Information & Security

J. Rajamäki & V. Katos
vol. 46, no. 2 (2020): 198-214
https://doi.org/10.11610/isij.4614

Published since 1998

ISSN 0861-5160 (print), ISSN 1314-2119 (online)

Research Article

Information Sharing Models for Early Warning Systems of Cybersecurity Intelligence

Jyri Rajamäki [a] (✉), *Vasilis Katos* [b]

[a] *Laurea University of Applied Sciences, Finland, https://www.laurea.fi/en/*

[b] *Bournemouth University, UK, https://www.bournemouth.ac.uk/*

ABSTRACT:

An Early Warning System (EWS) for cybersecurity intelligence will provide the capability to share information to provide up to date information to all constituents involved in the EWS. The development of EWSs will be rooted in a comprehensive review of information sharing and trust models from within the cyber domain as well as models from other domains. This article is the result of a qualitative multiple-case study analysis. It consists of theory development by systematic reviews of academic articles, seven case studies, and cross-case conclusions, from which a set of system requirements and features were established to support a model that promotes information sharing among partners, while also meeting regulatory requirements. Moreover, the final analysis includes the requirements for information sharing within and between partners across organisational boundaries as derived from multi-sector analysis. The study consists of a comprehensive review of information sharing and trust models from within the cyber domain (n > 50), as well as models from other domains, such as healthcare, maritime and critical infrastructure protection.

ARTICLE INFO:

RECEIVED: 08 MAY 2020
REVISED: 21 JULY 2020
ONLINE: 02 SEP 2020

KEYWORDS:

early warning system, ECHO, Information sharing, information sharing models, trust models, cybersecurity, case study

Introduction

An Early Warning System (EWS) for cyber intelligence aims at serving as a security operations support tool enabling the members of the network to coordinate and share information in near real-time. With EWS, stakeholders can retain

✉ E-mail: jyri.rajamaki@laurea.fi

their fully independent management of cyber-sensitive intelligence and related data management. EWS will work as a parallel part of other mechanisms in smart society. The development of EWS will be rooted in a comprehensive review of information sharing and trust models from within the cyber domain.

Fig. 1 shows how multiple case study research (MCSR) is applied in the creation of this study. The initial step in designing MCSR consists of theory development, and the next steps are case selection and definition of specific measures in the design and data collection process. Each individual case study consists of a whole study, and then conclusions of each case are considered to be the replication by other individual cases. The individual cases as well as the multiple result should be the focus of a summary report. For each individual case, the report should indicate how and why a particular result is demonstrated. Across cases, the report should present the extent of replication logic, including certain and contrasting results.[1]

According to Yin,[1] any use of multiple case design should follow a replication, not a sampling logic, and choosing of each case should be made carefully. In Fig. 1, the dashed-line feedback represents a discovery situation, where one of the cases does not suit the original multiple-case study design. This kind of a discovery stands for a need to reconsider the original theoretical foundations. This means redesign should take place before proceeding further, and in this view the replication approach represents a way of generalising that uses a type of test called falsification or refutation, which is the possibility that a theory or hypothesis may be proven wrong or falsified.[2]

The sources of evidence used in the individual case studies consist of documentation, archival records, interview, direct observations, participant-observation, and physical artefacts. From these, two to four multiple sources of evidence were used in every individual case study. Every individual case study has been reported separately as a conference paper and/or via ECHO SharePoint.

Cross-case conclusion were made via a document analysis exercise of the preceding sections and a selection of literature sources. The final analysis includes the requirements for information sharing within and between partners across organisational boundaries as derived from multi-sector analysis.

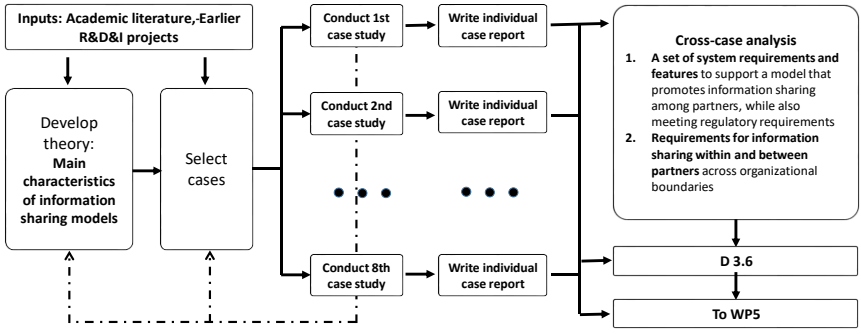

Figure 1: Multiple-case study method.

