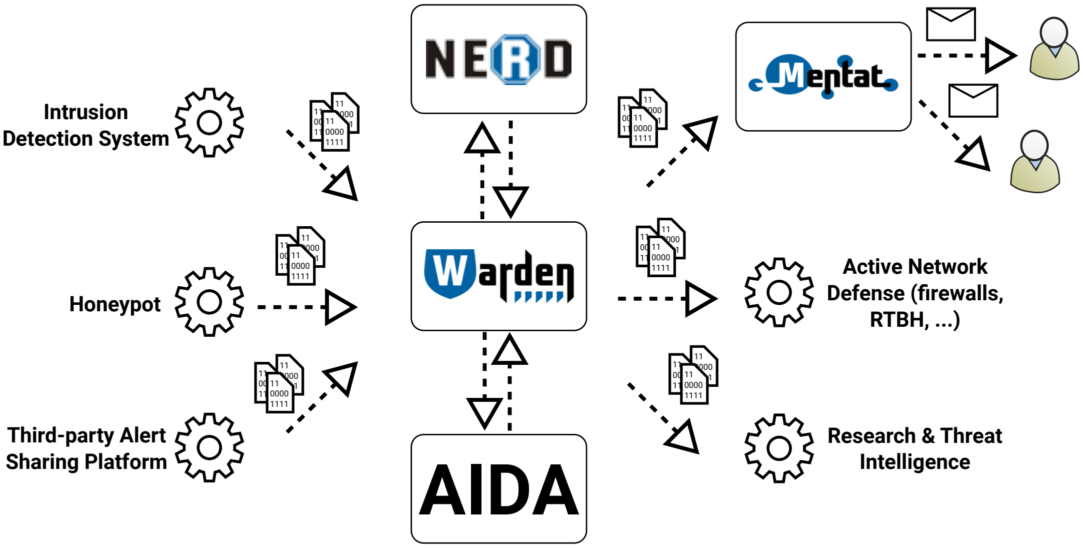
NERD
Mentat
Intrusion Detection System
Honeypot
Warden
Active Network Defense (firewalls, RTBH, ...)
Third-party Alert Sharing Platform
Research & Threat Intelligence
AIDA

Figure 1: Schema of SABU, the assessed intrusion detection alert sharing platform.

the alerts, adds the current score to the IP addresses in alerts, and provides other data enrichment, such as geolocation, etc. The score calculation is based mostly on the number and severity of alerts containing a given IP address and its presence on blacklists and other data sources.

The sharing platform uses Intrusion Detection Extensible Alert (IDEA)[6], a format of security alerts inspired by IDMEF [7], but customized to reflect operational needs [13], and enhanced with a taxonomy of security events [14]. Most interesting entries in IDEA message are *Category* with a list of tags describing an event according to the taxonomy, *Source* and *Target* that provide identifiers of attacker and victims in an event, and *Node* where the information on the sending connectors are stored. *Source*, *Target*, and *Node* all contain potentially sensitive information, such as IP and email addresses, URLs, organization name, and software used for the detection.

3.2 Security Alert Sharing Platforms and Personal Data

From GDPR's perspective, it is clear that personal data are processed within the security alert sharing platforms. The definition of personal data is according the Art. 4 para. 1 "any information about identified or identifiable person." The same article states that person may be directly or indirectly identified by, for example, ID number, location data, personal specifics, or an online identifier. As was shown in the previous section, the data shared by security alert sharing platforms may contain IP addresses, domain names, URLs, email addresses, or even pieces of transferred content. The important is, that even though we know that in some cases, these data are not sufficient to identify a specific person, we must treat them as personal data every time. The reasons are that they can be combined with other data, and then they can directly or indirectly identify a specific person. This approach supported by the Article 29 working party [1] and also by the Court of Justice of the European Union case laws (Scarlet case No. C-70-10 [5] and Breyer case No. C-582/14 [6]). Now we know that at least some of the data shared by security alert sharing platforms should be considered personal data. The tricky part is that it is technically impossible to separate personal data from other share data. Therefore, according to the Article 29 working party, all the shared data should be treated as personal data. We have already described in our previous work [18] how to make sharing of personal data in security alert sharing platforms legal. Therefore, we will not discuss this topic anymore in this article, and we will focus only on the issue of DPIA.

