D3.3

First implementation, test and validations of the CERT Pilot

WP3.3 – CERT Pilot

C3ISP

Collaborative and Confidential Information Sharing and Analysis for Cyber Protection

Due date of deliverable: <30/11/2018>
Actual submission date: <DD/MM/YEAR>

30/11/2018
Version 1.0

Responsible partner: ISCOM-MISE
Editor: Sandro Mari
E-mail address: sandro.mari@iscom.mise.it

Project co-funded by the European Commission within the Horizon 2020 Framework Programme		
Dissemination Level		
PU	Public	
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

The C3ISP Project is supported by funding under the Horizon 2020 Framework Program of the European Commission DS 2015-1, GA #700294

Authors: *Andrea Saracino (CNR) , Sandro Mari (ISCOM-MISE)*

Approved by: *Stefano Tranquillini (CHINO), Ali Sajjad (BT)*

Revision History

Version	Date	Name	Partner	Sections Affected / Comments
0.1	16/04/2018	Andrea Saracino	CNR	Initial ToC
0.2	15/08/2018	Andrea Saracino	CNR	Architecture and Deployment
0.3	12/09/2018	Andrea Saracino	CNR	Validation Strategy
0.4	30/09/2018	Andrea Saracino	CNR	GQM Validation definition
0.5	5/11/2018	Andrea Saracino, Sandro Mari	CNR, ISCOM-MISE	Validation results at M24.
0.6	11/11/2018	Andrea Saracino, Sandro Mari	CNR, ISCOM-MISE	Ready for internal review
0.8	22/11/2018	Andrea Saracino	CNR	Addressed comments from internal reviewers
1.0	23/11/2018	Andrea Saracino, Sandro Mari	CNR, ISCOM-MISE	Ready for submission

Executive Summary

This documents reports the progress related to the Work Package 3 of the C3ISP project. This document is particularly focused on the implementation and initial validation for the CERT pilot. The document will report a summary of the pilot overview and implementation, discussing the chosen architectural model and the actual deployment in the CERT premises. Furthermore, the document will describe the validation methodology, first at high level to present the involved actors of the validation process and the role of the C3ISP consortium. Hence, the document will report the full validation methodology, to be performed in accordance with the GQM validation, as agreed in the activities of WP6. Hence, we have reported the initial results related to the validation that has been performed at M24 with the current framework status. The validation has been performed with the CERT internal stakeholders, i.e. data collector, data analyser and data dispatcher and the results on the tests performed up to now are mostly positive. The performed tests have been sufficient to validate completely or partially 3 out of the 9 user stories defined in deliverable D3.1 [2]. The next deliverable will report the completed evaluation on all missing user stories.

Table of contents

1. Introduction

1.1. Purpose of the Document

The purpose of this document is to report the status of the implementation of the CERT pilot at Month 26, and the validation strategy designed for the CERT pilot, on the base of the requirements collected in deliverable D3.1. Hence, we will report the initial validation results, performed with the framework at the status of M24. This deliverable also provides input for both deliverable D6.3 and D6.5, related respectively to pilot implementation and pilot validation. Final validation results will be collected in the next months and will be reported in deliverable D3.4.

1.2. Scope of the Document

The scope of this document is the description of the current implementation status and the validation methodology and initial results. Considerations on the architectural choices and the pilot overview will be also reported.

1.3. Structure of the Document

The rest of the document is structured as follows. Section 2 will recall an overview of the pilot, describing the operative workflow, the actors and the objective of the pilot. Section 3 describes the pilot architecture, explaining how the C3ISP framework is deployed in the CERT workflow. Section 4 describes the current implementation of the pilot, reporting the interfaces and components specific for the CERT pilot and their current status. Section 5 describes the validation strategies and the initial validation results.

1.4. List of abbreviations

Term	Meaning
API	Application Program Interface
C3ISP	Collaborative and Confidential Information Sharing and Analysis for Cyber Protection
CERT	Computer Emergency Response Team
CVE	Common Vulnerabilities and Exposures
DMO	Data Manipulation Operations
DSA	Data Sharing Agreement
HE	Homomorphic Encryption
IAI	Information Analytics Infrastructure
ISI	Information Sharing Infrastructure
TLS	Transport Security Layer

2. CERT Pilot Overview

An overview of the pilot architecture is depicted in Figure 1, reporting the main actors, component and expected interactions.

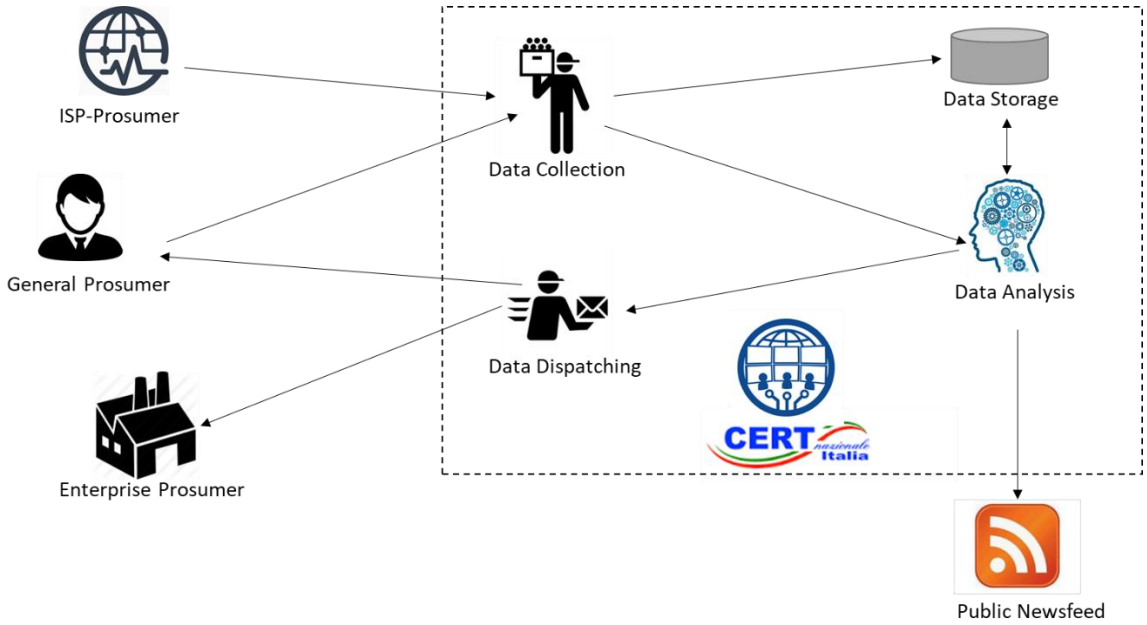

Figure 1: CERT Pilot Overview: main actors and functionalities

The CERT pilot is the one which better represents the C3ISP philosophy, including all the C3ISP functionalities in a single entity and having as prosumers a large set of stakeholders, which also includes the other pilots (i.e. ISP and Enterprise/SME). The CERT is a public entity which collects data related to cyber threats (CTI) from several prosumers (or providers), stores and categorize the collected information and exploits them to run analysis. In particular, the analysis can be requested from a specific prosumer, or issued by the CERT itself. Analysis results will be stored in the DPO as well, given that the data policy allows it, and/or are dispatched to interested prosumers. Furthermore, a set of information, collected or inferred will be made publicly available through the CERT website as newsfeed related to cyber threats of public interest.

The CERT has to consider the privacy requirements expressed by the data providers, which might apply to the data content itself or to computed results. The policies will specify which attributes can be made available in public access, a list of specific prosumers which can read data and derived analysis results and legal requirements on methodologies for data storage and processing. Moreover, this pilot requires that prosumers might be able to enforce some data policies on their premises, sanitizing data before sharing them with the CERT.

3. CERT Pilot Architecture

The CERT pilot architecture follows the hybrid model with On-Premises ISI with Centralised ISI and IAI, exactly as described in deliverable D7.2 [6], section 3.2.

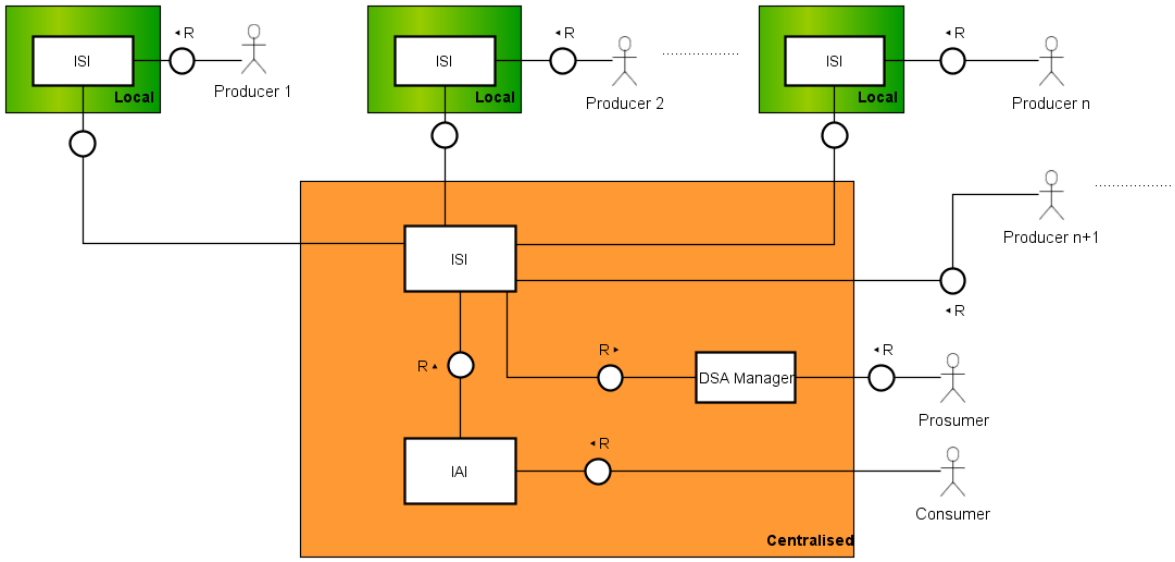

Figure 2: C3ISP Hybrid Architecture

Hence, the envisioned architecture considers the presence of a Local ISI on prosumer side, whilst the "centralized" architectural part entirely resides in CERT premises. The presence of the local ISI also allows to accommodate the requirement, also introduced in the former section, related to the sanitization of data on prosumers premises, before they are shared with the CERT.

