Sharing and Automation for
Privacy Preserving Attack Neutralization

(H2020 833418)

D5.8 Sharing Response Handling Information (M30)

Published by the SAPPAN Consortium

Dissemination Level: Public

H2020-SU-ICT-2018-2020 – Cybersecurity

Document control page

Document file: D5.8
Document version: 2.0
Document owner: Martin Zadnik (CESNET)

Work package: WP5
Task: T5.4 Sharing Response Handling Information
Deliverable type: Report
Delivery month: M30
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
1.0	Martin Zadnik (CESNET)	2020-01-31	Outline of the document
1.1	Martin Zadnik (CESNET)	2021-06-04	Initial discussion with CACAO
1.2	Martin Zadnik (CESNET)	2021-09-18	Draft version
1.3	Martin Zadnik (CESNET)	2021-10-22	Internal review version
1.4	Martin Zadnik (CESNET)	2021-10-28	Integrating reviewers' comments
2.0	Martin Zadnik (CESNET)	2021-10-29	Final version for submission

Internal review history:

Reviewed by	Date	Summary of comments
Lasse Nitz (FIT)	2021-10-25	Grammar and rephrasing
Mischa Obrecht (DL)	2021-10-26	Technical, story and grammar
Mehdi Akbari Gurabi (FIT)	2021-10-27	Technical and grammar

Executive Summary

Task T5.4 aimed at enabling sharing of response handling information. This is an updated version of the initial deliverable that describes our implementation activity toward sharing response handling information in the form of cybersecurity playbooks. The initial version already contained an overview of the existing work and development in this area. The initial version also collected general as well as playbook-specific requirements on the SAPPAN sharing system and its related components, functions and data model. The initial version outlined how cybersecurity playbooks can be shared in MISP. Since the first version, we got in touch with the Technical Committee (TC) of the OASIS specification Collaborative Automated Course of Action Operations (CACAO). We proposed several possible options for playbook sharing. We organized several joint teleconferences with the TC of CACAO to figure out how to best share the playbooks, what the use cases are for playbook sharing and whether they fit the proposed solution. The result is a data object template available in the official MISP repository, joint technical paper submitted and planned integration testing.

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 is dealt with. The documented systematic response enables the replication of the handling process, supports its monitoring and 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. 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 organizations 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, such as [3], 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 confidential details about an organization's handling procedures, tools, response capabilities, and infrastructure, as well as private or sensitive information (e.g. login and password).

1.1 Scope

The response handling information is 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 remainder 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 range from breaching some internal policy, finding a new vulnerability, detecting an indicator of a compromise to the reception of intelligence about a new threat or threat actor. The trigger also forms an initial hypothesis that something is wrong (or initial hypotheses but let us for the sake of simplicity consider only a single hypothesis at a time). 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, for example, 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 primary goal of the playbooks is to be executed while 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, task T4.1 developed a methodology for formalizing and modeling 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 modeling, such as the Resource Description Framework (RDF), RDF Schema (RDFS), and the Web Ontology Language (OWL). Alongside deliverable D5.7 (as the initial version of this deliverable), task T4.1 prepared (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. When the initial version of this deliverable was submitted we noticed the first draft of TC CACAO which was published in the same month (January 2021). This gave us a unique opportunity to consider both playbook formats and prepare the playbook data model to be ready for various formats.

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 modeling 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, namely the CACAO playbook. 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. [4]) but these playbooks are too abstract and may serve only as background material for the human operators or for constructing more detailed playbooks. However, there are efforts aiming at automatization of playbook execution in the Security Orchestration and Automation Response (SOAR) domain and we give a brief overview of these efforts.

Integrated Adaptive Cyber Defense (IACD) [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 and implement interfaces to other control and monitoring services and humans as is 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 not technology-agnostic 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 [5]. 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 of which focus on the cybersecurity domain and 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	D4.2
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 (links are provided in Table 2). These systems are promising candidates for sharing various pieces of information in SAPPAN. MISP is 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 straightforward 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. The attributes are of various categories and types, e.g. an attribute of type bank-account-nr belongs to the financial fraud category. 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 classifications such as MITRE [9]. MISP has a protocol used for synchronization between different MISP instances. The MISP platform allows defining the distribution of the cyber threat intelligence (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 representing 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 decided to 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 [10], RabbitMQ [11], GraphQL [12], redis [13]. It is developed as an open-source project. It focuses on processing, structuring and correlation of technical and non-technical CTI. Its schema is based on the STIX2.1 format. It can be integrated with other tools and projects such as MISP [14], The Hive [15], MITRE CTI [16]. OpenCTI aims at being a holistic tool for the integration of technical as well as non-technical information. OpenCTI supports the processing of structured and unstructured data, enabling inference of meaningful knowledge from the raw data. OpenCTI supports export and import into various formats such as CSV and STIX2 via its connectors.

