Sharing and Automation for
Privacy Preserving Attack Neutralization

(H2020 833418)

D5.7 Sharing Response Handling Information (M21)

Published by the SAPPAN Consortium

Dissemination Level: Public

H2020-SU-ICT-2018-2020 – Cybersecurity

Document control page

Document file:	D5.7
Document version:	1.0
Document owner:	Martin Zadnik (CESNET)
Work package:	WP5
Task:	T5.4 Sharing Response Handling Information
Deliverable type:	Report
Delivery month:	M21
Document status:	☒ approved by the document owner for internal review ☒ approved for submission to the EC

Document History:

Version	Author(s)	Date	Summary of changes made
0.1	Martin Zadnik (CESNET)	2020-12-04	Outline of the document
0.2	Martin Zadnik (CESNET)	2021-01-04	Preliminary version
0.3	Martin Zadnik (CESNET) Mischa Obrecht (DL) Tomas Plesnik (MU)	2021-01-18	Draft version
0.4	Martin Zadnik (CESNET)	2021-01-26	Internal review version
0.5	Martin Zadnik (CESNET)	2021-01-28	Integrating reviewers' comments
1.0	Martin Zadnik (CESNET)	2021-01-29	Final version for submission

Internal review history:

Reviewed by	Date	Summary of comments
Robert Rapp (USTUTT)	2021-01-27	Grammar and rephrasing
Tomas Jirsik (MU)	2021-01-26	Technical check

Executive Summary

The task T5.4 aims at enabling sharing of response handling information. This initial deliverable describes our implementation activity toward sharing response handling information in the form of playbooks. We summarize the existing work and development in this area. We also collect general as well as playbook-specific requirements on sharing system and its related components, functions and data model. We survey and assess promising sharing systems to select the best fit for SAPPAN. We introduce the MISP data model for playbook sharing and propose an interface between MISP and relevant SAPPAN components. Subsequently, we evaluate how the specific playbook-sharing requirements are met. We conclude with our plan for the next period which will be documented in the second version of this deliverable.

Table of Contents

1 Introduction

Response handling information documents how to respond systematically when a threat is discovered, an attack is detected or an incident dealt with. The documented systematic response enables the replication of the handling process, supports its gradual improvement and enables meeting the legal requirements. Response handling information covers several phases of the incident response process as defined by NIST [1] - preparation, detection, assessment, and handling. In either of these phases, the response handling information captures guidelines for the particular phase and the particular threat/attack/incident. The guidelines are often documented as playbooks that are high-level human-readable, written in plain text without structure.

The free form of the playbooks requires human operators to interpret and execute the playbooks. Given the need for a timely response as well as the need to utilize human capabilities efficiently, there is a demand for the automation of the response, but the high-level playbooks are too abstract to be automatized. Therefore, current efforts revolve around how to make the playbooks readable by machines.

The machine-readable playbooks have been introduced recently in the form of workflows [2] which elaborate the abstract playbooks into a greater level of technical detail. Sharing detailed response handling information enables organizations to prepare for a potential attack that has not yet reached its infrastructure. This allows organizations to prepare an appropriate response before an attack arrives, or even pro-actively prevent the threat from materializing and causing an impact. In addition, a single organization may only have a fragmented view on how to respond to the threat, attack, or incident. Sharing the pieces of response information and finding effective responses by correlating and combining them, can help an organization to recover from an attack more effectively.

Naturally, the sharing of response handling information faces obstacles. Since the extensive automation of response handling has been introduced recently, the standards are behind by the development of the cybersecurity vendors, resulting in multiple representations, toolsets, and approaches. Another issue is that organizations fear revealing confidential information when pieces of handling information leave an organization. E.g. while workflows support the automation, they may reveal what are the organization's handling procedures, what tools are being used, what are the organization's capabilities, what does the infrastructure consists of as well as private information.

1.1 Scope

The response handling information is a too general term covering everything from recommendations, guidelines, manuals over playbooks to workflows. SAPPAN focuses on playbooks, to bring them closer to the workflows, i.e. to enable automated interpretation of the playbooks and to enable sharing of such playbooks. In the rest of the deliverable, we use the term playbook to refer to anything between the abstract playbooks to detailed workflows.

The primary goal of the playbook is to lead humans or machines in the selection of proper actions to respond to a trigger. The trigger may vary from breaching some internal policy, finding a new vulnerability, detecting an indicator of a compromise, reception of intelligence about a new threat or threat actor. The trigger also forms an initial hypothesis (or initial hypotheses but let us for the sake of simplicity consider only a single hypothesis at a time) how to respond. The hypothesis is elaborated, confirmed, proved wrong, or rather considered more or less likely by the human or machine by executing the steps of the playbook. Even if the playbook does not contain any assessment steps but only a mitigation action, the successful result of the mitigation action proves the initial hypothesis correct. The less successful result proves the hypothesis less likely and should lead to improvement of the respective playbook. The secondary goals of the playbooks are to make the handling process repeatable, documented, quantifiable as well as to allow for automatization and improvement of the process.

An innovative goal set by the SAPPAN project proposal is to formalize the description of the playbooks. To this end, the task T4.1 developed a methodology for formalizing and modelling response and recovery actions and their triggers using semantic technologies. The deliverable D4.1 defined a formal methodology of a generic playbook for cybersecurity response and recovery actions using technologies from semantic knowledge modelling, such as the Resource Description Framework (RDF), RDF Schema (RDFS), and the Web Ontology Language (OWL). In parallel to this deliverable D5.7, T4.1 is preparing (in D4.2) vocabularies to standardize naming conventions which will help to map the playbooks into the realm of a particular organization and standardize naming across organizations.

1.1.1 Aim and outline

The aim of the T5.4 activity is to develop a proof of concept for sharing response handling information, taking into account the achievements of other SAPPAN Tasks focused on modelling and capturing response handling information in WP4. Nevertheless, we follow existing work regarding response handling information and the latest development in this area so that our proof-of-concept for sharing playbooks can also support other representations in the future. We provide a brief overview of the related work in Chapter 2. Since our aim is to share the playbooks, we collect general requirements on the sharing system as well as specific requirements on components, functions, and data involved in playbook sharing. The requirements are presented in Chapter 3. The general requirements help us to assess several sharing systems, and we select the sharing system based on the assessment in Chapter 4. Further, the deliverable introduces the data model specific to the selected sharing system in Chapter 5 and proposes an interface of the sharing system with the relevant SAPPAN components in Chapter 6. An example of the playbook captured in the MISP data model is presented in Chapter 7. Chapter 8 summarizes the deliverable and outlines our future plan.

2 Related work

We may see some playbooks being collected and made publicly available (e.g. [3]) but these playbooks are too abstract and may serve only as background material for the human operators or for constructing more detailed playbooks. In this deliverable, we consider the SAPPAN WP4 representation of playbooks. But while the outputs of WP4 serve as primary representation we also look for recent development in Security Orchestration and Automation Response (SOAR) domain.

Integrated Adaptive Cyber Defense [2] defines a strategy and framework to adopt an extensible, adaptive, commercial off-the-shelf (COTS)-based approach. Its goal is to change the timeline and effectiveness of cyber defense via integration, automation, orchestration, and sharing of machine-readable cyber threat information. IACD provides a strategy and a framework to enable organizations to increase the efficiency and effectiveness of their cybersecurity operations through the appropriate use of automation; leveraging orchestration of existing processes, products, and services; and engaging in threat sharing communities that support machine-based consumption and usage of intelligence to drive operational priorities and decisions. The effort is sponsored by the Department of Homeland Security (DHS) and the National Security Agency (NSA) in collaboration with the Johns Hopkins University Applied Physics Laboratory (JHU/APL).

IACD defines 3 levels of playbook abstractions - "playbooks", "workflows" and "local instances". IACD Playbooks represent general security processes in their most basic form and are meant to be shared between organizations. They provide a mapping to governance and regulatory requirements, describe industry best practices as available response and mitigation actions in human-readable form, and do not contain any conditional logic. Decisions are left to the analysts' discretion. IACD defines a minimal required taxonomy for IACD playbooks. Each IACD playbook must at least contain the following information [2]:

1. Initiating condition
2. Process steps
3. Best practices and local policies
4. End state
5. Relationship to governance or regulatory requirements

IACD Workflows are the machine-understandable codification of playbooks to enable automation of the procedures. Workflows are the technical steps and are repeatable and auditable. They can be tailored to organizations' needs and appetite for automation and are machine-readable. IACD proposes the usage of the business process modeling language (BPMN) to describe workflows. BPMN is a standard that allows for the representation of processes without requiring specific technologies. There are multiple free and non-free applications for editing and reading files in the BPMN format (e.g., Camunda Modeler, Flowable Modeler, etc.). BPMN is usually stored in an XML based format. Orchestration services execute workflows, interfacing with the other orchestration services and humans as necessary.