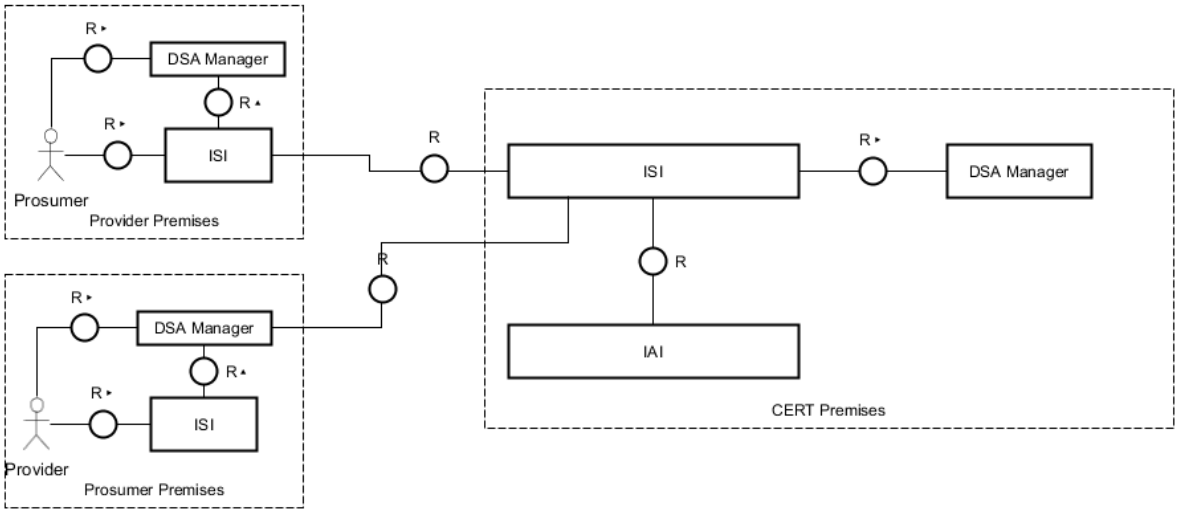

Figure 3: CERT Pilot Architecture

A high-level representation of the architecture is depicted in Figure 3, showing the components and their interconnections. As discussed, the ISI is present on both the Prosumer and CERT premises. In the following they will be addressed respectively as *Local ISI* and *Remote ISI*. An instance of the DSA manager is present both on provider/prosumer premises and in the CERT. The DSA manager local to prosumers is used, as in the other pilots, to define policies for the shared data. Part of these policies will be enforced by the local ISI, additional policies will be enforced instead by the remote ISI, in particular the one related to data analysis and to result

redistribution. The remote DSA manager is instead used to define additional constraints which are internal to the CERT organization, which, being a public organization has to implement standards related to data storage and maintenance. These policies are enforced by the remote ISI. The IAI is only present in the CERT premises, hence the prosumers are considered not able to run in house the analytics on their data and will demand the analysis directly to the CERT. Hence, the C3ISP analytics functionalities are all provided by the CERT, on prosumer request, or invoked directly by the CERT itself.

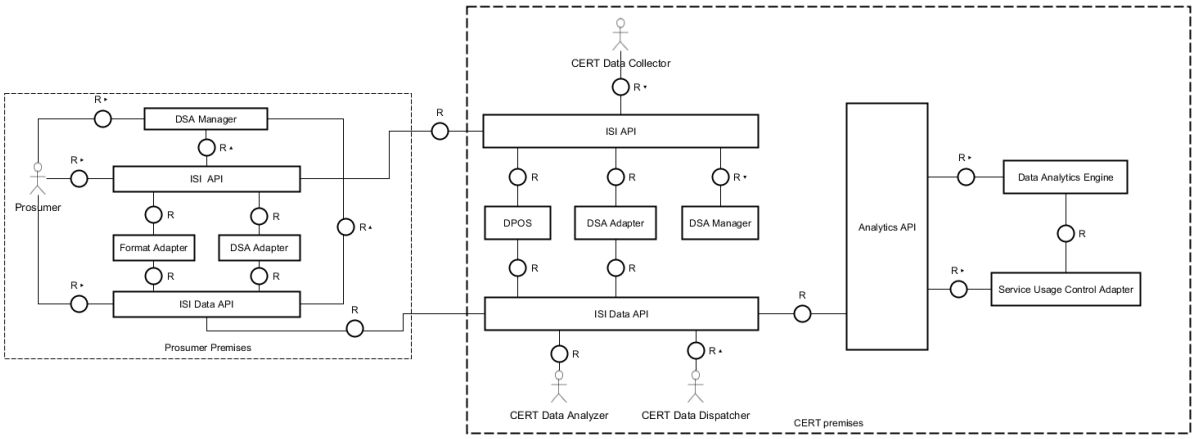

Figure 4: Detailed pilot architecture

The detailed pilot architecture, also showing the interaction with the main actors internal to the CERT, i.e. the data collector, data analyser and data dispatcher, as shown in deliverable D7.1 is presented in Figure 4 . Also, the pictures show more in details the fact that the Analytic components and the DPOS only reside on the CERT side. We will now discuss the single components in details.

3.1. Block Design

In the following we will report the block representation and the description of the main blocks of the CERT pilot architecture. In particular, we will describe the local architecture Prosumer/Provider side of the C3ISP infrastructure, and the C3ISP architecture for the CERT side.

3.1.1. The Prosumer local side

The operations and components related to the prosumer are detailed in Figure 5.

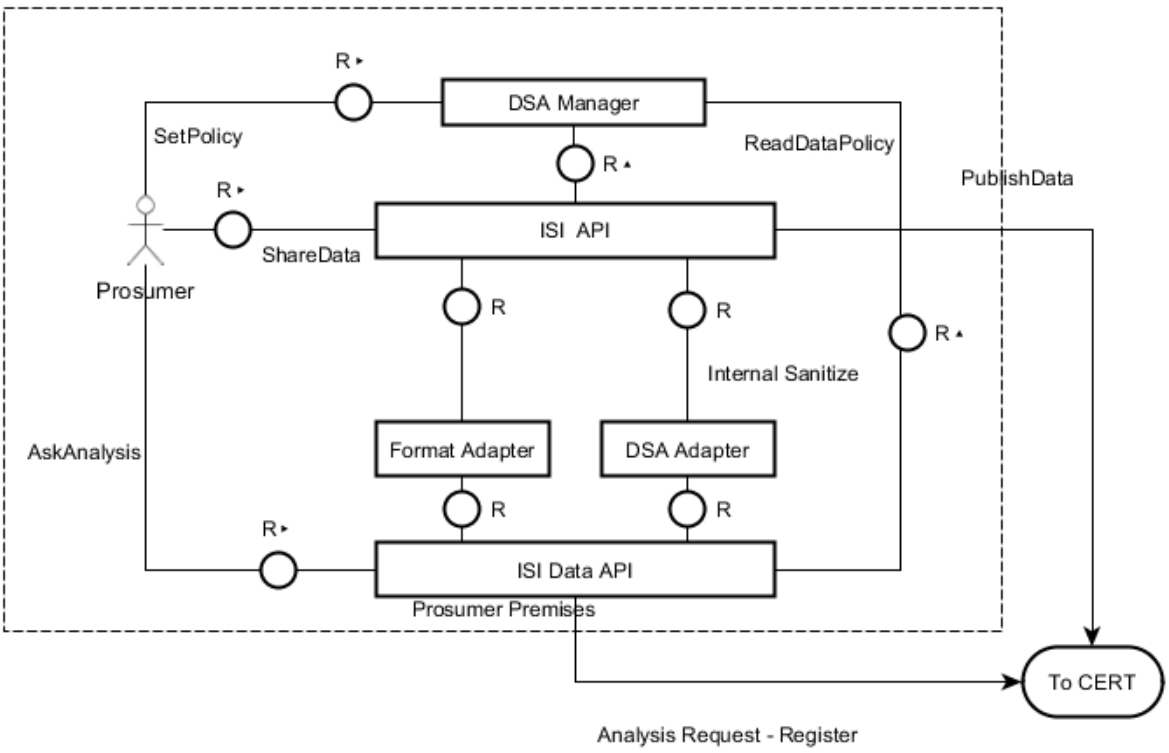

Figure 5: Prosumer side detail

In the CERT pilots the Prosumer will mainly do either:

i. asking the CERT for specific analytics
ii. or sharing CTIs with the CERT to let it compute additional knowledge about specific threats.

As shown, the prosumer interacts directly with the DSA manager to define policies for data to be shared, which might be partially enforced on the prosumer side by the DSA adapter, which will provide to sanitize the data before they are sent to the CERT.