Warden is a very simple sharing system built 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 the IDEA format. IDEA is a straightforward 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 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).

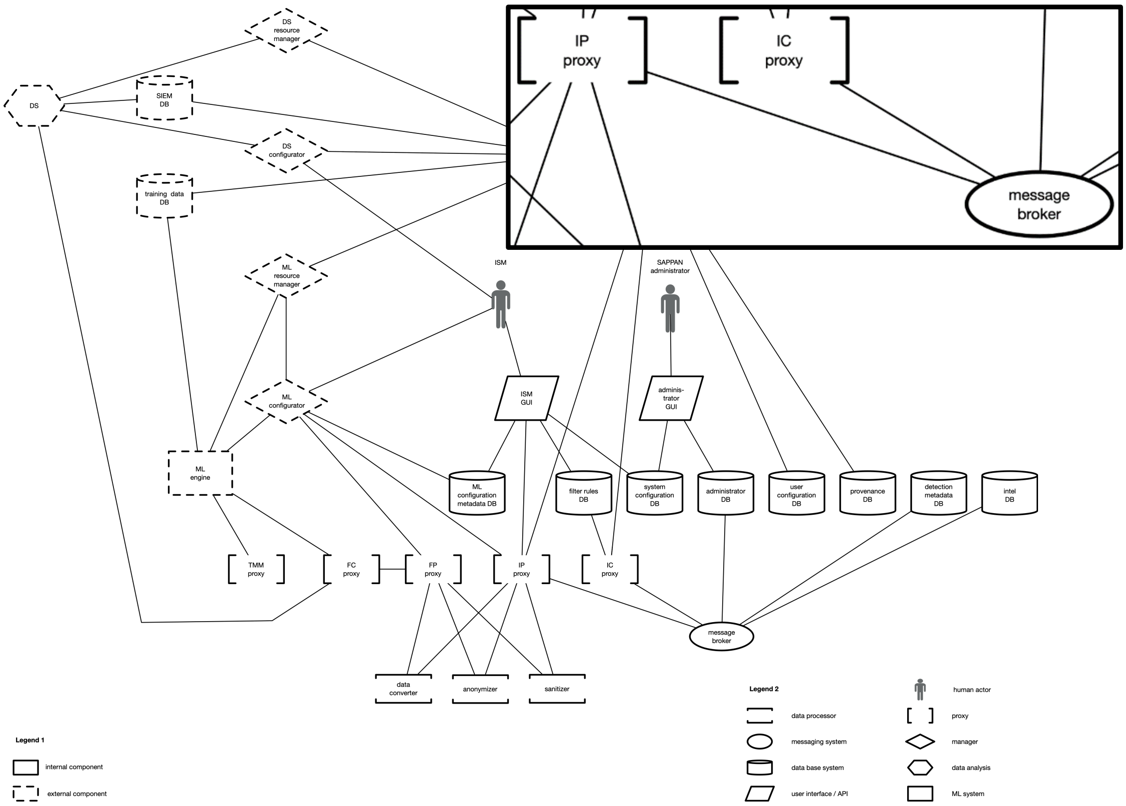

Figure 1: SAPPAN Technology Architecture Model diagram

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.