Local instances of IACD workflows are often thought of as runbooks or SOAR-playbooks. They incorporate technologies, products, and assets deployed in the local environment and respond to conditions or events that are occurring in that local environment. As opposed to playbooks and workflows, local instances are thus technologyagnostic anymore. Because of their hierarchical relationship to workflows and playbooks they are however still consistent with local policies, procedures, thresholds, and decision processes.

Another seminal work was published by the OASIS Collaborative Automated Course of Action Operations (CACAO) for Cyber Security Technical Committee [4]. It defines a standard for cybersecurity playbooks. The CACAO specification defines the schema and the taxonomy for cybersecurity playbooks. CACAO playbooks are made up of five parts; playbook metadata, the workflow logic, a list of targets, a list of extensions, and a list of data markings. The metadata and the data markings constitute an important element from the perspective of sharing the playbooks and we will consider them while designing the data model for sharing playbooks in the selected sharing system.

Besides IACD and CACAO there are other projects (see Table 1) that are focused on automating the processes into executable workflows some focus on cybersecurity, some are more general. Nevertheless, it is important to be aware of these projects, so that our sharing data model is prepared to support these formats in the future.

Table 1: Playbook/workflow automation projects

Project name	**Playbook representation**	**Link**
SAPPAN	JSON	D4.1
IACD	XML	https://www.iacdautomate.org/
COPS	YAML	https://github.com/demisto/COPS
OASIS Collaborative Automated Course of Action Operations (CACAO) for Cyber Security TC	JSON	https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=cacao
Act react	YAML	https://github.com/atc-project/atc-react
ThreatConnect	JSON	https://github.com/ThreatConnect-Inc/threatconnect-playbooks/tree/master/playbooks
Rapid7-InsightConnect	YAML	https://github.com/rapid7/insightconnect-workflows/tree/master/workflows
Apache Airflow	Python (DAG)	https://airflow.apache.org/docs/apache-airflow/stable/index.html
The Hive - Cortex	Python	https://github.com/TheHive-Project/Cortex-Analyzers

Further projects regarding automation of processes are aggregated on several GitHub pages [6], [7], [8].

We follow with a brief overview of the sharing systems: MISP, STIX/TAXII, OpenCTI, Warden (Table 2 provides references). These systems are promising candidates for sharing various pieces of information in SAPPAN. MISP itself is a community sharing platform that depends on the content which is generated by communities. MISP is an open-source, licensed under GPLv3. MISP is not only a software tool but also a series of data models created by the MISP community. MISP includes a practical and straight-forward information-sharing format expressed in JSON, which is the core format for the MISP platform itself. Its concept is based on objects, attributes, and taxonomies. MISP attributes contain the pieces of data themselves. MISP attributes are of various categories and types, e.g. an attribute of type bank-account-nr belongs to the financial fraud category. There are various categories and types. MISP objects allow building a collection of attributes. The objects are defined by a template that enumerates a set of attributes in the object. MISP also includes various existing taxonomies to classify events and attributes, such as CSIRTs/CERTs classifications, national classifications or threat model classification e.g. MITRE [5]. MISP has a protocol used for synchronization between different MISP instances. The MISP platform allows defining the distribution of the CTI records among organizations. MISP supports the export of records and attributes in different formats (e.g., OpenIOC, CVS, STIX in XML, and JSON) to allow integration with other tools.

STIX is a language and serialization format for exchanging cyber threat intelligence. STIX is developed by MITRE and OASIS, it is free and open source. It offers a structured, object-oriented approach with relationships represent threats, attacks, actors, malware and events. It integrates and links different standards (e.g. Cybox). STIX2.0 is using JSON to represent the data using multiple defined STIX Domain Objects (SDO) which consists of attributes to categorize the pieces of stored information. The relations between objects can be seen as a graph where SDOs are nodes and the relationships are edges. The community widely recognizes STIX as a comprehensive and well-developed taxonomy for the representation of cyber incidents. STIX includes its own taxonomies. Its development follows a formalized and open process but takes a significant amount of time to reach the next version. TAXII is an application layer protocol. Its purpose is the communication of cyber threat information over HTTPS. Primarily intended to be used with STIX, but can be used separately for other sharing systems. Since it is a protocol we rather should evaluate the tools implementing it such as commercial Eclectic-IQ or open-source OpenTAXII.

OpenCTI is an emerging system composed of modern technologies such as Elastic [9], RabbitMQ [10], GraphQL [11], redis [12]. It is developed as an open-source. It focuses on processing, structuring and correlation, and correlation of cyber threat intelligence of technical and non-technical. Its schema is based on the STIX2.1 format. It can be integrated with other tools and projects such as MISP, The Hive [13], MITRE CTI [14]. OpenCTI aims at being a holistic tool for the integration of technical as well as non-technical information. OpenCTI supports processing of structured and unstructured data enabling an inference of meaningful knowledge from the raw data. OpenCTI supports export and import into various formats such as CSV, STIX2 via its connectors.

Warden is a very simple sharing system build as a client-server model where clients send or receive shared events from the server. It is open-source based on GPL v3 license. Warden is using IDEA format. IDEA is a straight-forward format to describe alerts reported by network devices such as honeypots, IDS, NBA, firewalls, etc. IDEA is defined as at most a two-level deep tree of keys and values (JSON). That allows for just one basic level of indirection when represented in relational models (except arrays). It allows for referencing of other messages via different types of reference fields.

IDEA is extensible by any custom field which is ignored by a consumer if the consumer does not recognize the custom field. IDEA is easily machine-readable due to its low complexity and minimum possibilities for ambiguous representations of the defined event. Further details can be reached via websites of the projects in Table 2.

Table 2: References to the sharing systems assessed by SAPPAN

Project	Topic	URL
MISP	General	https://www.misp-project.org/features.html https://github.com/MISP/misp-book
	Core Format	https://github.com/MISP/misp-rfc/blob/master/misp-core-format/raw.md.txt
	Taxonomies	https://github.com/MISP/misp-rfc/blob/master/misp-taxonomy-format/raw.md.txt
	Detailed Documentation	https://www.circl.lu/doc/misp/book.pdf
STIX and TAXII	General	https://oasis-open.github.io/cti-documentation/ https://stixproject.github.io/about/
	STIX: Visualized Relationships	https://oasis-open.github.io/cti-documentation/examples/visualized-sdo-relationships
	Detailed Documentation	https://docs.google.com/document/d/1yvqWaPPnPW-2Ni-VCLqzRszcx91ffMowfT5MmE9Nsy_w/edit
	Documents, Schemas and Tools	https://oasis-open.github.io/cti-documentation/resources
OpenCTI	General	https://opencti.io
	User Guide	https://www.notion.so/OpenCTI-Public-Knowledge-Base-d411e5e477734c59887dad3649f20518
Warden	General	https://warden.cesnet.cz/en/about_project
	Architecture	https://warden.cesnet.cz/en/architecture
	IDEA Format	https://idea.cesnet.cz/en/index

3 Requirements

We collect the requirements to be able to assess the sharing systems, to be able to make an informed decision on system selection, to identify the needs and gaps of sharing the playbooks. We split the requirements into two groups - general requirements that are not specific to the playbooks but are related to the sharing system and requirements regarding sharing of the playbooks.

3.1 General requirements of SAPPAN architecture

The deliverable D2.4 Functional specification and architecture definition identified a component that is a prerequisite to the sharing of information among multiple entities. The component is the Message Broker with its related components IC and IP proxies (see Figure 1).

