Sharing Cyber Threat Intelligence under the General Data Protection Regulation

Adham Albakri[1,2], Eerke Boiten[2] , Rogério De Lemos[1]

[1] School of Computing, University of Kent, UK
{a.albakri | r.delemos}@kent.ac.uk
[2] School of Computer Science and Informatics, De Montfort University, UK
Eerke.Boiten@dmu.ac.uk

Abstract.
Sharing Cyber Threat Intelligence (CTI) is a key strategy for improving cyber defense, but there are risks of breaching regulations and laws regarding privacy. With regulations such as the General Data Protection Regulation (GDPR) that are designed to protect citizens' data privacy, the managers of CTI datasets need clear guidance on how and when it is legal to share such information. This paper defines the impact that GDPR legal aspects may have on the sharing of CTI. In addition, we define adequate protection levels for sharing CTI to ensure compliance with the GDPR. We also present a model for evaluating the legal requirements for supporting decision making when sharing CTI, which also includes advice on the required protection level. Finally, we evaluate our model using use cases of sharing CTI datasets between entities.

Keywords: Cyber Threat Intelligence, Information Sharing, General Data Protection Regulation GDPR, Legal evaluation.

1 Introduction

Sharing Cyber Threat Intelligence (CTI) between organizations is a good strategy for building better cyber defence [1]. It assists organizations in understanding existing cyber attacks, and helps them to react against those attacks efficiently and quickly. However, CTI potentially contains sensitive and identifying information, such as IP addresses, email addresses and existing vulnerabilities [2]. Therefore, we should establish proper safeguards before sharing CTI datasets with others. When sharing CTI datasets, organizations must ensure conformance with legal and regulatory requirements, such as those required by the state and federal level in the US [3], the Japanese Personal Information Protection Act (PIPA) [4], and the General Data Protection Regulation (GDPR) [5]. In the specific context of organizations being part of critical national infrastructure, the EU NIS Directive [6] mandates some level of CTI sharing. It requires all EU member states to establish national Computer Security Incident Response Teams (CSIRT), as a single point of contact, to report cyber incidents that affect critical infra-

structure and essential services. This is supported by the European Network and Information Security Agency (ENISA), which improves CSIRT capabilities by providing tools and methodologies to support network and information security [7].

In this paper, we investigate the legal aspects for sharing CTI datasets in the context of the GDPR [2] which is the principle law in the EU for regulating the processing of personal data in the EU. Personal data is defined as "any information relating to an identified or identifiable natural person ('data subject'). An identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person" (Art. 4(1), GDPR).

In this paper, we will present an approach for defining the required protection level on CTI datasets, if they contain personal data, as defined by the GDPR. Based on the GDPR rules, this approach would help to make the decision of sharing and processing personal information clear. Moreover, it helps to provide some practical and clear rules to build data sharing agreements between organizations, because during the evaluation phase, we establish the purpose of the sharing, the legal basis and security measures for compliance with the law. This paper has two main contributions. First, to provide a decision process about sharing CTI datasets containing personal data in the context of the GDPR. Second, to convert existing legal grounds into rules that help organizations share such data whilst being legally compliant with the GDPR. These rules establish an association between the CTI policy space and the defined protection levels.

The remainder of this paper is organized as follows. Section 2 describes the steps of the methodology to build the approach. Section 3 gives several use cases of sharing CTI datasets to validate our approach. Section 4 discusses related work and finally section 5 presents the conclusion and future research directions.

2 Methodology

This section presents the methodology we used to build an approach to evaluate the possibility of sharing personal data in the context of CTI datasets under the GDPR. The methodology consists of three main steps and is inspired by the DataTags project [8]. The first step is to define the possible levels of security requirements which agree with the principles considered by the GDPR when processing personal data in CTI datasets. The second step is to identify a policy space, i.e. a set of concepts, definitions, assertions and rules around the GDPR to describe the possible requirements for sharing CTI datasets. The last step is to build the decision graph, which defines the sequence of questions that should be traversed to establish and assess the legal requirements for CTI data sharing, represented with an outcome as so-called "tags". The DataTags project, developed by Latanya Sweeney's group at Harvard University, helps researchers and institutions to share their data with guarantees that releases of the data comply with the associated policy, including American health and educational legislation [9]. It consists of labelling a dataset with a specific tag based on a series of questions. Each question is created based on a set of assertions under the applicable policy.