4 DATA PROTECTION IMPACT ASSESSMENT

GDPR defines DPIA in its Art. 35. The controller shall carry out the Data Protection Impact Assessment (DPIA) when the processing is likely to result in a high risk to the rights and freedoms of natural persons. DPIA is a tool whose main purpose is to systematically analyze, identify, and minimize the risks of data processing for the rights of data subjects. In other words, it is a risk management tool that helps the controller to reach and demonstrate compliance with GDPR obligations. The main difference between classical risk management and DPIA is that assessing risks relates to a data subject's rights and not to the controller's. Now in the Article 29 working group guidelines, several criteria indicate that the processing of personal data may result in a high risk to the rights and freedoms of natural persons.

It follows from Article 35 that DPIA is only mandatory in certain cases and some of the lawyers would probably try to find the arguments why it is not necessary to conduct DPIA since it might appear to be a time-consuming exercise. That is not our case. As we stated in the intro, we conduct DPIA to distract the concerns arising from the possible violation of the GDPR by the processing of personal data using our security alert sharing platform. Also, DPIA not only allowed us to analyze the life cycle of the personal data in

[6] https://idea.cesnet.cz/en/index

the platform and identify possible weak spots but also serves as an effective tool to demonstrate compliance to the data subjects, data protection authorities and also users of the platform.

It is essential to point out that DPIA is not an individual activity to achieve GDPR compliance. On the contrary, GDPR requires a more general consideration of the privacy perspective. It is required to determine the purpose of data processing and its legal ground. This perspective was sufficiently covered in our previous work [18]. In this paper, we only focus on the DPIA process and its implications.

4.1 The DPIA Process in Brief

The process of DPIA is not strictly defined; the controller can develop its template or use template which was created by someone else. Since we wanted the template to meet our particular needs and context we developed our own, which is based on available templates published by various data protection authorities and EU institutions and simplified to meet our need for a quick and easy assessment tool. Our DPIA procedure consists of five steps.

In the first step, we try to get as much information about the assessed system, as possible. Firstly, we examine the system specification in detail. Secondly, we identify stakeholders within and outside of our institution who might provide relevant input about the functioning of the assessed information system, the data the system uses, and for which purposes. We often require the cooperation of the data protection officer, developers or suppliers, system operators, CSIRT team, users, management, and legal department.

The second step focuses on mapping and description of the assessed information system so we can understand how it works and handles personal data. In this step, we describe personal data life cycle within the system and conduct decomposition of the system into functions and components. The purpose of this is to analyze what functions and purposes are the personal data needed for, and which components are processing it, and then implement targeted controls that ensure data minimization and protection.

In the third step, we use the collected information about the system to identify the risks that may arise from the processing to the data subject and to evaluate these risks in terms of their probability of occurrence and their severity.

In the fourth step, we identify the controls that are already implemented in the system or which should be implemented in the future to reach an acceptable residual risk.

Finally, the fifth step consists of documentation of the whole process of assessment and planning of the review schedule. We usually expect the DPIA to be reviewed once a year to be able to identify new risks that may be caused by changes in the processing or the information system itself.

4.2 Getting Information About the Assessed System

We started our DPIA with the examination of the specification of the sharing platform system. In our case, we were lucky because the sharing platform has its requirements well documented, and therefore, we quickly got a basic understanding of the purpose and functioning of the system. Then, we focused on consultations with stakeholders within and outside of our institution. During the internal consultations, we interviewed platform administrators and developers. The external consultations concerned members of the sharing community and the National Cyber and Information Security Agency (NCISA) of the Czech Republic, where the sharing platform is operated. Members of the sharing community are data producers, i.e., operators of intrusion detection systems and other data sources, and data consumers, e.g., security incident response teams (CSIRT) and cybersecurity researchers. NCISA encourages cooperation within the cybersecurity community in the Czech Republic, which includes development and operations of alert sharing platform.