Hence, the prosumer operations are the following:

- *Ask Analysis - Register*: The prosumer queries the CERT for specific data and/or for an analysis (classification, clustering etc.) on dataset already stored in the CERT DPOS. The AskAnalysis might also result in a Registration to a specific topic or set of information.
- *PublishData*: The prosumer introduces a piece of data, already structured according to the CTI format standard, protected through a data bundle with the attached DSA. According to the DSA, data can be sanitized before they are sent to the CERT, which will store the bundle in the DPOS.
- *InternalSanitize*: Is the operation of sanitization performed by the DSA adapter local to the prosumer.

We assume thus that in the CERT pilot the prosumer is not able/willing to run in house analytics and will demand this task to the CERT itself.

3.1.2. CERT Remote Side

The CERT side includes the majority of the components of the pilots. In Figure 6 is reported a more detailed view of the CERT infrastructure in the C3ISP pilot,

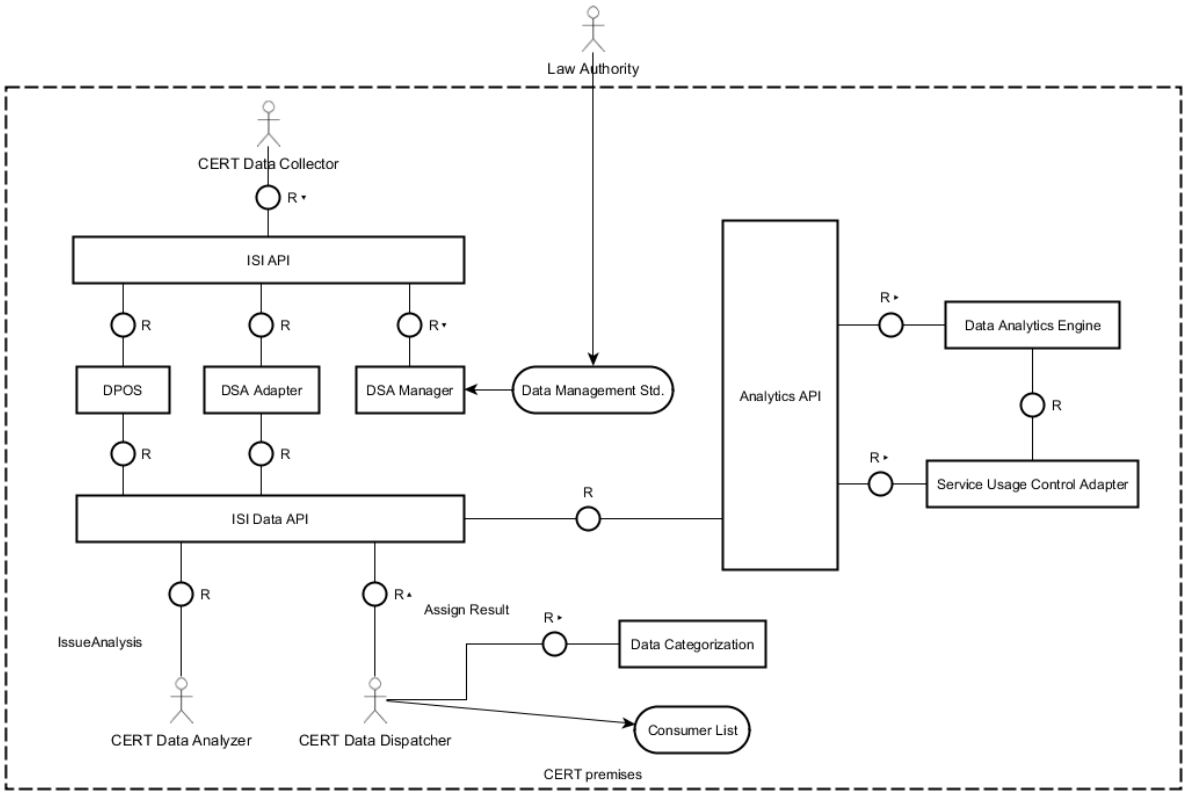

Figure 6: Detailed architecture of the CERT side

The three main actors on the CERT side are the Data Collector, the Data Analyzer and the Data Dispatcher. The Data Collector directly interacts with the ISI API, managing thus the data storage operations acting as interconnection between the Local and Remote ISI. The Data Collector might either passively act by receiving and storing data from prosumers in the DPOS, or actively act by requesting specific information or data streams from prosumers. The Data Analyser issues the analysis operations, either when receiving requests from prosumers, or acting as a consumer itself, generally to infer information of public interest. The Data Dispatcher interacts with the ISI via API for receiving the analysis results which have been extracted by the IAI. It also receives the registration requests issued by prosumers, storing them in a Consumer List. Hence, through the CERT internal Data Categorization component, the dispatcher matches the analysis results with the consumer registrations, sending automatically interesting results to them, without the necessity for them to issue an analysis. It is worth noting that such a process is done automatically, providing results coming from collaborative analysis without explicit user analysis request. Finally, it is worth noting the presence of a set of Data Management Standards, which are used to define additional policies for data, enforced on the CERT side. These standards are defined by law authorities and might regard national or international regulations for data storage, management of classified information, etc.

The functions to be considered on the CERT side are:

- *IssueAnalysis*: Invoked by the Data Analyzer by itself, or as a response to a prosumer request for a specific analytic.
- *AssignResult*: This function is invoked by the data dispatcher to match a new information, received by a prosumer or computed through the IAI, with interested

consumers. The information dispatching phase is however handled by the ISI, to ensure that result is delivered in accordance with DSA policies.

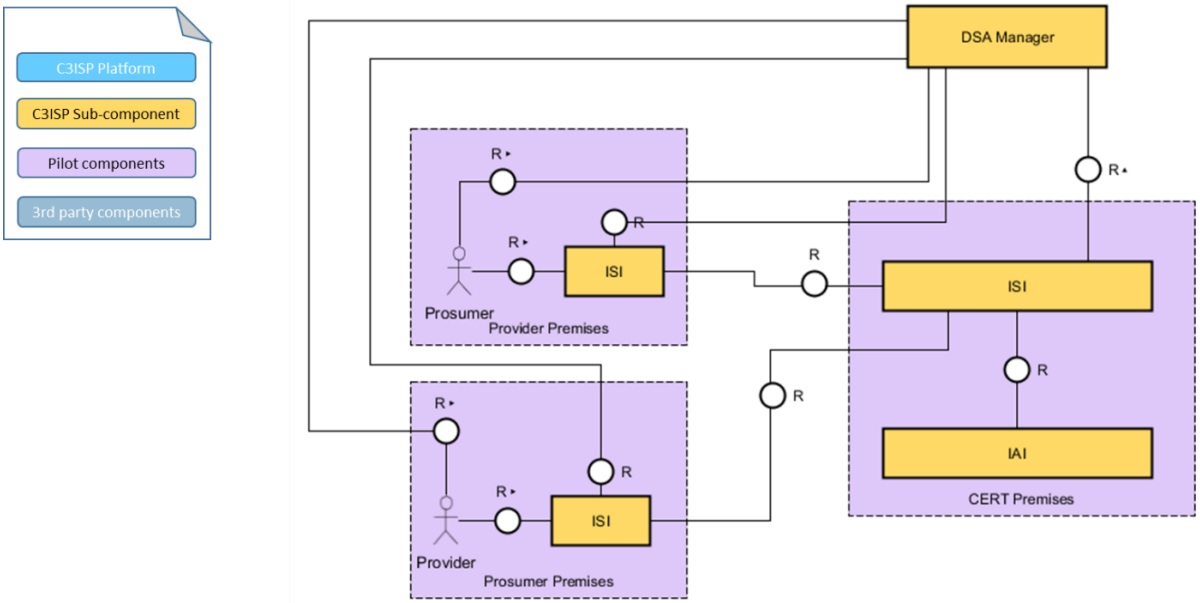

Figure 7: Pilot implementation

Figure 7 reports the implementation of the CERT pilot. As shown, the Local ISI is deployed in the prosumer premises, which are the external stakeholders of the CERT and considered part of the pilot. The Prosumers publish data and ask for analysis via the local ISI and IAI API. The centralized ISI, together with the IAI is instead the core of the CERT pilot, which is installed in the CERT premises. The current configuration of the system where the virtual machine has been installed is shown in the following Figures.

