D5.1
Pilot Deployment and Evaluation Report

Grant Agreement number:	740723
Project acronym:	CS-AWARE
Project title:	A cybersecurity situational awareness and information sharing solution for local public administrations based on advanced big data analysis
Principal author:	Jerry Andriessen, Wise & Munro jerry@wisemunro.eu
Contributors	Thomas Schaberreiter, University of Vienna Christopher Wills, Caris Research Kim Gammelgaard, Rheasoft John Forrester, Manuel Leiva, Massimo DellaValentino, Cesviter Consulting Laurentiu Vasiliu, Peracton
Internal Reviewers	Stefania Tola, Thomas Schaberreiter, Chris Wills
Document version:	1.0

Executive Summary

This Deliverable provides a report of the deployment and evaluation activities and outcomes during the final year of the CS-AWARE project. The main objective of the CS-AWARE project (https://cs-aware.eu/) is to provide a cybersecurity situational awareness software solution for Local Public Administrations (LPA's), small- to medium-sized enterprises (SME's) and financial institutions. This solution enables the detection, classification and visualisation of cybersecurity incidents in real-time, supporting the prevention or mitigation of cyber-attacks. It is intended that this solution will provide an intermediary step towards automation of cyber incident detection, classification and visualisation, that impacts on cybersecurity awareness of users. User involvement is crucial in all steps of design, implementation, deployment and evaluation.

Deployment can be defined as locally implementing the system and making sure that the system functions in the two LPA contexts. Like everything else in this project, the nature of this work is socio-technical. This means that ongoing feedback of users is a crucial aspect, as well as testing the technology in the authentic user context. We collected user feedback and tested the technology during three cycles of three months. For each cycle, we implemented user feedback as an ongoing process until the very end of the project, this is called agile programming and software development. At the start, we established the user needs and ambitions in a deployment scenario. This served to shape our deployment and evaluation activities.

For evaluation, the main qualitative input came from frequent usability testing and interactions with users from the pilot deployment teams. Through these flexible methods, we gathered valuable information about using the system and refinement of its functionalities. In combination with the agile approach, feedback could be implemented quickly. Through the use of questionnaires, we got informed about how users rated the system, in terms of general characteristics, usability, their awareness of cybersecurity, and the impact on the organisation. Our attempts to get responses from a fifth questionnaire, for potential new users in other municipalities failed to receive any response, most probably because these users were focusing on the corona pandemic rather than on cybersecurity. Together, the questionnaire scores were very positive, which confirmed the KPIs set for our system.

At the end of deployment, users successfully completed the four exercises that were part of the final usability test. This means that our users, and very probably also users in other municipalities and professional contexts, with similar motivation and experience, can work with the CS-AWARE system to perceive, comprehend, project and mitigate cybersecurity threats. Moreover, we have shown that working with CS-AWARE increases awareness of cybersecurity, this applies to individual users, such as system administrators but also to greater awareness of cybersecurity within their organisation. This is the basis for realising the objectives that users formulated in the deployment scenario: efficient detection of cyberthreats, and better communication and collaboration between departments, between system administrators and managers, and between departments and citizen users of services. Better collaboration, learning, and increased reputation can be built on this basis.

Table of contents

Deployment and Evaluation Report for the pilots in Roma Capitale and Larissa

Jerry Andriessen[*], Thomas Schaberreiter[**], Chris Wills[○] & Kim Gammelgaard[Δ]

[*] (main author, usability and awareness evaluation) Wise & Munro, Learning Research, NL
[**] (reviewer, technical coordinator, technical level evaluation) University of Vienna, AT
[○] (reviewer, SSM workshops, organisational level evaluation) Caris Research, Fowey, UK
[Δ] (usability) Rheasoft, Aarhus, DK

WP5 Objectives - *The main objective of this work package is the deployment of the CS-AWARE solution at each of the user sites as pilot case studies and to collect and evaluate the results and success of the solution through end user validation. This objective is achieved through local pilot-specific development and learning. This means that we will evaluate the specific dynamics of each pilot to see to what extent this case realises the goals that have been set by the public administration and stakeholders. Furthermore, through assessment or comparative benchmarking, the results from the different pilots will be compared to evaluate similarities and differences. This type of evaluation is needed to draw more generalisable conclusions about cybersecurity solutions. Finally, the results of the pilots will enable lessons to be learned and conclusions to be drawn from each pilot to help form the final shape of the CS-AWARE solution, evaluating the usefulness of the solution.*

I. Introduction

This deliverable provides a stepwise walkthrough for Experts for understanding the deployment and evaluation activities of the CS-AWARE cybersecurity awareness solution[1] at the municipality of Roma Capitale, and at the municipality of Larissa, during three cycles of design, collecting feedback, testing and revision (see table 1 below).

Cycle	Period	Months	Section in this report
1	Oct-Dec 2019	M26-28	V & VI
2	Jan-March 2020	M29-31	VII, VIII, IX
3	April-June 2020	M32-34	X, XI

Table 1: Three Deployment cycles

This report addresses the following evaluation questions*:* 1) To what extent is the technical implementation effective? 2) To what extent is the CS-AWARE system usable by expected target users? 3) To what extent is the awareness of users affected by discussing and using the system during deployment? 4) To what extent does using the CS-AWARE system have an impact on cybersecurity awareness at the organisational level? 5) To what extent can other municipalities be involved in our approach?
In this report we first summarise (in section II) the activities that involved users in building the CS-AWARE system. We then (in section III) present the deployment scenario, and our evaluation methodology (section IV). In subsequent sections, deployment and evaluation outcomes for each of the cycles are described, summarised and main conclusions will be derived. In our conclusions, we will also focus on differences between the two municipalities, the lessons learnt, and on methodological issues. Final conclusions are presented in section XII.
The technological developments of the system, partly inspired by user feedback, will be reported for cycles 2 and 3, as part of deployment activity. For these cycles, we will explicitly refer to requirements and KPI's for evaluation outcomes.

II. The CS-AWARE Project

The main objective of the CS-AWARE project (https://cs-aware.eu/) is to provide a cybersecurity situational awareness software solution for Local Public Administrations (LPA's), small- to medium-sized enterprises (SME's) and financial institutions. This solution enables the detection, classification and visualisation of cybersecurity incidents in real-time, supporting the prevention or mitigation of cyber-attacks. It is intended that this solution will provide an intermediary step towards automation of cyber incident detection, classification and visualisation, that impacts on cybersecurity awareness of users. User involvement is crucial in all steps of design, implementation, deployment and evaluation. In this section, we briefly revisit the three iterations of *workshops* for building the application during the first 25 months of the project. We do this because these workshops have strong user involvement and therefore contribute to awareness of cybersecurity in

[1] The CS-AWARE project was funded by Horizon 2020 research and innovation programme, under grant agreement No: 740723. It started on September 1st, 2017 and is due to end on August 31st, 2020.

the LPAs. After this period, we distinguish three deployment *cycles*, each of about three months, during which we systematically tested and revised the CS-AWARE application with the two pilots. The technical work until month 28 is reported extensively in WP3 and 4 deliverables. Technical development after month 28 is reported the current deliverable.

Workshop 1: The method exploited in three workshops is part of the analysis according to the soft systems methodology (SSM) that was chosen for CS-AWARE. In CS-AWARE the analysis and identification of assets, dependencies and monitoring points of the existing and organically grown complex socio-technological systems found in all larger organizations - like Local Public Administrations (LPAs) - is an integral part of the proposed cybersecurity awareness solution. We argue that in complex systems good cybersecurity awareness can only be provided if the relevant relations between the mission critical aspects of the system are understood, and relevant case specific monitoring points can be utilized. The *first workshop* focused on getting an overview analysis of cybersecurity related aspects, relevant to CS-AWARE in the context of LPAs. The analysis is based on three thematic focus points: an initial threat assessment for LPAs, an analysis of external information sources that may be relevant to CS-AWARE, and an analysis of the pilot scenarios (relevant work flows) collected from the users during the first round of workshops. We have investigated potential monitoring points at 4 different levels that allow to identify suspicious behaviour related to data operations: the database level, the application/service level, the network level and the security appliance level.

Workshop 2: A second workshop was organized to refine our understanding in the three main thematic focus points covered in the first iteration, as well as to assess a fourth focus point, the definition of CS-AWARE use cases, based on our understanding and results of the first three topics. This allowed us to identify the critical processes of the services that are used for CS-AWARE piloting, and the associated information flows through the system that those processes create in day-to-day operations of both system users and administrators. To achieve this understanding, the workshops were organized in two parts: the system and dependency focused workshop to refine the understanding of systems and processes already started in the first round of workshops, as well as a more end-user focused story telling workshop to determine the cybersecurity related problems users and administrators alike face on a daily basis, and the procedures and processes used to solve those problems. The outcomes of this second workshop and how these were exploited is further described in section VII on awareness.
Four specific *use cases* could be identified: (1) Vulnerability awareness (map classified vulnerabilities to specific LPA systems/components); (2) Behaviour analysis (identify suspicious behaviour and if possible, classify according to data received from threat intelligence). (3) General security warnings (informing about general and/or currently ongoing security events that may become relevant to the specific context of each LPA); and (4) an analysis of connections originating from or going to specific IP/DNS entries that are classified as malicious by relevant communities can be conducted. We assume that the list of use case scenarios covers all aspects relating to cybersecurity awareness that can be covered considering the data provided by the various communities.

Workshop 3: In the third workshop the goal was to define, together with the LPA users and building upon the first two iterations, normal and abnormal behaviour within the identified business processes (pilot scenarios), and how this behaviour is reflected within the data sources collected from the LPA systems on the database, service, security appliance and network level. The behaviour of system elements during day-to-day operations according to the identified business processes, and how this reflects in the data sources CS-AWARE collects, is a crucial input for the definition accurate and relevant monitoring patterns. The resulting patterns were validated through the consent of CS-AWARE security and data analysis experts as well as the employees of the Municipalities (users, administrators, managers) who are the ones the cybersecurity awareness system is intended for. Similarly, self-healing policies have been defined that allow mitigation of events detected by cybersecurity patterns in an automated way.

During the first 25 months, the system was developed, and the outcomes of the workshops implemented. Then, deployment and evaluation started.

III. THE DEPLOYMENT SCENARIO

Deployment can be defined as locally implementing the system and making sure that the system functions in the two LPA contexts. Like everything else in this project, the nature of this work is socio-technical.
This means that ongoing feedback of users is a crucial aspect, as well as testing the technology in the authentic user context.