2.1 Defining DataTags related to cybersecurity information sharing

The first step to achieving our goal is to define the tags that will be the possible decisions reached after a series of questions that interrogate CTI datasets for GDPR requirements. The legal requirements of the GDPR indicate in the first instance whether we can share or not. However, when the answer is positive, additional obligations for such sharing arise out of the principles and articles of the GDPR, in particular: the principle of data minimization; the requirement that personal data must be processed securely; and that the data must not be retained when no longer relevant. Hence, the decision process also leads to conclusions on how sharing can take place by translating these constraints into technical requirements. All of this is represented in the "data tags" of the leaves of our decision graph. The organizations that are sharing CTI datasets should ensure that the receiving organization understands the sensitivity of this information and receives clear instructions on what they are allowed to do with the information, e.g. potential on-sharing. We will follow the Traffic Light Protocol (TLP) [10] levels as a springboard, and expand them by adding security measures for each level in order to address the GDPR requirements of processing personal data when sharing CTI datasets. TLP was created to facilitate the sharing of information by tagging the information with a specific color. TLP has four colors, indicating different levels of acceptable distribution of data, namely [10]:

- WHITE - Unlimited.
- GREEN - Community Wide.
- AMBER - Limited Distribution.
- RED - Personal for Named Recipients Only.

This protocol records whether recipients may share this information with others. We have extended this protocol by adding appropriate security measures that are required for the legality of CTI sharing. To increase the trustworthiness between the entities and encourage entities to share CTI, we require the receiving organization to apply these security measures whilst keeping in mind that, in general, organizations use different approaches and levels of security practices. However, enforcing the receiver to apply these security measure is a challenge in itself and is beyond the scope of this paper. Table 1 shows the levels that we are going to use in order to label the shared datasets. Cells in columns "Tag type", "Description", and "Examples" are taken from the TLP description [11]. The values in columns "Security Measures" and "Transfer/Storage" are our proposals to meet the legislative requirements for securely sharing this data. We have proposed technical methods that would help organizations to achieve what the GDPR mandates as a technical requirements to ensure confidentiality and protect data subjects (Article 32). When proposing the security measures, we had to take into consideration with whom we are going to share CTI datasets and their trustworthiness because recipients who cannot be relied upon to protect the shared information need to be eliminated from further sharing.

We combine the notion of privacy preservation of the data with the trust level of the recipient organization, and because of that, we recommend the use of the Attribute-

based Encryption (ABE) technique [12] [13]. For encryption, ABE can use any combination of a set of attributes as a public encryption key. Decryption privileges of the data in this type of encryption are not restricted to a particular identity but to entities with a set of attributes which may represent items such as business type and location. For example, an organization chooses to grant access to an encrypted log of its internet traffic, but restricts this to a specific range of IP addresses. Traditional encryption techniques would automatically disclose the log file in case the secret decryption key is released.

Table 1 lists example values of some attributes in the data. The first attribute is the location of the organization. Due to the different legal systems associated with international transfer information exchange, we will consider three levels: National, EU and International. The second attribute is the sector of the organization, because of the similarity of the working processes and procedures and likely similar threat models. The value might contain energy, health, education, finance and so on. Finally, the size of the organization may be relevant because the number of employees has been empirically related to the number of threats [14]. To use ABE, before sharing the data with other organizations and in case it is not shared to the public, the Setup Key Authority generates a master secret key along with a public key. It publishes the public key so everyone has access to it. The key authority uses the master secret key to generate a specific secret key for the participating organization in the sharing community. For example, there might be an organization called "Alpha" which gets a specific secret key from the key generator authority. "Alpha" is an organization operating at the national level in the telecom sector. Before sharing any dataset with "Alpha", the user will encrypt the dataset that has its own specific access policy. Hence, this user encrypts the dataset such that anyone at the national level working with the telecom business will be able to decrypt it. The organization sharing CTI datasets generates ciphertext with this policy. As a result, the organization "Alpha" will be able to decrypt the dataset.

Table 1. ABE attribute

Attribute	**Value**
Location	National, EU, Global
Organization Sector / Similarity of Business	Central Authority, similar business, connected groups, …
Organization Size	Small, Medium, Big

At all levels, Green, Amber and Red, data will be encrypted using the ABE method. In addition, we need to consider the data minimisation principle as defined in GDPR Art. 5(1)(c) "1. Personal data shall be: (c) adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed (data minimisation)". Hence, sharing should be designed to provide only the required data to successfully achieve a specific goal. This implies that we should use the minimum amount of identifiable information to decrease any privacy risk on individuals whose personal data might be included. Doing so will reduce the risks of the following potential privacy attacks on the data:

Identity disclosure [15] [16]: this threat occurs when the attacker is able to connect a data subject with their record in a CTI dataset. For example, an attacker might identify

a victim because the dataset contains direct identifying information such as an email address, IP address or credential information.

Membership disclosure [17]: this threat occurs when an attacker can derive that a specific data subject exists in the dataset. For example, the dataset contains information about specific malware victims. Any person established to be in the dataset reveals that this victim has been hacked by this malware.

Attribute disclosure [18]: This threat occurs when data subjects are linked with information about their sensitive attributes such as biometric data that is used to uniquely identify an individual. Some personal information is more sensitive and defined as a special category under the GDPR. The GDPR (Art. 9) defines special categories that need extra protection and prohibits processing this type of data unless certain conditions are applied.