Figure 8: CPU Specifics for CERT machine with C3ISP ISI and IAI

The machine installs an Intel i7 with two CPUs, with a speed of 3.4 GHz. The amount of dedicated memory is 4GB and the installed OS is Ubuntu 18.04.

```
c3ispmanager@ubuntu:~/Desktop/spam_sample$ cat /proc/meminfo
MemTotal:        4015684 kB
MemFree:          471480 kB
MemAvailable:    3268476 kB
Buffers:          335580 kB
Cached:          2441088 kB
SwapCached:            0 kB
Active:          2140596 kB
Inactive:         927468 kB
Active(anon):     165516 kB
Inactive(anon):   127768 kB
Active(file):    1975080 kB
Inactive(file):   799700 kB
Unevictable:           0 kB
Mlocked:               0 kB
SwapTotal:       2097148 kB
SwapFree:        2097148 kB
Dirty:                20 kB
Writeback:             0 kB
AnonPages:        289276 kB
Mapped:           194608 kB
Shmem:              1888 kB
Slab:             366820 kB
SReclaimable:     311640 kB
SUnreclaim:        55180 kB
KernelStack:        7696 kB
PageTables:        18960 kB
NFS_Unstable:          0 kB
Bounce:                0 kB
WritebackTmp:          0 kB
CommitLimit:     4104988 kB
Committed_AS:    1815072 kB
VmallocTotal:   34359738367 kB
VmallocUsed:           0 kB
VmallocChunk:          0 kB
HardwareCorrupted:     0 kB
AnonHugePages:         0 kB
ShmemHugePages:        0 kB
ShmemPmdMapped:        0 kB
CmaTotal:              0 kB
CmaFree:               0 kB
HugePages_Total:       0
HugePages_Free:        0
HugePages_Rsvd:        0
HugePages_Surp:        0
Hugepagesize:       2048 kB
DirectMap4k:      165696 kB
DirectMap2M:     4028416 kB
DirectMap1G:     2097152 kB
c3ispmanager@ubuntu:~/Desktop/spam_sample$
```

Figure 9: Memory specifications for the CERT machine for C3ISP ISI and IAI.

```
c3ispmanager@ubuntu:~/Desktop/spam_sample$ screenfetch
                           ./+o+-       c3ispmanager@ubuntu
                   yyyyy- -yyyyyy+      OS: Ubuntu 18.04 bionic
                ://+//////-yyyyyyo      Kernel: x86_64 Linux 4.15.0-24-generic
            .++ .:/++++++/-.+sss/`      Uptime: 71d 4h 41m
          .:++o:  /++++++++/:--:/-      Packages: 1608
         o:+o+:++.`..```.-/oo+++++/     Shell: bash 4.4.19
        .:+o:+o/.          `+sssoo+/    CPU: Intel Core i7-6700 @ 2x 3.408GHz
   .++/+:+oo+o:`             /sssooo.   GPU: svgadrmfb
  /+++//+:`oo+o               /::--:.   RAM: 448MiB / 3921MiB
  \+/+o+++`o++o               ++////.
   .++.o+++oo+:`             /dddhhh.
        .+.o+oo:.            `oddhhhh+
         \+.++o+o``-````  .:ohdhhhhh+
          `:o+++ `ohhhhhhhhyo++os:
            .o:`.syhhhhhhh/.oo++o`
                /osyyyyyyyo++ooo+++/
                    ````` +oo+++o\:
                           `oo++.
```

**Figure 10: OS and general specifics for the CERT machine for C3ISP ISI and IAI**

The ISCOM MISE will provide in the next months a dedicated server for supporting heavy load of operations issued by several prosumers.
```

The DSA Manager is the only component that is not integrated in the CERT architecture, since it is offered by HPE as a remote component.

3.2. Components Implementation

The CERT pilot does not require a set of pilot specific components. The CERT, in facts, install and offers the C3ISP services and components as they are provided by the WP7 activities, as described in [7] and [8].

The component which is specific for the CERT pilot is the *User Interface* used to access the ISI (local and remote) and the IAI. The C3ISP framework is accessed for the CERT pilot via a user interface, which wraps the ISI and IAI functionalities. In the current implementation, this APIs are offered through a set of Java functions that enable the users to publish files via the ISI, to search and assign DSAs from those defined through the DSA manager. In the following we report a code snippet used to invoke the ISI API to publish a new file on the DPOS.

```
private String uploadSingleFile(MultipartFile file, Metadata metadata,
      RestTemplate rest) throws Exception {
    LinkedMultiValueMap<String, Object> toUpload = new
LinkedMultiValueMap<>();
    toUpload.add(FILE_SUBMIT_FORM, file);
    toUpload.add(METADATA_FORM, metadata);
    RequestEntity<MultiValueMap<String, Object>> requestEntity;
    requestEntity = RequestEntity.post(new URI(URI))
        .contentType(MediaType.MULTIPART_FORM_DATA).body(toUpload);
```

The full set of functions available for the CERT pilot can be found at the following address.

Link to code: https://devc3isp.iit.cnr.it:8443/wp2wp3/pilot-interface

All the functions of the CERT interface can be invoked at M24 via command line with the following command:

```
java –jar [method].jar [parameters]
```

In these months the User Interface is being extended and completed by wrapping it in a web service accessible thus via web interface. As for the other components of the C3ISP framework this User Interface is thus provided as a WAR file to be installed via Tomcat and accessed via a Web Interface. The component will present a GUI that makes accessible all the functionalities of ISI and IAI.

3.2.1. Defined Data Sharing Agreements

In the following we report the data sharing agreements defined up to now for the CERT pilot. The process of definition of DSAs has also been considered a part of the validation, in particular for what concerns the CERT-AT-1, as will be detailed in the following.

The DSAs defined at month 26 concerns the management of email files, used as test data for the validation of CERT-US-1 and CERT-US-2. The first DSA shown in Figure 11 includes two policies, the first is an authorization to publish email files, given to a specific Subject with Role “Producer”. The second policy is a prohibition that denies the possibility to invoke the analytic “spamEmailAnalysis” on this data piece, if the Subject issuing the action belongs to a specific company.

Figure 11: DSA for email allowing publishing and forbidding spam analysis operation to a specific organization

The second DSA reported in Figure 12 expresses an authorization for email publication followed by an obligation, which forces the system to anonymize the email by suppressing the destination address field.

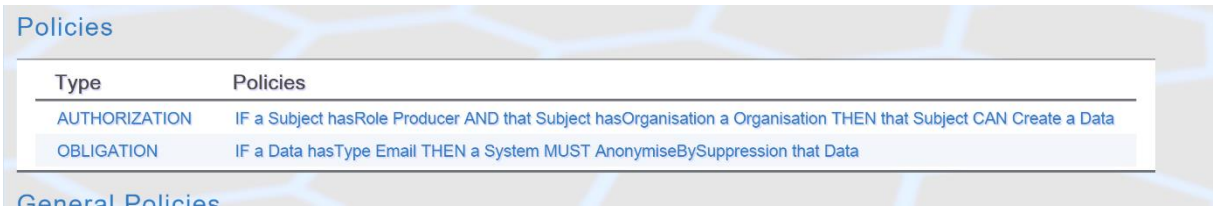

Figure 12: DSA for email allowing publishing with obligation of anonymizing the destination address.

Additional DSAs will be defined in the next months, in parallel with the enrichment of the dictionary and the interaction that will be performed with external stakeholders for implementing their requirements.

4. Testing and Validation Strategy

4.1. Testing and Validation Methodology

The testing and validation strategy for the CERT pilot aims at verifying that the functionalities are working properly in all the steps of the operative workflows and to measure the user acceptance, verifying that their requirements have been met. The validation strategy for the CERT pilot follows the guidelines provided in deliverable D3.1, by performing the Acceptance Tests for the various functionalities as they are described. More specifically, the performed acceptance test will involve the various actors of the CERT pilot, in particular Data Collector, Data Analyzer, Data Dispatcher and CERT customer, who will be asked to answer questionnaires and perform specific tests on the provided platform.

According to the user stories defined in D3.1, we can divide the requirements extracted from them in two macro-sets: (i) requirements from CERT operators and (ii) requirements from CERT customers. The CERT operators are interested in improving their operative workflows by having improved performances, simplified procedures, improved result accuracy, and minimized risk of non-legal compliance. To assess these improvements, we are going to provide a set of questions to be answered after that the C3ISP platform has been used. Moreover, some practical validation tests will be performed with CERT operators to verify the functionalities related to specific acceptance tests.

For what concerns the set (ii) of requirement from customers, they will be verified through questionnaires and feedback that will be provided directly to the CERT.

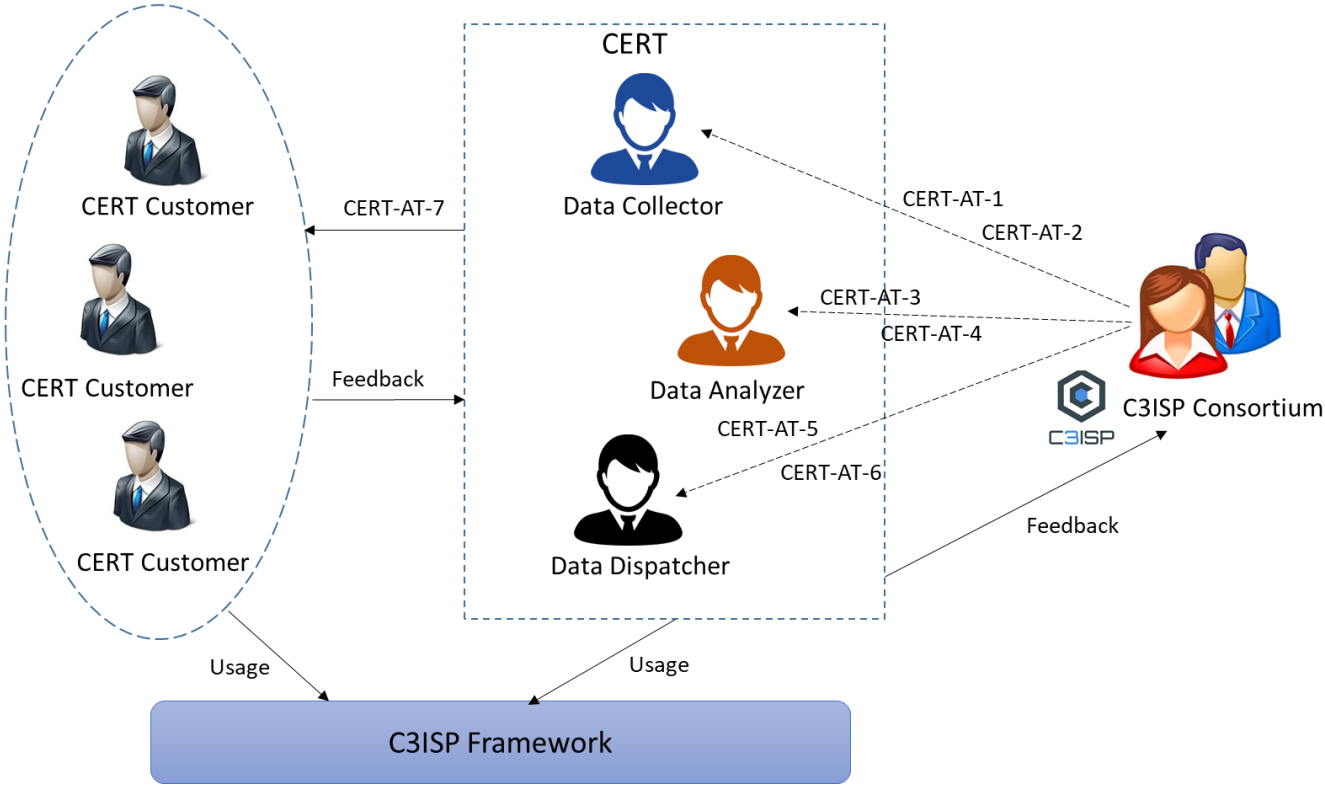

Figure 13: CERT validation strategy

The workflow of the validation strategy for the CERT pilot is depicted in Figure 13, together with the main actors and the interactions.

In accordance with the other pilots, the validation will be performed according to the GQM validation methodology.

4.2. Test Data

Data to be used for validation are for the majority real data coming either from public sources or provided by CERT stakeholders. In particular the ISCOM-MISE has provided a set of real emails to be analysed for spam analysis, together with a set of malware collected through internal honeypots.

The data formats used in the CERT pilot are the following:

4.2.1. Email Files

Emails are used for Provided in `eml` format or as raw text files including only the email header information. The email files are converted by the Format Adapter in STIX format before being stored in the DPOS. Here it follows an example of email file.