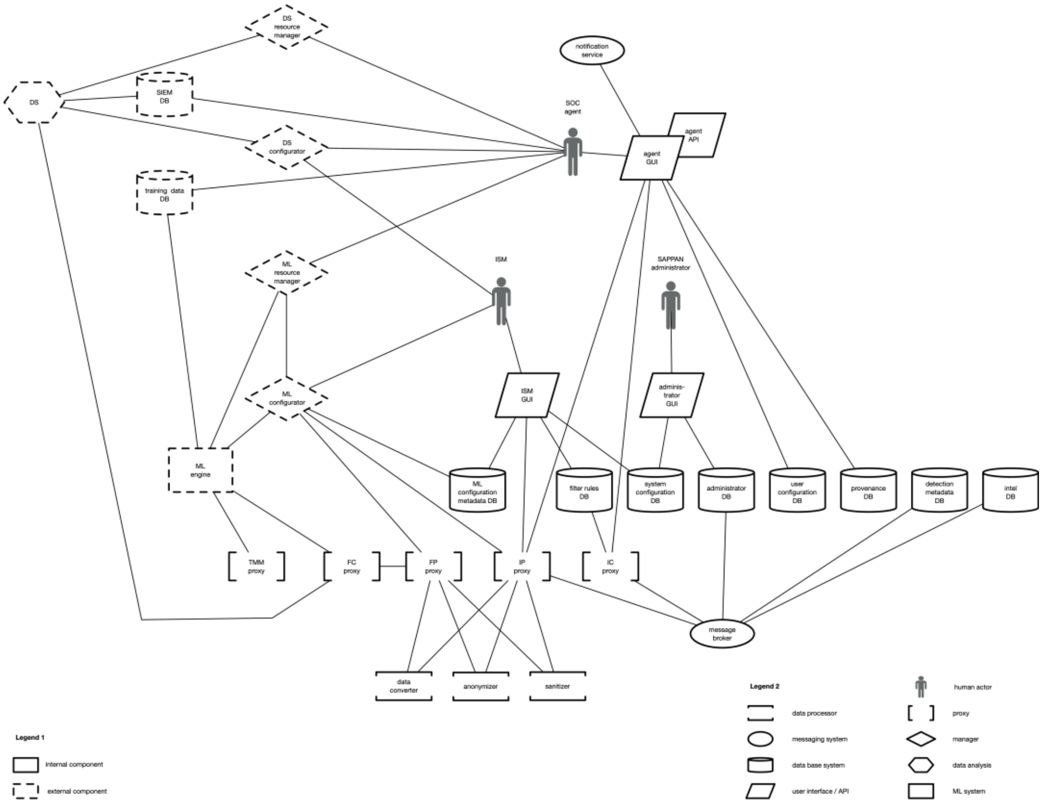

Figure 1: SAPPAN Technology Architecture Model diagram

The Message Broker is primarily responsible for the distribution of intelligence. It receives converted, anonymized, and sanitized data from the IP proxy and forwards it to the IC proxy. Additionally, it deals with authorization by checking client privileges in the Administration Database. It can further store the received intelligence in the Intelligence Database and retrieve it for future distribution. The Message Broker also manages the Detection Metadata Database, where data about collaborative learning tasks are stored.

As the whole SAPPAN concept involves sharing various types of data (see Figure 2) in various use cases we must consider requirements that are more general and will not limit the sharing system hosting these use cases and data.

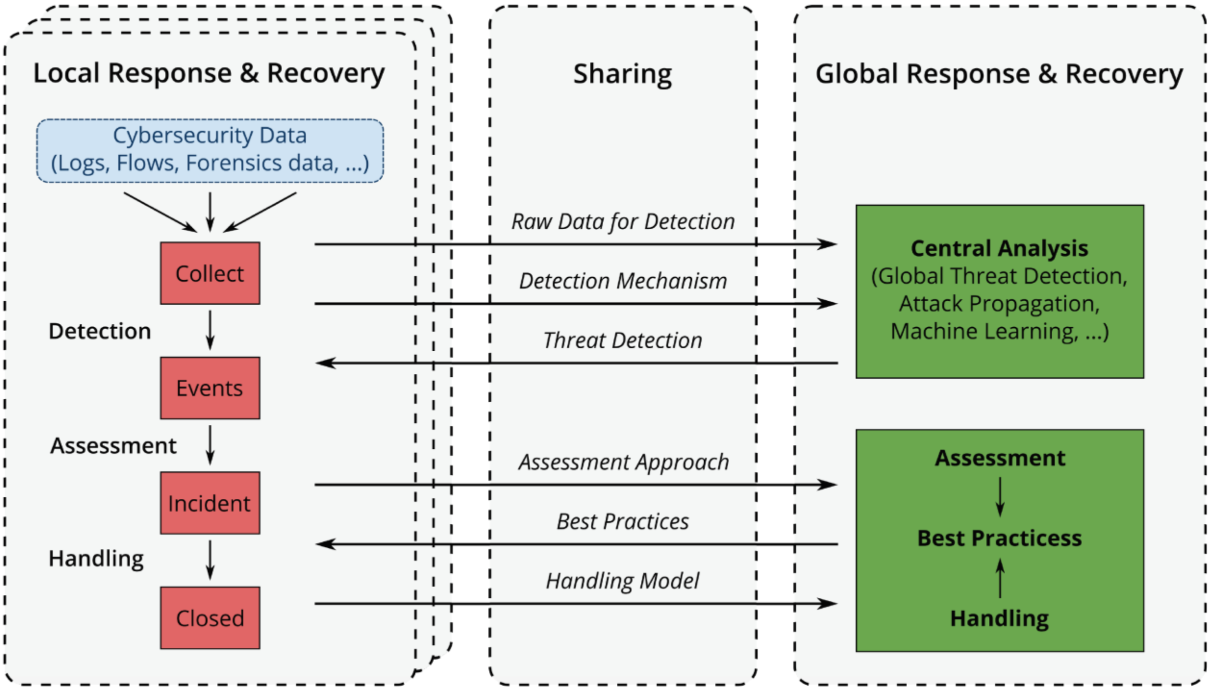

Figure 2: Sharing of various data in SAPPAN

The Table 3 captures the SAPPAN sharing requirements from an architecture and functional perspective.

Table 3: General sharing requirements

ID	Name	Description
ShRq-1	History	The message broker must be able to store and provide data for a long time, e.g. an organization that joins the sharing community can retrieve data submitted in the message broker earlier. The minimal required time is a year for threat intelligence data and permanent availability for glossaries. The expected amount of data per day is in the order of megabytes.
ShRq-2	Data updates	The message broker must be capable to invalidate/delete uploaded data by the author. Optionally, the message broker provides a capability to update/edit the stored data with new information, e.g. given feedback, by the author as well as other participants.
ShRq-3	Open-source	The message broker should be available as open-source software that can be modified/extended freely. The more open the better. Ideally, without obligations.

ShRq-4	SAPPAN proprietary data	The message broker and its data model/format must support an object (or similar abstraction) expressing proprietary blob with proprietary metadata. This means that the data model/format already offers a convenient object or that the model/format can be extended with a proprietary object carrying the metadata and blob.
ShRq-5	User management	The message broker must have user management (register, delete user, password recovery, etc.).
ShRq-6	Authorization and access to data	The message broker must be able to authorize a user. In case of a user with write and admin permission a two-factor authentication can be considered. Based on user roles and permissions in ShRq-12 the message broker allows the user only adequate actions and adequate access to data.
ShRq-7	Storage of diverse data	The message broker should be able to store not only threat intelligence but also static data such as glossaries, overlaps with ShRq-8.
ShRq-8	Glossary	The message broker must provide a glossary that describes metadata.
ShRq-9	Clients/proxies	The message broker may have clients/libraries providing the capability to connect to the server/to each other if the p2p model is considered. This means that there is an existing reference implementation of a client and connects to the API of the message broker.
ShRq-10	Filtering	The API of the message broker may have the ability to filter data that are being shared.
ShRq-11	Auditing	The message broker should provide logging capability and monitoring of the log for specific patterns.
ShRq-12	User roles/permissions.	The message broker must have a concept of user roles or permissions. Such as the admin of the message broker has permissions to manage admins per organization and to freely access and modify any data in the message broker. Admin per organization (or there might be specific organization account) has permission to add users per organization and grant them various permissions, e. g. read-only access or edit access.
ShRq-13	Encryption at transport	Encryption of data at transport between client and server is a must.
SqRq-14	Encryption at rest	Encryption of data in the internal storage would be nice to have.

SqRq-15	Maturity	The message broker must be stable and well supported by its developers. For the deployment in a productional environment, it must allow minimal effort to maintain it during its lifetime.
SqRq-16	Community	The message broker must be widely accepted and deployed in the cybersecurity community.

3.2 Playbook requirements

Protecting privacy is essential when sharing any piece of information outside of the group of people who are entitled to work with given personally identifiable information (PII). The sharing of playbooks is not different. In general, playbooks do not contain personal information but in case they do, e.g. send an email to john.doe@administrator.com, such information must not make it into the playbook that is shared. On the other hand, when sharing the playbook within the entitled persons in a single organization, the privacy is maintained by default and privacy-protective actions should not be applied. We identify two privacy requirements in Table 4.

Table 4: Privacy requirements

ID	Name	Priority	Description
PrRq-1	Limit the recipients	MUST	The producer of a playbook wants to define who can receive/view the given playbook. The set of recipients can be either only users from the same organization as is the producer, or only enumerated organizations or all the users participating in this sharing community.
PrRq-2	No personal information	MUST	The playbook shared outside of an organization must not contain personally identifiable information unless the producer gives explicit consent. For example, a managed service provider shares its playbooks with its customers and deliberately specifies a step that contains an email address with a Managed Service Provider (MSP) member in a network operation center (NOC), in such a case the step must be marked as deliberately containing PII that should not be removed. Otherwise, PII must be either removed or anonymized.

Leaking confidential information from the cybersecurity perspective is the main concern of organizations when asked to share their playbooks. The confidential playbooks which could be leaving the sharing community are a potential threat exploitable by attackers as well as by aggressive business competitors. Therefore, SAPPAN must seek to abstract confidential information. On the other hand, the playbook cannot be too abstract as it would hamper its automation and lower its value from the perspective of how to respond properly. We collect the related requirements in Table 5.