We suggest never to skip the consultations, even though if your system is well documented. The inputs from consultations may give you whole different points of view on risks and controls because every stakeholder has different expertise and interests.

4.3 Mapping and Description of the Assessed System

In this step, we conduct the decomposition of the sharing platform alert sharing platform to enumerate its functions and components. Subsequently, we describe information flows of security alerts in the system to infer life-cycle of personal data in the system. Please, recall Figure 1 for a scheme of the decomposed system.

The functional components identified in the sharing platform are as follows:

(1) Sending connectors are interfaces between data sources (intrusion detection systems, honeypots, etc.) and the sharing platform. They are operated by data producers within the community.
(2) Warden is a central hub for exchanging all the alerts operated by administrators of the sharing platform. The data there are available to all receiving connectors and analytical tools.
(3) NERD is a reputation database that provides additional, potentially sensitive information to the alerts and identifiers in them.
(4) iABU is an analytical component operated by platform administrators. The community may only access outputs of analyses (statistics, etc.) and newly generated data, such as alerts of predicted events or meta-alerts (compositions of multiple related alerts).
(5) Receiving connectors are operated by data consumers within the community and consume the shared alerts. Actions, such as blocking or blacklisting, may be executed by receiving connectors in peers' networks.
(6) Mentat is a special type of receiving connector that provides periodical reports of alerts to interested parties, not necessarily members of the community. Typical reports summarize alerts related to a receiver's constituency (IP range, domain name, etc.).

Each component runs on a dedicated host, sending and receiving connectors are distributed among the networks of community members. To secure communication over the network, all the connectors (including iABU and Mentat) must be registered, provide a valid certificate, and communicate from a registered domain name. Communication is encrypted.

The information flows, i.e., the life-cycle of the data goes as follows. The alerts are collected by sending connectors and sent to the central hub. As of early 2019, the central hub receives approximately 1.5 million alerts per day, which contain slightly less than 1.5 million IP addresses, thousands of hostnames, and a few URLs. Although email addresses and user identifiers can also be shared, there are currently no data producers of such data involved. Further, around 100,000 alerts contain attachments with additional data, such as system logs, network traffic samples (NetFlow, PCAP), and malware binaries. The IP addresses, in most cases, refer to attackers outside of constituency of sharing community. Target IP addresses are in some cases anonymized, namely when the target is a honeypot. The central hub is not persistent storage, but a queue of recent alerts; old alerts are dismissed approximately after one month.

The receiving connectors, including iABU and Mentat, download the alerts from Warden. The connectors may access all the alerts and keep them. However, filters are usually applied, and alert storage is discouraged due to the amount of data. Mentat is persistent storage and keeps the history of alerts for two months except for alerts, in which the event source ("attacker") is in the constituency of the community, e.g., network range of CESNET, national research and education network, that operates the platform. Such alerts are kept indefinitely, but they pose only 1-2 % of all the alerts.

Mentat uses NERD to enrich the data, i.e., the identifiers in alerts (IP addresses, hostnames, etc.) are enriched with contextual information, such as geolocation, hostname lookup, calculated reputation score based on the history of security alerts, presence on blacklists, scans by Shodan, etc. Further, iABU receives and processes the alerts, and may generate new ones if it predicts an event or folds a meta-alert of several alerts. Thus, it appears as sending and receiving connector simultaneously. However, no alerts are stored for analysis; only the statistics and other outputs.

4.4 Risks Identification and Evaluation

From the first two steps, we have enough information to identify the relevant risks for the data subjects. We describe every risk, and then we qualitatively evaluate the severity of its impact and likelihood of occurrence. The mechanism of evaluation is clear from the Table 1.