Theory Development

The first step in MCSR is the development of a rich theoretical framework that needs to state the conditions under which a particular phenomenon is likely to be found.[1] The theoretical framework of this study forms by way of a systematic review. Collected and analysed materials consist of scientific literatures, research articles and official publications. The research question of the literature review was "What are the main characteristics/features of cyber information sharing and trust models?" In order to capture a reasonably full range of the literature concerning the main features of cyber exchange models, following scientific data-bases have been used: Database of the JYKDOK library at the University of Jyväskylä (wide database concerning cybersecurity and it provides access e.g. to the IEEE Xplore). The IEEE Xplore library (provides web access to more than 4.5 million documents from publications in computer science and ca. 200 journals and ca. 1 700 conference proceedings), Springer link (Database area of engineering contains 17 000 books) and AI tool called IRIS which search engine based on 100 entered keywords. Also, several studies were based on public sources. Ana-lysed information sharing and trust models are listed in Annex 1. The qualitative analysis was made by using traditional half-manual processing and Glue (Orange3) Python to explore collected database. As a result, the main characteristics of cyber information sharing models were defined. These characteristics were used as embedded units of analysis in the individual case studies.

The literature review indicates that "cyber security information sharing" is not precisely defined in the area of cyber security. As mentioned above, the structures of the information sharing models are generally very sector-specific and created in different environments. There is a need for a common early warning solution. Usually a word "warning" means also preventive functions as US intelligence ser-vices operates. The fight against hybrid threats means not only preventing cyber-attacks, but also identifying, tracing and prosecuting a criminal / criminal group. This means an even deeper integration of government systems in the future.

Relevant information from the site of major hybrid incident must be directly shared to the national participant's e.g. cyber security centres. To determinate discrepancies of limits is relevant to allocate additional reliable data. Combining pieces of information to ensure the correct and reliable information to be shared is primary importance. The essential information should process to the desired shape for the participants. In the future cyber-defence operations are more integrated and automated according to local capabilities, authorities and mission needs. Shared common operational picture means that real time communication link from local level to nation and EU level exist. A common cyber situational awareness is needed for both operating CPS and for emergency and crisis management. There should be the connection between cyber situational awareness and emergency management.

When developing an EWS at the EU level, three requirements exist: 1) The possibility that some EU Member State may leave an early warning system; 2) engaging participants in the values of western world; and 3) the possibility of

joining Cyber Threat Warning System to NATO Cyber Situational Awareness Solutions. These factors have a direct link to sharing confidential information.

It is important to take into account how national Cyber Security Centres cooperate with other organisations within critical infrastructure in national level. The states departments of the United States work closely together in the fight against threats in the field of cyber security. The organisations of public administration in European Union work together more formally. This is important noticed when cyber security expertise is being strengthened. The fundamental problems of the European community must be solved before permanent solutions can be built. However, this does not prevent the development of operating models, but this factor must be taken into account when developing new systems. Firstly, confidence between member states must be on a stable basis.

What are those fundamental differences of administrative functions between European Union and The United States? Mainly there are more similarities than differences. Legislation and regulation between USA and EU are coming closer with each other. NIS directive in EU will help to develop next generation early warning systems. USA and EU have made quite fundamental agreements to generate a common base for fluent information sharing.

As Ilves et.al [3] mentioned, there is no crucial barriers to increase collaboration concerning early warning solutions between US, NATO and EU. US's Cyber security sharing act and Europe´s directive on Network and Information Security (NIS) have similar goals. In addition to this, EU and NATO signed a technical arrangement in 2016 to increase information sharing between the NATO Computer Incident Response Capability and EU Computer Emergency Response Team.[3]

Public safety actors like European law enforcement agencies need common shared situational picture for the cross-boarding tasks in a way that operational co-operation will be based on reliable platform.