Table 5: Confidentiality requirements

ID	Name	Priority	Description
CoRq-1	No sensitive identifiers leak	MUST	The shared playbook should not contain identifiers that leak information that an organization considers confidential.
CoRq-2	No tools leak	MUST	A more specific requirement to the CoRq-1. The producer of the playbook does not want to reveal which particular tools are being used by its organization.
CoRq-3	No infrastructure element leak	MUST	A more specific requirement to the CoRq-1. The producer of the playbook does not want to reveal which infrastructure elements are being used in its organization or how the infrastructure looks like, e.g. its topology.
CoRq-4	Detail to abstract	COULD	One of the sanitization as well as anonymization techniques is abstraction using vocabularies to translate from a playbook, specific to an organization, into an abstract playbook.
CoRq-5	Abstract to detail	COULD	This requirement is reverse to the Detail to abstract. The point of this requirement is to enable consumers of the playbook to map abstract identifiers onto identifiers that are relevant to the specific of the consumer's organization.
CoRq-6	Confidential step	MUST	If the step is marked confidential then it must be either left blank even it does not contain sensitive or private information.
CoRq-7	Anonymity of producer	COULD	Producer could choose to remain anonymous when pushing the playbook into the sharing system to prevent affiliation of its organization with the shared playbook.

Cooperation between organizations is vital to improve the overall cybersecurity on the Internet. To this end, there must be a mutual understanding of the information that is being shared, possibly not only between humans but between machines as well to support automation (see Table 6 for the requirements).

Table 6: Automation requirements

ID	Name	Priority	Description
AuRq-1	Human-readable	MUST	The playbook must be human-readable.
AuRq-2	Machine-readable	MUST	The playbook must be machine-readable.
AuRq-3	Machine interpretable	SHOULD	The playbook format should support the interpretation of the playbook by a machine, i.e. if the playbook contains a description detailed enough to be executed by a machine, the playbook format should support it. In other words, the format should support various levels of playbook abstraction (high-level = interpretable by human, low-level = interpretable by machine).
AuRq-4	Lists of defined values	SHOULD	Wherever possible the playbooks, as well as any additional data related to the playbook, should utilize predefined values to describe actions, techniques, tools, etc.
AuRq-5	Mapping	WOULD	Two different organizations utilize different tools during the incident handling process. As a part of the sharing, it may be supported mapping of various tools with the same functionality so that the consuming organization can replace the tools in the playbook that are not utilized in the consuming organization.
AuRq-6	Visualization	WOULD	As a part of the sharing, it would be useful to support visualization of the playbooks so that the human operator can fast assess the playbook she or he is interested in.
AuRq-7	Filtering	SHOULD	Besides sharing the playbook itself, it should be shared with additional metadata which can be used to filter the playbooks of interest by an operator of the consuming organization.

The data model of a message containing a playbook influences the possibilities of a consumer to operate with the received piece of information and how to interpret it. We identify requirements specific to data model of playbook in Table 7.

Table 7: Data model requirements

ID	Name		Description
DaRq-1	Format identification	MUST	When sharing the playbook, the consuming party must be able to read the arriving playbook. To this end, the format and its version must be clearly identified.
DaRq-2	Support of various playbook formats	SHOULD	As it is the case even with sharing formats, there are various communities, vendors, and standardization bodies, each introducing their format. Therefore the sharing itself should not prohibit the utilization of various playbook formats.
DaRq-3	Parameters to filter	SHOULD	The metadata should serve attributes to filter a particular playbook of interest.
DaRq-4	Extensible	SHOULD	The structure of the shared information should be extensible to allow additional metadata as well as different representation of the playbook.

4 Sharing system selection

We assess how various systems meet the requirements defined from the perspective of the whole SAPPAN architecture. Please note that in the case of STIX/TAXII we assess the tools implementing these standards where applicable. This work has in fact been done within the scope of WP6 T6.2 but we present it here already to draw the complete picture of how we want to approach the sharing of the playbooks. The Table 8 captures how the requirements are met.

Table 8: Assessment of sharing systems

ID	Name of sharing requirement	MISP	OpenCTI	STIX/TAXII implementations	Warden
ShRq-1	History	MISP stores all the data since the start of the instance, the only limit is the disk space. There are no automatic cleanup routines, old data can be cleaned up via MISP API if needed.	OpenCTI stores all the data into a database (Elasticsearch) and storage (MinIO) since the start of the instance, the only limit is the disk space. Automatic cleanup routines are not implemented yet	STIX/TAXII are based on immutable history. The internal storage is specific to a particular TAXII server implementation. The available implementations use standard persistence stores which are usually limited only by disk space.	Warden stores alerts up to a certain threshold which is given during configuration. When this threshold is reached old alerts are removed.

ShRq-2	Data updates	Editing and deletion of an event are possible by its author (and other users in the same organization). Others can create a proposal to change the event, which the author can accept.	Editing is most likely possible but it is not possible to fully assess this feature based on documentation. Due to the recent appearance of this system, there was no time for experiments in this direction.	STIX/TAXII are based on object versioning and revocation as a mechanism for updating and deleting the data.	Warden does not allow editing of alerts. Warden does not allow to delete an alert. Warden allows to invalidate an alert but this is not in its standard interface.
ShRq-3	Open-source	AGPLv3 license (i.e. open-source, extensions are possible, but must be released under the same license).	Apache 2.0	STIX/TAXII are open standards. Several prototypes of open-source implementations are available, e.g. https://medallion.readthedocs.io/, https://opentaxii.readthedocs.io/. OpenCTI is an open-source platform based on STIX.	3-clause BSD license.

ShRq-4	SAPPAN proprietary data	MISP core format allows extending the set of already predefined MISP object templates by custom MISP objects. Every object template describes a set of attributes, which the object uses. Some of the attributes may be mandatory, others are optional. By using a custom object template it should be possible to describe any needed data and metadata and there should not be any problem with future extensions or updates of the object.	The schema is built on STIX2.1 and that allows for extensions.	STIX standard provides mechanisms for its flexible extension and customization. However, representing completely new data is difficult and needs to be officially standardized in the next version.	Warden and its native format IDEA (JSON) allow for any proprietary extension, i.e. to define proprietary tags of string or encoded-binary type. Moreover, additional metadata can also be stored in tag Note (string) and binary data in tag Content-X (MIME encoding).

ShRq-5	User management	Complete user management is available. Each user belongs to an Organization. Administrator of a MISP instance can create and edit Organizations and assign admin accounts. Organization admins can create new users, edit users (password reset and information update), delete users, and display all detailed information about a user (includes the last login, if the user has PGP key set, subscription to auto alerts, etc.).	OpenCTI supports several authentication providers via integrating several authentication strategies that provide (local or external) user management.	Access control to an instance of the TAXII API is specific to the sharing community, vendor, or product and is not defined by the specification.	Warden does not have a concept of users. Warden has the only concept of clients. Client management is not user-friendly.

ShRq-6	Authorization and access to data	Authorization and different access levels are supported. Access level groups can be modified and new ones can be created if needed. Every user has assigned a certain role, which allows him to do certain actions (read, write and publish data, sync MISP instances, edit templates, etc.). Moreover, every user is assigned to an Organization, which also influences his access rights to display or modify certain events and attributes in the MISP instance. Two-factor authentication is not available.	Full control of data access management using groups with permissions based on granular markings on both entities and relationships. Two-factor authentication is not available.	Access control to an instance of the TAXII API is specific to the sharing community, vendor, or product and is not defined by the specification.	Warden does not have a concept of roles. Warden provides everything to everybody. Two-factor authentication is not available. The client authenticates via a certificate.

ShRq-7	Storage of diverse data	MISP can store various types of data, the closest one to a 'glossary' are probably MISP Galaxies – lists of data, such as description of malware types, attack patterns or known threat actors. Each galaxy is basically a mapping of a term/value (with optional list of synonyms) to a description. Each 'event' can be linked to some of the terms from galaxies.	Supports various glossaries such as MITRE.	STIX standard provides mechanisms for its flexible extension and customization.	Native format IDEA is dedicated to bear alerts. Other type of messages are allowed but do not have taxonomy or any further technical support.
ShRq-8	Glossary	Such a glossary can be set up as a new MISP Galaxy (see above).	The system support custom markings.	The Glossary is defined by the STIX standard itself and the possible extensions and customizations.	No support of glossaries or permanent object.

ShRq-9	Clients/proxies	MISP platform supports API, which offers almost all actions, which can be done via web GUI. It includes search, creating, updating, and deleting events, attributes, user management, etc. All the API tasks can be automated with the python module PyMISP which supports all API actions. Every user has its own authentication key, which is used for accessing the API. Peer2peer synchronization between MISP instances is a core functionality of MISP server.	OpenCTI offers such API via Python or GO client.	There are several open-source STIX and TAXII libraries. TAXII is designed to operate in publish-subscribe mode between its peers.	Warden client allows to filter alerts based on categories, tags. It allows positive as well as negative filtration rules.

ShRq-10	Filtering	Sharing of data between MISP instances is governed by the ‘distribution level’ metadata attribute, which is assigned to every piece of data (events, misp-attributes, misp-objects). There are 5 levels – Your organisation only, This community only, Connected Communities, All communities, Sharing group. For example, if an event or an attribute has a distribution level set to 'This community only', then all users in the same MISP instance can view it, but it is not shared to any other instance. Also, no event is shared outside of the instance until it is ‘published’. Besides this, TLP tags (“traffic light protocol”, red, amber, green, white) are used to mark distribution possibilities of information outside of MISP.	It is not possible to fully assess this feature based on documentation. Due to the recent appearance of this system, there was no time for experiments.	The TAXII standard provides limited support for filtering.	Warden can share data with other communities. It does not support any distribution levels. Simply everything is shared.