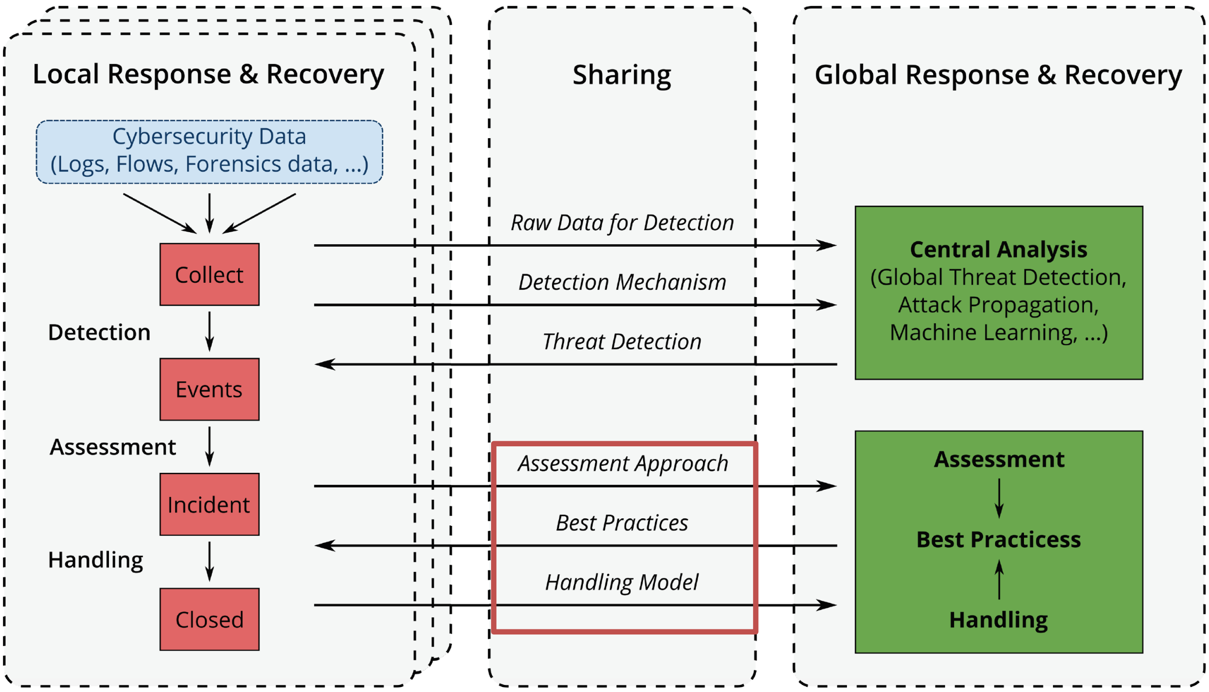

Figure 2: Sharing of various data in SAPPAN

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 to 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 participants who are entitled to work with given personally identifiable information (PII). In this regard, sharing of playbooks is no exception. 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 be preserved in the playbook that is shared. On the other hand, when sharing the playbook within the group of entitled persons in a single organization, privacy is maintained inherently and privacy-protective actions should not be applied as the private information is very likely relevant information for others. 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) employee 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 an abstraction, e.g. using vocabularies to translate a playbook from an organization-specific into an abstract playbook.
CoRq-5	Abstract to detail	COULD	This requirement is a 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 specifics of the consumer's organization.
CoRq-6	Confidential step	MUST	If the step is marked confidential then it must be left blank, even if it does not contain sensitive or private information.
CoRq-7	Anonymity of producer	COULD	A 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 interpre-table	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 part of sharing, the mapping of various tools with the same functionality may be supported such 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 quickly 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 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 the 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 below captures how the requirements are met. 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 is 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 is based on object versioning and revocation as a mechanism for updating and deleting the data.	Warden does not allow the 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	The 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, which allows for extensions.	The 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. An 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 (including the last login if the user has PGP key set, subscription to auto alerts, etc.).	OpenCTI supports several authentication providers by 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. It only knows the concept of a 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.	The STIX standard provides mechanisms for its flexible extension and customization.	The native format IDEA is dedicated to bare alerts. Other types 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 objects.

ShRq-9	Clients/proxies	The MISP platform supports an API, which offers almost all actions, which can be done via the 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 a respective 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.	The Warden client allows filtering 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 an attribute creation or update. The web GUI interface allows some basic search and filtering of audit logs. It does not support monitoring of the log for specific patterns.	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.	Logging is available, but no advanced options are supported.
ShRq-12	User roles/permissions	The 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.	The OpenCTI platform supports different access levels and roles. The default role is admin, which has all access rights.	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 log in to see everything that 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 for 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 in its components. But so far the community is small.	STIX is being widely used and deployed, and well adopted by commercial organizations.	Warden is only used by a small community.