If the risk is high, there is a need to implement such measures, that ensure a sufficient reduction in the risk and a sufficient level of protection of rights and freedoms of data subjects, while preserving the possibility of implementing the intended processing. If such measures are unavailable, the controller should consider a different kind of processing with the same purpose. If the risk is a medium risk, the controller should consider the implementation of measures that strengthen the protection of data subjects rights and freedoms.

4.4.1 The Inability of Data Subject to Exercise Rights. We have found, based on our legal analysis, that personal data in the system can be processed for the purpose of protection of infrastructure even without the explicit consent of the data subjects [18]. The reason here is that the protection of infrastructures is legitimate interest not only of us as the operator of the sharing service, but also of the peers. Our analysis also suggests that processing implemented within the sharing platform causes very little interference with the rights and freedoms of the data subject, whereas it greatly improves the ability of the peers to protect their infrastructures. Therefore, we conclude that interference with the rights of the data subject is proportionate to the intended purpose of processing and the legitimate interest of the operator and peers.

However, we still need to provide the data subjects with the necessary information to allow them to exercise their rights derived from the GDPR. There are a couple of issues here. The sharing platform processes personal data collected by many different operators/peers, who are collecting various identifiers that may be related not only to their customers but also to various third parties. Also, these identifiers may be shared within the sharing platform to other peers who therefore become controllers.

Given nature and of the personal data that is processed within the sharing platform, we concluded, that the severity of the impact on rights and freedoms of the data subjects caused by their inability to exercise their rights related to such data is very low. The reason is that we collect and process only general identifiers, which contain only very little information about the data subjects. Likelihood of occurrence of such impact is also improbable since we provide information in compliance with Art. 14 of the GDPR and the data subject who learns about the processing of their personal data within the sharing platform, will be provided with sufficient information to exercise their rights under the Regulation.

4.4.2 Loss of Control over the Use of Personal Data, Loss of Confidentiality. Not all the data the sharing platform collects is considered personal data, in fact, most of it are just identifiers of networked devices, some of these identifiers could, however, be connected to a specific natural person, like attacker or user. Also, the system processes other data connected to these identifiers like hostnames, URLs, logs, flow data, etc., which are being used for the analysis of individual incidents and also as a big data resource for machine learning implemented in the sharing platform. Both identifiers and related incident data are provided by peers, who also determine the structure and level of detail of the data they choose to provide.

For the purpose of effective mitigation of incidents by peers, we need to share all of this data within the platform. Therefore all peers and their designated operators are allowed to access the data. That brings a risk of misuse of the data that is processed within the sharing platform.

The severity of possible impact caused by disclosure or misuse of the data should be minimal or very small since we are doing our best to store only data relevant to the incidents, which generally indicate nothing about the characteristics and the minimum about the behavior of the data subject. The same goes for the likelihood of occurrence, which is improbable. The data is not publicly available, logged access to it is provided only to authenticated operators, all the communication channels within the sharing platform are encrypted, and we also ensure deletion of legacy and unusable data within individual components of the system.

4.4.3 Inability to Access Services or Opportunities. Another risk for the data subjects is the inability to access services or opportunities due to restrictive measures taken by receiving peers to protect their networks and information systems. Technically, receiving peers may create blacklists and set firewall rules using the identifiers in the shared alerts. Further, the reputation database, which is a

Severity of impact	Serious harm	Low risk	High risk	High risk
	Some impact	Low risk	Medium risk	High risk
	Minimal impact	Low risk	Low risk	Low risk
		Improbable	Reasonably likely	More likely than not
		Likelihood of occurence		

Table 1: Mechanism of risk evaluation.

component of the sharing platform, uses machine learning technologies to analyze similarities and correlations in the collected data and calculate the score for each identifier [3]. Such a score is then provided to the peers. The issue here is that this kind of scoring may be considered profiling in the meaning of the GDPR, and also that based on the score, some peers may decide to limit or otherwise alter the access of the identifier to a specific resource. Technically speaking, a network entity may receive a score because it is located in a notorious network segment, autonomous system, or country. Using the score based on this information could lead to limitation of rights and freedom of natural persons.