		The client API also supports filters to search for specific data.			
ShRq-11	Auditing	Logging is available, the server logs almost every action, from user actions (like login) to event or attributes creation or update. Web GUI interface allows some basic search and filtering of audit logs. It does not support monitoring of the log for specific patterns.	Yes, logging is available in different log levels: debug info, warning, error. It does not support monitoring of the log for specific patterns.	The used auditing mechanism is specific to a particular TAXII server implementation.	Yes, logging is available but no further support.
ShRq-12	User roles/permissions	MISP platform supports different access levels. By default 'admin', which has all access rights, 'org admin' which maintains organization users, 'read-only', etc. Access level groups can be created and modified if needed.	OpenCTI platform supports different access levels and roles. By default 'admin', which has all access right.	Access control to an instance of the TAXII API is specific to the sharing community, vendor, or product and is not defined by the specification.	Warden has its admin and multiple clients. It allows everybody who can login to see everything what is inside.

ShRq-13	Encryption at transport	Both the API and the synchronization protocol are based on HTTP/HTTPS, so all the communication is encrypted if HTTPS is always used (which is recommended). Emails sent by MISP instance can be encrypted using GnuPG or SMIME if the user has set a corresponding key/cert in his/her profile (an instance can be configured to enforce encryption and never send unencrypted emails).	The communication protocol between a client and the server (web and API service) is based on HTTPS.	Encryption is supported via HTTPS transport.	The communication protocol between a client and the server is based on HTTPS.
SqRq-14	Encryption at rest	Data are stored in a MySQL database, encryption is not supported by MISP.	Data are stored unencrypted in the database at the server. OpenCTI use ElasticSearch (Database), redis (Events Stream) and MinIO (storage).	The internal storage is specific to a particular TAXII server implementation. The available implementations use standard persistence stores which usually support encryption.	Data are stored unencrypted in the database at the server.

SqRq-15	Maturity	MISP is mature enough capable of productional deployment. Moreover, it has significant development support.	Under heavy development.	Mature tools, as well as mature products (commercial), are available.	Warden is very stable being used productionally since 2012.
SqRq-16	Community	A large community of users as well as developers. CIRCL.LU offers training on MISP development.	Promising due to modern technologies used as its components. But so far the community is small	STIX is being widely used and deployed, well adopted by commercial organizations.	Only small community.

Although all the presented platforms provide multiple features for data sharing, storing, and analysis, in terms of sharing capabilities, the flexibility of the format, maturity, and adoption is the MISP platform prevalent. It provides an external API for the automation between machines and other platforms, also allowing to import and export data using various formats, including STIX standard. MISP includes a real-time publish-subscribe channel that enables entities to automatically obtain new events, indicators, sightings, or tagging. For this purpose, MISP uses the ZeroMQ and Apache Kafka capabilities. In particular, the ZeroMQ plugin of MISP operates at a global level, thus publishing all activities within the ZeroMQ pub-sub channels. MISP allows defining the distribution for each event before sharing it, thus sharing data with entities according to sharing levels and Traffic Light Protocol (TLP) Neither of the platforms meets the optional requirement for two-factor authentication.

5 Data model for sharing

To derive the data model for the sharing of incident handling the information we follow the methodology proposed in [15]. The methodology collects the pieces of information into a structure and offers a mapping mechanism onto MISP existing attributes, objects, taxonomies, or galaxies as well as it identifies the gaps where new objects, taxonomies, and galaxies should be introduced. The table 9 captures the important aspects of what should be shared.

Table 9: Methodology table with playbook sharing use case

Collection
Use case
In summary, the goal is to share response handling information. The details of the use case are described in the Introduction section of this deliverable and by the playbook requirements.
Data points
Playbook - playbook describes individual steps which are executed to handle an incident. Playbooks standard - identification of the playbook standard used to describe the playbook. Attack tactic (contextual, optional) - attack tactic of an adversary that the playbook reacts to. Attack technique (contextual, optional) - tool/technique an adversary uses to perform tactic, the playbook reacts on using the particular technique. Malware - (contextual, optional) - particular malware the playbook reacts to. IOC (contextual, optional) - indicator of compromise (e.g. hash of a file) the playbook reacts to. CVE (contextual, optional) - common vulnerability exposure the playbooks reacts to or fixes. Person role (contextual, optional) - roles of a users, that should be involved in the playbook. Organization type (contextual, optional) - type of an organization that the playbook is intended for. Confidentiality (optional) - indication if the playbook contains confidential information that should not be shared. Distribution (optional) - the distribution level defines who can receive the playbook. Playbook impact (optional) - how safe it is to execute the playbook from the perspective of breaking down or influencing infrastructure, services, users. Playbook abstraction (contextual, optional) - indicator of how abstract the playbook is. Whether the playbook is detailed enough to be automatically executed by a machine or requires mapping on the specifics of an organization, up to the playbook does not contain any input for automation and must be purely executed by a human operator. Playbook type (contextual, optional) - identifies type of actions in the playbook and what phases according to [1] the playbook type relates to. Created (optional) - date and time when the playbook was created by the author.
Sets

There are several candidate sets of data points that logically belongs together.

PlaybookSet = {Playbook, Attack tactic, Attack technique, IOC, CVE, Person role, Organization type, Confidentiality, Distribution, Playbook impact, Playbook abstraction, Playbook type}

PlaybookMetaData = {Attack tactic, Attack technique, IOC, CVE, Person role, Organization type, Confidentiality, Distribution, Playbook impact, Playbook abstraction, Playbook type}

PlaybookTriggers = {Attack tactic, Attack technique, IOC, CVE}

Relationships

Data point/set	*relationship*	*Data point/set*
PlaybookTriggers	trigger	Playbook
PlaybookMetadata	characterize	Playbook

Individually, the data points in PlaybookMetaData are related to the Playbook data point as well.

Mapping

Types and categories

Assign types and categories to data points.

Playbook - missing type in MISP - string containing the serialized playbook, we propose new type of attribute called playbook. The playbook attribute itself will be represented in prevalent cases either in JSON or XML format as a string. For example the JSON scheme captures the playbook expressed as OWL graph based on the D4.1 as well as CACAO playbooks while the XML captures playbooks defined by IACD.

Playbook-standard - a taxonomy of playbook standard formats is missing due to missing standard until recently published [4].

Playbook-type - the type of the playbook should be expressed as a machine tag using taxonomy.

Attack tactic - candidate for galaxy.

Attack technique - candidate for galaxy.

IOC - MISP contains plenty of types that can represent an IOC. The particular IOC type should be selected to represent given IOC and added as an attribute.

CVE - identifier of CVE - missing type in MISP

Person-role - missing in MISP - (management, analysis, development, incident handling, legal support, communication, network administration, other)

Organization type - prevalent type of an organization (critical infrastructure, governmental, enterprise, education, research, corporate, SME, ISP, MSP), candidate for galaxy.

Distribution - maps on distribution levels native to MISP

Playbook-impact - optional attribute defined as integer from 0 to 100 according to [4], or magnitude of impact on organization's mission/business defined as taxonomy - levels: none, low - executing the playbook will impact with low chance or will impact low number of users or non-critical services, medium - impact with medium chance or number of users or internal services, high - impact with high chance or high number of users or some business services, critical - severe impact on large number of users or on business critical service.

Published - this maps on the timestamps that are inherent to each event in MISP.

Playbook-abstraction - is a candidate for taxonomy.

Created - created timestamp does not map on any timestamp inherent in MISP event therefore dedicated attribute of datetime type will be added.

Objects

There are no objects dedicated to playbook sharing in MISP.

Tags, taxonomies, galaxies

There are several relevant galaxy and taxonomies already available in MISP:

- Galaxy 6 Attack pattern
- Galaxy 46 Techniques
- Galaxy 20 Malpedia
- Galaxy 28 Course of Action
- Galaxy 13 Sector
- Taxonomy file-type
- Taxonomy tlp
- Taxonomy iep

Attack tactic - missing Galaxy in MISP - to represent https://attack.mitre.org/tactics/enterprise/

Attack technique - maps on Galaxy 6 - Attack pattern (by MITRE) and Galaxy 46 - Techniques (by MITRE, newly available in MISP since January 2021).

The file-type taxonomy can capture the format of the playbook. This taxonomy already defines XML type but the taxonomy should be extended with JSON type.

Malware - maps on Galaxy 20 - Malpedia.

Confidentiality - maps on TLP or IEP (IEP2) taxonomies available in MISP.

Playbook-type - missing taxonomy in MISP. The taxonomy can be based on [4] can serve as a basis for the taxonomy can be considered.

Organization-type - maps on Galaxy 6 but some sectors might be missing, also this galaxy should be put inline with [4].