Individual Case Studies

This MCSR is made up of following individual case studies:

- *Taxonomies for cyber information sharing* is based on analysis results from the ECHO partners' research, development and innovation work in earlier projects. It provides a definition of taxonomies as used in the cyber domain for cyber information sharing model for collaborative incident response.
- *Health information sharing* was selected as an example of sensitive information sharing models from other than cyber domain. This case study analyses Health Information Exchange (HIE) methods and models, and studies, for example, how to share and analyse the detected physiological profiles.
- The third case study published in [4] analyses the *information sharing models applied in maritime domain.* The main research question is "how can cyber in-formation sharing models be understood in maritime domain?"
- The fourth case study publish in [5] analyses *inter-sector cyber information sharing models in critical infrastructure protection*. It studies how the cyber situational awareness of an organisation can be developed; how do the

organizations exchange their cyber security related information; and how an organisation's cybersecurity capability can be utilised more extensively?

- The fifth case study publish in [6] analyses *cyber threat prevention mechanisms in Finland.* It finds out the pros and cons of the national HAVARO system, and studies what are the factors (requirements), which effect for implementing national EWS system to common early warning ecosystem in EU lev-el. Every EU member country has its own system for monitoring and protecting cyber domain among vital functions.
- The sixth case study published in [7] compares *information sharing between US and EU* emphasising cyber information sharing models in US. In addition, it handles legislative factors, organisational factors and features of the models.
- The seventh case studies effects of national fundamental risks to the international trust warning system and information sharing policy. These are crucial factors within smart societies. Political decision makers are elective, and also many of highest authorities are chosen based on political selection criteria. Hybrid or cyber influencing can create instability to the society in many ways, one key aim is to influence political decision-making. In practice, this means that there is a need to integrate organisational, administrative and operative functions. A trust model with cyber information sharing in CIP is a part of the preventive early warning solution. Secure national and international decision making needs a trust model
- The E-EWS and E-FCR are two of the four vital technologies developed within the ECHO project. Both can exploit each other in order to maximise their capabilities and offerings to the users. The eight case study deals with the synergies of information sharing needs with E-EWS and E-FCR. According to it, the following are the three most relevant use cases where data exchange is re-quired:
  - The Early Warning System can be a part of an exercise or a training that runs on one of the cyber ranges which is also connected to the E-FCR. The data and incident reports that are produced from the exercise/training can be fed into the EWS which will then make an analysis of this. This analysis can be used to by the organisers of the exercise or training to maximise the impact of the exercise/training on the participants.
  - The EWS can collect threat intelligence data from the realistic simulation environment running on E-FCR. This in turn can be used as input by the EWS for alarms or any further analysis done on the EWS. Potentially a digital twin can be set up for the E-FCR where various simulations and scenarios can be run and tested. The EWS in turn can use this input for analysis.
  - EWS can share quarterly data and analysis with the E-FCR's Content Providers in order to allow them to design training and exercise scenarios based on real world needs.

Cross-case Conclusions and System Requirements

This section presents the recommendations following a document analysis exercise of the preceding sections and a selection of literature sources.

Context

At the kernel of information sharing lies the intelligence data item (IDI). In the context of ECHO, an intelligence data item is defined as any piece of data that potentially contains actionable information relating to cyber security. Appreciating the enormous value of information and its potential, an information sharing framework is required in order to appropriately manage the lifecycle of the corresponding data items, from their generation, processing, dissemination all the way to their destruction. ECHO envisages the creation of a community of a large pool of stakeholders who will engage in joint intelligence activities and reliably share information and collaborate in handling security incidents in an effective and timely manner. As such, establishing and ensuring trust is a key factor for the successful adoption of the EWS.

ECHO's information sharing and its instantiation as the E-EWS will adopt the joint intelligence process comprising of the 6 operations (planning and direction; collection; processing and exploitation; analysis and production; dissemination and integration; evaluation and feedback) and adapt and extend - if necessary - the MISP taxonomies.

Characteristics of intelligence Data Items

At a first level of discrimination, IDIs can be structured, semi-structured or unstructured. Typically unstructured data refer to primary sources of information that are normally processed by automated or human means for extracting the necessary information. This process would generate structured IDIs that would allow automated processing. It should be noted though that there can be primary sources ingested into the EWS that are structured (e.g. log files).

IDIs can also be distinguished as reference information or operational information. Reference information refers to the IDIs that contribute in achieving situational awareness, allowing the beneficiary to make informed judgements on the cyber risks of the organisation. Operational information relates to those IDIs that support the actual decision making, handling incidents and so forth.

The IDIs should be accompanied by metadata that will contextualise the contained information but also enable the EWS to implement and enforce authorisation and access control mechanisms. Common identifiers and enumerations should be used whenever possible.