For better understanding the socio-technical aspects of deployment of the system in the municipal contexts, we developed a systematic co-creation approach, called a *deployment* scenario (Figure 1). Similar to the joint efforts in the SSD workshops, co-creation of the elements of a deployment scenario (by researchers and stakeholders at the municipalities) can be seen as part of a trajectory towards more awareness of cybersecurity of public administrations

	System administrators	Managers	Local users involved in services
Objectives	**(1) Easy treat identification and classification** **(2) Reduction of time for threat understanding** **(3) Mitigation suggestions**	**(1) More effective relation with service providers in handling cybersecurity** **(2) Better informed about threats and mitigation** **(3) Quality of services** **(4) More satisfied citizens**	**(1) Efficiency of work** **(2) Service reliability** **(3) Personal data protection** **(4) Increased trust**
System artefacts	**(1) Trouble ticket on mobile device** **(2) Real time info on number, resolution time, types of attacks**	**(1) Weekly incident reports** **(2) Monthly trend reports**	**Users indicated they did not want to be notified of threats *before* they were resolved**
Desired Behaviour	**(1) Using the tool on a daily basis** **(2) Reflection on past problems and solutions** **(3) Regular communication with technical team and internal users**	**(1) Improved knowledge** **(2) Policy making** **(3) Proactive approach to users *and* senior management**	**(1) Follow the guidelines** **(2) Acquire information** **(3) Communication with citizens**

Table 2: Some negotiated elements of the deployment scenario for Roma Capitale

(Schønheyder & Nordby, 2018)[i]. This makes deployment also a learning trajectory, with greater awareness of cybersecurity as the outcome. Our approach is similar to what is called project articulation, which starts with developing a joint vision rather than with a list of things to do (Strauss, 1988)[ii]. To support our municipalities in developing clear goals and expectations, and in articulating these together, we employed a scenario template, to guide the discussion with the pilots. This template was based on earlier experiences with this procedure in another project[iii].

The template was applied during the same meeting as the third SSM workshop. The template was revisited with the users but they did not feel the need to make any changes until the end of deployment.
The 9 elements of the scenario are as follows: (1) *The Deployment Team*: At the pilot sites, a local deployment team is established. The role of the deployment team is to design, prepare and monitor execution of the deployment scenarios in coherence with the requirements. In Rome, the deployment team consisted of system administrators (n=6), managers (n=5), and local users from the municipality (n=2). In Larissa, the team comprised a manager/owner, 3 system administrators and their unit manager, and two service users from the municipality. Members of the deployment team are also involved in evaluation activities. (2) *The Deployment Objectives*: These are negotiated local objectives for deploying the system, made explicit for the different stakeholders. (3) *Desired Artefacts*: This is the tangible output from the CS-AWARE system that users desire for their context. This means the concrete output of the system: a report about incidents, an overview of what was shared, what healing operations succeeded, etc. (4) *Desired User Behaviour*: What will users do to make sure that the objectives will be realised? We are interested in visible, manifest activity: what does effective (and not effective) stakeholder activity look like? How can we observe things are going well, as desired, or not? (5) *Participants*: we discuss within the deployment team who will participate in the various meetings and tests in using the system. (6) *Impact*: Impact of deployment is what we expect to change in the organisation when deployment objectives will be realised, in a long-term perspective. Thinking about desired impact allows us to understand 'how far we are' – what can be realistically achieved, in the eyes of the users. (7) *Evaluation:* Evaluation will be carried out at five levels, corresponding to the five main research questions (section I): (a) The technical

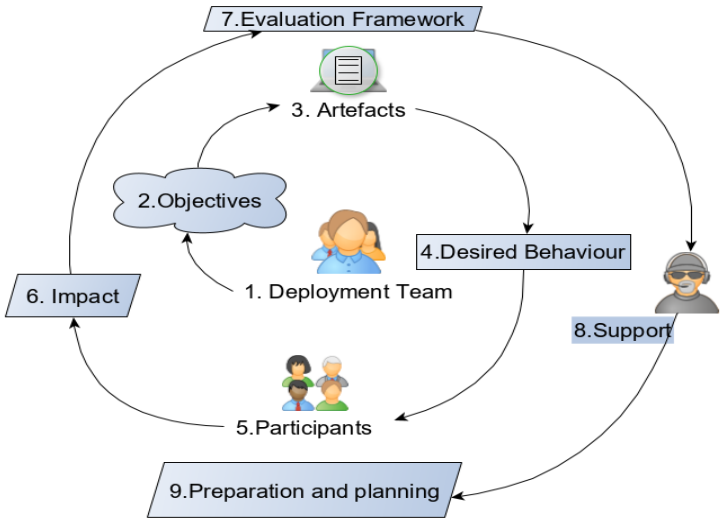


Figure 1: The deployment scenario

	System administrators	Managers	Local users involved in services
Objectives	**(1) Timely detection of treats (2) No false alarms (3) Up-to-date information (4) Trustworthy information (5) Room for making your own decisions**	**(1) No service down time (2) No extra costs or resources (3) No reputation damages**	**(1) No additional burden (2) Feel safe and protected (3) Timely, clear, concise information (4) Not being watched**
System artefacts	**(1) Ranking of threats by frequency (monthly) (2) List of system affected nodes (weekly) (3) Report of information shared by other LPA's (monthly)**	**(1) Weekly incident reports (2) Weekly report of services affected**	**Weekly report of threat sources**
Desired Behaviour	**(1) Assigning monitoring roles on a daily basis (2) Actively learn from the system (3) Collaboratively discussing solutions**	**(1) Reading the reports (2) Increased trust in system administration**	**Better following the guidelines**

Table 3: Some negotiated elements of the deployment scenario for Larissa

Deployment Scenario Element	Is used for…..
Team	**• Contacts & feedback**
Objectives	**• Views on using the system • Evaluation (c): User Awareness**
Artefacts	**• Information from CS-AWARE system**
Behaviour	**• Evaluation (c-d): Awareness**
Impact	**• Long term expectations • Evaluation (d): organisational awareness**
Support	**• Immediate needs, manuals**
Planning	**• Preparation and planning**

Table 4: Use of information from the deployment scenario

level of system functioning; (b) Usability of the interface; (c) User awareness; (d) Organisational awareness; and (e) marketability. Evaluation is further explained in the next section (IV). (8) *Support*: Intended here is the support foreseen during deployment, within the organisation and/or by the deployment team: is there someone watching over the users, technical support, are there manuals or training. (9) *Preparation and planning*: this involved all preparatory activities for the deployment: technical preparation, implementation and testing, recruiting users, communication with users beforehand, introductions, manuals, preparatory meetings. In Annex 1 the full versions of the deployment scenarios can be found, for both LPAs. Tables 2 and 3 describe particular highlights focus points for Rome and Larissa, respectively. After the deployment scenario was established, we implemented regular meetings with the local deployment team. During these meetings, we further discussed the status of the deployment scenario, and collected feedback on various aspects of deployment.

Table 4 shows how we have used the various elements of the deployment scenario: for further implementation of the technology, for understanding and user feedback, and for evaluation.

IV. EVALUATION

Evaluation of the CS-AWARE system is part of deployment. During deployment, aspects of this system are still being modified, as a result of internal discussions in the team, and of user feedback. Agile development of technology, combined with soft systems methodology, does not enable the clear prediction of outcomes. Instead, what they afford is a solution that is closely tied to user needs and local circumstances. Therefore, in line with this approach to design, we developed an appropriate approach to evaluation, which is design-based, and evaluation is seen as a formative intervention (Berliner, 2003[iv]; The Design-Based Research Collective, 2003[v];Paavola et al., 2011[vi]; Christensen & West, 2017[vii]). The following *methodological principles* for evaluation (quoted and adapted from Engeström, 2011[viii]) apply to our situation: (1) participants face a problematic and contradictory object, i.e. cybersecurity, embedded in their vital life activity, which they analyse and expand by constructing a novel concept, cybersecurity awareness, the contents of which are not known ahead of time to participants nor to the researchers; (2) the contents and course of the intervention, using the CS-AWARE system, are subject to negotiation and the shape of the intervention is eventually up to the participants, who thereby gain agency (from awareness) and take charge of the process (ownership); (3) the aim is to generate new concepts that may be used in other settings as frames for the design of locally appropriate new solutions; (4) the researcher aims at provoking and sustaining an expansive transformation process led and owned by the practitioners. In more practical terms, this means *a cyclical approach*: (a) we will be collecting evidence from the users: their ideas on cybersecurity and their activities with the CS-AWARE system; and (b) adapt the design (of one or more CS-AWARE system components) based on that evidence, and (c) revisit how this affects the ideas and

#	Requirements	F	NF	EU
	Technical System Requirements			
S1	Provide cybersecurity awareness	X		X
S2	Allow information sharing	X		X
S3	Enable system self-healing	X		X
S4	Enable data collection from internal LPA and external cyber security information sources	X		
S5	Allow pre-processing to bring data into a unified format	X		
S6	Enable data analysis by setting external and internal data into context	X		
S7	Ensure international usability of the system by providing multiple languages	X		
S8	Identifying relevant internal and external sources	X		
	General System Requirements			
S9	Usability (The usability of the CS-AWARE system, as determined by the end users)		X	X
S10	Compliance (Compliance to LPA regulations, policies and procedures)		X	X
S11	Integratability (Integratability of CS-AWARE system into LPA work flows)		X	
S12	Open Source (How much of the CS-AWARE components can be open sourced and how much is kept proprietary)		X	
S13	Internationalization (Integration into different cultural and language contexts)		X	
S14	Cost/Marketability of CS-AWARE solution		X	

Table 5: System Requirements (F: Functional; NF: Non-Functional; EU: End user based

activities of the users; etc. A researcher should therefore not impose understanding of the main concepts, it is up to the users how they interpret those concepts and use the technology.

The hybrid approach: In addition to the formative intervention methodology, we decided to also collect quantitative evidence to evaluate the CS-AWARE system. To this end, we developed a hybrid approach to evaluation that therefore involves two sources of evidence (A) *Qualitative evidence*, derived from (1) the deployment scenarios designed at the beginning and discussed at the end of deployment, (2) user verbalisation and actions during usability testing of the system, and (3) feedback from users collected at regular intervals during meetings; and (B) *Quantitative evidence*, collected through 5 questionnaires, one for each level of evaluation, and administered twice during deployment.

The hybrid approach serves to provide answers to the following questions: 1) *Technical Level*: to what extent is the technical implementation effective? 2) *Usability Level:* To what extent is the CS-AWARE system usable by foreseen target users? 3) *Awareness Level:* To what extent is the awareness of users affected by discussing and using the system during deployment? 4) *Organisational Level:* To what extent does using the CS-AWARE system have an impact on cybersecurity awareness at the organisational level? 5) *Business Level:* To what extent can other municipalities be involved in our approach?
Qualitative evidence is used for understanding and evaluating how users work with the system, as well as how their awareness (qualitative understanding) of cybersecurity evolves during deployment. Quantitative evidence provides evidence on the attainment of key performance indicators (KPIs), related to requirements for each level, as specified in the next section.