```
Delivered-To: bruce@untroubled.org
Received: (fqmail 31297 invoked from network); 01 Aug 2017 05:47:25 -0000
Received: from mx10.futurequest.net (mx10.futurequest.net [69.5.6.182])
  by 10.170.1.170 ([10.170.1.170])
  with FQDP via TCP; 01 Aug 2017 05:47:25 -0000
Received: (qmail 12819 invoked from network); 1 Aug 2017 05:47:24 -0000
Received: from implement.virtualvictorygroup.com
(implement.virtualvictorygroup.com [174.138.191.55])
  by mx10.futurequest.net ([69.5.6.182])
  with ESMTP via TCP; 01 Aug 2017 05:47:24 -0000
DKIM-Signature: v=1; a=rsa-sha256; c=relaxed/relaxed;
      d=virtualvictorygroup.com; s=dkim; t=1501566444; q=dns/txt;
      bh=D2vhgqE+0O1uDLCcs7aIwaa0THKf4rcklHL4QvM9H/4=;
      h=mime-version:date:message-id:subject:from:to:content-type;
      b=SII5J5xT8h7Jl3j7edi4DgwaGpN9mBvowR7F7PiT2KQcm/MhaOKpNpP3SzEX/HYsl
       KGit4V7RgSBwQPz3P2ZHuI4D/J3KKqRv+0oTTYfPAQKSCNAatrv5PM2YbM/VVJFwgA
       7nQMUMaNj9aMeqQyXy0UEflnY+jm10wlKAi5/LeE=
MIME-Version: 1.0
Date: Tue, 01 Aug 2017 01:47:24 -0400
Message-ID: <1787ab02bdf63b28024829b_ec158059@virtualvictorygroup.com>
Subject: Make your LinkedIn Leadgen do the work for you
From: "Dana Bailey"<linked-university@virtualvictorygroup.com
To: bruce@untroubled.org <bruce@untroubled.org>
Content-Type: text/html
List-Unsubscribe: <mailto:unsubscribe-
1787ab02bdf63b28024829b_ec158059@virtualvictorygroup.com>,
<http://www.virtualvictorygroup.com/1787ab02bdf63b28024829b_ec158059/U/>
Content-Length: 15746

<!DOCTYPE html>
<html lang="en"
xmlns="http://www.virtualvictorygroup.com/1c8381473ab63b28024829b4661d_ec1
```

Figure 14: Excerpt of an email file in eml format

As anticipated, the format adapter converts the email in a JSON format compatible to be integrated in a STIX record. The converted format expresses in each JSON tag one of the property of the email file, such as Subject, Sender Address, Recipient Address, Body, etc. The following JSON schema can be integrated in a STIX record, rep

```
{
  "type": "schema",
  "id": "schema--94377c156735b39dfa4ac607234cb87c",
  "name": "emailschema",
  "description": "A general schema for describing an email",
  "created": "2017-06-09T01:32:37.459000Z",
  "modified": "2017-06-09T12:44:54.459000Z",
  "version": "2",
  "object": {
    "object_type": "email",
    "email_attributes": {
      "email_format": "mixed",
      "body": {"<!DOCTYPE html> <html lang=\"en\"
xmlns=\"http://www.virtualvictorygroup.com/1c8381473ab63b28024829b4661d_ec15
8059-0101090e0005/C/\" xmlns:v= […]
"email_language": "English",
      "email_size": "20480",
      "subject": {
        "type": "Make your LinkedIn Leadgen do the work for you",
        "format": "text",
        "language": "English",
        "characters": 46
      },
      "recipient_data": {
        "recipient_number": "1",
            "items": {
              "recipient": {
              "type": "individual",
              "id": "individual--4de6805f83a7d4fa0069fcac6a7ffbe5",
              "ip": "10.170.1.170",
              "name": ["bruce"],
              "address": "bruce@untroubled.org",
              "recipient_category": "To"
          } […]
        }
       },
       "sender": {
         "type": "individual",
         "ip": "174.138.191.55",
         "name": ["Dana", "Bailey"],
         "address": "linked-university@virtualvictorygroup.com"
         }
       }
```


4.3. Pilot specific analytics

The analytic for Spam e-mail analysis has been the first analytic made available through the IAI API. At M24 the Spam E-mail Analysis is made by two API call available through the IAI APIs fully implemented and integrated with the CERT pilot. A third analytic has been added and is currently under development. The three analytics are reported in Figure 15.

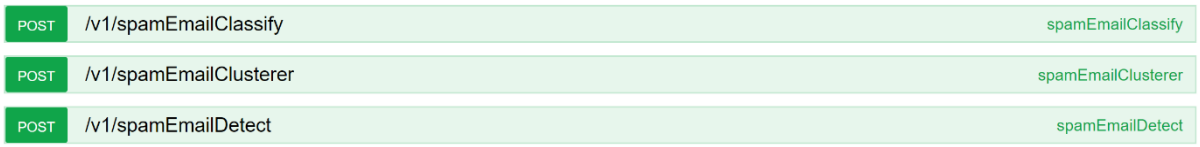

Figure 15: IAI API calls of analytics for SPAM e-mail analysis

- The *spamEmailClassify* takes as input a set of *eml* files and assigns to them a class based on the spammer goal. The analytics extract first a set of structural features converted in a vector of numerical features to be used as input for the classifier. The full list of features is reported in Figure 16. The classifier will assign to each vector one of the following classes: (i) Advertisement: E-mail used for unsolicited advertisement of products; (ii) Phishing: E-mail used for attempting to steal user credentials on some services; (iii) Malware: E-mail used as a vector for a malicious software to infect the receiving matching; (iv) Portal: E-mail generally showing a single link that redirects the user on a portal of different categories of products for sale. It is similar to advertisement but makes it harder to assign responsibilities; (v) Confidential Trick: E-mails that attempt to trick the user into paying an amount of money in exchange of a fake service. This function has been fully implemented and tested on an initial set of emails in the format described in the previous section. A more detailed description of this analytic can be found in [6] and the full description of classes and performance can be found in [4].

Attribute	Description
RecipientNumber	Number of recipients addresses.
NumberOfLinks	Total links in email text.
NumberOfIPBasedLinks	Links shown as an IP address.
NumberOfMismatchingLinks	Links with a text different from the real link.
NumberOfDomainsInLinks	Number of domains in links.
AvgDotsPerLink	Average number of dots in link in text.
NumberOfLinksWithAt	Number of links containing “@”.
NumberOfLinksWithHex	Number of links containing hex chars.
SubjectLanguage	Language of the subject.
NumberOfNonAsciiLinks	Number of links with non-ASCII chars.
IsHtml	True if the mail contains html tags.
EmailSize	The email size, including attachments.
Language	Email language.
AttachmentNumber	Number of attachments.
AttachmentSize	Total size of email attachments.
AttachmentType	File type of the biggest attachment.
WordsInSubject	Number of words in subject.
CharsInSubject	Number of chars in subject.
ReOrFwdInSubject	True if subject contains “Re” or “Fwd”.
NonAsciiCharsInSubject	Number of non ASCII chars in subject.
ImagesNumber	Number of images in the email text.

Figure 16: Features for e-mail analysis

- The *spamEmailClusterer* exploits the CCTree **Error! Reference source not found.** algorithm to separate spam e-mails in campaigns, exploiting structural similarity. Campaigns are a set of spam e-mails with a similar structure, generally generated by the same spammer. The input is a set of *eml* files from which are extracted the same features exploited by the *spamEmailClassify* analytic. This analytic at M24 is fully implemented and integrated in the CERT pilot.
- The *spamEmailDetect* analytic takes as input a set of e-mails and separates them into genuine (ham) e-mails and unsolicited ones (spam). The analytic exploits a deep learning classifier, namely a Recursive Convolutional Neural Network, trained with a

set of Ham and Spam emails coming from public datasets. The classifier is able to generalize to different datasets, showing an excellent accuracy. The output of this function is a JSON reporting the main elements of the email (subject, sender, body…), the classification “spam/not spam” and the confidence of the decision.