Fig. 2 shows the key components and benefits and goals of the ECHO intelligence information sharing approach.

Table 1 presents an initial list of the categories of information and their expressions as IDIs. IDIs that potentially contain Personal Information will need to also meet the privacy requirements (see subsection below). The categories will be further expanded and refined following the requirements elicitation and specification of the E-EWS (WP5).

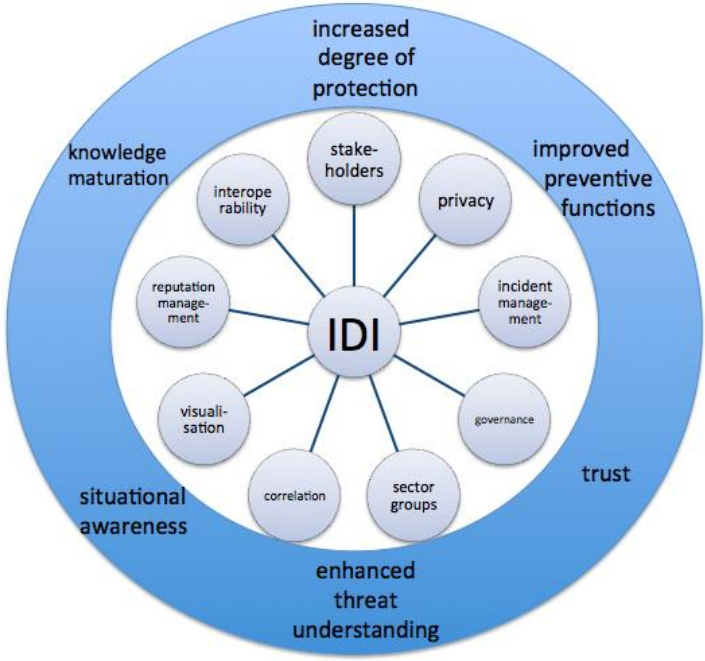

Figure 2: ECHO's information sharing at a glance.

Table 1. Intelligence items.

Information category	IDI	structured/ unstructured	reference/ operational	Personal Infor-mation
Technical threat indica-tor	IOC (email, IP ad-dress, file hash, mutex, domain)	S	R	
Intrusion at-tempt	Threat Actor	S	O	X
	IOC (atomic, composite, be-havioural)	S	O	
Security alert	Ticket	Semi	O	
	Readiness level	S	R/O	
Vulnerability information	CVE	S	R/O	
	CVSS	S	R/O	
	Threat identifica-tion	Semi	O	
			R	
	Geopolitical	U	R	

	Exploitability	S		
Vulnerability report	Vulnerability scanning report	S	R	
Incident report	Report	U	O	?
TTP	ATT&CK STIX object	S S	R/O R/O	
Remediation actions	Operating procedure Playbook	U U	O O	
Asset	CPE to describe system platforms CCE (common configuration enumeration)	S S	R/O R/O	
Discussion	Discussion item	U	R/O	?
Blog post	Reference	U	R/O	
Poll	Poll item	U	R/O	
Raw data	Log-file Netflow Packet capture RAM image dump Malware sample VM Image File Email	S S Semi Semi Semi U U U	O O O O O O R/O R/O	X X ? X X

Information sharing model assumptions

Against all the above, the proposed ECHO information sharing model is based on the following assumptions or premises:

- There will be a clear and concise governance model for the intelligence data items, where each item will be described by a comprehensive list of contextual information (metadata) to allow fine-grained decision making on the management and handling of the data.
- There will be a clear process for on-boarding and off-boarding of participating organisations.

- It is expected that it would be easier for organisations being in the same sector or having similar goals and purpose to form easier clusters for sharing threat intelligence information, as they are more likely to have established and mature exchange arrangements; therefore they are more likely to reach consensus. On the contrary, organisations that operate in orthogonal industries (i.e. where their respective industries have virtually nothing in common) is expected that would be less forthcoming in sharing information.
- Stakeholders and participants are expected to join pre-defined and ad hoc groups.
- Trust will be delivered through technical, organisational and human means.
- Due to the nature and diversity of sectors, in order for information sharing to provide a meaningful and accurate services, the scope of the data items should be extended to encompass Cyber Physical Systems; indicatively, this can consider the practices found in the Maritime Sector where there is a clear distinction between cyber (e.g. IT networks) and Physical (e.g. Operational Technology networks) highlighting the existence and interdependencies between the physical and cyber plane.
- Translation and normalisation services will allow the standardisation of intelligence data items. The underlying taxonomies and schemas should cater for the verticals by including optional fields.
- Existing standards for information processing and sharing will be adopted wherever possible.