IV.1 Requirements, KPI's, Questionnaires

Requirements: Within the context of the CS-Aware framework and the project objectives, we formulated two sets of requirements that the system is supposed to meet (table 5). The first set were *technical requirements*, of which the first three match the general objectives of the project, and the others refer to technical system components. These requirements were functional: they pertained to functionalities that the system should offer. The other set involves non-functional *general system requirements*, which are requirements for the system as a whole. These general requirements (except S12, 'open source') are to be tested with end users. The three main system requirements (S1, S2, S3) are also subjected to end user evaluation. The degree to which our approach is supposed meets the requirements is expressed in terms of *key performance indicators (KPIs)*, which are scores on a scale that maps test results to specific requirements, in terms of the five research questions. For defining KPIs in the next subsections, we take the name of the subscale, not the score (which always has to be 60% or higher). As figure 2 shows, the KPIs are coupled to the five research questions, and provide therefore important evidence for the success of the CS-AWARE system on the five different levels.
For each evaluation level, we developed a questionnaire. Questionnaires 1-4 will be answered at the pilot sites, questionnaire 5 is for potential users outside our pilot sites. The questionnaire items allow users to indicate the degree to which they agree with a statement (Likert Scale 1-5, a well-accepted scientific approach: Allen & Seaman, 2007[ix]). All questionnaires start (or end) with a section where the participants can indicate their experience with CS-AWARE, and their role in their organisation. We

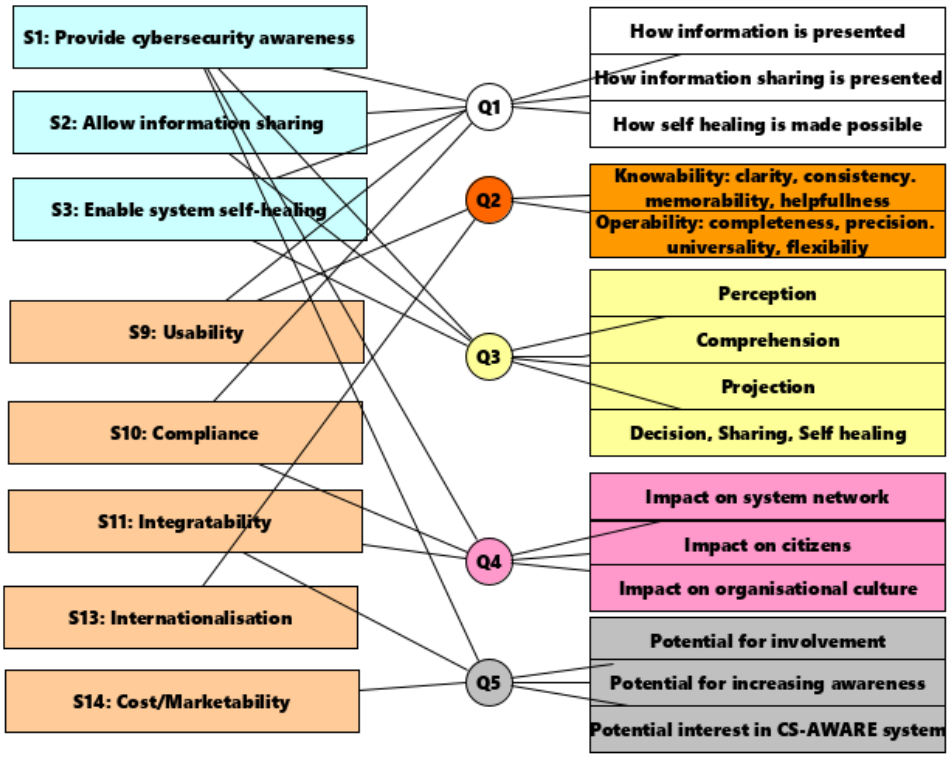


Figure 2: System requirements (column on the left), questionnaire numbers, and KPIs (column on the right); Q1: System effectiveness; Q2: Usability; Q3: User awareness; Q4: Organisational impact; Q5: Business

have set a baseline score per KPI at 60% appreciation, as this reflects the accepted level of satisfaction. Annex 2 provides all questionnaires. Figure 2 provides an overview of the relationship between system requirements, questionnaires, and KPI's. In this section, we briefly explain how the requirements for each level are addressed.

***Level 1, Technical**:* Evaluation at this level addresses (1) technical validation and (2) the extent to which users assess the technology of CS-AWARE as sound (i.e.: it works as it is supposed to).
(1) The technical validation focuses on testing the functional requirements for each individual CS-AWARE component (system and dependency analysis - GraphingWiki, data collection and storage, data pre-processing, data analysis and pattern recognition, multi-language support, visualization, information sharing and self-healing), as well CS-AWARE integration testing. The component testing is based on the functional requirements defined for each component in Table 5, and in more detail in D2.2. Some CS-AWARE system requirements (in particular the functional requirements, except S9-S11) are also evaluated based on technical criteria, that do not involve users. The functional requirements mentioned are evaluated by requiring passing the functional tests defined and reported in CS-AWARE deliverable D4.3.
(2) User technical evaluation is done through *Questionnaire 1*. This questionnaire addresses the main requirements S1, S2, S3, S9, and S10. The KPI's that we link to these requirements for technical evaluation are the following:

- For S1, providing cybersecurity awareness, we inquire the extent to which the users (including new users, who have seen the system, as a demo) appreciate how information is presented in the various overviews and detailed views that the CS-AWARE system provides.
- For S2, provide information sharing, we evaluate the extent to which users appreciate the information and the technical implementation.
- For S3, system self-healing, we ask about user appreciation of the relevant information and the technical implementation of self-healing.
- For S9, usability, we have a general question about user appreciation. More detailed information is investigated in Q2, but only for users who have actually worked with the system.
- Finally, for S10, compliance, we ask the extent to which the user thinks the system complies with internal regulations.

The 11 questions for level 1are formulated in a general manner, e.g. *I am happy with how information is presented in dartboard overview* (addresses S1), or *I do not like how information is presented in system overview* (addresses S2), and *I think the CS-AWARE technology complies with our regulations, policies and procedures* (addresses S10). The general formulation of the questions allows 'new users', i.e. users who have not worked with the technology, or only incidentally, and users from outside the LPAs, to fill in the questionnaire, for example after a demonstration session at a conference.

***Level 2, Usability**:* Usability addresses requirement S9, that we will split up into more detailed requirements, and S13, the use of language features.
We adopted the taxonomy of usability provided by Alonso-Rios et al. (2009). For our specific purposes,

we formulated usability requirements for *Knowability* (user can understand, learn and remember how to use the system) and *Operability* (the system provides the user with the necessary functionalities). Both are split up into more detailed KPIs for usability. For Knowability:

- U1 (*clarity*) addresses the extent to which the elements on the screen (options, colours, used classifications) are clearly represented, easy to find, and have a clear function.
- U2 (*consistency*) addresses the extent to which the elements are consistently used throughout the application, and have a consistent structure and function.
- U3 (*memorability*) addresses the extent to which the elements are remembered, including their structure and function.
- U4 (*helpfulness*) is about the extent to which documentation and online supportive elements provide sufficient support.

For Operability, we have the following KPIs:

- U5 (*completeness*): addresses the degree to which the user thinks the information provided by the CS-AWARE system is complete, reliable, and helps in deciding what to do.
- U6 (*precision*): the extent to which the user thinks the information is what is needed to understand and resolve a specific threat
- U7 (*universality*): the extent to which the translations that are provided by the system are clear. This also addresses the requirement S13 (internationalisation).
- U8 (*flexibility*): the extent to which the system provides various types of users (system administrators, service provides, managers) the options they need for their role.

The 18 questions of Questionnaire 2 result in scores for each of these requirements for user evaluation of usability.

Level 3, Awareness: The specific requirements for awareness that we test are based on the phases described in Hibshi et al. (2016)[x]: *Perception* (user perceives a threat), *Comprehension* (User understands the threat, its characteristics and information provided by the system), *Projection* (user foresees consequences of actions), and *Decision* (User makes a decision). To those, we add awareness of *sharing* and *self-healing*. The requirements for awareness link to the general requirements for cybersecurity awareness (S1), sharing (S2), and self-healing (S3). The requirements, combined with the phases, lead to a set of KPIs for awareness. For each of these, we indicated in what way the CS-Aware dashboard affords it:

- A1 (*perception*): which involves the affordances of the opening screen of the interface: (a) detection of a threat; (b) estimation of its possible impact, on a dimension very harmful – innocent; and (c) table of main features of the threat: date, type, part of the system network involved. Users will now be immediately aware of a threat and its main characteristics. To what extent will users process and understand this information?
- A2 (*comprehension*) of the cybersecurity threat, a process for which the console provides additional information about the threat, from reliable sources. Please note, that the main goal of our approach is cybersecurity awareness, which goes beyond the resolution of a threat, therefore consultation of this information is a crucial part of this. awareness requirement. The comprehension process (understanding the threat before action is undertaken) is supported at the console by the threat information screen. Providing necessary and useful information is a main research question for evaluation.
- A3 (*projection*), understanding future events or consequences, including potential, foreseeable attacks, or failures that result from poor security is a crucial aspect of cybersecurity awareness, which goes beyond detection and repair of threats. This awareness process is facilitated by the system and dependency graph, which is an interactive visualisation of the system network, showing information about the network components when the user clicks on them, and also displaying the components that are linked, or functionally interdependent (i.e. being part of the same pilot scenario, such as salary administration). We ask users about their use of this visualisation, and the extent to which they are aware of the dangers of the threat to nodes in the system network.
- A4 (*decision*) involves the possibility of indicating resolution of the threat (resolution itself happens outside of CS-AWARE), applying self-healing (if this is available for this threat, and the user has the rights, it can be automatically applied), and also deciding to share the information about the threat with the relevant expert authorities or communities. There is some difference between the complexity of decision-making in Rome, where several people may be involved in threat resolution, and in Larissa, where one or two experts are sufficient. For this part of evaluation, we ask the users about their communications with others, and understanding (sometimes: being told) if a threat is resolved.
- A5 (*sharing*): here we ask users about the extent to which they understand the need to share information and about the local policies for sharing. We expect no high scores here until management policies have been considered.
- A6 (*self-healing*): for awareness, it does not only matter if the user can apply self-healing, more important is it when this happens with full understanding of the consequences.

Because awareness is a rich concept, Questionnaire 3 involves 36 questions. This questionnaire will be applied in cycles 2 and 3.
For cycle 1, the phases that we described will also be leading in determining the baseline level of awareness, to be able to see if awareness levels increase in cycle 2 and 3.
This *baseline level* for awareness will be constructed by interpreting the outcomes from user story workshops in month 14. For example, for *perception*, we infer from the user stories that the users only perceive the threat when an employee comes to them with an issue (not immediate), they have to research what caused the issue (no immediate diagnosis) and what kind of threat is causing the issue. This will give no points for immediate awareness of a threat. For comprehension, some baseline understanding may be assumed, once the threat has been diagnosed, but not all details will be known. For projection, the measures will be adequate, but probably too rigorous, even at the level of not allowing certain applications to be used. For decision, some interactions and communication has been implemented, but not always, and not in every case. Similar conjectures apply to sharing and self-healing. We will look into this more precisely when we set the baseline for the two municipalities.