According to our evaluation, the severity of impact and the likelihood of its occurrence is medium. It is very likely that the consumers of shared alerts and records in a reputation database will use this information to limit access to their resources for given entities. Peers in the community are using the alerts and the reputation database to create blacklists of IP addresses, set firewall rules, etc. However, such measures are taken only in the constituency of the peer, i.e., its network or information system, so that the assortment of services with limited access may be very low. Moreover, the restriction is typically applied in a limited time window, a few weeks at maximum, as the information are becoming obsolete, and the capacity of restrictive appliances is limited.

4.4.4 Other Economic or Social Disadvantage. The last risk which we have identified is a risk related to social or economic interests of the subject. The substance of this risk lies basically in damage to the reputation of a data subject.

Concerns about reputational damage are particularly relevant to information sharing in the decision-making process of data producers whether they will voluntarily proactively share information or not [16]. To natural persons, the risk for reputation is rather a hypothetical thought exercise than a real problem. Arguments based on the construction that the social or economic interests of the data subject were harmed due to the inclusion of its IP address or another identifier in a security alert or negative rating in a reputation database like the one deployed in the assessed platform do not sound very solid. More specifically, the potential of shared identifiers to damage the reputation of the data subject is very small, given their minimum information value. Finally, if such a situation occurs, the subject is not necessarily perceived negatively, but rather as a victim, as the subject's resources might have been abused for a cyber attack. However, certain subjects, e.g., financial institutions, may feel harmed even in such a situation. Also, they are not shared with the entire world but only with registered peers, so the likelihood of occurrence is generally low. To sum it up, we assess this risk as very low.

4.5 Controls Identification

The controls to mitigate the risks can be split into two groups, existing and newly proposed. The existing were mostly covered in previous sections. However, for the sake of completeness, they include encryption of the network communication, deleting the obsolete data from the system, and data access management and accounting. Further, the sharing platform applies compliance with requirements regarding transparency and information provided to the data subject, agreement or other legal document defining duties and responsibilities of peers, and non-disclosure agreements with the operators and user.

The controls proposed to mitigate the risks identified in the previous step goes as follows. First, the risk of the inability of data subject to exercise rights was identified as very low. From our perspective, it would be probably sufficient to publish the information on the processing as required in the Art. 14 of the GDPR and provide the link to the peers, we decided to mitigate the risk further and strengthen data subjects rights by making the information as accessible to the data subjects as possible. That's why we also asked all the peers to publish or otherwise make available the information to their users and customers. We also asked peers to identify in their privacy policies or other relevant legal documents security as one of the purposes for collection of personal data processed by the security alert sharing platform. This way, we achieved an acceptable level of residual risk.

The low risk was also identified in case of loss of control over the use of personal data and the loss of confidentiality. However, in this case, we decided to implement several measures to reduce the risk further. The first measure is that in the service policies, we will require the peers to allow access to the system only to operators with signed a non-disclosure agreement (NDA). The second is that we will develop guidelines for the peers, detailing what incident data is relevant and should be shared within the sharing platform. This would mitigate the risk of peer sharing too detailed personal data within the system. By implementing these measures, we achieved negligible residual risk.

To reduce the risk of inability to access services and opportunities, we propose setting the recommendations for using the alerts and reputation database. Users of the data should be aware of the origins of the data. For example, using the identifiers contained in alerts has a lower risk, because the alerts are generally trustworthy, and the entities were involved in security events. The data from reputation database, however, should not be used directly to restrict access to any services, but rather as a supporting feature in intrusion detection and similar use cases, where the reputation is combined with an observed activity of the entity.