Although all the presented platforms provide to some extent 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 allows to import and export data using various formats, including the 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 the 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 information we follow the methodology proposed in [17]. The methodology collects the pieces of information in a structured way. It identifies the data points and offers a mapping mechanism onto the existing MISP attributes, objects, taxonomies, or galaxies. Also, it detects the gaps where new objects, taxonomies, and galaxies should be introduced. The table below captures the important aspects of what should be shared or taken into account during the sharing process. 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. The use case further considers the formats enumerated in the related work.
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 an attack, the playbook reacts to. 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 disrupting down or influencing infrastructure, services, users. Playbook abstraction (contextual, optional) - indicator of how abstract is the playbook. The abstraction ranges from the playbook that is detailed enough to be automatically executed by a machine, 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.

5.1 **Possible mappings**

There are several possible ways of mapping playbooks into the MISP data model. The mapping varies based on the use case. We take into consideration these possible mappings and discuss their pros and cons.

5.1.1 MISP galaxy

First, we discuss using MISP galaxies for playbook sharing. Indeed, there already exists a MISP galaxy named Course of Action galaxy (COA galaxy). The COA galaxy is very interesting as it references MITRE ATT&CK Mitigation techniques - a well-established framework. The mitigation actions have a unique ID and their plain-text description in the COA galaxy. Based on our observation, the COA galaxy is used to tag MISP events containing IoC about malware. The COA galaxy defines the actions that are straightforward (there is no conditional branching, no workflow, just a single step action referring to a MITRE ATT&CK Mitigation technique). An event can be tagged with multiple galaxy tags, but such a tagging still does not represent a workflow. Therefore the COA galaxy is not equivalent to playbooks and cannot be used to represent them. It can, however, be used as additional context information.

Rather than using the COA galaxy, we may specify our own galaxy containing predefined playbooks. But in that case, if there is just a small modification in such a galaxy needed then a new version of the whole galaxy must be issued. The MISP galaxies cannot be edited, they are static. A possible solution to address the update problem would be to version the playbook via an external repository and the content of the galaxy would be just a link to the repository and basic metadata about the playbook in the repository. However, that would require another access management to an external repository. Only if the playbooks are shared publicly then the requirement for access management can be dropped.

In summary, there is a significant downside when new playbooks are created or the existing playbooks are modified. It would either result in a new version of a galaxy or a galaxy with links into an external repository and a need for additional access management when the playbooks are shared within a closed community or between a customer and a vendor. The advantage would be a straightforward affiliation with an IoC by tagging the event with the playbook galaxy tag.

5.1.2 MISP object

Another representation is to create a custom MISP object that would contain an attribute that carries the entire playbook (as a JSON text or as a file attachment) as well as other attributes and tags that would represent the playbook metadata.

In fact, the current MISP contains an object for representing a course of action (COA, https://www.misp-project.org/objects.html#_course_of_action). The MISP COA object stores the textual description of the course of action and additional pieces of information such as name, the impact of the action, its cost, type and the stage of the threat management lifecycle. The object is used when there is a need to describe a new procedure on how to handle an incident. Unfortunately, the MISP COA object representation is designed to host human-readable playbooks and as such is not a good fit to host CACAO playbooks.

Additionally, the possibility of creating multiple objects that would correspond to the objects in the CACAO standard (e.g. each step can be a single object) was discussed in mutual teleconferences between the TC of CACAO and members of SAPPAN. Such a mapping in the MISP data model would achieve higher granularity of the playbook, e.g. even the individual steps. This would allow MISP to index the steps and their attributes, making them searchable and ready for correlation. However, representing the graph-based structure of the workflow in MISP is problematic. Further granularity in the form of multiple different MISP objects would bring more issues than benefits. For example, the concrete formats of individual playbooks may vary in the future, making the maintenance of many MISP objects rather problematic. Also, the mutual conversations with TC CACAO also revealed that the multi-object playbook would compromise the signature of the original playbook and would require a complex process to restore the signature correctly.