```
[
  {
    "language": "english",
    "sender": [
      [
        "Kholl Dzheyms",
        "info@prilivochka.ru"
      ]
    ],
    "spamProbability": "0.99968326",
    "body": "Hey an interesting offers To qualify click on the link below --- mail_boundary --- <A
href=https://drive.google.com/file/d/1XYrhijaI9HdUndb4vdLIzpzGgrSMW7_5/preview> Hey     an interesting offers To qualify click on the link below ",
    "receiver": [
      [
        "",
        "andrea.saracino@iit.cnr.it"
      ]
    ],
    "filename": "There is  an amazing  offers.eml",
    "isSpam": true,
    "mail_original": "Return-Path: <info@prilivochka.ru>\r\nDelivered-To: andrea.saracino@iit.cnr.it\r\nReceived: from smtp.iit.cnr.it ([192.168.1.151])\r\n\tby mx1 (Dovecot) with LMTP id
qUgMF6/J9VuR6gEAbw8qlA\r\n\tfor <andrea.saracino@iit.cnr.it>; Wed, 21 Nov 2018 22:12:03 +0100\r\nReceived: from localhost (localhost [127.0.0.1])\r\n\tby smtp.iit.cnr.it (Postfix) with ESMTP
id 4F5E2B8026A\r\n\tfor <andrea.saracino@iit.cnr.it>; Wed, 21 Nov 2018 22:12:03 +0100 (CET)\r\nX-Virus-Scanned: Debian amavisd-new at mx4.iit.cnr.it\r\nAuthentication-Results: mx4.iit.cnr.it
(amavisd-new); dkim=pass (1024-bit key)\r\n\theader.d=prilivochka.ru; domainkeys=pass (1024-bit key)\r\n\theader.from=info@prilivochka.ru header.d=prilivochka.ru\r\nReceived: from
smtp.iit.cnr.it ([127.0.0.1])\r\n\tby localhost (mx4.iit.cnr.it [127.0.0.1]) (amavisd-new, port 10024)\r\n\twith ESMTP id rEi7ugU5k8sw for <andrea.saracino@iit.cnr.it>;\r\n\tWed, 21 Nov 2018
22:12:01 +0100 (CET)\r\nReceived-SPF: Pass (sender SPF authorized) identity=mailfrom; client-ip=37.228.118.159; helo=prilivochka.ru; envelope-from=info@prilivochka.ru;
receiver=andrea.saracino@iit.cnr.it \r\nDMARC-Filter: OpenDMARC Filter v1.3.1 smtp.iit.cnr.it 3B8D7B80739\r\nAuthentication-Results: smtp.iit.cnr.it; dmarc=pass
header.from=prilivochka.ru\n\nX-SMTP-Auth: no\n\nReceived: from prilivochka.ru (prilivochka.ru [37.228.118.159])\n\n\tby smtp.iit.cnr.it (Postfix) with ESMTP id 3B8D7B80739\n\n\tfor
```


Figure 17: Result of the spamEmailDetect analytic on a spam email of type portal

This analytic can also be used as a pre-filter to separate ham and spam emails, passing then the resulting spam email to the SpamEmailClassify and SpamEmailCluster analyitics.

5. Prototype for the CERT Pilot

The development of the CERT pilot prototype is currently ongoing, still already offers the possibility to run a complete workflow with data and analytics relevant for the pilot.

5.1. Prototype Development Status

In Figure 18 we report the current implementation, highlighting the components related to the CERT pilot, colouring them based on their completeness.

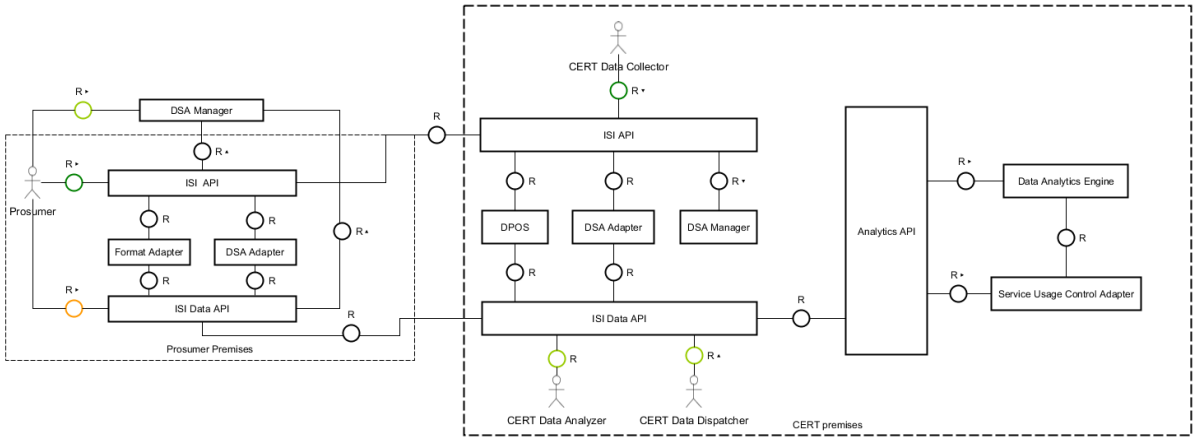

Figure 18: Prototype implementation status

The interface to the ISI API has been completed. The CERT prosumer has in fact now a GUI and a set of APIs, which can be invoked through a command line or programmatically. The GUI also links to the DSA manager components, which allows the prosumer to define his own DSAs. DSA vocabulary needed to enforce all the policies needed by the CERT pilot, will be completed in the next months.

The interface to the IAI API is currently under development, adding the possibility to invoke the required APIs as they are added to the IAI API. Also in this case, the interface available to the prosumer is a GUI and a set of API invoked via command line and programmatically. Similar functionalities are also provided to the users CERT Data Collector, Data Dispatcher and Data Analyzer.

5.2. Prototype Implementation

The GUI for interacting with the ISI API and IAI API components has been developed in Java 8, using the Bootstrap framework. The communication with the IAI API and ISI API exploits instead the Springboot framework for communication through REST interface.

5.3. Prototype Deployment

The prototype has been deployed in compliance with all other C3ISP components, by exploiting the Jenkins tool for automatic deployment and integration with the framework. The C3ISP infrastructure has been deployed in the CERT premises via installation of a virtual machine mounting Ubuntu 18.04.

6. Prototype Testing and Validation

Category	User Stories	Requirements	Validation Questions	CERT AT
CTI Collection	CERT-US-1	RVQ-COL1	Can the data collector collect CTI data?	CERT-AT-1: Verify that the data collector is able to receive data from external stakeholders.
		RVQ- COL2	Can the data collector filter the CTI data that is collected?	CERT-AT-1: Verify that the data collector is able to define policies on data.
		RVQ- COL3	Is the filtering of CTI data sufficient for the relevant stakeholder?	CERT-AT-2: Verification of rejection of non-compliant data.
CTI Processing	CERT-US-5 CERT-US-6	RVQ-PRO1	Can the CTI data be encrypted before it is shared?	CERT-AT-6 and CERT-AT-8 for verification of data protection
		RVQ- PRO2	Can the user pseudo-anonymise the CTI data before it is shared?	CERT-AT-6 and CERT-AT-8 for pseudo-anonymization conditions
		RVQ- PRO3	Can the user anonymise the CTI data before it is shared?	CERT-AT-6 and CERT-AT-8 for anonymization conditions
		RVQ- PRO4	Is the CTI processing functionality sufficient or not for the relevant stakeholder?	CERT-AT-7 and CERT-AT-9 to verify correctness and utility of analytics results.
CTI Sharing	CERT-US-3 CERT-US-5 CERT-US-6	RVQ-SHA1	Can the Data Dispatcher share data with the required partners?	CERT-AT-5 to verify the capability of the data dispatcher to return results to interested parties only.
		RVQ- SHA2	Can the relevant stakeholder prohibit specific entities from sharing the CTI data?	CERT-AT-6 and CERT-AT-8 for preventing the CERT to share

				CTIs with unwanted entities
		RVQ- SHA3	Is the CTI data sharing functionality sufficient for the relevant stakeholder?	CERT-AT-5, CERT-AT-7 and CERT-AT 11 for relevance of received results.
CTI Analysis and Results	CERT-US-2 to 9	RVQ-ARE1	Can the Data Analyser analyse the shared CTI data?	CERT-AT-3 and CERT-AT-4 for verifying functionalities of analytics.
		RVQ- ARE2	Are analysis functions sufficient for the relevant stakeholder?	CERT-AT-7, CERT-AT-9, CERT-AT-11, CERT-AT-13
		RVQ- ARE3	Can the Data Dispatcher retrieve and distribute the results of the analysis?	CERT-AT-5
		RVQ- ARE4	Can the user control who has access to the analysis results?	CERT-AT-1, CERT-AT-6, CERT-AT-8, CERT-AT-10, CERT-AT-12
		RVQ- ARE5	Is access control of the results sufficient for the user?	CERT-AT-6, CERT-AT-8, CERT-AT-10, CERT-AT-12
		RVQ- ARE6	Can the analysis and results collection be performed asynchronously	CERT-AT-3, CERT-AT-4, CERT-AT-5, CERT-AT-7, CERT-AT-9, CERT-AT-11, CERT-AT-13
		RVQ- ARE7	Can the analysis and results collection be performed synchronously	CERT-AT-3, CERT-AT-4, CERT-AT-5, CERT-AT-7, CERT-AT-9, CERT-AT-11, CERT-AT-13
Non-functional Requirements	CERT-NFR-1 to 4	RVQ- NFR1	Can the terms and conditions for using the C3ISP infrastructure be viewed and accepted/rejected?	
		RVQ- NFR1	How useful is the process of CTI data collection?	CERT-AT-6, CERT-AT-8, CERT-AT-10, CERT-AT-12

		RVQ- NFR1	How useful is the process of CTI data processing?	CERT-AT-6, CERT-AT-8, CERT-AT-10, CERT-AT-12
		RVQ- NFR1	How useful is the process of CTI data sharing?	CERT-AT-5, CERT-AT-6, CERT-AT-8, CERT-AT-10, CERT-AT-12
		RVQ- NFR1	How useful is the process of CTI data analysis?	CERT-AT-3, CERT-AT-4, CERT-AT-5, CERT-AT-7, CERT-AT-9, CERT-AT-11, CERT-AT-13
		RVQ- NFR1	How useful is the process of collecting CTI data analysis results?	CERT-AT-3, CERT-AT-4, CERT-AT-5, CERT-AT-7, CERT-AT-9, CERT-AT-11, CERT-AT-13
		RVQ- NFR4	What is the perceived security of C3ISP framework?	
		RVQ- NFR1	What is the performance of the C3ISP framework?	
		RVQ- NFR4	What are the remarks regarding C3ISP framework security features?	

6.1. Pilot's User Stories

The ISCOM-MISE embeds the CERT for private citizens, enterprises and small companies, providing a public service to consistent number of stakeholders. When defining User Stories in D3.1, we have identified the internal and external stakeholders of the C3ISP infrastructure. The internal stakeholders are those users internal to the ISCOM-MISE that perform the operations of data collection, analysis and distribution, named respectively Data Collector, Data Analyzer and Data Dispatcher. Furthermore we have considered 3 user stories coming from the domains of external stakeholders, i.e. those users that actually exploit the services offered by the CERT.