Playbook abstraction - missing taxonomy in MISP. The taxonomy can be partially based on [4] but there are only two levels of playbook-template or playbook. We envision a more fine-grained definition (guideline, playbook template, playbook, partial workflow, full workflow, fully scripted - a script interpretable e.g. by Python).

Relationships

There is no need for the explicit expression of relationships.

Other relevant features

Enabling MISP correlation of the IOC or CVE attributes will link the particular playbook event with other MISP events containing the same IOC, CVE. Thus, the person/machine who is handling the IOC/CVE event immediately can have a look how to deal with it. But if the correlation is enabled for these attributes all the events with the same IOC or CVE will be correlated with each other. This may result in a large set of correlated events where the playbook event will be lost in the number of other events. Therefore, the correlation upon IOC or CVE should be enabled only in such MISP instances where the data contain versatile IOCs or CVEs. Or the same IOC/CVE are specifically structured in such a way that they map the playbook event and only the sighting is increased.

Also, it is worth considering creating a dedicated MISP object for playbooks which would allow for better structure and embedding into other events. The experiments with sharing the playbooks will provide us enough feedback to assess typical usage and decide on the structure of the object for the playbooks.

6 Sharing interface

There will be two interfaces, API and GUI, how the information about response handling information can be uploaded into the MISP, or consumed from MISP.

From the perspective of the architecture defined in D2.4, there are two components communicating with MISP - Intelligence Consumer (IC) and Intelligence Provider (IP) proxy. The proxies de facto implements the application programming interface between MISP and other specific SAPPAN components.

The IC proxy:

- connects to MISP,
- listens for the new events in MISP,
- filters the events according to the defined ruleset,
- reads and stores the events locally.

The IP proxy:

- connects to MISP,
- loads the event from local storage,.
- applies sensitivity and privacy functions,
- sends the resulting playbook into MISP.

The IC and IP proxies will be implemented using the libraries and methods that are supported by MISP. The IP proxy will connect to certain MISP instance using the MISP Automation API key. The IP proxy will pick the playbook relevant for sharing and applies sanitization and anonymization functions to remove or modify confidential or private information. The IP proxy then pushes the playbook via PyMISP into the MISP instance. The IC proxy will connect to certain MISP instance using the MISP Automation API key. It will register to its ZeroMQ notifications about newly published events. The IC proxy will filter notifications based on JSONpath query language, so that only the events containing response handling information are considered. The IC proxy will store each event in a separate file in a directory that will serve as a repository of the shared playbooks. The concept is depicted in Figure 3.

When the sharing of a playbook is performed manually via GUI then the sanitization and anonymization functions must be applied manually at the side of the playbook producer before the playbook is inserted in the GUI. The producer just calls the respective sanitization and anonymization functions as scripts that read the playbook and return the modified version of the playbook with sanitized and anonymized information.

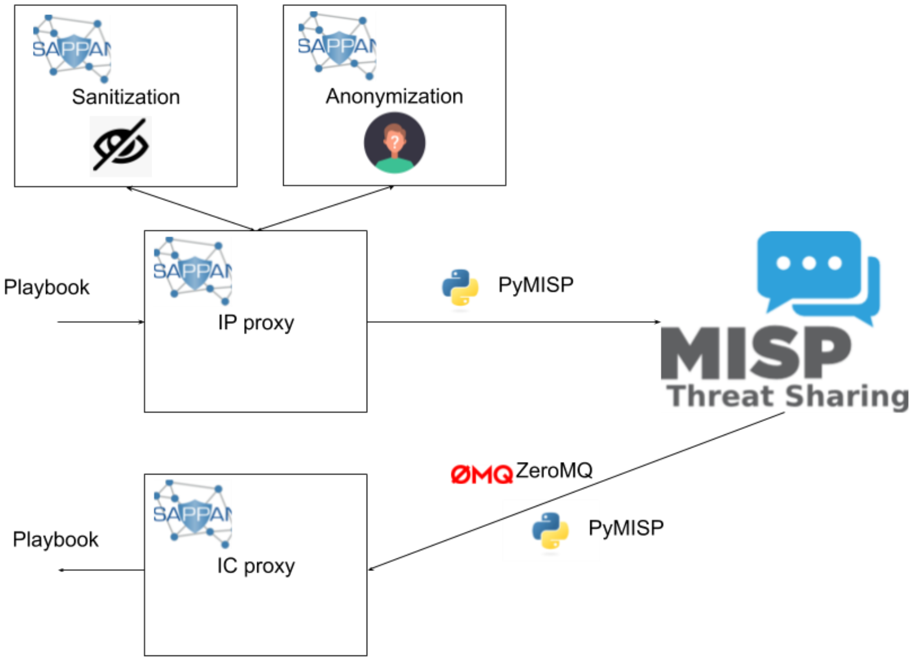

Figure 3: Basic concept of communication between producer/consumer and MISP

7 Meeting playbook sharing requirements

In Table 10, we document how the requirements, defined in Tables 4, 5, 6, 7, are met.

Table 10: Meeting the playbook sharing requirements

ID	Name	Meeting the requirement
PrRq-1	Limit the recipients	MISP implements distribution levels to define if a playbook can reach a particular group of consumers or individual consumers. The distribution level natively supports these groups: only users from the same organization as is the producer, or only enumerated organizations or particular group, or all the users registered in the given MISP instance. At the producer part, i.e. at the IP proxy, the playbook events must be marked with tlp or iep tag to indicate the consumers how to further disseminate, if at all, the playbooks.
PrRq-2	No personal information	The anonymization function applied to the playbook takes care of removing or anonymizing (as defined by the input parameter) PII. The function omits anonymization only in cases when the step is labeled as do not anonymize. The function looks for typical PII identifiers and values such as email address, telephone number, login.
CoRq-1	No sensitive identifiers leak	The sanitization function applied to the playbook takes care of removing or abstracting (as defined by the input parameter) sensitive identifiers. The sanitization function is applied if the confidentiality level of the playbook is labeled at least as partially confidential. Based on the level of sanitization. In such a case it must be sanitized before it is shared.
CoRq-2	No tools leak	This is a more specific requirement to the CoRq-1. The sanitization function removes or abstracts all the specific tools referenced in the playbook.
CoRq-3	No infrastructure elements leak	This is a more specific requirement to the CoRq-1. The sanitization function removes or abstracts all the specific infrastructural elements referenced in the playbook.
CoRq-4	Detail to abstract	The sanitization function should use a vocabulary defined in D4.2 to enable translation from local/detailed level into more abstract. For example, tcpdump → packet analyzer → analyze packet capture → investigate network traffic.
CoRq-5	Abstract to detail	This requirement is reverse to the detail to abstract and should serve for consumers to map the abstract identifiers/actions to the local policies.

CoRq-6	Confidential step	The sanitization function removes or deletes content (based on parameters) if the step is marked confidential.
CoRq-7	Anonymity of producer	The sanitization function must remove the author/organization identifying fields from the playbook description, moreover, the MISP platform does not support anonymous sharing as such. Partially, this requirement can be met by using pseudo-anonymization which can be achieved by delegating publishing of an event to another organization as documented by MISP https://www.misp-project.org/features.html or https://www.misp-project.org/compliance/ISO-IEC-27010/.
AuRq-1	Human readable	The MISP event containing the playbook is machine-readable due to its structure. Within the SAPPAN project, we will use the model of the playbook defined in D4.1, this model can be represented both as JSON and XML (but we choose JSON as our primary format). The JSON format is to some extent human-readable but definitely not convenient. To this end, SAPPAN has developed a tool to convert playbooks represented as JSON into a visual graph-based representation.
AuRq-3	Machine-readable	This requirement is met by using a structured approach to share playbooks via MISP and representing the playbooks in JSON format which can be easily processed by a machine.
AuRq-4	Machine interpretable	The JSON format is flexible enough to support various levels of details of playbooks. Moreover, using the vocabularies defined in D4.2 it is in some cases possible to map the pre-defined values on organization-specific procedures, tools, and infrastructure elements. The level of the playbook helps the consumer to estimate how well the playbook can be interpreted by the machine.
AuRq-5	Lists of defined values	The SAPPAN playbooks are using pre-defined values from D4.2 whenever possible.
AuRq-6	Mapping	Mapping of different tools with the same functionality should be to some extent possible using the vocabularies as well as to perform detail to abstract and back to detail translation using the vocabularies.
AuRq-7	Visualization	SAPPAN provides a tool for visualization of playbooks represented in JSON format defined in D4.1.
AuRq-8	Filtering	MISP provides the capability to filter the events based on all the attributes, tags, galaxies both in GUI as well as in the API, where we add special support to be able to filter notifications using JSON path query language.