Level 4, Organisation: To what extent is the system used in the organisation, are the managers aware of this use, and what is its impact?
The requirements for organisational awareness link to the general requirement of the CS-AWARE system: to provide cybersecurity awareness (S1), as well as to the system requirements set for compliance (S10) and Integratability (S11). We distinguish awareness of impact on the municipality network, on services for citizens, and on organisational culture (Anttila, 2006)[xi]:

- O1 (*Impact on the system)*: one aspect of organisational awareness is related to increased security of municipality data, which is their greatest asset. Also, we ask about compliance with local regulations, and about improved security and resilience in general. These KPIs are related to S10 and S11.
- O2 (*Impact on citizens)*: this concerns the extent to which more effective delivery of services is made possible by using CS-AWARE. It concerns S1 and S11.
- O3 (*Impact on culture*): we inquire about the extent to which senior management is more aware of cybersecurity, and the impact on general organisational culture and security awareness. This involves requirements S1, S10 and S11.

We will administer Questionnaire 4 to the managers within the organisation. These will be managers of the system department, but also managers from the departments involved in the pilot scenarios that have been selected for CS-AWARE.

Level 5, Marketing/Business: Here, we address the interested users and potential early adopters. These are either participants to workshops organised by Cesviter and OTS, or new users that we send the questionnaire. We are interested in decision makers, mayors and managers of departments. From these people, we want to know the degree to which the CS-AWARE system addresses their cybersecurity needs and organisational structure. Three main business requirements are addressed:

- B1 (*Potential for organisational Involvement*): We ask potential users if there is sufficient awareness and ownership of cybersecurity. It is important if someone in the top policy group (that is, the city council) is specifically charged with overseeing cybersecurity), and with operational oversight (Someone with authority on the administrative staff should be charged with managing cybersecurity activities). Links to system requirements S11 and S14.
- B2 (*Potential for increasing awareness*): to what extent is cybersecurity an important (or crucial) and recurrent issue for this potential user? Links to system requirement S1.
- B3 (*Potential interest in system*): we ask (after a demo) about the extent to which there is sufficient administrative and policy interest in a system that provides awareness, self-healing and sharing of information. These business KPIs link to the system requirements of Marketability (S14) and Integratability (S11) and to the general awareness requirement (S1).

IV.2 Other Instruments and Procedures per Level

In the previous section we addressed how requirements and subsequent KPI's are addressed in the questionnaires for each level. In this section we will describe other evaluation instruments and the qualitative ways adopted in some of the levels, as a result of our hybrid evaluation approach. Furthermore, an overview of the procedures for all evaluation instruments per level is provided.

For *Level 1* (Technical evaluation) we developed functional tests for each component of the CS-AWARE system, setting criteria that were all or none (passed/not passed) with respect to the technical requirements S1-S8 and S12 (Appendix 3). When a component did not pass a test, the component was examined and adapted until it passed the test. Tests and results are explained in D4.3.
Procedure Questionnaire 1: this questionnaire was sent to the two LPAs twice, once at the end of cycle 2, once at the end of cycle 3. The request was to hand it out to the members of the deployment team, and to any other user in the municipality who would be interested in CS-AWARE.

	Cycle 1	Cycle 2	Cycle 3: Final Test
Technical Level	**Formal tests (D4.3)**	**Questionnaire 1**	**Questionnaire 1**
Usability level	**Usability Test**	**Usability Test, Questionnaire 2**	**Usability Test, Questionnaire 2**
User Awareness Level	**Baseline setting Qualitative analysis**	**Questionnaire 3**	**Questionnaire 3**
Organisation Level	=	**Questionnaire 4**	**Questionnaire 4**
Business Level	=	**Questionnaire 5**	**Questionnaire 5**

Table 6: Overview of hybrid evaluation methods during deployment, underlined methods are qualitative

Level 2, Usability: The purpose of usability testing is to get hands-on information on how foreseen users handle the CS-AWARE interface. We were interested in collecting their feedback, in order to improve various aspects of the interface. For each cycle a specific usability test was designed. These usability tests were run in a test environment which presented a number of cybersecurity incidents that were typical instances of the use cases (section II, workshop 2). The users are asked to express aloud their thoughts, feelings and opinions on any aspect of the system or prototype[xii]. *Procedure usability testing:* Users work on the usability test individually. After a brief introduction, users perform the tasks without much explanation. They are asked to think aloud and comment on what they are thinking and doing. The verbalisations of the users and their corresponding usage of the interface were all recorded and transcribed for analysis. User consent was obtained beforehand.
Procedure Questionnaire 2: this questionnaire was administered immediately after the usability test, in cycle 2 and 3. Participants were requested to fill in the questionnaire keeping in mind their recent experience with usability testing.

Level 3, Awareness, qualitative analysis: early development of awareness, i.e., during cycle 1, was interpreted by comparing the base-level (set by interpreting the story workshop) with the deployment scenarios that were constructed by the deployment teams of Larissa and Rome (discussed in section III, constructed in month 26 of the project).
Procedure Questionnaire 3: this questionnaire was administered to participants immediately after the usability test in cycle 2 and 3.

Level 4, Organisational: No other qualitative instruments were used for level 4.
Procedure questionnaire 4: this questionnaire was sent to the managers of the two LPAs. This was done twice; at the end of cycle 2 and at the end of cycle 3.

Level 5: Business: Our partner Cesviter organised two workshops, with participants from several municipalities in the regions of Cagliari and Puglia. The workshops are described as a dissemination activity in WP6. Further workshops were planned for cycle 2 and 3. For evaluation purposes, we prepared questionnaire 5 to be sent to participants of these workshops as a follow-up activity. Our partner OTS planned to involve their network of Greek municipalities after February 2020. This is a campaign to all 353 municipalities in Greece, consisting of an initial informative mail containing information about the project (and Questionnaire 5), to all the majors and vice majors responsible for IT & e-governance. The next step is to involve OTS-support teams to start calling in the appropriate offices, making sure that the material sent reached the right people. Additionally, a request for an appointment for a face-to-face presentation will be made. In the next phase and once appointments will be scheduled, OTS sales team will be involved visiting the municipalities across Greece.
Procedure questionnaire 5: this questionnaire is sent to participants of the workshops and of individual meetings as a follow-up activity.
Unfortunately, the business level suffered from the difficulties imposed by changed priorities of potential clients as a result of the corona-crisis, that happened during cycles 2 and 3. This means that the Italian workshop participants were less inclined to respond to our requests to answer the questionnaires. Also, the OTS campaign started three months later than originally planned.

Differences between cycles: We should note that the three cycles of deployment involved different deployment activities, and, consequently, they were not identical in terms of evaluation. For cycle 1, we performed technical and qualitative usability evaluation, and compared user awareness between what was revealed in the user stories (month 14, baseline) and the deployment scenario (month 26). This will be reported in the outcomes section for cycle 1 below.
For cycles 2 and 3, all levels were formally tested by administering questionnaires, and qualitatively by organising two more round of usability testing.
Table 6 provides a summary of the methods and instruments used for evaluation in this project.

V. OUTCOMES FOR CYCLE 1

The first deployment cycle took place between months 26 and 28 of the project (Oct-Dec 2019). The goal of this cycle was to implement and test the system in the LPA-contexts, to plan feedback procedures with the users, and to collect initial measures for evaluation.

Because during cycle1, implementation of the system in the local network was still ongoing, we focused on the first three evaluation levels: technical, usability, and awareness. Awareness was approached through the interpretation of the user stories (baseline setting) and the interpretation of thinking revealed in the deployment scenarios (qualitative analysis of awareness). Table 7 shows the activities undertaken during deployment cycle 1. We first discuss our deployment activities, and then report the outcomes for evaluation.

V.0 Deployment activity in cycle 1

Deployment Preparation phase: Deployment at the CS-AWARE project level started in May 2019. We formed a Pilot Working Group, which consisted of the leaders of WP2, 4, and 5. This PWG had weekly meetings. From December 2019 onwards we were joined by people from Cesviter. A template for a deployment scenario was proposed and discussed. We further specified the evaluation plan that was proposed to the commission after the first review. Within the context of WP4, we worked on interface usability testing by engaging a number of colleagues from the CS-AWARE project in a number of small tests. A brief report can be found in D4.3.
During the project General Assembly in Fowey (at the beginning of September 2019), a workshop was organised to discuss expectations from the CS-AWARE partners about deployment, use, evaluation, and impact. We discussed the feedback we would be collecting from the users. Furthermore, we addressed issues at the organisational level, both from the user point of view (how would experiences with the CS-AWARE solution be disseminated inside the municipality?) and from the project point of view (can the system generate information that can be exploited by the organisation? How can this be disseminated and exploited?). We also discussed the possibility of the changing role of the system administrator, when
awareness is increased (e.g. more or different responsibilities), and, accordingly, more accurate and effective services that can be provided. All these points of discussion fed into finalising the deployment scenario template, and also in our understanding of the current level of awareness of the future users of the CS-AWARE system in Rome and Larissa who were present at this meeting and participated at the workshop.

Deployment scenarios have been constructed in October 2019, in two workshops together *with the users from both municipalities.* The report is in Annex 1 to this Deliverable.

Deployment activities with users: Two deployment scenarios were established, which provided insight on the objectives of the users at the two LPA pilot sites. Annex 1 presents the full outcomes, our summary is displayed in tables 2 and 3. After the deployment scenario was developed, we had three meetings with the deployment teams in Larissa and Rome. We further discussed the deployment scenario, and many details of the CS-AWARE interface. In addition, there were the discussions about technical implementation, taking place on a weekly basis, and through frequent email exchanges. We organised formal usability tests for evaluation in December in which three users (all members of the deployment teams) participated at each location.
It should be stressed that all documents and comments were discussed with Larissa and Rome extensively and are published with their full consent.

CYCLE 1	M26	M27	M28
Deployment			
Technical	**System Development**		
Users	**Deployment team and scenario (L)**	**Deployment team and scenario (R)**	**Deployment team meetings**
Evaluation			
Level 1: Technical			**System Test**
Level 2: Usability			**Testing (L-R)**
Level 3: Awareness		**Qualitative baseline setting / analysis of awareness (L-R)**	
Level 5: Business	**Cagliari workshop**		

Table 7: Activities for deployment and evaluation during cycle 1 (October-December 2019)

Deployment activities, technical: Technical developments during this period was reported in D4.3: CS-AWARE solution validation and testing report.