With the current version of the prototype, we were able to validate the requirements related to the user stories from 1 to 3, i.e. the ones of internal stakeholders. In fact, the current prototype has been installed in the CERT premises and validated by the internal operators for what concerns the operations of data collection, requirement specification (privacy, anonymization), data analysis and result categorization. Validation with external stakeholder will be performed at month 33 when the final version of the prototype will be available and can be presented as a full-fledged service provided by the CERT. One user story has been removed from the validation, in particular CERT-US-9. This is due to the fact that the ISCOM-MISE does not

provide any more services to the public administration, hence governmental organizations are not currently CERT stakeholders.

6.1.1. CERT-US-1

The following Table reports the validation methodology related to the user story CERT-US-1, for the user CERT Data Collector. The level of acceptance will be measured through a set of Yes or No (Y/N) questions and a 1-10 scale for measuring the level of improvement brought by the introduction of C3ISP.

Goal	CERT – US- 1	As a CERT data collector, I want to receive information about incident and vulnerabilities which affected or might affect my stakeholders,so that I can promptly list and communicate them.
Acceptance Test ID	**Questions**	**Metrics – (Answer)**
CERT-AT-1	Were you able to automatize through C3ISP the process of data collection related to vulnerabilities?	Y/N – (TBD)
	Were you able to automatically assign DSAs to data in a specific folder?	Y/N – (Y)
	Were you able to store and retrieve the correct information from the C3ISP ISI APIs?	Y/N – (Y)
	Rate from 1 to 10 the procedure of defining a DSA through the DSA-Manager web tool.	Numeric scale 1-10 – (7)
CERT-AT-2	Have you been notified of the non-compliance issue with standard regulations of specific files?	Y/N – (TBD)

The first question of CERT-AT-1 requires interaction with an external stakeholder. In the current version of the framework has not been possible to validate it.

CERT-AT-2 requires the evaluation of complex policies not yet available with the current prototype.

6.1.2. CERT-US-2

The following Table reports the validation methodology related to the user story CERT-US-2, for the user CERT Data Analyzer. The level of acceptance will be measured through a set of Yes or No (Y/N) questions, aimed at evaluating the functionality of the C3ISP framework.

Goal	CERT – US- 2	As a CERT data analyser of MSS data, I want to infer automatically useful information about incident and vulnerabilities from large amounts of unorganized data, so that I

		can reduce the amount of work and the time needed to detect and communicate a vulnerability.
Acceptance Test ID	**Questions**	**Metrics – (Answer)**
CERT-AT-3	Did the analytics brought useful results on the data you provided?	Y/N – (Y)
CERT-AT-4	When requiring access to data whose policy denies access in specific conditions, such as time interval, where you able to access those data when condition was not met?	Y/N – (N)
	Where you able to identify the original sender of an email after anonymization?	Y/N – (N)

6.1.3. CERT-US-3

The following Table reports the validation methodology related to the user story CERT-US-3, for the user CERT Data Dispatcher. The level of acceptance will be measured through two questions where the improvement in the workflow should be reported in numerical terms. Furthermore, a more objective test will be performed by giving a set of processed data to the system and verifying that they are delivered to the correct recipient. False Negatives and False Positives will be used as evaluation index.

Goal	CERT – US- 3	As a Vulnerability info dispatcher, I want to Automatically categorize information stakeholders So that Vulnerabilities are communicated easily and automatically
Acceptance Test ID	**Questions**	**Metrics – (Answer)**
CERT-AT-5	Rate from 1 to 10 the improvement in simplicity brought by the C3ISP framework to the process of notifying customers about new vulnerabilities.	Numeric scale 1-10 – (8)
	Rate from 1 to 10 the improvement in timeliness brought by the C3ISP framework to the process of notifying customers about new vulnerabilities.	Numeric scale 1-10 – (10)
	After the test execution, did the recipient received any unwanted information?	FNR – Percentage – (TBD)
	After the test execution, did the recipient missed any desired information?	FPR – Percentage – (TBD)

The CERT-AT-6 has not been performed yet, since it is necessary to have a complete version of the framework to correctly run the tests with external stakeholders.

6.1.4. CERT-US-4

In the following Table we report the GQM validation methodology for the CERT-US-4, i.e. the one for the Enterprise user. This test is related to the retrieval of information related to possible vulnerabilities relevant for the company. The use case will consider that the Enterprise is sharing information related to its systems and this should be done in a privacy preserving way. The level of acceptance will be measured through two Y/N answers and the percentage of correct results received through the vulnerability search analytic.

Goal	CERT – US- 4	As an Enterprise I want to be informed about major threat and vulnerabilities related to my sector, so that I can take countermeasures and protect my systems, employee and customers.
Acceptance Test ID	**Questions**	**Metrics**
CERT-AT-6	Where you able to define the privacy conditions needed to preserve eventual shared information about your system?	Y/N (N/A)
	Where you able to define the party that CAN and/or CANNOT view the information you are sharing?	Y/N (N/A)
CERT-AT-7	Where you able to receive vulnerability information relevant for your system?	FNR – Percentage (N/A)

6.1.5. CERT-US-5

The following Table reports the questionnaire for the validation of the CERT-US-5. This user story is related to the analysis of spam emails requested by an Enterprise to avoid malware infection. This user story implies that the Enterprise is willing to share (be a producer) of email files, with different privacy requirements. Hence, the enterprise will expect from the C3ISP framework to know which email is actually a malware vector and if there is the possibility to identify one or more spam campaigns (different emails sent by the same attacker) out of the shared email files. The validation is performed by a set of Y/N questions to assess if the functionalities work as expected.

Goal	CERT – US- 5	As an Enterprise, I want to be protected from malware which might be received through spam email and recognize email attempts to trick my users in giving private information via email, so that I can avoid damages to my company and my employees
Acceptance Test ID	**Questions**	**Metrics**
CERT-AT-8	Where you able to define the privacy conditions needed to preserve the privacy of the email files you are sharing for analysis?	Y/N (N/A)
	Where you able to associate the correct level of anonymization to each party that	Y/N (N/A)

	CAN and/or CANNOT view the email files you are sharing?	
CERT-AT-9	Where the available analytics for spam email satisfactory?	Y/N (N/A)
	Did you detect through the framework any attempt to infect your systems through spam emails?	Y/N (N/A)
	Where you able to find correlations among the spam email received, possibly identifying a single attacker?	Y/N (N/A)

6.1.6. CERT-US-6

The following Table describes the acceptance tests for the user story CERT-US-6, again involving an Enterprise as a customer. This user story assumes that the Enterprise is willing to share its network logs for analysis, expecting to know if it is possible to identify DDoS attack traces. The acceptance test is aimed at verifying that the Enterprise user is able to define, with the C3ISP framework, security policies to anonymize the network logs, and to receive the expected information.

Goal	CERT – US- 6	As an Enterprise, I want to be protected from Denial of Service attacks so that I can avoid unavailability of my services and failures of my IT system.
Acceptance Test ID	**Questions**	**Metrics**
CERT-AT-10	Where you able to define the privacy conditions needed to anonymize your network logs?	Y/N (N/A)
	Where you able to define the stakeholders that CAN and CANNOT access and use your data for analytics?	Y/N (N/A)
	Where you able to define the accepted provenance for data that can be used for analytics involving your data?	Y/N (N/A)
CERT-AT-11	Where the available analytics for traffic analysis satisfactory?	Y/N (N/A)

6.1.7. CERT-US-7

The following Table reports the validation for the CERT-US-7 which involves an SME as user. This user story assumes that the user would like to receive information about threats that might be able to affect its system. Sharing information for this user story is not mandatory. The acceptance is performed through two Y/N questions related respectively to data sharing and data analysis.