DaRq-1	Format identification	SAPPAN will tag the playbook with its format type, i.e. JSON. The playbook itself must contain identification of the schema that was used to create the playbook. As the playbook will be extracted and stored from the MISP event it must be self-descriptive.
DaRq-2	Support of various playbook formats	The attribute playbook of the MISP event can store any playbook expressed as a text string. I.e. it supports all formats that can be serialized into a text string. The particular format type expressed by the taxonomy file-type is either XML or JSON but the taxonomy can be extended or changed to a different taxonomy. Also, the additional attributes with metadata are optional and therefore do not prevent using a different playbook format.
DaRq-3	Parameters to filter	MISP GUI and API provides sufficient filter capabilities to select playbooks relevant to the user based on the attributes in the event.
DaRq-4	Extensible	The MISP data model allows for extensions by default. It does not enforce the rigid structure of the event.

8 Example

We take one of our playbooks for dealing with DGA detection and we draft a MISP event based on the proposed data model. The playbook is represented in Figure 3.

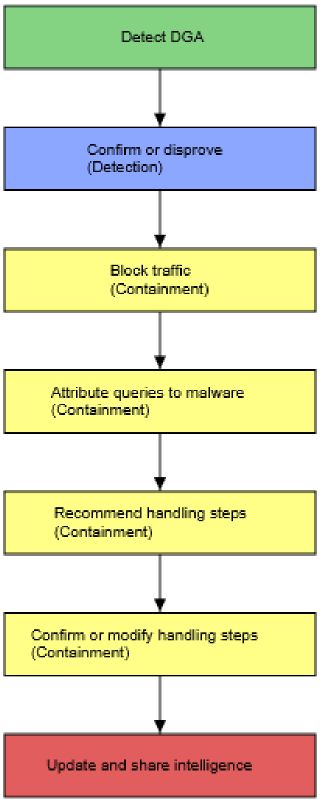
Detect DGA
Confirm or disprove (Detection)
Block traffic (Containment)
Attribute queries to malware (Containment)
Recommend handling steps (Containment)
Confirm or modify handling steps (Containment)
Update and share intelligence

Figure 3: Simple DGA playbook.

The DGA plabyook is expressed in a shortened version as the MISP event below, the full MISP event is captured in the Attachment chapter.

```
"Event":
   {
      "id": "12345",
      "orgc_id": "1",
      "org_id": "1",
      "date": "2021-01-07",
      "threat_level_id": "4",
      "info": "General guideline when DGA is detected",
      "published": true,
      "uuid": "20fca1d2-a2e1-45c9-8d90-be7cabd41916",
      "attribute_count": "34",
      "analysis": "2",
"timestamp": "1610019001",
"distribution": "1",
"proposal_email_lock": false,
"locked": false,
"publish_timestamp": "1610019004",
"sharing_group_id": "0",
"disable_correlation": false,
"extends_uuid": "",
"event_creator_email": "author@company.com",
"Org": {
    "id": "1",
    "name": "COMPANY",
    "uuid": "2d216963-3dc8-4132-b747-5fbde7aa501f",
    "local": true
},
"Orgc": {
    "id": "1",
    "name": "COMPANY",
    "uuid": "2d216963-3dc8-4132-b747-5fbde7aa501f",
    "local": true
},
"Attribute": [
    {
        "id": "865",
        "type": "date-time",
        "category": "Other",
        "...
        "value": "2021-01-07T02:53:01CET",
        "Galaxy": [],
        "ShadowAttribute": []
    },
    {
        "id": "866",
        "type": "playbook",
        "category": "Other",
        "to_ids": true,
        "uuid": "f6c5f7e9-cbfb-4b9c-8953-43a6ca52cb71",
        "event_id": "12345",
        "distribution": "5",
        "timestamp": "1610019002",
        "comment": "",
        "sharing_group_id": "0",
        "deleted": false,
        "disable_correlation": false,
        "object_id": "0",
        "object_relation": null,
        "first_seen": null,
        "last_seen": null,
        "value": "{\"Steps\":\
                    [{\"0\":[{\"Name\":\"Detect DGA\"},{\"Category\":\"InitialStep\"},\
                     ….
                     {\"ConfidentialityLevel\":\"Public\"},{\"Reporter\":\"\"}]}",
        "Galaxy": [],
        "ShadowAttribute": []
      },
      {
        "id": "867",
        "type": "playbook-impact",
        "category": "Other",
        "to_ids": true,
        "uuid": "920df7ec-67ef-4eb0-b718-e4a9e264a98b",
        "event_id": "60",
        "distribution": "5",
        "timestamp": "1610019002",
        ...
        "value": "40",
        "Galaxy": [],
        "ShadowAttribute": []
      }
    ],
    "ShadowAttribute": [],
    "RelatedEvent": [],
        "EventReport": [],
        "Tag": [
            {
                "id": "2678",
                "name": "playbooks:standard=\"SAPPAN playbook format version 1.0\"",
                ...
            },
            {
                ....
                "name": "playbook:type=\"Investigation\"",
                ...
            },
...
    }
```

9 Summary and plans

This deliverable documented our implementation activity toward sharing response handling information in the form of playbooks. It summarized the existing work regarding response handling information and the latest development in this area. The deliverable collected general as well as specific requirements on sharing system and its related components, functions and data model. The deliverable proceeded with a short survey and assessment of promising sharing systems to select the best fit for SAPPAN use cases. Further, the deliverable introduced the MISP data model for playbook sharing and proposed an interface MISP and relevant SAPPAN components. In the end, the deliverable evaluated how the specific playbook-sharing requirements are met by MISP, by the designed interface between MISP and SAPPAN components, and the proposed data model. The deliverable presented also an example of a playbook expressed in the proposed data model for MISP.

We believe that the integration of playbook sharing into MISP is the right choice not only as a proof-of-concept for the SAPPAN project but also in a long run beyond the end of the project due to the wide adoption of MISP. Moreover, the standardization of the playbooks indicates that this may become a commercially interesting field, e.g. similarly to IDS ruleset we envision that there will be open playbooks as well as commercial playbooks where MISP can serve both of them as a vehicle to deliver the playbooks to users.

Our plan in T5.4 is to further evolve the integration into MISP and also to evolve sanitization and anonymization capabilities that are proprietary to playbooks. Also, the recent publication (in January 2021) of a standard for playbook description by OASIS [4] and the presentation - Cybersecurity standardization within the "Horizon” (by Dr. Vasileios Mavroeidis) during [16] highlighted the need for pushing the standard and building a community. To this end, we plan to join OASIS CACAO Technical Committee to help to form the standard towards sharing as well as to contribute with our proof-of-concept playbook sharing integration experience into MISP. To disseminate the results of T5.4 we plan to introduce our work to MISP developers and to propose the merging of the perspective results into the official MISP repository. We will document our achievements in the next version of this deliverable which is planned in M30.

10 References

[1] Cichonski, Paul & Millar, Tom & Grance, Tim & Scarfone, Karen. (2012). NIST Special Publication 800-61 Revision 2, Computer Security Incident Handling Guide. http://dx.doi.org/10.6028/NIST.SP.800-61r2

[2] Integrated adaptive cyber defense, 2021, available online : https://www.iacdautomate.org/

[3] Incident Response Consortium. Playbooks, 2021, available online: https://www.incidentresponse.com/playbooks/

[4] OASIS Collaborative Automated Course of Action Operations (CACAO) for Cyber Security TC. CACAO Security playbooks specification v1.0. Available online: https://docs.oasis-open.org/cacao/security-playbooks/v1.0/cs01/security-playbooks-v1.0-cs01.html

[5] Blake E. Strom, Andy Applebaum, Doug P. Miller, Kathryn C. Nickels, Adam G. Pennington, Cody B. Thomas. MITRE ATT&CK: Design and philosophy, 2018.

[6] Awesome SOAR, available online: https://github.com/correlatedsecurity/Awesome-SOAR

[7] Awesome workflow engines, available online: https://github.com/meirwah/awesome-workflow-engines

[8] Awesome pipeline, available online: https://github.com/pditommaso/awesome-pipeline

[9] ElasticSearch, available online: https://www.elastic.co/

[10] RabbitMQ, available online: https://www.rabbitmq.com/

[11] GraphQL, available online: https://graphql.org/

[12] Redis, available online: https://redis.io/

[13] The Hive, available online: https://github.com/TheHive-Project/TheHive

[14] MITRE CTI, available online: https://github.com/mitre/cti

[15] D4.1: Cybersecurity threat intelligence common data model, available online: https://www.sparta.eu/assets/deliverables/SPARTA-D4.1-Cybersecurity-threat-intelligence-common-data-model-PU-M18.pdf

[16] CyberSANE: Joint standardization workshop Dynamic Countering of Cyber-attacks projects in January 2021