There are methods to remove personal information from an individual's record in a way that decreases the possibility of all these attacks. Some of these methods that we can use are k-anonymity [15] which uses suppression and generalization as the main techniques, l-diversity [19] [18] which is an extension of k-anonymity to protect the shared data against background knowledge and Homogeneity Attacks, and t-closeness [20] which is another extension of l-diversity that decreases the granularity and makes the distribution of the sensitive attribute close to the distribution of the entire attribute.

Table 2. Proposed DataTags relating to four proposed classes of access

Type	**Description**	**Examples**	**Security Measures**	**Transfer / Storage**
WHITE	Information does not contain any personal data or sensitive information so it can be shared publicly.	Sharing public reports and notifications that give a better understanding of existing vulnerability.	Anonymization (Identity disclosure, Membership disclosure, Attribute disclosure).	Clear
GREEN	Information shared with community or a group of organizations but not shared publicly.	Sharing cybersecurity information within a close community. For example, sharing e-mail with malware link targeting specific sector.	Anonymization (Identity disclosure) Attribute-Based Encryption (ABE)	Encrypted
AMBER	Share information with a specific organization; sharing confined within the organization to take effective action based on it.	Sharing cybersecurity information that contains indicators of compromise, course of action to a specific community or sector e.g. financial sector.	Anonymization (Identity disclosure) Attribute-Based Encryption (ABE)	Encrypted
RED	Information exclusively and directly given to Central Authority. Sharing outside is not legitimate.	Sharing that you have been attacked or notifying central authority about an incident.	Attribute-Based Encryption (ABE). Data Minimization to share only relevant data.	Encrypted

2.2 Policy Space

We build the policy space of our model as a set of assertions using the context of the CTI dataset. The evaluation of cases will be based on the defined assertions. The assertions will contain the legal grounds under which personal data can be processed, in this case for the purpose of ensuring network and information security. For instance, assertions for sharing CTI information with other parties are based on both the purpose of sharing which is "GDPR Recital 49 - ensuring network and information security" such as the prevention of any access to the critical system after credentials leaks, and the related legal basis which is "GDPR Art 6.1(c) - processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party". These steps offer a clear, practical framework, justifying the sharing of cyber threat Intelligence. The tagged data which meets the rules based on applicable assertions will be derived from the decision graph. In order to build the CTI policy space, we use a JSON file maintained by Computer Incident Response Center Luxembourg CIRCL [21] for the related context of use of data by CSIRTs. The goal of the file is to track processing personal information activities and support automation. Many assertions refer to the GDPR Art.30 which prescribes all the recordable details of processing activities. The main categories of the assertions contain:

- Purpose: "The purpose of the processing. Ref GDPR Art. 30(1)(b)"
- Legal ground: "Lawfulness/grounds for the processing activity. Ref GDPR Art. 6 & 5(a)."
- Data subjects: "Categories of the data subjects. Reference GDPR Art. 30 (1)(c)."
- Personal data: "Personal data processed. Reference GDPR Art. 30 (1)(c)."
- Recipients: GDPR Art. 30 (1)(d).
- International transfer: "Whether any personal data in this processing activity is transferred to a third country or an international organization. Reference GDPR Art. 30(1)(e)."
- Retention period: "Retention schedule/storage limitation. Reference GDPR Art. 30(1)(f) and Art. 5(e)."
- Security measures: "Security measures & Integrity & Confidentiality. Security measures can be technical and/or organizational. Reference GDPR Art. 30(1)(e), 32(1) and Art. 5(f)."

Based on the previous assertion list, we need to extract the relevant assertions categories specifically related to CTI sharing. We will consider only those assertions that are directly related to CTI sharing. In the GDPR the purpose of processing personal data should be precise and for that the GDPR offers clear recognition of "ensuring network and information security" GDPR Recital 49 as the purpose of processing personal data for actors such as public authorities and CSIRTs. The legal grounds for processing personal data are provided in GDPR Art. 6 & 5 (a). CIRCL has published a discussion [22] of the legal grounds of information leak analysis and the GDPR context of collection, analysis and sharing information leaks. The legal grounds relevant in our context are "processing is necessary for the compliance with a legal obligation to which the con-