V.1 Evaluation Level 1: Technical validation in cycle 1

The technical validation focuses on testing the functional requirements for each individual CS-AWARE component (system and dependency analysis - GraphingWiki, data collection and storage, data pre-processing, data analysis and pattern recognition, multi-language support, visualization, information sharing and self-healing), as well based on the functional requirements defined for each component in Table 4 of CS-AWARE deliverable D2.2. The integration testing is based on the eight technical CS-AWARE system requirements as defined in Table 5. It is the responsibility of the technical partner developing/integrating the components to provide testing environment capable of testing the functional requirements

Figure 3: Simplified interface for threat detection. Old version on the left, revised version on the right

and conducting the relevant tests. Annex 1 of deliverable 4.3 contains functional testing reports for each CS-AWARE component and the integrated CS-AWARE system. The reports define the testing environment/procedure and report on the success of the individual tests. It should be noted that all tests are expected to report a positive result, since in case of failure the component is further developed until a positive result is achieved. Furthermore, the test dates reported represent the date of a successful test run, but is should be noted that functional tests based on the presented test environments/procedures are conducted on a continuous basis if components are developed further, to ensure new changes do not break existing functionality.
The test results show that all basic technical functional requirements defined by CS-AWARE have been fulfilled. Work continues on refining the technical basis based on input derived from CS-AWARE piloting.

V.2 Evaluation Level 2: Usability outcomes for cycle 1

Feedback on usability was a very productive line of investigation and resulted in many important improvements of the CS-AWARE system.
Developments on the CS-AWARE console (i.e. the interface for users to the CS-AWARE system) were (up until cycle 3) subject to an ongoing sequence of revision and feedback, based on comments, from users, but also from the technical developers and other project team members. The results that we report here as outcomes of cycle 1, therefore pertain to a version of the interface that
no longer exists, although the structure of the CS-AWARE interface is still the same.
Table 8 shows the main features of the usability tests in both Rome and Larissa. As can be seen, there were 3 participants in the usability testing at each LPA. In Larissa, participants were tested individually, in Rome, the participants were tested with the active presence of their colleagues and a translator from CS-AWARE. We have explained in the evaluation section (IV) that we approached threat detection, comprehension and resolution as a decision-making process, comprising 4 phases. A schematic overview

Summary	**Usability study, Rome & Larissa, cycle 1**
use cases	**1) A vulnerability use case, the network has been infected. Tasks: locate the source, find the infected parts, and undertake appropriate action, including deciding on self-healing. 2) A general security warning, such as a DDoS attack. 3) A malicious IP address has been detected. 4) An attack against LPA data. For example: unusual behaviour in security log, network traffic, database, at different moments in time.**
Components	**All components of the system are involved**
Participants (Rome)	**2 system administrators (SUET), 1 manager (Head of the Data Center of Roma Capitale)**
Participants (Larissa)	**3 system administrators, specialised in applications or in database services**
method	**Cognitive Walkthrough**
outcomes	**Table in Annex 3.3 listing improvements; impact of expertise (role) and organisational complexity**

Table 8: Main features of usability tests in cycle 1

of the phases and the appropriate screens of the console can be seen in figure 4. Here, we will discuss the main changes made based on user feedback, for each of the four phases of the decision-making process.

Phase 1, perception: This refers to the opening screen showing a dartboard shape with threats, in different colours. These colours were simplified for the sake of clarity, (Figure 3). The opening screen also displays a list of threats, with some crucial characteristics. This was much appreciated by the users, and some of the details were improved, for clarity of description and easier identification of threats in the list.

Phase 2, Comprehension: The threat information window (see figure 3) is revealed when the user clicks on a threat in the opening screen. Graphics were improved, and a comprehensive list of technical details of the threat for system administrators was added.

Phase 3, Projection: In the network visualisation. (see figure 3), more details of the system nodes were provided at the request of the users. We added the possibility of threat filtering (by group, person, location) and the possibility of focusing on a single flow

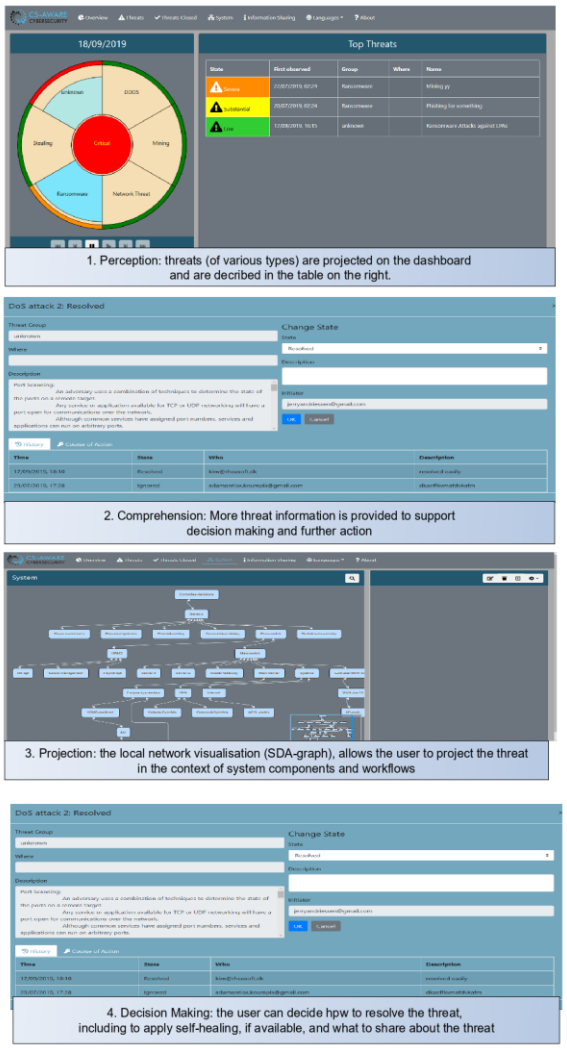


Figure 4: Process steps(phases)in cybersecurity decision-making linked to features of the interface.

of information (e.g. for Finance). Zooming was improved, search functions were added, and the shapes of the nodes in the visualisation more consistency matched their function in the network.

Phase 4, Decision-making: the most important addition was a ticketing system for assigning roles, which was useful for a large municipality. This was a consequence of the differences in user roles and expertise of participants that we observed in the usability tests.
Moreover, the interface for information sharing was enhanced allowing more detailed sharing options.

Differences between municipalities: The thinking aloud approach revealed important differences in the ways that the individual participants handled cybersecurity issues. This confirmed what we already understood from the workshops, and allowed us to make these differences in approach explicit, more particularly with respect to the phase of decision-making.

In Larissa, we worked with experienced system administrators, who had sufficient autonomy to decide if a threat was resolved or not. During the full process, they were already used to consulting their colleagues and manager, depending on the kind of threat. We noted that system administrators differ in their expertise: specialisation in database administration, or in handling issues with applications. This difference in expertise was reflected in decision-making, where the appropriate expert was always consulted before a decision.

In Rome, the same difference appeared between expertise in database management, and in other services, but there was an additional layer of decision-making. A new threat was always handed over to the expert for the affected part of the system network. The first pass of handling an issue therefore involved a manager, deciding what nodes in the system network were in danger, and
then handing over the issue to the dedicated expert. The manager was very knowledgeable, did not resolve the issues, but was responsible for deciding if a threat was resolved or not.

V.3.LEVEL 3: AWARENESS

Awareness baseline setting: For a baseline understanding how the two municipal organisations handle and think about cybersecurity, we organised a story-telling workshop[xiii] . We think stories are a good way to capture personal views and experiences. However, in our workshop, we added the collaborative dimension.
In our view, a collaborative story (a story written together) can be taken as joint understanding and a coherent integration of personal views and experiences. The stories therefore can be interpreted as providing a picture of how the organisation (and its employees) look at cybersecurity, in other words, on their initial level of awareness. The stories and their interpretations are reported in detail in D2.2.

Here, we first summarise the main characteristics of the two LPA contexts, as they were revealed by the stories. In **Larissa**, there is one system administration department for the municipality (including smaller municipalities around Larissa), where all cybersecurity issues are handled. System administrators in Larissa have different expertise, are aware of each other's competences, and regularly share information. Managers from other departments seem or need to be less involved. In **Rome**, there are many departments responsible for many services, making for a complex structure. Departments handling data, or particular applications, or citizen internet services, do not necessarily share information on cybersecurity. It is unclear how and when communication between the various departments takes place. For

awareness, this means it is highly *distributed*, and often unshared.
Developing knowledge on cybersecurity is a local affair, and this is fine in Larissa, but less so in Rome. Application of safety rules is monitored in Larissa by the system administration, but it is less clear if other department managers share this attitude. In Rome, if and how safety regulations are monitored, is highly dependent on the department managers. The role of system administrators is to help users, sometimes even help them avoiding regulations.

From this brief analysis we learnt that increasing awareness of cybersecurity issues and their resolution within system administration involves: a) having access to the knowledge of threats (from various sources); b) having an interest not only in threat resolution, but also on the implications for the system network of the organisation; c) having an interest in safety of the department and the organisation; and d) having an interest in communicating and sharing experiences with other municipalities, CERTs and NIS authorities. This is not the same as saying that system administrators should be aware of all of these things, it means that these are the crucial aspects of cybersecurity awareness in any organisation.
The *rules and norms* that the IT-Department maintains underlying the regulations for monitoring safety behaviour, and providing services to the users, are different in the two municipalities. In Larissa, we see that the system administration expects citizens (internal users) to stick to the rules and regulations. Conversely, the citizens expect an immaculate and timely resolution of their cyber-issues. This means there is always a tension: trust may grow but also decrease. In Rome, there seems to be a general norm that the system as a whole is too complex, and we should accept imperfection to the extent that not all policy regulations are created in the interest of maximal cybersecurity. It looks like the various IT-departments are still able to provide quick services to users with problems. These services are provided on an individual basis, for a user who detects an issue, there is no follow-up, and there may be no spread of information to other users in case of a security issue. For awareness this means, that there is a lack of awareness of threats, especially in terms of understanding the threat and its impact on the system network. In Rome, this lack of awareness is related to the distributed knowledge of the complex system network: there is no single individual who completely oversees the whole network.
Concerning *sharing of information within the organisation*, in case of Rome, there is the issue of information sharing and knowledge management: in Rome, knowledge is highly distributed and less shared, system users seem to have no concern for each other's cybersecurity issues, and the culture favours general regulations but individual solutions (not shared). For awareness, this creates a problem: it does not evolve nor spread. Also, in Larissa, while employees within the system department are aware of the cyberincidents that have occurred, this knowledge remains within the department, and seems of less concern for others, although the other employees that we met seemed to have some interest.
Although the situation in Rome reflects a more complex organisation for handling cybersecurity, underlining the importance for managing and sharing information, the main characteristics of awareness for Rome and Larissa are (perhaps surprisingly) similar.

Awareness requirements, baseline level: For interpreting the baseline level of awareness, which is the state of cybersecurity awareness at the beginning of the project, we combine the steps of cybersecurity decision-making with inferences from the user stories. In our discussion, we focus on system administrators (level 3 of evaluation), but also on relevant aspects of the organisation (level 4).