Goal	CERT – US-7	As a SME, I want to be protected from malware which might be received through different channels, so that I can implement suggested counter-strategies and recovery best practices.
Acceptance Test ID	**Questions**	**Metrics**
CERT-AT-12	Where you able to define privacy rules for eventual information that you shared with the CERT?	Y/N (N/A)
CERT-AT-13	Where you able to receive specific information about malware that can affect your systems and the appropriate countermeasure?	Y/N (N/A)

6.1.8. CERT-US-8

The following Table reports the validation for the CERT-US-8 which involves an ISP as user. This user story assumes that the user would like to receive information about threats that might be able to affect the IP blocks it is managing. The acceptance is performed through two Y/N questions related respectively to data sharing and data analysis.

Goal	CERT – US-7	As an ISP, I want to receive automatically any information related to *incidents* and vulnerabilities involving my IP blocks and systems, so that I can take immediate action on the interested IPs and systems.
Acceptance Test ID	**Questions**	**Metrics**
CERT-AT-14	Where you able to define the parties that can access and use the data you provide?	Y/N (N/A)
CERT-AT-15	Where you able to receive specific information about threats that can affect your IP blocks?	Y/N (N/A)

6.2. Non Functional Requirements

The non-functional requirements for the CERT pilot, as defined in [2] are the following:

CERT-NFR-1 Communication between the provider and CERT should be protected through the C3ISP framework. The compliance with this requirement is automatically ensured by the C3ISP framework since all communication with the C3ISP framework are protected via TLS, with mandatory https. Formally the validation of this requirement should be completed by testing communication between the CERT and external stakeholder.

CERT-NFR-2 Received information should match a standard format. This requirement is validated since all the information exchanged in C3ISP are encoded in STIX format.

CERT-NFR-3 The CERT analyser might not be allowed to see some data to be analysed. This requirement has been validated by locally testing the implementation of two anonymization

algorithm to anonymize the mail recipient address and to directly communicate the numerical feature vector of Figure 16.

CERT-NFR-4 Communication between the CERT and data recipient should be protected. This requirement is dual to CERT-NFR-1 and has the same validation procedure.

6.3. Bug and feature tracking

The bug tracking has been done via support from the C3ISP consortium. Non relevant bugs have been observed in this initial test phase, since the components have been largely tested in CNR premises before being deployed in the CERT premises.

7. Conclusions and Future Work

At month 26 the CERT pilot has experienced a first deployment of the C3ISP architecture and has been capable of validating the first functionalities validating part of the user stories for the internal CERT stakeholders and defining an initial set of policies. Also the CERT has provided an initial set of data used to test relevant analytics and validate the related use case. In next months the WP3 will look forward to the full deployment of the complete C3ISP framework, starting to provide the project services to external stakeholders and validating the remaining use cases. A summary of the acceptance test is reported in the following:

Test	Brief Description	Responsible Component or Use Case	Result / Status
CERT-AT-1	Collect MSS data from C3ISP	CERT-US-1	In progress
CERT-AT-2	Conformance to legal constraints	CERT-US-1	In progress
CERT-AT-3	Collaborative analysis brings useful results	CERT-US-2	Validated
CERT-AT-4	CTI Data is sanitised or analysis is forbidden	CERT-US-2	Validated
CERT-AT-5	C3ISP delivers stakeholder-specific reports that can be validated	CERT-US-3	Validated
CERT-AT-6	Enterprise-specific events	CERT-US-4	In progress
CERT-AT-7	Enterprise analysis insight	CERT-US-4	In progress
CERT-AT-8	Receive Enterprise email CTI	CERT-US-5	In progress
CERT-AT-9	Classify email malware CTI	CERT-US-5	In progress
CERT-AT-10	Receive Enterprise network CTI	CERT-US-6	To be done
CERT-AT-11	Detect Enterprise DoS attacks	CERT-US-6	To be done
CERT-AT-12	Receive SME threat information	CERT-US-7	In progress
CERT-AT-13	Recommend SME threat removal	CERT-US-7	In progress
CERT-AT-14	Receive ISP CTI alerts	CERT-US-8	To be done
CERT-AT-15	ISP analysis insight	CERT-US-8	To be done

CERT-NFR-1	Communication between the provider and CERT should be protected through the C3ISP framework.	C3ISP front-end	In progress
CERT-NFR-2	Received information should match a standard format.	ISI API	Validated
CERT-NFR-3	The CERT analyser might not be allowed to see some data to be analysed	ISI	Validated
CERT-NFR-4	Communication between the CERT and data recipient should be protected	C3ISP front-end	In progress

The following deliverable, i.e. D6.4, will report the results of the completed validation and full deployment of the C3ISP architecture in the CERT premises.

8. References

[1] A. Saracino, F. Martinelli, S.Mari, CERT Pilot Architecture, C3ISP deliverable D3.2.

[2] A. Saracino, F. Martinelli, S.Mari, Requirements for the CERT Pilot, C3ISP deliverable D3.1.

[3] M. Sheikhalishahi, A. Saracino, M. Mejri, N. Tawbi, F. Martinelli, Fast and Effective Clustering of Spam Emails based on Structural Similarity, FPS 2015.

[4] M. Sheikhalishahi, A. Saracino, M. Mejri, N. Tawbi, F. Martinelli, Digital Waste Sorting: A Goal-Based, Self-Learning Approach to Label Spam Email Campaigns, STM 2015.

[5] Y. Lecun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, Proceedings of the IEEE 1998.

[6] T.H. Nguyen et al, C3ISP Data Sharing, Analytics and Crypto Technology Maturation, C3ISP deliverable D8.2

[7] M. Manea et al, First version of C3ISP Architecture, C3ISP deliverable D7.2

[8] M. Manea et al, First version of the C3ISP platform and test bed, C3ISP deliverable D7.3

9. Appendix A – Installation Guide for CERT Pilot

The components of the CERT pilot can be easy installed both in the CERT premises and in external stakeholder system. Following the design pattern used throughout C3ISP, in fact, all functionalities are provided as web services. The general requirement for installing a C3ISP component is thus the presence of a web server. As detailed in [8], the webserver used for testing the C3ISP framework is the Apache Tomcat web server. We will now detail the components to be installed and their requirements:

IAI

Location: CERT Systems

Comes as: Set of dependent WAR files

Installation: IAI components are released as a Web Archive (.war) files; installation is straightforward and requires to upload the war files onto Tomcat. The Tomcat version shipped with Ubuntu Linux requires the war files to be copied to /var/lib/tomcat8/webapps.

Table 1 – IAI components

Component	Path and file
IAI API	/var/lib/tomcat8/webapps/iai-api.war
FHE Analytics	/var/lib/tomcat8/webapps/fhe-conn-malicious-host.war
C3ISP Analytics Engine	/var/lib/tomcat8/webapps/monitoring-dga.war

ISI

Location: CERT Systems

Comes as: Set of dependent WAR files and NodeJS files.

Installation: The following ISI components are released as a Web Archive (.war) files; installation is straightforward and requires to upload the war files onto Tomcat. The Tomcat version shipped with Ubuntu Linux requires the war files to be copied to /var/lib/tomcat8/webapps.

Table 2 – ISI components (1)

Component	Path and file
DSA Adapter Front End	/var/lib/tomcat8/webapps/dsa-adapter-frontend.war
Event Handler	/var/lib/tomcat8/webapps/event-handler.war
Continuous Authorization Engine	/var/lib/tomcat8/webapps/multi-resource-handler.war /var/lib/tomcat8/webapps/UsageControlFramework.war
Obligation Engine	/var/lib/tomcat8/webapps/trigger-engine.war /var/lib/tomcat8/webapps/action-engine.war
DMO Engine	/var/lib/tomcat8/webapps/dmo-engine.war
Bundle Manager	/var/lib/tomcat8/webapps/bundle-manager.war
Buffer Manager	/var/lib/tomcat8/webapps/buffer-manager.war
Data Protected Object Storage	/var/lib/tomcat8/webapps/dpos-api.war
ISI API	/var/lib/tomcat8/webapps/isi-api.war

The following ISI component runs on NodeJS. It is required to deploy this on top of NodeJS; we use PM2 (Process Manager 2[1], a nodejs application runner):

Table 3 – ISI components (2)

Component	Path and file
Format Adapter	/home/nodejs/format-adapter/converter.js

Local ISI

Location: Customer premises.

Comes as: Set of Dependent WAR

Installation: Same as for ISI.

User Interface

Location: Customer premises

Comes as: WAR file

Installation: The installation is straightforward and requires to upload the war file onto Tomcat. The Tomcat version shipped with Ubuntu Linux requires the war file to be copied to /var/lib/tomcat8/webapps.

[1] http://pm2.keymetrics.io/