11 Attachment

Complete MISP Event capturing basic DGA response playbook.

```
  "Event":
   {
      "id": "12345",
      "orgc_id": "1",
      "org_id": "1",
      "date": "2021-01-07",
      "threat_level_id": "4",
      "info": "General guideline when DGA is detected",
      "published": true,
      "uuid": "20fca1d2-a2e1-45c9-8d90-be7cabd41916",
      "attribute_count": "34",
      "analysis": "2",
      "timestamp": "1610019001",
      "distribution": "1",
      "proposal_email_lock": false,
      "locked": false,
      "publish_timestamp": "1610019004",
      "sharing_group_id": "0",
      "disable_correlation": false,
      "extends_uuid": "",
      "event_creator_email": "author@company.com",
      "Org": {
         "id": "1",
         "name": "COMPANY",
         "uuid": "2d216963-3dc8-4132-b747-5fbde7aa501f",
         "local": true
      },
      "Orgc": {
         "id": "1",
         "name": "COMPANY",
         "uuid": "2d216963-3dc8-4132-b747-5fbde7aa501f",
         "local": true
      },
"Attribute": [
  {
    "id": "865",
    "type": "date-time",
    "category": "Other",
    "to_ids": false,
    "uuid": "2efd0ab9-4441-46b4-a5cd-41c0a4d980f9",
    "event_id": "12345",
    "distribution": "5",
    "timestamp": "1610019002",
    "comment": "",
    "sharing_group_id": "0",
    "deleted": false,
    "disable_correlation": false,
    "object_id": "0",
    "object_relation": null,
    "first_seen": null,
    "last_seen": null,
    "value": "2021-01-07T02:53:01CET",
    "Galaxy": [],
    "ShadowAttribute": []
  },
  {
    "id": "866",
    "type": "playbook",
    "category": "Other",
    "to_ids": true,
    "uuid": "f6c5f7e9-cbfb-4b9c-8953-43a6ca52cb71",
    "event_id": "12345",
    "distribution": "5",
    "timestamp": "1610019002",
    "comment": "",
    "sharing_group_id": "0",
    "deleted": false,
    "disable_correlation": false,
    "object_id": "0",
        "object_relation": null,
        "first_seen": null,
        "last_seen": null,
        "value": "{\"Steps\":\
[{\"0\":[{\"Name\":\"Detect DGA\"},{\"Category\":\"InitialStep\"},\
{\"Description\":\"The classifiers detect an algorithmically generated domain name\"},\
{\"ConfidentialDescription\":\"\"},{\"Mean\":\"None\"},{\"ConfidentialityLevel\":\"Public\"}{\"AllowedUsers\":\"\"}]},\
{\"1\":[{\"Name\":\"Confirm or disprove\"},{\"Category\":\"IntermediateStep\"},\
{\"Description\":\"The analyst confirms or disproves the incident. The triage can be omitted in case of high-confidence classification.\"},\
{\"ConfidentialDescription\":\"\"},{\"Mean\":\"Detection\"},{\"IndicatorOfCompromise\":\"\"},\
{\"ConfidentialityLevel\":\"Public\"},{\"AllowedUsers\":\"\"}]},\
{\"2\":[{\"Name\":\"Block traffic\"},{\"Category\":\"IntermediateStep\"},\
{\"Description\":\"Network traffic for the infected host is blocked until malware is removed.\"},\
{\"ConfidentialDescription\":\"\"},{\"Mean\":\"Containment\"},{\"IndicatorOfCompromise\":\"\"},\
{\"ConfidentialityLevel\":\"Public\"},{\"AllowedUsers\":\"\"}]},\
{\"3\":[{\"Name\":\"Attribute queries to malware\"},{\"Category\":\"IntermediateStep\"},\
{\"Description\":\"The malicious DNS queries are attributed to a malware by another classifier.\"},\
{\"ConfidentialDescription\":\"\"},{\"Mean\":\"Containment\"},{\"IndicatorOfCompromise\":\"\"},\
{\"ConfidentialityLevel\":\"Public\"},{\"AllowedUsers\":\"\"}]},\
{\"4\":[{\"Name\":\"Recommend handling steps\"},{\"Category\":\"IntermediateStep\"},\
{\"Description\":\"Based on the detected malware, handling steps are recommended.\"},\
{\"ConfidentialDescription\":\"\"},{\"Mean\":\"Containment\"},{\"IndicatorOfCompromise\":\"\"},\
{\"ConfidentialityLevel\":\"Public\"},{\"AllowedUsers\":\"\"}]},\
{\"5\":[{\"Name\":\"Confirm or modify handling steps\"},{\"Category\":\"IntermediateStep\"},\
{\"Description\":\"The analyst confirms or modifies the handling steps.\"},\
{\"ConfidentialDescription\":\"\"},{\"Mean\":\"Containment\"},{\"IndicatorOfCompromise\":\"\"},\
{\"ConfidentialityLevel\":\"Public\"},{\"AllowedUsers\":\"\"}]},\
{\"6\":[{\"Name\":\"Update and share intelligence\"},{\"Category\":\"FinalStep\"},\
{\"Description\":\"Intelligence of known malicious domain names and DGAs (e.g. DGArchive) is updated and shared\"},\
{\"ConfidentialDescription\":\"\"},{\"Mean\":\"None\"},\
{\"ConfidentialityLevel\":\"Public\"},{\"AllowedUsers\":\"\"}]}],\
\"Connections\":[{\"0\":[{\"From\":\"Detect DGA\"},{\"To\":\"Confirm or disprove\"}]},\
{\"1\":[{\"From\":\"Confirm or disprove\"},{\"To\":\"Block traffic\"}]},\
{\"2\":[{\"From\":\"Block traffic\"},{\"To\":\"Attribute queries to malware\"}]},\
{\"3\":[{\"From\":\"Attribute queries to malware\"},{\"To\":\"Recommend handling steps\"}]},\
{\"4\":[{\"From\":\"Recommend handling steps\"},{\"To\":\"Confirm or modify handling steps\"}]},\
{\"5\":[{\"From\":\"Confirm or modify handling steps\"},{\"To\":\"Update and share intelligence\"}]}]},\
\"Playbook\":[{\"Incident\":\"\"},{\"AttackCategory\":\"Phishing\"},{\"Vulnerability\":\"\"},\
{\"IndicatorOfCompromise\":\"\"},{\"ConfidentialityLevel\":\"Public\"},{\"Reporter\":\"\"}]}",
             "Galaxy": [],
            "ShadowAttribute": []
        },
        {
            "id": "867",
            "type": "playbook-impact",
            "category": "Other",
            "to_ids": true,
            "uuid": "920df7ec-67ef-4eb0-b718-e4a9e264a98b",
            "event_id": "60",
            "distribution": "5",
            "timestamp": "1610019002",
            "comment": "",
            "sharing_group_id": "0",
            "deleted": false,
            "disable_correlation": false,
            "object_id": "0",
            "object_relation": null,
            "first_seen": null,
            "last_seen": null,
            "value": "40",
        "Galaxy": [],
        "ShadowAttribute": []
    }
],
"ShadowAttribute": [],
"RelatedEvent": [],
"EventReport": [],
"Tag": [
    {
        "id": "2678",
        "name": "playbooks:standard=\"SAPPAN playbook format version 1.0\"",
        "colour": "#00bdbd",
        "exportable": true,
        "user_id": "0",
        "hide_tag": false,
        "numerical_value": null,
        "is_galaxy": false,
        "is_custom_galaxy": false,
        "local": 0
    },
    {
        "id": "2004",
        "name": "playbook:type=\"Investigation\"",
        "colour": "#285c00",
        "exportable": true,
        "user_id": "0",
        "hide_tag": false,
        "numerical_value": null,
        "is_galaxy": false,
        "is_custom_galaxy": false,
        "local": 0
    },
    {
        "id": "2123",
        "name": "playbook:type=\"Mitigation\"",
        "colour": "#285c00",
        "exportable": true,
        "user_id": "0",
        "hide_tag": false,
        "numerical_value": null,
        "is_galaxy": false,
        "is_custom_galaxy": false,
        "local": 0
    },
    {
        "id": "2746",
        "name": "misp-galaxy:Sector=\"IT\"",
        "colour": "#0088cc",
        "exportable": true,
        "user_id": "0",
        "hide_tag": false,
        "numerical_value": null,
        "is_galaxy": true,
        "is_custom_galaxy": false,
        "local": 0
    },
    {
        "id": "2847",
        "name": "misp-galaxy:mitre-attack-pattern=\"Domain Generation Algorithms - T1520\"",
        "colour": "#0088cc",
        "exportable": true,
        "user_id": "0",
        "hide_tag": false,
        "numerical_value": null,
        "is_galaxy": true,
        "is_custom_galaxy": false,
        "local": 0
    },
    {
        "id": "2343",
        "name": "file-type:type=\"JSON\"",
            "colour": "#0088cc",
            "exportable": true,
            "user_id": "0",
            "hide_tag": false,
            "numerical_value": null,
            "is_galaxy": false,
            "is_custom_galaxy": false,
            "local": 0
        },
        {
            "id": "2133",
            "name": "tlp:white",
            "colour": "#0088cc",
            "exportable": true,
            "user_id": "0",
            "hide_tag": false,
            "numerical_value": null,
            "is_galaxy": false,
            "is_custom_galaxy": false,
            "local": 0
        }
        {
            "id": "2133",
            "name": "playbook:abstraction=\"Guideline\"",
            "colour": "#0088cc",
            "exportable": true,
            "user_id": "0",
            "hide_tag": false,
            "numerical_value": null,
            "is_galaxy": false,
            "is_custom_galaxy": false,
            "local": 0
        }
    ]
}
```