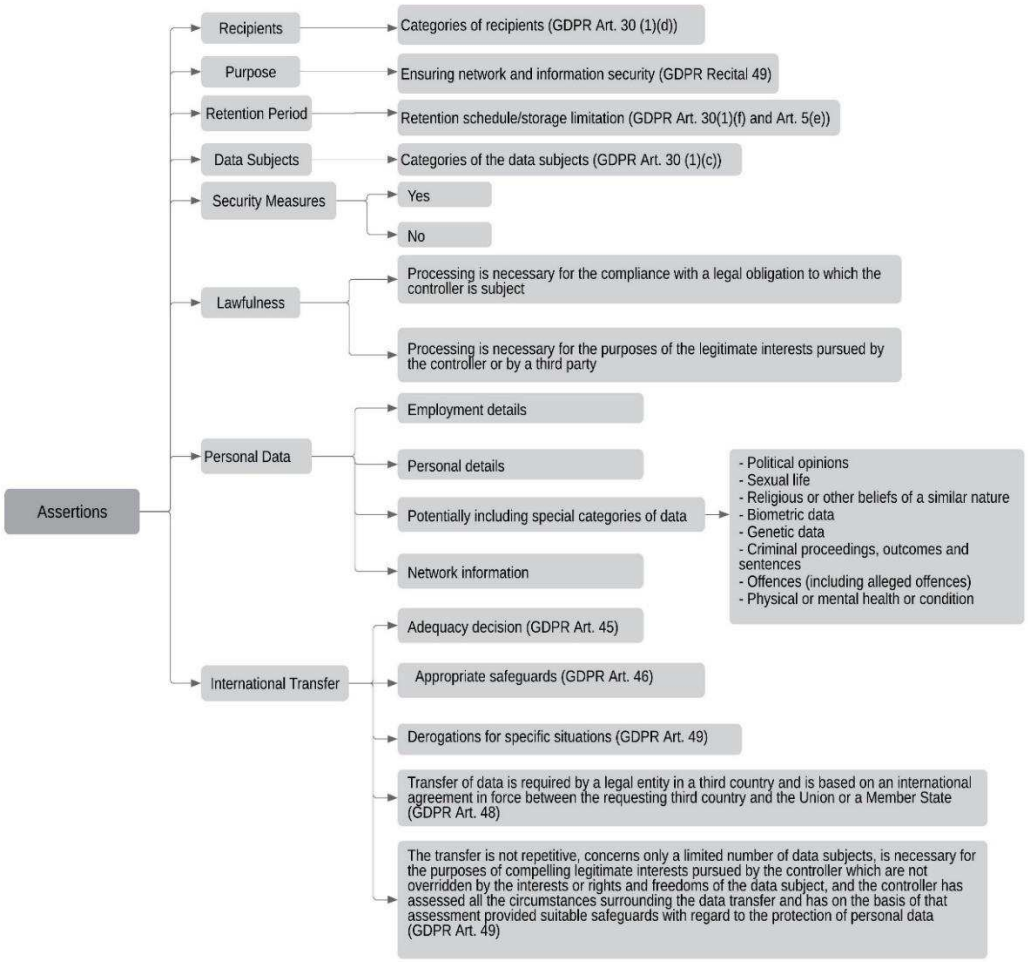
troller is subject" where it applies to CSIRTs and data protection authorities and "processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party" otherwise. In the "legitimate interest" sharing CTI information will enable organizations to better detect and prevent attacks by, for example, identifying the IP address of a malware communications and control hub. We do not consider "consent" GDPR Art. 6 (1)(a) a credible legal basis for processing personal data in the context of sharing cyber threat Intelligence. This is because it is very hard to get consent of data subjects especially when dealing with huge amounts of data [22] (e.g. 1bn Yahoo accounts were compromised from a 2013 hack [23]) or when personal data such as IP addresses concerns the perpetrator of a cyber-attack. Also, the vital interest Art.6 (1) (d) is not feasible to be used to justify sharing and processing CTI. The rationale is most likely there is no personal data in CTI datasets which would relate to a threat to life. However, the public interest Art.6(1)(e) would be the justification to process personal data in the case of acting under specific authorization from an official authority to check that the cyber incident could affect the public interest. The description of the personal data that pertains directly to the GDPR is described in Art.30 (1) (c). The conditions under which personal data can be transferred to third countries or an international organization are described in GDPR Art. 30(1)(e). As a result, the CTI policy space is described in Fig. 1.

2.3 Decision graph

In this step, we propose an assessment based on the previous assertions. This assessment contains a set of questions, and the answer to each question will lead to different questions or a final decision and as a result, we will assign a specific tag to the CTI dataset or even in some cases, the decision would be to not share. This assessment is not definitive, but it gives a chance to reflect on our understanding of sharing CTI datasets under the GDPR. Fig 2 shows the decision graph for sharing CTI datasets under the GDPR. Some of the decisions in the graph still require human judgement, so we make no claims of the process being fully automatable.

The process first establishes whether the proposed data sharing falls within the scope of the GDPR. Then it establishes the legal basis for any special category data included. This is likely to be rare in CTI datasets, but we could imagine biometric data following an attack that included a physical breach. Next, it establishes the legal basis for the overall processing. Then, it checks and selects appropriate retention and security protections. We assume the "trust level" node's result has been determined based on previous knowledge of the trustworthiness of the entity that we are sharing with. The outcome matches one of the TLP tags as described in the previous section. Of course, the CTI datasets are also likely to contain "sensitive" information about the infected asset and the exploitable vulnerability that should be protected. The outcome reflects concerns for the data protection angle only; included information that is sensitive in a different dimension might require strengthening of the security measures.