Another possibility is to share only the workflow (without metadata) in a single attribute and store metadata, that were originally in the playbook, as additional attributes. But this approach is problematic because when such an object is downloaded from MISP. Hence the playbook would have to be reconstructed from the workflow and the attributes. Therefore, it is better to share the whole playbook in an attachment and not just the workflow despite the fact that the playbook metadata will essentially be duplicated within the new playbook object.

In Table 10, we identify all possible metadata about the playbook and that will be required in terms of sharing. Some of them can map as attributes and some of them as tags from taxonomies or galaxies.

Table 10: Mapping of data points onto MISP data structures

Mapping
Types and categories
Our chosen approach is to create one MISP object representing a playbook with one attribute (JSON text or attachment) and additional attributes representing important metadata about the playbook. The data points presented here are mapped onto attributes defined in table below. Playbook - missing type in MISP - a string containing the serialized playbook, we propose a 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 RDF graph based on 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 [5]. Playbook-type - the type of the playbook should be expressed as a machine tag using a taxonomy. Attack tactic - candidate for the galaxy. Attack technique - candidate for the galaxy. IOC - MISP contains plenty of types that can represent an IOC. The particular IOC type should be selected to represent the given IOC and added as an attribute. CVE - identifier of CVE - missing type in MISP Person-role - missing in MISP (possible values, e.g., management, analysis, development, incident handling, legal support, communication, network administration, other) Organization type - a prevalent type of an organization (critical infrastructure, governmental, enterprise, education, research, corporate, SME, ISP, MSP), candidate for the galaxy. Distribution - maps on distribution levels native to MISP Playbook-impact - optional attribute defined as an integer from 0 to 100 according to [5], or magnitude of impact on organization's mission/business defined as taxonomy. Possible levels: none, low: executing the playbook will impact with low chance or will impact a low number of users or non-critical services, medium: impact with medium chance or the number of users or internal services, high: impact with high chance or a high number of users or some business services, critical: severe impact on a 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 [5]. Organization-type - maps on Galaxy 6 but some sectors might be missing, also this galaxy should be put in line with [5]. Playbook abstraction - missing taxonomy in MISP. The taxonomy can be partially based on [5] 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.

Based on the mapping table we create the final CACAO-compatible playbook object. The object may host not only CACAO playbooks but also other specifications of the playbooks such as the one proposed in the SAPPAN deliverable D4.2. The Table 11 defines a basic list of attributes that have been identified as necessary for search and filtering purposes in the MISP platform and form attributes of the CACAO MISP object.

Table 11: MISP playbook object

Custom name of attribute/property	**Attribute category**	**Attribute type**	**Attribute description**	**Correlation**	**Required**
playbook	payload delivery	attachment	Content of the whole playbook.	True	True
playbook-standard	other	text	Identification of the playbook standard used to describe the playbook.	False	True
playbook-type	other	text	Identifies types of actions in the playbook.	False	True
person-role	person	text	Roles of users, that should be involved in the playbook.	False	True
organization-type	other	text	Type of an organization, that the playbook is intended for.	False	False
playbook-impact	other	counter	How safe it is to execute the playbook from perspective of breaking down or influencing infrastructure from 0 to 100.	False	False
playbook-abstraction	other	text	E.g. playbook template, playbook, partial workflow, full workflow, ...	False	False

creator	other	text	Creator organization of the playbook.	False	False
description	other	text	Description of the playbook.	False	False

We build the MISP object template described in JSON and, after consultation with the MISP developers, we pushed it into the official MISP repository. The template is also included in the attachment of this deliverable together with an example of a playbook.

6 Sharing interface

here will be two interfaces, API and GUI which allow information about response handling information to be uploaded into MISP, or to be consumed from MISP. The GUI and the API provide the MISP platform itself but we wrap the API to simplify its usage. 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 implement 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 the MISP instance using the MISP Automation API key. The IP proxy picks the playbook relevant for sharing and applies sanitization and anonymization functions to remove or modify confidential or private information. The sanitization and anonymization functions are optional and their implementation is outside of the scope of T5.4. The IP proxy then pushes the playbook via PyMISP into the MISP instance. The IC proxy will connect to the MISP instance using the MISP Automation API key. It will register to MISP 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 through a form as depicted in the following Figure 4.