Finally, the risk of other economic or social disadvantages was identified as very low and, thus, we did not propose any further controls in addition to those discussed above.

4.6 Documentation and Planning of Review Schedule

The documentation is crucial for the demonstration of compliance to the data subjects, data protection authorities, and also the users of the platform. We conduct DPIA to distract the concerns arising from the possible violation of the GDPR. Therefore, we will use the documentation as an input in our communication activities to encourage potential peers to join the platform and share their data.

The GDPR does not prescribe any obligatory form of DPIA process documentation. We perceive the following structure as suitable:

(1) Identification of purposes for which was DPIA conducted.
(2) Description of getting information about the assessed system – what documents were examined, which stakeholders were consulted, and what were the lessons learned.
(3) Description of the information system, information flows, and mechanism of personal data processing.
(4) Identification and evaluation of risks for rights and freedoms of data subjects.
(5) Identification of existing controls and newly proposed controls.
(6) Plan for the implementation of newly proposed controls.
(7) Conclusion of DPIA results and planning the review schedule.

Except for the plan for implementation of newly proposed controls, which is out of the scope of this article, everything else from the proposed structure of the documentation is covered in this article.

In relation to planning the review schedule, it is important to highlight the importance of precise set up of its process. Due to the dynamic nature of the cyber environment, we will in our case conduct the "smaller scale" DPIA yearly to identify and evaluate new risks and newly available controls. We also implemented a rule that after any significant changes in the system design or the mechanism of the personal data processing is necessary to conduct DPIA.

5 CONCLUSION

To conclude our paper, we discussed the issues GDPR brought into cybersecurity operations and software development, namely the automated exchange of information and knowledge. Cybersecurity community operates a plethora of tools and platforms to exchange alerts of security events and other information that may contain private data. GDPR's obligations triggered fear of information sharing among the community that we addressed in our work. Although preliminary work did not find substantial problems with information sharing for cybersecurity purposes[16, 18], the fears and doubts remained. We introduced an alert sharing platform the sharing platform as an example of alert sharing platform and conducted DPIA, thorough legal analyses, to dismiss any remaining concerns and answer frequent questions. The DPIA is not mandatory in our case; more simple legal measures and argumentation could have been taken. However, given the aspect of the unwillingness of the community to share, we believe that the work gave us strong arguments in favor of information sharing in cyber security under the novel legal framework.

Although we could say, with regards to the results of our analysis, that the hype around GDPR within the cybersecurity community was "much ado about nothing," we consider it important as it raised many interesting questions and implications. It helped us identify potentially sensitive information that we were working with daily, not knowing about its sensitivity. The information flows related to intrusion detection alert sharing were mapped, so now we are aware who does what with the data. Finally, alerts sharing platforms in many cases improved their own security. We may only argue if all the work would be done if it was not for GDPR; probably yes, but in the much longer time frame. We hope that our work addresses all the issues discussed in the community, so that the information sharing may be further conducted and developed.

ACKNOWLEDGMENTS

This research was supported by ERDF "CyberSecurity, CyberCrime and Critical Information Infrastructures Center of Excellence" (No. CZ.02.1.01/0.0/0.0/16_019/0000822).

REFERENCES

[1] Article 29 Working party. 2007. Opinion 4/2007 on the concept of personal data. http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2007/wp136_en.pdf.

[2] Article 29 Working party. 2017. Guidelines on Data Protection Impact Assessment (DPIA). https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611236.

[3] Václav Bartoš, Martin Žádník, Sheikh Mahbub Habib, and Emmanouil Vasilomanolakis. 2019. Network entity characterization and attack prediction. *Future Generation Computer Systems* 97 (2019), 674 – 686.

[4] Jaspreet Bhatia, Travis D. Breaux, Liora Friedberg, Hanan Hibshi, and Daniel Smullen. 2016. Privacy Risk in Cybersecurity Data Sharing. In *Proceedings of the 2016 ACM on Workshop on Information Sharing and Collaborative Security (WISCS '16)*. ACM, New York, NY, USA, 57–64.