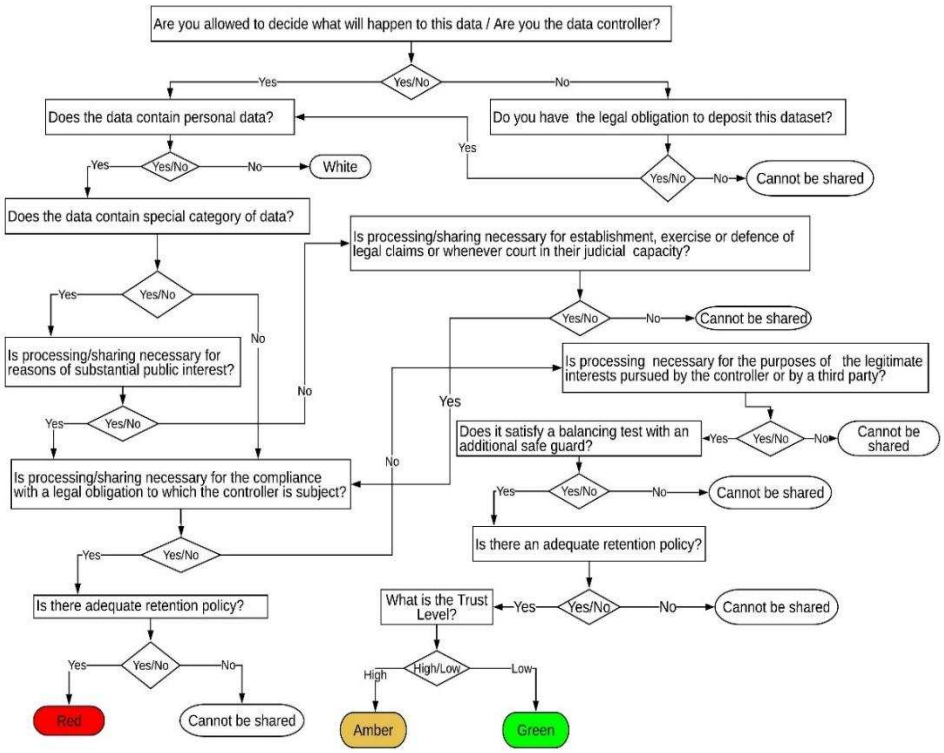
Fig. 1. CTI Policy Space



3 Use cases

Sharing information regarding current or ongoing attacks including information on threat actors, attack vectors, victims and impact of the attack is an essential scenario of sharing cyber threat Intelligence. In order to see how to apply the tags on CTI datasets two different use cases were developed. In the first use case, the organization that is the victim informs a central authority about the attack. In the second use case, an organization informs another organization about a recent attack that affects the availably, confidentiality or integrity of services.

Fig. 2. Decision Graph



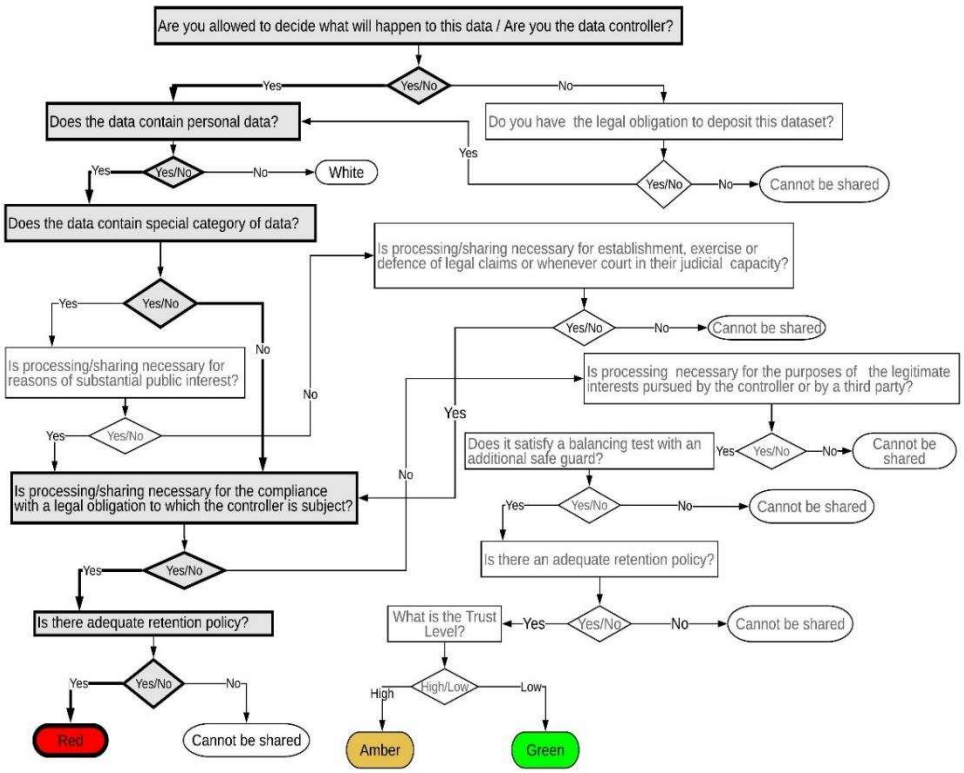
Use Case 1: informing central authority

This case study consists of two organizations, A and C (Central Authority) where an organization A wants to report an incident to organization C about a remote access tool (RAT) used by different threat actors. Before sharing the information, the reporter wants to be sure that sharing it is legitimate under the GDPR.