A1, Perception (The 'A' in 'A1' signifies 'awareness'): Threat perception at baseline is not immediate, as it usually depends on the system administrator (the *user*) being informed of an incident (by a service user, or a system alert). Then, the user has to identify the threat, discover its possible impact, such as inferring when the issue probably appeared, in what part of the municipality network, or with which employees. Our impression is that the focus in both municipalities is on threat identification, and less on the other processes (comprehension, projection, decision-making). Base-level perception is not immediate, identification can be time-consuming, focuses on individual service users, and is local (part of the system), not (whole) system oriented.

A2, Comprehension: For further understanding of a threat, if it is not already well known, a system administrator has to access the same resources as for identification. Searching for and reading relevant information may take time, depending on the experience of the user, and often does not happen at all. The user may have to research log-files of the system, and maybe of particular services, to figure out what exactly is the problem. System administrators may do this, if they are responsible for the network or a part of it. If their role is to manage the database of a particular service, they may not be interested in such log files. In that case, communication with another specialist is needed, which requires information about the threat to be collected and understood.
For awareness this means that for a system administrator resolving a threat, full comprehension is not always necessary, and sharing information is only needed sometimes. Therefore, comprehension is limited, at the individual and at the organisational levels.

A3, Projection: If the system administrator wants to understand the risks imposed by a threat to a system node, as well as to the other nodes, and which services will be in danger, a thorough understanding of the network as well as of the process in which the nodes are involved is a requirement. As awareness often is distributed, these aspects of resolution are a problem for users, especially in larger organisations. Users may be very knowledgeable for their own service, but may not always oversee the possible implications for the rest of the network.

For awareness, this means that projection is local, and users may lack the knowledge to research their network. In Larissa, communicating with colleagues solved the problem. In Rome, such communication may be more complicated, and require the intervention of management.

A4: Decision: The final phase is called decision-making, and this involves several different actions. In the baseline situation, it usually means *resolving* the threat, for example by applying a patch or update, or blocking a user or part of the network. In a complex organisation, it may mean referring resolution to the expert responsible for that part or service in the network. In all of these situations, experts, colleagues, including managers, may be informed or consulted. This requires sufficient comprehension. Also, the user may need to ascertain that a threat is actually resolved. Most awareness in the base-level situation depends on communication between different stakeholders. In other words, awareness requires the collaborative attitude of equality of contributions, desire for sharing, and careful consideration of contributions by all[xiv].

Qualitative analysis of Awareness: The deployment scenarios were obtained at the beginning of the third year of the project, hence one year after the story workshop. Compared to the user stories collected in year 2, we can see (see appendix 1) already increased user awareness of cybersecurity, reflected as (1) *Perception*: expectations for timely and accurate detection of cybersecurity threats; (2) *Comprehension*: awareness of the importance of good quality information and good quality reporting; (3) *Projection*: interest in their own system network, its assets and vulnerabilities, and the additional possibility of reflection on past problems and solutions; (4) *Decision*: the possibility of mitigation reports and user responsibility in making decisions; (5) Interest in impact on the organisation and communication with internal and external service users.

Nevertheless, we also see aspects where awareness gains are quite possible: (1) The focus is on detection rather than on comprehension; (2) Especially in Larissa, trust in the CS-AWARE system still needs to evolve; (3) Users did not discuss sharing and self-healing; (4) Efficiency rather than awareness is the main focus.

Also, we note the following differences between Rome and Larissa, some of which may be related to differences in the complexity of their organisations: (1) Department-oriented objectives (L) versus organisation oriented objectives (R); (2) Orientation towards greater efficiency and trust (L) versus collaboration between managers and between the different departments (R); (3) Service users are interested (L) versus not interested (R) in security; (4) System administrators stress the importance of being alerted (R) versus being involved in active learning and discussion (L), and, finally, (5) higher involvement and expectations by either managers (R) or by system administrators (L). One comment about (2) orientation is, that this orientation signifies aspects that are not yet realised in the department or organisation.

Awareness development, cycle 1: The two tables (8a and 8b) summarise the differences noted between the awareness of cybersecurity at the beginning of the project (month 14) and one year later (month 26), for system administrators in Rome and Larissa. It should be noted that when system administrators or managers discuss some issue in a deployment team, which may signify awareness, this does not imply that this awareness is always applied when resolving a threat. For example, the most efficient use of the CS-AWARE system is possible by simply confirming current practice: nothing changes, all remains the same. In that case, although detection of a threat is improved with CS- AWARE (automatic notification and diagnosis of threats), the user can follow the advice from the system until the threat is resolved, with only minimal reflection (greater perception awareness, but not much greater comprehension).

In terms of increased awareness of cybersecurity threats, the notion of (better) threat comprehension is therefore very important. In both LPAs, initially (in month 14) comprehension was handled as efficiently as possible. In month 26 users in Rome indicate they see the potential for (increased) efficiency. Greater efficiency in Rome is an organisational desire, which involves cybersecurity awareness, but is also a more general issue. The managers in Rome were clearly interested in reflection on how the organisation handles cybersecurity. In Larissa the users want to learn and discuss, but only if they have developed sufficient trust in the system.

Concerning projection, although none of the users in Rome referred to the need for understanding the system network, all agreed that more effective communication was needed. In Larissa, the users expressed the desire to have overviews of what system components were threatened, and in collecting evidence for the weaker components in their network.

Rome (sysadmin)	Month 14	Month 26 (Cycle1)
Threat Perception	Not immediate, local, individual	Real time alerts, efficiency
Threat Comprehension	Distributed expertise, distributed rights & roles	Efficiency
System Projection	Service (not network) oriented, requires management	Service (not network) oriented
Decision-making	Distributed	Trouble ticket, more effective relations
Need for Sharing	Manager distributes tasks	Better communication with internal users
Self-healing	=	=
General	Individual help, no follow-up	Reflection

Table8a: Awareness baseline, and at the start of deployment (Rome, system administrators)

Larissa (sysadmin)	Month 14	Month 26 (Cycle 1)
Threat Perception	Not immediate, local, time consuming	Immediate, no false alarms
Threat Comprehension	Only if needed, collaboration	Up to date, if trusted then collaboration
System Projection	Collaborative	Proactive, Reports, Collaborative
Decision making	Collaborative	Informed, self-paced
Need for Sharing	Internal	Internal
Self-healing	=	=
General	Trust between sysadmin and user is an issue	Learning

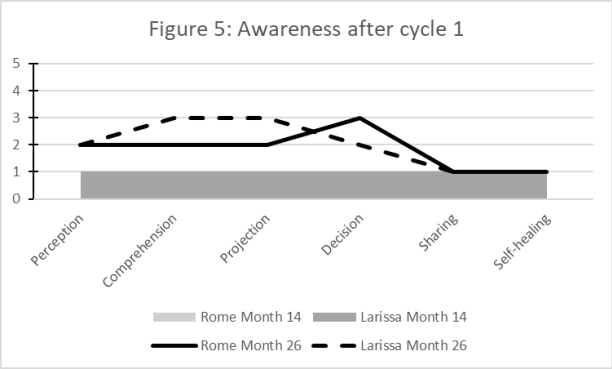
Table8b: Awareness baseline, and at the start of deployment (Larissa, system administrators).

As for decision-making, the users in Larissa indicated they expected this to be better informed with the CS-AWARE system, which would also support internal collaboration. In Rome, high expectations were formulated for better communication with service providers, and more effective management of the various users involved in threat resolution processes.

The need for sharing only was mentioned as an LPA-internal possibility, not with external communities. Self-healing was not mentioned at all.

In our qualitative approach we interpreted the meaning of user stories and deployment scenarios for awareness. Clearly, we are interpreting and evaluating a developmental trajectory for awareness, not a fixed state. Our conclusion is that, for awareness, we could already see clear developments of awareness, on both sites, between month 14 and month 26. For Rome, this means increased awareness of the importance of CS-AWARE for threat perception and comprehension, and especially, for decision-making.

For Larissa, increased awareness is particularly about comprehension, but also about perception, projection and decision-making. Figure 5 displays the (arbitrary) scores that we have assigned to these qualifications, for Rome (straight line) and Larissa (dashed line). For the sake of quantification, explained earlier in the section on hybrid approach to evaluation, we have set the base level (Month 14) as one, on a five-point scale ranging from low awareness (1) to high awareness (5).



VI. CONCLUSIONS FOR CYCLE 1

Achievements in cycle 1: Table 9 summarises the outcomes for cycle 1. The CS-AWARE system was finalised and deployment in the LPA system networks started at the end of the cycle. The deployment teams were established, which meant we had installed a group practice for feedback and testing. The deployment scenarios made clear what our participants expected and the state they were in concerning cybersecurity awareness. Usability testing was a very productive source of information, not only for the system, but also for understanding its use by the participants. We discovered many differences in use, linked to the expertise and role of the participants in the organisation. We observed a considerable increase in awareness of cybersecurity during months 14 and 26, which could be linked to participation in the design of the CS-AWARE system, and the workshops (section II) in particular.

Lessons learnt from cycle 1: In this project, we are designing a bespoke solution for cybersecurity awareness for the two LPAs. The assets of this solution are clear: the system works in the local context, and for the local users. The limitations are clear as well: the approach takes time, and generalisation from two pilots seems difficult. Nevertheless, our qualitative approach to deployment has given us information on the development of awareness of cybersecurity at the two municipalities.

We have proposed, on the basis of the user stories, the following definition of cybersecurity awareness that we will further exploit during cycles 2 and 3:

Awareness of cybersecurity includes both knowledge and agency.

Knowledge pertains to:
(1) Cybersecurity threats
(2) The system networks
(3) The organisation and its users and
(4) The cybersecurity community.

Agency is about the ability and willingness to act:
(5) When a threat is imminent
(6) When there is no threat

We suppose that the categories (each to a certain degree) qualify user-awareness, its role during detection, but also when there is no threat to resolve. We do not test detailed knowledge of cybersecurity threats directly, but we will test, during action to resolve a threat (the usability test), and after that action (questionnaires 2 and 3) the extent to which the user specifies that a particular threat is understood, including the implications for the system network, and what to share with the cybersecurity community. In addition, we ask about (in questionnaire 4) the implications for the organisation: the system network, the cybersecurity culture, and the impact on citizens.

The CS-AWARE system provides *knowledge* about threats, about the system network, and the components and service processes implied in the threat. This may lead to increased awareness of cybersecurity in the organisation. CS-AWARE supports the user agency when resolving a threat, and this may facilitate awareness and reflection also when there is no immediate threat. The CS-AWARE system asks users about what they want to share about a threat with cybersecurity communities. This requires a dedicated policy for sharing at the municipality.
So, as a general expectation, all aspects of awareness may improve, but not automatically. In that respect, we note the important concept of sharing that will be the focus for the next phase.