Information Sharing Architecture

Information sharing is highly dependent upon and influenced by the regulatory frameworks as well as the cultural norms both within a sector and the organisation itself. In academia for example, barriers to sharing are expected to be lower than the other sectors, due to the culture of freedom of academic expression and an academic citizen mentality of peer review and dissemination of research output. On the other hand, in critical infrastructure type of sectors such as Energy, or in banking, information sharing is more intensely regulated, and this also is reflected in the respective organisational cultures. This creates a tessellation of regulatory frameworks and cultural antecedents on the following levels:

- Intra-Organisational, influenced by specific internal policies and procedures;
- Intra-Sector, imposed by the respective sector;
- National-governmental, governed by the respective strategic decisions on a national level;
- Transnational, through the international agreements, treaties and EU legislation and directives, in the case of the organisation operating within the

EU. This may include frameworks for information sharing with Law Enforcement entities.

The above are also complemented by horizontal legislation such as the GDPR that cuts across all sectors.

Provided that:

- The ECHO pilot is part of the EU initiative on establishing a network of competency centres, and
- ECHO aims to support information sharing among and between a multitude of sectors with Healthcare, Energy and Maritime being initially considered,

a modified hybrid model architecture is recommended as this appears to best fit the requirements following the cross-case analysis. In essence, the hybrid approach will allow to maintain a basic form of hierarchy, and at the same time it will allow the connection of different hubs, forming a higher-level peer to peer. This is also in accordance to how CERTs operate and share information, which is done on a peer to peer basis but also within their level of operation (e.g. national, organisational, etc.). Allowing some degree of centralisation will also enable centralised decision making and support the emergence of Coordination Centres. A hub could represent a variety of communities, such as a specific sector, an interest group or a national point. It is recommended that each hub will refer to organisations of common characteristics, goals or sector, simplifying its management, internal governance and deployment complexity. This would be inline with the E-EWS architecture supporting tenants allowing also seamless integration through the sharing API capability that will connect EWS instances.

From a governance perspective, the immediate consequence of this would be to have trust realms, two tiers of cross organisational boundaries, as shown in Fig. 3.

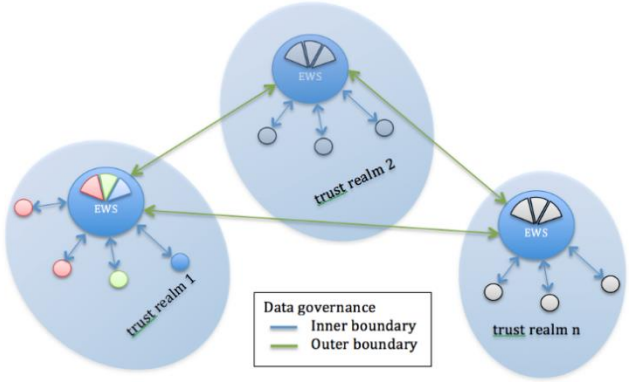

Figure 3: Information sharing architecture.

In the figure above, three trust realms are presented. Each realm can correspond to any type of organisational cluster, e.g. realm 1 could be academic

CERTs, realm 2 national cyber security competency centres and realms 3 maritime sector. Every trust realm can have more than one EWS instances, for scalability and resilience purposes. The governance model could refer to policies and security certification requirements for deploying an E-EWS instance.

The first step for an organisation or individual joining the E-EWS ecosystem is to complete the on-boarding process. Upon successful application, the organisation is allocated a tenant slice. This will host all information provided by the participating organisation. Organisation boundaries can be crossed within a given trust realm and these are specified through the inner boundary data governance. It is expected that these will be the first to be formed, upon the emergence of the E-EWS.

Inter-realm information sharing is controlled by the outer boundary data governance models. These are expected to be more complex and diverse and will require a longer maturity period. It should be noted that not all trust realms will necessarily connect to each other; such configurations imply that some realms will emerge to be more authoritative and trustworthy than others, but should also indicate that transitive trust should not be guaranteed or offered.