The incident report contains personal information such as contact information of the reporter and credential information. Therefore, sharing and processing of such personal data would need to be legitimate under the GDPR. In order to decide how to share this information, the reporter needs to run an evaluation. The organization A is the owner of this dataset and has the right to process this information, hence in this scenario the organization A is considered the controller. Although the incident information contains personal data, it does not contain any special category data, such as, biometrics or political opinion, religious or philosophical beliefs, etc. In order to share this information with a Computer Security Incident Reporting Team CSIRT or the central authority, the reporter can rely on GDPR Art. 6(1)(c) where the legal ground states “processing is necessary for the compliance with a legal obligation to which the controller is subject”. Organization A has a retention policy in place. The security measures that should be

applied to reduce the risk of harm to data subjects before sharing this dataset are: encrypted storage associated with a secure protocol to transmit this information. Moreover, the data will be encrypted by using ABE techniques with the properties (National, CA, Big) so as a result the final tag for this data will be RED. Fig. 3 shows a sample questionnaire covering this case study.

Fig. 3. Use case 1 Assessment Graph



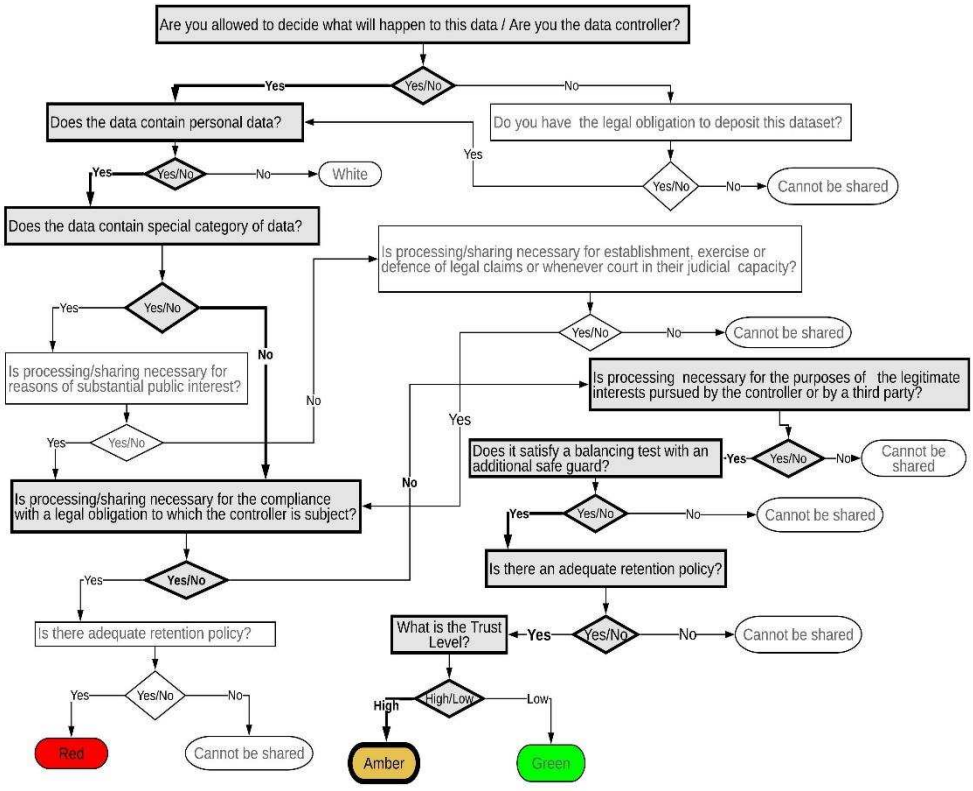
Use Case 2: Sharing information about port scanning for incident prevention.

Suppose an organization O1 in the energy sector detects port scanning from a specific IP address for port range 0-1023 which is considered a potential threat. For incident prevention purposes, they may want to share information containing the source IP address, port range, the time of the incident, signs of the incident, and the course of action such as improve monitoring on these ports.

The personal information in this scenario consists of the reporter information along with that of the individual who has made the observations. Organization O1 is the controller of this data and needs to share this information with trusted company O2. Because the dataset contains personal information, sharing needs to be legitimate under the GDPR. The dataset does not contain any special category data so we can continue

and check the purpose of this sharing, which is the GDPR Recital 49 – "ensuring network and information security". The reporter can rely on the GDPR Art. 6(1)(f). The legal ground for sharing this information is "processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party". Presumably there is a retention policy in place. The security measures that will be associated before sharing this dataset are: encrypted storage associated with a secure protocol to transmit this information, Anonymization against any Identity disclosure and the data will be encrypted by using ABE techniques associated with the properties (EU, Energy sector, Medium). The trust level based on an assumed external calculation is high so as a result the final tag for this data will be AMBER. Fig. 4 shows a sample questionnaire covering this case study.