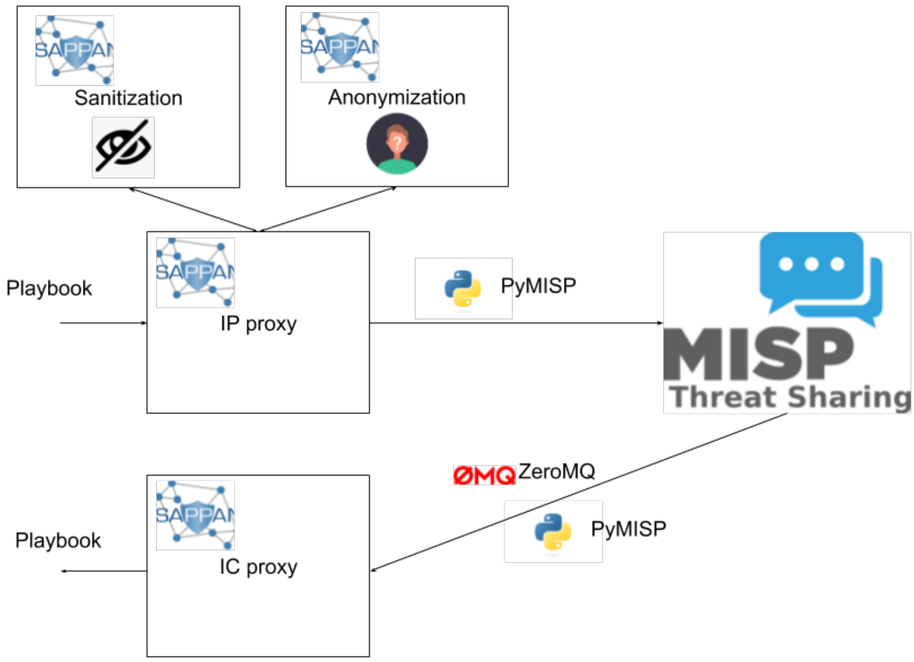

Figure 3: Basic concept of communication between producer/consumer and MISP

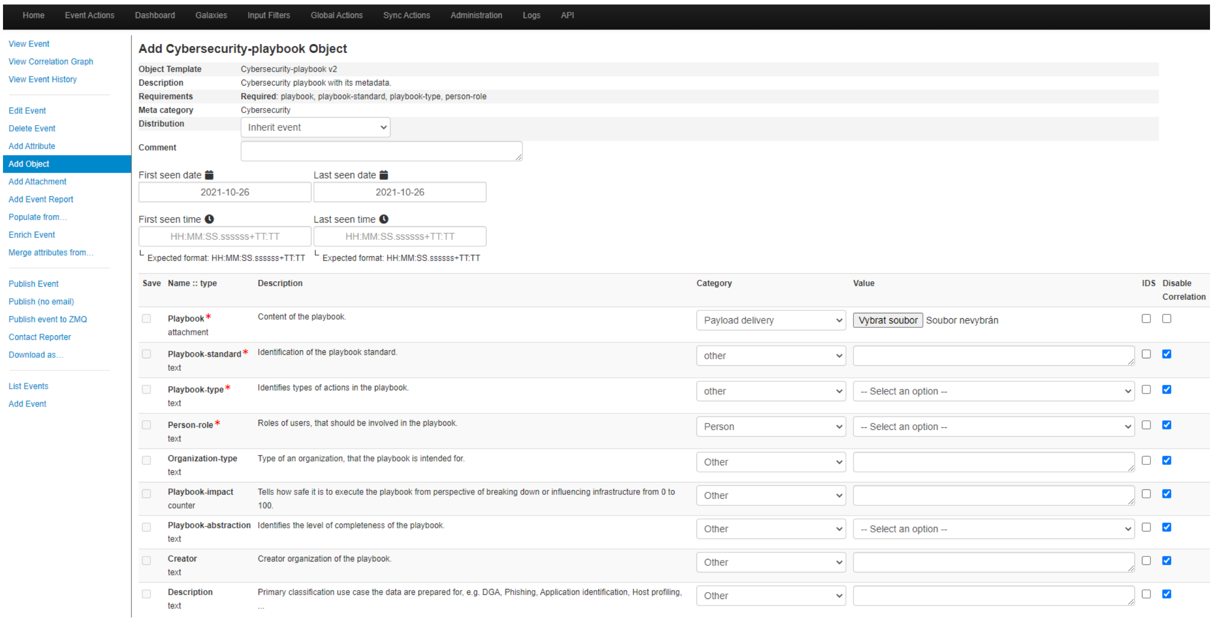

Figure 4: MISP GUI with a playbook form

7 Meeting playbook sharing requirements

In Table 12, we document how the requirements, defined in Tables 4, 5, 6, 7, are met.