IDIs containing personal information will be go through anonymization and redaction layers prior to leaving a tenant's area. For structured IDIs, automated processes would seamlessly and efficiently implement the underlying privacy policy. Information classification schemes will be enforced at the organisational boundaries (coarse grained access control) as well as internally (fine grained).

As the organisation participation and connectivity between the hubs increases, the value of the network is expected also to increase, in accordance to Metcalfe's Law. However, as this increase is very likely to result to generation of large volumes of data, the perceived usefulness is expected to decrease. In order to compensate for this, information sharing should not only be limited by access control criteria, but additional contextual features to enable effective filtering of non-relevant information (noise). A representative feature for this task is asset information. For example, by using the Common Platform Enumeration (CPE) convention, an organisation can describe their assets in a standardised way. By doing this it would be possible to quickly filter out attacks and vulnerabilities that are not applicable to a particular organisation's attack surface.

Stakeholders

Stakeholders operate at different levels, having potentially diverse agendas and priorities:

CERTS/CSIRTS:

- National/EU CERTS – protect national and critical infrastructure
- ISP CERTS – protect Internet related services and backbone
- Organisational CERTS – protect organisation
- ICT Vendor CERTS – protect products
- Academic CERTS – protect academic infrastructure, facilitate research and innovation.

Public safety organisations:

- Law enforcement Agencies. These are secondary users that may be involved when handling incidents. As such, the E-EWS will allow the collection and preservation of evidence in a forensically sound manner.

Information sharing entities:

- Individuals
- Researchers
- Organisations
- National
- Private
- Critical infrastructure
- Research

Features of the Information Sharing System

A modular approach for the E-EWS is considered. The core EWS should comprised of a ticketing system supporting distributed workflow among a number of different partners and organisations. The EWS should allow the enrichment and contextualisation of the introduced and ingress information. As such, a standard description and an expandable information taxonomy should be considered.

An initial list of features of the perspective E-EWS is presented below:

- *A suitable confidentiality model, such as the traffic light protocol*
  All intelligence items will need to be assigned with a designation to ensure that the sensitive information is shared with the appropriate audience. TLP is recommended because it is less formal, does not really require NDAs, etc., it is more of a "gentlemen's agreement" and allows a faster communication of incident data. TLP will of course run in conjunction with the standard system's access control mechanisms, such as RBAC. For the E-EWS system inparticular and upon a joint decision, FIRST's TLP definition is adopted to support future interoperability and standardisation with all pilots. Moreover, the confidentiality model – due to the nature of the EWS – should include introduction of information by protecting source attribution (Chatham House rule), in order to facilitate the submission of any information that can be vital when handling security incidents. A direct consequence of this is the consideration of the reliability of the data, defined further below.
- *An access control scheme, capable of making fine-grained access control decisions*
  The audience accessing intelligence items shall be controlled through access classifiers such as organisations, groups, and roles.
- *Support of multiple taxonomies and standards for intelligence sharing.*
  This will allow the hosting of organisations belonging in different sectors.

- *Capabilities for a structured sharing of intelligence data*
  e.g. use of Structured Threat Information eXpression (STIX)
- *The system should facilitate the exchange of intelligence between CERTS/ CSIRTS and LEAs*
  Terminologies used in the two communities are sometimes different. ENISA recommends using the 'Common Taxonomy for Law Enforcement and The National Network of CSIRTs'.
- *Common data and document formats support*
  Use of common formats e.g. Word, PDF, and CSV facilitate intelligence sharing where the use of specialised formats is not an option.
- *Capability to evaluate the reliability of the source of an intelligence data item*
  All information sources should be assessed for reliability based on a technical assessment of their capability, or in the case of human intelligence source, their history.
- *Assessment of the credibility of an intelligence data item based on likelihood and levels of corroboration by other sources*
  An EWS allowing a quick turnaround and fast decision making requires that the ingress information is trusted. The system should have mechanisms to assess the credibility of the information and include fake news protection mechanisms.
- *A shared workflow management system for incident handling*
  This is one of the main purposes and core functionalities of the E-EWS, allowing also to monitor the effectiveness and efficiency of the system.
- *Trust-boosting security technologies*
  Supporting the creation of closed communities and encrypted peer to peer communication.
- *Data redaction capabilities, for privacy compliance*
  The system will need to redact personal information for data items marked to contain PI when exporting them to other EWS instances based on a privacy protection policy. For structured data, this can be done automatically. For un-structured data, this can be done semi-automatically, but may require human inspection and approval.
- *Attribution capabilities, identification of the origins of the source of information*
  For traceability, disseminated information shall contain appropriate origin describing meta-data.
- *Anonymous sharing of information*
  Despite the attribution requirements, it is advised that the system would still al-low anonymous information, however, these items will need to be clearly marked as anonymous and is expected to have an impact on the reliability of the information.
- *Customisable exchange of intelligence data*

Customisation may be in accordance with internal (originating organisation) or external requirements.