Fig. 4. Use case 2 Assessment Graph



As a result, we present two use cases for sharing CTI datasets between different entities. The datasets have been evaluated based on the decision graph built in section 2. The decision is positive in both use cases, but it associated with different protection

levels based on the flow of the assertions. Hence, our approach can give any organization intends to share CTI datasets the ability to determine that they are legally compliant with the GDPR.

4 Related Work

Many papers have addressed issues related to terms and rules extracted from regulations and policies for protecting personal data. K. Fatema, Chadwick, and Van Alsenoy [24] converted the precursor of the GDPR, the 1995 EU Data Protection Directive [25] into executable rules to support access control policies. The authors presented a system to automate legal access control policy to make automated decision concerning authorization rights and obligations based on the related legal requirements. Doorn and Thomas [26]developed a specialized tool for privacy control based on the GDPR to share sensitive research datasets. The authors defined the security measures of the data tags levels based on the DANS EASY repository [27]. The authors focused on datasets managed by researchers in a general context. Breaux and Antón [28] [29] worked to extract data access rights from a legal test of the US Health Insurance Portability and Accountability Act (HIPAA). They used ontology to classify legal rules of privacy requirements from regulations to give a decision to grant or deny the access right. In [30] Schweighofer, Kieseberg and Kieseberg, a privacy by design solution to exchanging cyber security incident information between CSIRTs is presented. This solution focused only on sharing information between closed user circles such as the CSIRTs. The authors aimed to illustrate the legal requirements about sharing CTI datasets which contain personal information between the CSIRTs without giving a systematic way to help the CTI datasets manager to check the legality of sharing such information. In our work, we aim to build a set of sharing requirements that CTI datasets managers will check to provide a decision about sharing CTI dataset(s) under the GDPR.

5 Conclusion and future work

In this work, we have presented an approach that can help different entities to make a decision compliant with the GDPR when sharing CTI datasets. We have suggested adequate privacy preserving methods that should be applied when sharing CTI datasets. Then we have defined the policy space that related to the CTI in the context of the GDPR and finally built the decision graph that checks the legal requirements and provides a decision on how to share this information.

There are limitations in our approach. In complex use cases, the decisions in the assessment graph may still be very demanding, such as whether the Recital 49 objective justifies any privacy impacts on the data subject. Furthermore, including additional regulations or local policies besides the way they will interact with the GDPR requirements would make the decision graph more complex. Additional legal and technical requirements might make the data tag collection harder to structure and manage, as well as complicating the decision process.

In our previous work [2], we have identified the associated threats of disclosing CTI. Here we have specifically addressed the legal risks associated with sharing CTI datasets. Our overall work aims to mitigate all threats associated with sharing CTI datasets and improve the sharing process. As future work, we will extend the current model to evaluate the trust level and the associated risks in more detail. In addition, we intend to study the tradeoff between the privacy preservation and utility of processing CTI datasets.

6 Acknowledgment

This work has received funding from the European Union Framework Programme for Research and Innovation Horizon 2020 under grant agreement No 675320.

7 References

[1] F. Skopik, G. Settanni, and R. Fiedler, "A problem shared is a problem halved: A survey on the dimensions of collective cyber defense through security information sharing," in Computers and Security, 2016, vol. 60, pp. 154–176.

[2] A. Albakri, E. Boiten, and R. De Lemos, "Risks of Sharing Cyber Incident Information," in Proceedings of the 13th International Conference on Availability, Reliability and Security - ARES 2018, 2018, pp. 1–10.

[3] Latanya Sweeney, "Operationalizing American Jurisprudence for Data Sharing," Technical report, 2013.

[4] J. Personal Information Protection Commission, "Amended Act on the Protection of Personal Information," 2016. [Online]. Available: https://www.ppc.go.jp/files/pdf/Act_on_the_Protection_of_Personal_Information.pdf. [Accessed: 03-Jan-2019].

[5] European Union, "Regulation 2016/679 of the European parliament and the Council of the European Union of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/," Official Journal of the European Communities, vol. 59, no. May. pp. 1–88, 2016.

[6] ENISA, "Directive on security of network and information systems," 2017. [Online]. Available: https://ec.europa.eu/digital-single-market/en/network-and-information-security-nis-directive. [Accessed: 13-Dec-2018].

[7] ENISA, "information Security Agency,'A step-by-step approach on how to set up a csirt,' no," WP2006/5.1, 86 (2006).

[8] M. Bar-Sinai, L. Sweeney, and M. Crosas, "DataTags, Data Handling Policy Spaces and the Tags Language," in Proceedings - 2016 IEEE Symposium on Security and Privacy Workshops, SPW 2016, 2016, pp. 1–8.