Implications for next phase: During the next phase, we will focus on the implementation and testing of information sharing of threats. Sharing such information to CRTs and other stakeholders does not have a high priority for our LPAs at the moment. Furthermore, we feel safe now to build on our understanding of cybersecurity awareness at the pilot sites to undertake a quantitative evaluation in the form of questionnaires.

VII. CYCLE 2: DEPLOYMENT

The goals of cycle 2 were achieving a full implementation of the CS-AWARE system at the user sites, collecting further user feedback, getting results on the questionnaires for all levels, and to perform further usability tests. Table 10 shows the activities in cycle 2.

Technical deployment: Based on the already existing (and tested) functionalities, work continued to be able to test 3 main use cases in the deployments in Larissa and Rome. This consisted of work to instantiate the specific use cases, and refine the components involved in the cases along the way.

(a) Instantiation of Rome/Larissa *specific monitoring patterns*: The involved components in this use case are system and dependency analysis, data collection, data pre-processing, data analysis and visualization. The purpose of this use case was to instantiate the monitoring patterns that were derived during the third system and dependency analysis workshops in Larissa and Rome, to be able to test realistic

Deployment	Rome	Larissa
Technical achievements	Data Center requirements gathered; Enterprise docker set-up; IAM integration	Cloud requirements gathered; full deployment on AWS
Deployment Team	13 participants	6 participants
Objectives	Efficient detection, better collaboration	Effective detection
Foreseen Impact	Improved Communication with service providers	Learning, reputation
Desired	Alerting and reporting	Reporting
Evaluation	**Rome**	**Larissa**
1) Technical		
method	Functional component and integration test set	
outcomes	Full set of functional tests (component and integration) passed (Annex 1 to Annex 9 of CS-AWARE deliverable D4.3)	
2) Usability		
participants	3 (1 Data Center manager and 2 sysadmin)	3 (system admin)
method	Cognitive walkthrough	Cognitive walkthrough
outcomes	List of recommendations	List of recommendations
processes	Distributed	Collaborative
3) Awareness		
method	Story telling	Story telling
participants	15, 5 stories	13, 6 stories
awareness baseline	Local, individual	Local, individual
participants	13	6
method	Co-creation	Co-creation
more aware of	Threats, network, organisation	Threats, need for trust
Not more aware of	Sharing, self-healing	Sharing, self-healing

Table 9: Summary of outcomes of cycle 1 for the two LPAs

Timing	M29	M30	M31
Deployment			
Technical	**System Development**		
Users	**Project meeting**	**Review meeting**	**Deployment team meetings**
Evaluation			
Level 1: Technical			**Questionnaire 1**
Level 2: Usability			**Usability test Questionnaire 2**
Level 3: Awareness			**Questionnaire 3**
Level 4: Organisation			**Questionnaire 4**
Level 5: Business	**Puglia workshop**		

Table 10: Activities for deployment and evaluation during cycle 2 (January-March 2020)

threat events in the context of the "suspicious behaviour monitoring" use case. Following work on the individual components was required to provide this use case:
System and dependency analysis: It became clear that the asset and dependency graph format required, in addition to be able to model asset and dependency information, needed additions to be able to also model information relating to log files and monitoring patterns. The resulting protocol can be seen in Annex 3.1.

- Data collection and data pre-processing: The log files for Rome and Larissa have already been collected and pre-processed in earlier phases. Small refinements and bug fixes to individual log file parameter processing have been implemented on a per-case basis.
- Data analysis: The Rome/Larissa monitoring patterns, as reported in CS-AWARE deliverable D2.3, have been implemented by the data analysis component.

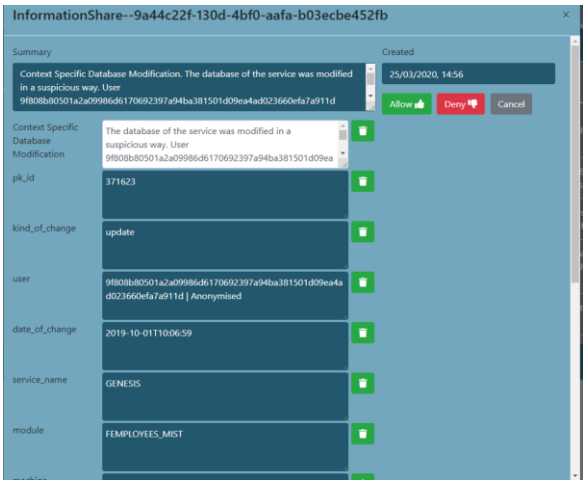


Figure 7: Information sharing details

- Data visualization: Minor refinements and bug fixes have been implemented in the process of visualizing the Rome/Larissa specific patterns

(b) *Information sharing*: Based on the information provided by the Rome/Larissa specific monitoring patterns, information sharing functionality has been refined in cycle 2. The involved components have been information sharing and visualization. The main effort has been to rework the communication protocol between the information sharing and visualization components, based on the experiences with the data from the real Rome/Larissa use case patterns. The resulting updated communication protocol between information sharing and self-healing is reported in Annex 3 (Section 3.2). Figures 6 and 7 show examples of the main screens for information sharing.

(c) *Social media monitoring* (general security warnings use case): A concrete implementation of the general security warnings use case was provided in cycle 2. It was decided to allow the user to monitor

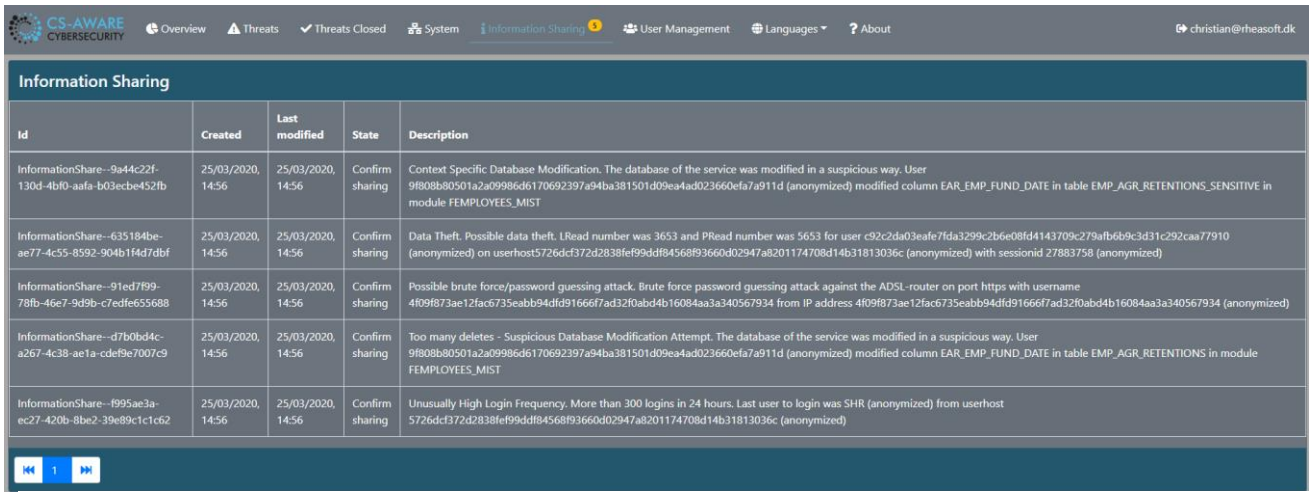


Figure 6: Information Sharing overview

social media messages for specific keywords on a per asset basis. If a keyword is detected, it is displayed as a "social media report" with default "low" threat level in the CS-AWARE overview, and visualized at the asset for which the keyword was specified. The involved components are system and dependency analysis, data pre-processing, data analysis and data visualization:

- System and dependency analysis: Keywords are modelled in the asset and dependency graph in the already foreseen "categories" parameter. No changes were required.
- Data pre-processing: A specific pre-processing task was added that allows data analysis to trigger a keyword-based search in a specified set collected information form information sources (like e.g. social media), based on the keywords defined in the asset and dependency graph.
- Data analysis: The ability to compile social media reports from messages that were retrieved from keyword-based search, in line with the previous protocol (where threat type is "report" and priority is "low" by default), was implemented. No changes to the exchange protocol were required.
- Data visualization: Representation of social media threats is in line with the visualization concept, no changes were required - except for minor additions to the code to allow social media messages to be displayed on a per-asset basis in the systems overview.

Deployment, users: During cycle 2, we collected comments from users on versions of the three additional functionalities mentioned in the previous section. It should be noted that cycle 2 was a turbulent period, because of the coronavirus, which halted communication with Larissa during the month of March 2020, which was the third month of cycle 2. The users of both municipalities often had to work from their homes, and monitoring the CS-AWARE system was therefore more difficult, especially for the users in Larissa. During the second month, the preparation and digestion of the second project review took some time, but this did not hamper deployment.

VIII. CYCLE 2: EVALUATION OUTCOMES

During cycle 2, we administered the 5 questionnaires in both municipalities.

Figure 8 shows the outcomes for each requirement, collapsed over all questionnaires. For example, the score for S2 (allow information sharing) is the average of scores for S2 from questionnaires Q1 (system) and Q3 (awareness). As can be seen, for all requirements, the KPIs are above threshold (.6), so the scores seem to confirm the positive outcome of the CS-AWARE system deployment. Below, we will look at the results for each questionnaire in more detail.

Impact Coronavirus: Questionnaire 5 was meant for potential new users at municipalities in Italy and Greece, and also for potentially interested visitors at two events that could potentially draw new interest. Due to the corona virus, these new users could not be reached, and the two foreseen events were cancelled. Therefore, questionnaire 5 was not administered during cycle 2. This implies that S14 (Marketability) remains untested in cycle 2.

VIII.1 EVALUATION LEVEL 1: TECHNICAL VALIDATION IN CYCLE 2

Questionnaire 1 was administered to users in Larissa and Rome. Table 11 shows the outcomes for each municipality, separated for type of user and the requirements that were tested.

The following observations seem relevant, where we take a difference in scores of more than 20% as noteworthy:

- The KPIs for all requirements overall are 60% or more, which is satisfactory. This means that, at this stage, the key performance indicators confirm that the requirements for S1 (provide cybersecurity awareness), S2 (allow information sharing), S3 (provide self-healing), S9 (usability), and S10 (compliance) are met.
- For Rome, there is a relatively high representation of management in the participants who filled out questionnaire 1. These managers generally give higher scores than their system administrators (on average 15% higher) on all requirements. At the same time, both managers and system administrators indicate they have not spent much time in working with the system yet.
- For Rome, the ratings by system administrators are, in spite of their lack of experience with the system, lower than the ratings by their managers, for all requirements (see table 11), with the exception of S1 (provide cybersecurity aware
- ness). Their comments on features of sharing information, self-healing, and usability in general will be collected in the usability test.
- For Larissa, we see highly experienced system administrators, who give very high scores to each of the requirements, with the exception of S2 (allow information sharing). We take the combination of high scores and high experience with CS-AWARE as a very encouraging result
- Requirement S2 receives low ratings (below 60%) from system administrators from both municipalities. We will explain this low score in the usability section.