- *Predefined criteria for data dissemination*
  This relates to both the originator of the information (e.g. the criteria a set in accordance with audience, trust realms etc.) and the consumer of the information (e.g. data versions and revisions, severity, etc.)
- *Data normalisation*
  The system shall normalise all ingress data under a common format, or data model. This will enable compatibility, interoperability and other functions (correlation)
- *A flexible data model*
  Expansion of the data model is a prerequisite to allow E-EWS to grow across different domains and verticals. The system can allow custom creation of tags and the enrichment of existing IDIs. This could be automatic or manual. For example an IDI may be enriched by external information from OSINT activities.
- *Correlation capabilities*
  At a minimum level, the system should automatically link newly imported IDIs with existing IDIs.
- *Data items curation*
  The system shall curate and de-duplicate IDIs imported from different sources and datasets. This is for ensuring that the integrity and accuracy of analytics is offered.
- *Advanced data analytics*
  Situational awareness will be considerably supported from data analytics techniques (e.g. clustering and classification). This could include production of trends over time related data to support predictive analytics.
- *Visual analytics*
  The system should provide visual analytics through a dynamic, interactive UI.
- *Pivoting capabilities*
  In order to support the analytics processes and allow complex correlations and analytics, the system should offer pivoting capabilities over data.
- *Data exporting formats*
  The system shall support exporting of data in different formats e.g. STIX, OpenIOC, CSV, Yara, sigma, etc.
- *Filtering capabilities*
  The system should support filtering of information across a number of parameters and features. This also includes both whitelisting, blacklisting, to filter out benign activity and to pin down suspicious/malicious events.
- *Triaging*

The system should provide a high level overview of the data so that the analyst can quickly get a "gist" of what they contain. For example, for numerical data, the basic statistical information should be presented.

- *Alerting and communication*
  This feature is required to improve the response times to incidents. This involves capabilities to match asset configuration with vulnerability information (for example describing assets as CPE and pairing with CVE and CVSS items) and sending a message to a designated contact point if a criticality level of an event exceeds some threshold. For example, this can be done if an asset de-scribed through a configuration is detected to be vulnerable to an exploit with a CVSS score.
- *Intelligence report generation*
  The information shared should be available to the stakeholders in an appropriate format and level of detail.

Privacy Requirements

In order to identify the personal information to be managed and processed by ECHO, the consortium carried out a detailed analysis of the different categories of personal information to be processed and its lifecycle. This analysis is described in the Data Protection Impact Analysis Report.[8]

ECHO is underpinned by a series of privacy statements. These comply with the General Data Protection Regulation (GDPR) and are managed and overseen by the Data Controller and Data Protection Officer (DPO) for the project, RHEA System SA (RHEA).

In addition to these statements, each consortium member will liaise with the DPO to establish a Data Protection Impact Assessment (DPIA) must be conducted prior to any data collection or processing taking place. This decision will be reviewed whenever the data category, type or the nature and/or scope of the processing changes.

Data Processing refers to any handling of data whether this is capturing, creating, modifying, adding, deleting, sharing or otherwise handling of data. Therefore, any and all data captured/to be captured and processed, whether manually or by automation as part of this project will be processed in some way and potentially fall within the remit of the General Data Protection (GDPR).

Collecting IDIs having personal information such as a threat actor, log-file, RAM image dump, etc. (see Table 1) will be processed and stored in accordance with the following privacy requirements:

P1. Lawful basis for processing:

a. the lawful basis for processing data will be specified, recorded and justified;
b. the data will be classified in accordance with sensitivity as either:
   I. Personally Identifying (PI)
   II. Non-personal (N)
   III. Other (O) (meaning the classification is to be confirmed pending discussion with project lead, privacy officer or DPO)

P2. Purpose Limitation: personal data should only be processed for needed and specified purpose; no personal data should be reused without informed consent first being obtained. Informed consent templates are provided within the ECHO project documentation (Reference Materials, documents folder);

P3. Data Minimisation: only necessary data for the specified purpose will be processed;

P4. Accuracy: the data will be accurate and kept up to date;

P5. Storage Limitation: data will be pseudonymised or anonymised as soon as practicable and kept for no longer than absolutely necessary ('the data life'). At the end of the data life, data will be securely deleted and/or destroyed.