Table 12: 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 a tlp or an iep tag to indicate to 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, 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 requirement is met by using a vocabulary defined in D4.2 to enable translation from local/detailed level into more abstract.
CoRq-5	Abstract to detail	Same to the above.
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, we developed as part of the SAPPAN dashboard in T6.1 a user interface to access playbooks in a visual graph-based representation based on business process model and notation (BPMN). BPMN is a graphical notation developed primarily to communicate processes between analysts, technical developers and business managers.
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 provide 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 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 the 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 MISP as the best fit for SAPPAN use cases. Further, the deliverable introduced the MISP data model for playbook sharing which was thoroughly discussed with the TC CACAO and is now publicly available in the official MISP repository (https://github.com/MISP/misp-objects/pull/324#issue-1009464958). We proposed and implemented MISP interface components to communicate with the SAPPAN environment. In the end, the deliverable evaluated how the specific playbook-sharing requirements are met by the proposed sharing environment and the playbook data model. The deliverable presents an example of a playbook expressed in the proposed data model for MISP in the attachments chapters.

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 TC CACAO is composed of commercial partners who are interested in utilizing the standard and the set of tools developed around the standard to deliver the playbooks to their customers, e.g. similarly to IDS ruleset, the vision is 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.

To further support the effort and disseminate the SAPPAN results, CESNET joined 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. We disseminate the playbook effort through the Twitter accounts of project partners, SAPPAN blog post as well as we co-authored a technical paper about the sharing of course of actions and submitted it to BigData 21 conference for review.

9 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] FortinetSOAR, 2021, available online: https://www.fortinet.com/content/dam/fortinet/assets/solution-guides/sb-accelerate-with-fortisoar.pdf

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

[5] 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

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

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

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

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

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

[14] MISP - Open Source Threat Intelligence Platform & Open Standards For Threat Information Sharing, available online: https://www.misp-project.org