[9] Sweeney L, Crosas M, Bar-Sinai M. Sharing Sensitive Data with Confidence: The Datatags System. Technology Science. 2015101601. October 16, 2015. https://techscience.org/a/2015101601

[10] I. FIRST.ORG, "Traffic Light Protocol (TLP)," 2001. [Online]. Available: https://www.first.org/tlp/. [Accessed: 14-Aug-2018].

[11] CIRCL, "Traffic Light Protocol (TLP) - Classification and Sharing of Sensitive Information," 2018. [Online]. Available: https://www.circl.lu/pub/traffic-light-protocol/. [Accessed: 29-Sep-2018].

[12] V. Goyal, O. Pandey, A. Sahai, and B. Waters, “Attribute-based encryption for fine-grained access control of encrypted data,” in Proceedings of the 13th ACM conference on Computer and communications security - CCS ’06, 2006, pp. 89–98.
[13] J. Bethencourt, A. Sahai, and B. Waters, “Ciphertext-policy attribute-based encryption,” in Proceedings - IEEE Symposium on Security and Privacy, 2007, pp. 321–334.
[14] M. E. Johnson, “Information Risk of Inadvertent Disclosure: An Analysis of File-Sharing Risk in the Financial Supply Chain,” J. Manag. Inf. Syst., vol. 25, no. 2, pp. 97–124, 2008.
[15] L. Sweeney, “k-ANONYMITY: A MODEL FOR PROTECTING PRIVACY,” Int. J. Uncertainty, Fuzziness Knowledge-Based Syst., vol. 10, no. 05, pp. 557–570, 2002.
[16] X. Xiao and Y. Tao, “Personalized privacy preservation,” in Proceedings of the 2006 ACM SIGMOD international conference on Management of data - SIGMOD ’06, 2006, pp. 229–240.
[17] M. E. Nergiz, M. Atzori, and C. Clifton, “Hiding the presence of individuals from shared databases,” in Proceedings of the 2007 ACM SIGMOD international conference on Management of data - SIGMOD ’07, 2007, pp. 665–676.
[18] D. Kifer, “l-Diversity : Privacy Beyond k -Anonymity,” in Proceedings of the 22nd International Conference on Data Engineering, 2006, pp. 24–36.
[19] A. Machanavajjhala, J. Gehrke, D. Kifer, and M. Venkitasubramaniam, “ℓ-Diversity: Privacy beyond k-anonymity,” in Proceedings - International Conference on Data Engineering, 2006, p. 24.
[20] L. Ninghui, L. Tiancheng, and S. Venkatasubramanian, “t-Closeness: Privacy beyond k-anonymity and ℓ-diversity,” Proc. - Int. Conf. Data Eng., no. 2, pp. 106–115, 2007.
[21] CIRCL, “Legal compliance and CSIRT activities,” 2018. [Online]. Available: https://github.com/CIRCL/compliance. [Accessed: 29-Sep-2018].
[22] CIRCL, “AIL information leaks analysis and the GDPR in the context of collection, analysis and sharing information leaks,” 2018. [Online]. Available: https://www.circl.lu/assets/files/information-leaks-analysis-and-gdpr.pdf. [Accessed: 20-Dec-2018].
[23] S. Thielman, “Yahoo hack: 1bn accounts compromised by biggest data breach in history,” The Guardian (UK), 2016. [Online]. Available: https://www.theguardian.com/technology/2016/dec/14/yahoo-hack-security-of-one-billion-accounts-breached. [Accessed: 24-Oct-2018].
[24] K. Fatema, D. W. Chadwick, and B. Van Alsenoy, “Extracting access control and conflict resolution policies from European data protection law,” in IFIP Advances in Information and Communication Technology, 2012, vol. 375 AICT, pp. 59–72.
[25] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. .
[26] P. Doorn and E. Thomas, “Tagging Privacy-Sensitive Data According to the New European Privacy Legislation: GDPR DataTags - a Prototype,” 2017. [Online]. Available: https://dans.knaw.nl/en/current/first-gdpr-datatags-results-presented-in-workshop. [Accessed: 20-Dec-2018].
[27] H. Tjalsma, “DANS Data Archiving and Networked Services,” 2012. [Online]. Available: https://easy.dans.knaw.nl/. [Accessed: 24-Oct-2018].
[28] T. D. Breaux and A. I. Antón, “Analyzing regulatory rules for privacy and security requirements,” IEEE Trans. Softw. Eng., vol. 34, no. 1, pp. 5–20, 2008.
[29] T. D. Breaux and A. I. Antón, “A Systematic Method for Acquiring Regulatory Requirements: A Frame-Based Approach,” in RHAS-6, 2007.
[30] E. Schweighofer, V. Heussler, and P. Kieseberg, “Privacy by design data exchange between CSIRTs,” in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2017, vol. 10518 LNCS, pp. 104–119.