P6. Integrity and Confidentiality:

a. Confidentiality: Ensuring data is only accessible to authorised stakeholders
b. Integrity: Ensuring non-repudiation and reliability for each piece of data, i.e. processing correct, authentic, and unmodified data.
c. Availability: Ensuring data is usable on demand and accessible to authorised stakeholders
d. Unlinkability: Ensuring data cannot be sufficiently distinguished to linked across platforms or domains with similar context or purposes
e. Unobservability/ Undetectability: Ensuring data is anonymised so that the anonymity and undectability of the individual is preserved
f. Anonymity: Obfuscating links between data and identity i.e. the ability to distinguish any one individual from the data
g. Pseudonymity: Replacing identifying data with pseudonyms ensuring any links to original data cannot be made by unauthorised parties

P7. Intervenability: Enabling data subject access and/or supervisory authority access to affect action on the records (e.g. request modification and/or deletion). In that way it can be seen as a safeguarding measure that must be included within any process or system involving personal data

a. Transparency: Openness – Providing assurance, accountability and traceability for internal and external stakeholders.

P8. Proportionality: Proportionality requires that any limitation on the rights of the individual have to be justified. For example, making sure that the measure(s) taken in processing the data do not disproportionally limit the rights of the individual whose data is being processed. A pre-condition is that the measure(s) taken in processing or safeguarding are sufficient to achieve the objective while only relevant personal data for the purposes of the processing is collected and processed.

These privacy goals comply with the GDPR and are based on the privacy principles of GDPR (P1-7) and those in the Privacy Lifecycle PLAN (i-ix), that forms part of the Privacy and Compliance framework (PACT).[9]

Conclusions

This study consists of a comprehensive review of information sharing and trust models from within the cyber domain, as well as models from other domains,

such as, healthcare information sharing. From these models a set of system requirements and features is established to support a model that promotes information sharing among partners, while also meeting regulatory requirements. The content of the paper is based on results of analysis of eight case studies carried out in the ECHO project, and cross-case conclusions of them. Moreover, the final analysis includes the requirements for information sharing within and between partners across organisational boundaries as derived from multi-sector analysis.

Acknowledgement

This work was supported by the ECHO project which has received funding from the European Union's Horizon 2020 research and innovation programme under the grant agreement no.830943.

References

1 Robert K. Yin, *Case Study Research and Applications: Design and Methods*, Sixth edition (Los Angeles: SAGE Publications, Inc., 2017).

2 Karl Popper, *Conjectures and Refutations: The Growth of Scientific Knowledge* (London: Routledge Classics, 2009).

3 Luukas Ilves, Timothy Evans, Frank Cilluffo, and Alec Nadeau, "European Union and NATO Global Cybersecurity Challenges: A Way Forward, *PRISM* 6, no. 2 (2016).

4 Jyri Rajamäki, Ilkka Tikanmäki, and Jari Räsänen, "CISE as a Tool for Sharing Sensitive Cyber Information in Maritime Domain," *Information & Security: An International Journal* 43, no. 2 (2019): 215-235.

5 Jouni Pöyhönen, Viivi Nuojua, Martti Lehto and Jyri Rajamäki, "Cyber Situational Awareness and Information Sharing in Critical Infrastructure Organizations," *Information & Security: An International Journal* 43, no. 2 (2019): 236-256.

6 Jussi Simola and Martti J. Lehto, "National Cyber Threat Prevention Mechanism as a part of the E-EWS," *International Conference on Cyber Warfare and Security*, 2020, pp. 539-XV.

7 Jussi Simola, "Comparative Research of Cybersecurity Information Sharing Models: The Common Cyber Ecosystem of ECHO," *Information & Security: An International Journal* 43, no. 2 (2019): 175-195.

8 ECHO, "DPIA pre-assessment Report, v0.2," 2019.

9 Jane Henriksen-Bulmer, *Incorporating Contextual Integrity into Privacy Decision making: A Risk Based Approach* (Bournemouth University, 2019).