[5] Court of Justice of the European Union. 2011. Judgement in Case C-70/10 Scarlet Extended SA v Société belge des auteurs, compositeurs et éditeurs SCRL (SABAM).

[6] Court of Justice of the European Union. 2016. Judgement in Case C-582/14 Patrick Breyer v Bundesrepublik Deutschland,.

[7] H. Debar, D. Curry, and B. Feinstein. 2007. The Intrusion Detection Message Exchange Format (IDMEF). RFC 4765 (Experimental). http://www.ietf.org/rfc/rfc4765.txt

[8] ENISA. 2011. A flair for sharing - encouraging information exchange between CERTs. https://www.enisa.europa.eu/publications/legal-information-sharing-1/at_download/fullReport

[9] ENISA. 2013. Detect, SHARE, Protect – Solutions for Improving Threat Data Exchange among CERTs. https://www.enisa.europa.eu/activities/cert/support/information-sharing/detect-share-protect-solutions-for-improving-threat-data-exchange-among-certs/at_download/fullReport.

[10] ENISA. 2014. Standards and tools for exchange and processing of actionable information. https://www.enisa.europa.eu/publications/standards-and-tools-for-exchange-and-processing-of-actionable-information/at_download/fullReport.

[11] Otto Hellwig, Gerald Quirchmayr, Walter Hötzendorfer, Christof Tschohl, Edith Huber, Franz Vock, Florian Nentwich, Bettina Pospisil, Matthias Gusenbauer, and Gregor Langner. 2018. A GDPR Compliance Module for Supporting the Exchange of Information Between CERTs. In *Proceedings of the 13th International Conference on Availability, Reliability and Security (ARES 2018)*. ACM, 3:1–3:7.

[12] Martin Husák and Jaroslav Kašpar. 2019. AIDA Framework: Real-Time Correlation and Prediction of Intrusion Detection Alerts. In *Proceedings of the 14th International Conference on Availability, Reliability and Security (ARES 2019)*. ACM.

[13] Pavel Kácha. 2013. IDEA: Designing the Data Model for Security Event Exchange. In *17th International Conference on Computers: Recent Advances in Computer Science*.

[14] Pavel Kácha. 2014. IDEA: Security Event Taxonomy Mapping. In *18th International Conference on Circuits, Systems, Communications and Computers*.

[15] Oscar Serrano, Luc Dandurand, and Sarah Brown. 2014. On the Design of a Cyber Security Data Sharing System. In *Proceedings of the 2014 ACM Workshop on Information Sharing & Collaborative Security (WISCS '14)*. ACM, New York, NY, USA, 61–69.

[16] Florian Skopik. 2017. *Collaborative cyber threat intelligence: detecting and responding to advanced cyber attacks at the national level.* CRC Press.

[17] Pavol Sokol, Jakub Míšek, and Martin Husák. 2017. Honeypots and honeynets: issues of privacy. *EURASIP Journal on Information Security* 2017, 1 (2017), 4.

[18] Václav Stupka, Martin Horák, and Martin Husák. 2017. Protection of personal data in security alert sharing platforms. In *Proceedings of the 12th International Conference on Availability, Reliability and Security*. ACM, Reggio Calabria, "65:1–65:8".

[19] Fredrik Valeur, Giovanni Vigna, Christopher Kruegel, and Richard A. Kemmerer. 2004. Comprehensive approach to intrusion detection alert correlation. *IEEE Transactions on Dependable and Secure Computing* 1, 3 (July 2004), 146–169.

[20] Emmanouil Vasilomanolakis, Shankar Karuppayah, Max Mühlhäuser, and Mathias Fischer. 2015. Taxonomy and Survey of Collaborative Intrusion Detection. *ACM Comput. Surv.* 47, 4, Article 55 (May 2015), 33 pages.