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

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

[17] 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

10 Attachment 1 – MISP playbook object


```
{
  "attributes": {
    "created": {
      "categories": [
        "Other"
      ],
      "description": "The time at which the playbook was originally created.",
      "disable_correlation": true,
      "misp-attribute": "datetime",
      "ui-priority": 1
    },
    "creator": {
      "categories": [
        "Other"
      ],
      "description": "Creator organization of the playbook.",
      "disable_correlation": true,
      "misp-attribute": "text",
      "ui-priority": 1
    },
    "description": {
      "categories": [
        "Other"
      ],
       "description": "Primary classification use case the data are prepared for,
e.g. DGA, Phishing, Application identification, Host profiling, ...",
      "disable_correlation": true,
      "misp-attribute": "text",
      "ui-priority": 1
    },
    "id": {
      "categories": [
        "other"
      ],
      "description": "A value that uniquely identifies the playbook.",
      "disable_correlation": false,
      "misp-attribute": "text",
      "ui-priority": 1
    },
    "impact": {
      "categories": [
        "Other"
      ],
       "description": "A positive integer that represents the impact the playbook
has on the organization from 0 to 100.",
      "disable_correlation": true,
      "misp-attribute": "counter",
      "ui-priority": 1
    },
    "label": {
      "categories": [
        "Other"
      ],
      "description": "An optional set of terms, labels or tags associated with this
playbook.",
      "disable_correlation": true,
      "misp-attribute": "text",
      "multiple": true,
```

```
      "ui-priority": 1
    },
    "modified": {
      "categories": [
        "Other"
      ],
      "description": "The time that this particular version of the playbook was last
modified.",
      "disable_correlation": true,
      "misp-attribute": "datetime",
      "ui-priority": 1
    },
    "organization-type": {
      "categories": [
        "Other"
      ],
      "description": "Type of an organization, that the playbook is intended for.",
      "disable_correlation": true,
      "misp-attribute": "text",
      "ui-priority": 1
    },
    "playbook": {
      "categories": [
        "Payload delivery"
      ],
      "description": "Content of the whole playbook.",
      "misp-attribute": "attachment",
      "ui-priority": 1
    },
    "playbook-abstraction": {
      "categories": [
        "Other"
      ],
      "description": "Identifies the level of completeness of the playbook.",
      "disable_correlation": true,
      "misp-attribute": "text",
      "ui-priority": 1,
      "values_list": [
        "guideline",
        "playbook template",
        "playbook",
        "partial workflow",
        "full workflow",
        "fully scripted"
      ]
    },
    "playbook-standard": {
      "categories": [
        "other"
      ],
      "description": "Identification of the playbook standard.",
      "disable_correlation": true,
      "misp-attribute": "text",
      "ui-priority": 1
    },
    "playbook-type": {
      "categories": [
        "other"
      ],
      "description": "Identifies types of actions in the playbook.",
```

```
      "disable_correlation": true,
      "misp-attribute": "text",
      "multiple": true,
      "ui-priority": 1,
      "values_list": [
        "notification playbook",
        "detection playbook",
        "investigation playbook",
        "prevention playbook",
        "mitigation playbook",
        "remediation playbook",
        "attack playbook"
      ]
    },
    "priority": {
      "categories": [
        "Other"
      ],
       "description": "A positive integer that represents the priority of this
playbook relative to other defined playbooks.",
      "disable_correlation": true,
      "misp-attribute": "counter",
      "ui-priority": 1
    },
    "revoked": {
      "categories": [
        "Other"
      ],
      "description": "A boolean that identifies if the playbook creator deems that
this playbook is no longer valid.",
      "disable_correlation": true,
      "misp-attribute": "boolean",
      "ui-priority": 1
    },
    "severity": {
      "categories": [
        "Other"
      ],
      "description": "A positive integer that represents the seriousness of the
conditions that this playbook addresses.",
      "disable_correlation": true,
      "misp-attribute": "counter",
      "ui-priority": 1
    },
    "valid-from": {
      "categories": [
        "Other"
      ],
      "description": "The time from which the playbook is considered valid and the
steps that it contains can be executed.",
      "disable_correlation": true,
      "misp-attribute": "datetime",
      "ui-priority": 1
    },
    "valid-until": {
      "categories": [
        "Other"
      ],
      "description": "The time at which this playbook should no longer be considered
a valid playbook to be executed.",
```

```
      "disable_correlation": true,
      "misp-attribute": "datetime",
      "ui-priority": 1
    }
  },
  "description": "Security playbook with its metadata for executing course of action
in cyberspace defense.",
  "meta-category": "misc",
  "name": "security-playbook",
  "required": [
    "playbook",
    "playbook-standard",
    "playbook-type",
    "person-role"
  ],
  "uuid": "48894c92-447b-4abe-b093-360c4d823e9d",
  "version": 1
}
```

11 Attachment 2 - 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 4.

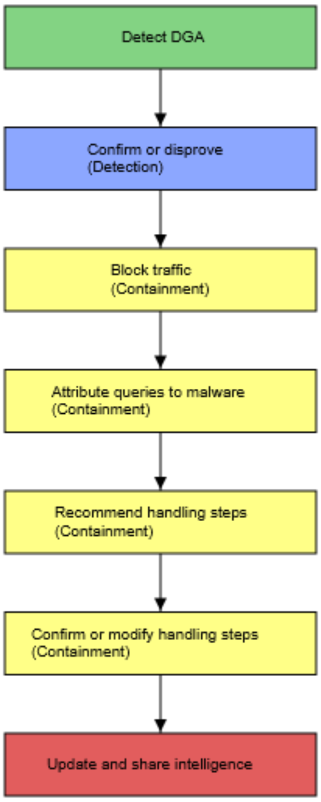

Figure 4: Simple DGA playbook.

The DGA plabyook is expressed in a shortened version as the MISP event below.

```
"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\"},{\"Cate-
gory\":\"InitialStep\"},\
                         ….
                         {\"ConfidentialityLevel\":\"Public\"},{\"Re-
porter\":\"\"}]}",
            "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\"",
                ...
            },
...
    }
```