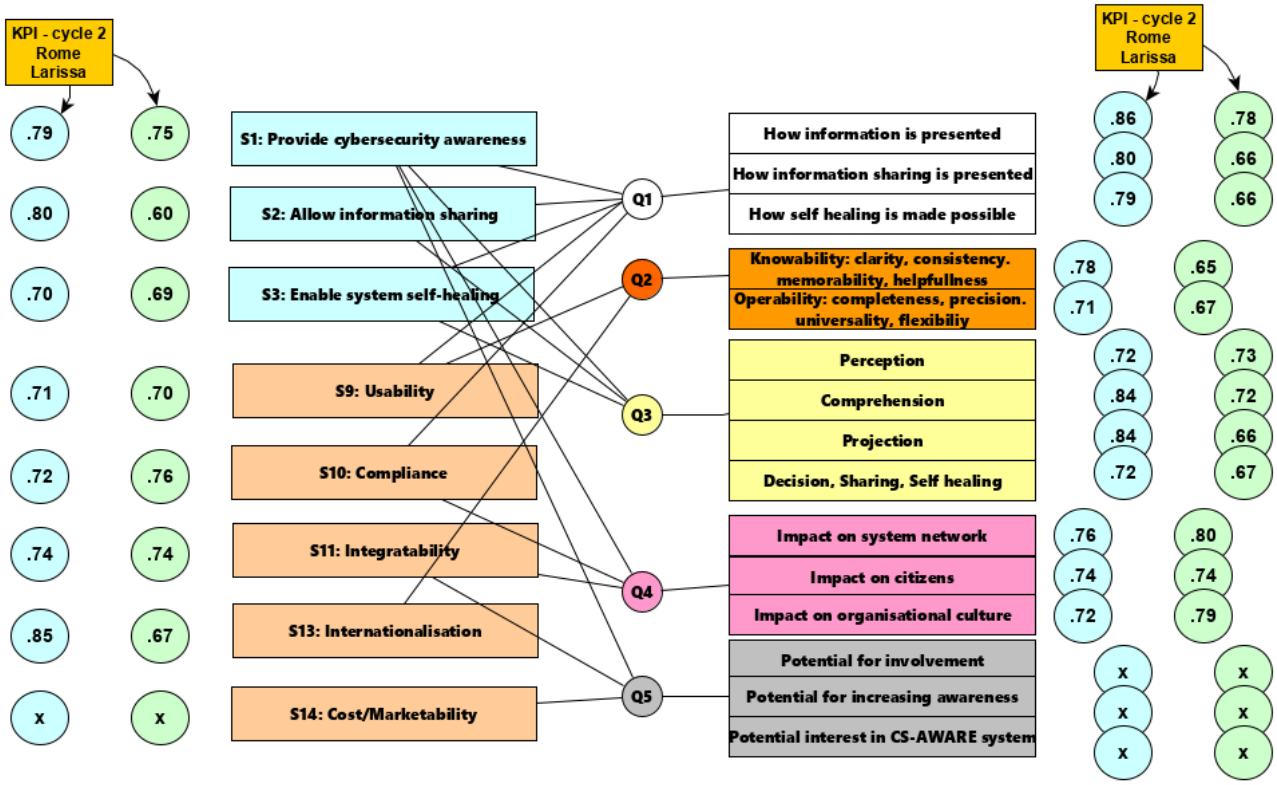


Figure 8: Cycle 2, requirements and KPIs

- Service users at both LPAs evaluate the system quite positively, but they do not have hands-on experience, these scores should be taken as the result of them participating in workshops and witnessing a demo. For them, CS-AWARE 'looks good'.

Q1-CYCLE 2	Sysadmin	Managers	Service users	Mean
Rome				
N	**4**	**10**	**6**	**20**
Exp	**2**	**1.5**	**1.8**	**1.7**
S1	**0.74**	**0.82**	**0.75**	**0.78**
S2	**0.57**	**0.72**	**0.61**	**0.66**
S3	**0.6**	**0.72**	**0.59**	**0.66**
S9	**0.6**	**0.78**	**0.73**	**0.73**
S10	**0.6**	**0.70**	**0.67**	**0,72**
Mean	**0.64**	**0.76**	**0.67**	**0.71**
Larissa				
N	**3**	**0**	**3**	**6**
Exp	**5**		**2**	**4.67**
S1	**0.82**		**0.9**	**0.86**
S2	**0.56**		**0.91**	**0.69**
S3	**0.8**		**0.87**	**0.84**
S9	**0.8**		**0.93**	**0.87**
S10	**0.8**		**0.87**	**0.83**
Mean	**0.74**		**0.9**	**0.81**

Table 11: Outcomes for Questionnaire 1 (cycle 2), for each requirement, in KPI (0-1). Exp means experience of the user (1-5).

The outcomes of Q1 are positive overall, although there are differences between managers (more positive) and system administrators in Rome, the system administrators in Larissa have more experience and are more positive, and it is worthwhile to inspect the process of sharing information (at the interface) more closely.

VIII.2A Evaluation Level 2: Usability, Qualitative

Impact Coronavirus: In Larissa, due to the situation of citizens/workers being locked at home, the users worked in their own time on testing the interface. They provided a set of written comments, that were further discussed with them in a deployment team session.

In Rome, we organised a number of usability tests, involving four different users. In the first three tests, we noted that it was impossible for one system administrator, for example responsible for maintain the SUET-database, to make decisions about how to resolve a threat, as that would require a manager with overview over the municipal network. We therefore organised an additional session, in which the manager and the system administrator together resolved a use case.

We analysed in more detail a single use case, with respect to the three new features (discussed above, under technical deployment in cycle 2): a specific monitoring pattern triggered a critical warning, and

Summary	Usability study, Rome & Larissa, CYCLE 2
use cases	A general security warning: a DDoS attack.
components	All components of the system are involved, we had specific interest for LPA specific monitoring patterns, details of information sharing, and keyword search.
Participants (Rome)	2 system administrators (SUET), 1 manager (Head of the Data Center of Roma Capitale), 1 system administrator
Participants (Larissa)	3 system administrators, specialised in applications or in database services
method	Cognitive Walkthrough
outcomes	Table in Annex 3.4 listing main improvements;

Table 12: Main features of usability tests in cycle 2

in addition to resolving the threat, users were requested to comment on the detailed pattern information, the options for sharing, and, as an additional feature, the output from key word search on social media that the system had performed on user provided keywords.
The fictional threat was 'injected' into the deployed CS-AWARE system. Participants in Rome were recorded when they were thinking out loud in their attempts to resolve the threat.

The main comments from the usability tests in Rome and the testing at home from Larissa, and the resolutions by the CS-AWARE team are in Annex 3.4.
In general, we noted that the sharing of information about the threat posed some problems for certain system administrators. Firstly, they were unaware of with whom the information would be shared. This was an unsettled issue indeed, and we suggested that the information that was marked as shared would end up in a database that could be accessed by people, internal and external, still to be designated, and with specific permission. Secondly, the information that was shared, was highly technical, to be understood by system administrators with the relevant knowledge. However, the decision to share or not was a management decision, even required formal policies for sharing. Neither municipality had such policies installed, although they were willing to undertake the effort. Thirdly, all users were bothered by the anonymity of the IP-Address, which, they thought, would be the most critical information to share.
To further understand if the information provided by the system for sharing was useful, in Rome we undertook another treat simulation with the CS-AWARE database manager and one system administrator together. This session revealed the central position of the manager, including his prerogative for decision-making. The system administrator (responsible for a particular application, for example) was assigned particular tasks by the manager and had to get back to him as soon as the task was accomplished. The manager decided if enough if done to resolve the threat, after communication with the administrators for all network components that could be affected by the threat. From the system, this required the extended possibility to send multiple messages to several administrators. It should be noted that the current medium for handling communication about threat resolution in Rome is a ticketing platform. Implementing CS-AWARE would require either integrating this platform in the communication process, or replacing its functionalities. The role of the manager is crucial for the success of the process.

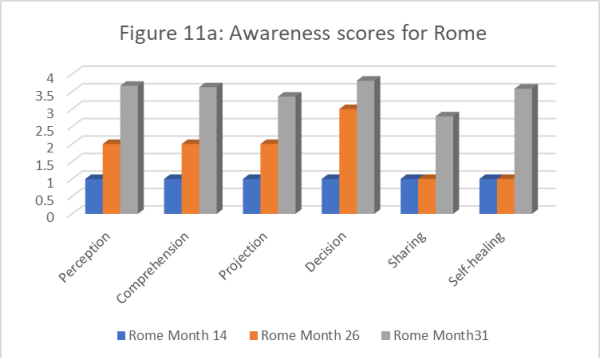
Usability, cycle 2	Rome	Larissa
N	3	4
Exp	2	5
S9: Knowability		
Clarity	0.67	0.85
Consistency	0.76	0.85
Memorability	0.69	0.72
Helpfulness	0.5	0.70
S9: Operability		
Completeness	0.78	0.73
Precision	0.78	0.72
Internationalisation (S13)	0.67	0.85
Flexibility	0.47	0.55
Mean	0.66	0.75

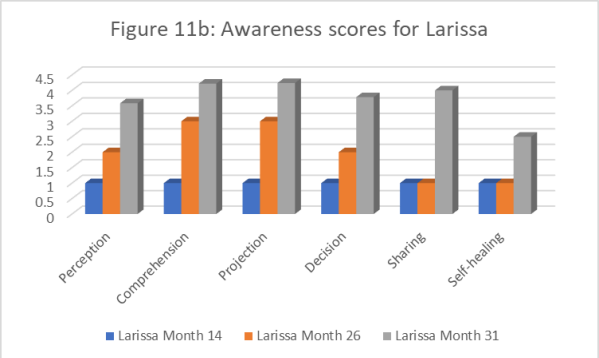
Table 13: Outcomes for Questionnaire 2 (cycle 2), for each requirement, in KPI (0-1). Exp means experience of the user (1-5).

VIII.2B Evaluation Level 2: Usability, Quantitative

The users in Rome filled in questionnaire 2 immediately after the usability test, and by the users in Larissa during the last week of April, after individual testing of the interface. Table 13 displays the outcomes. The following comments can be made:

- Again, there is a difference in experience between the participants in Rome (low experience) and those in Larissa (high experience). These differences reflect the difference in complexity of the organisations, rather than underlying motivations
- Again, all scores a well above threshold (.6), except the score for Helpfulness in Rome, and the score for Flexibility in both municipalities. The mean score overall can be taken as evidence that the users evaluate the usability of the system as above average.
- The low score for Helpfulness in Rome means that the users indicate they needed more help than the system was providing. Our observations during usability support the explanation that system administrators need support for: a) deciding that a threat is resolved (because more





components could be involved); 2) interpreting information for sharing (because this requires a management policy); and 3) understanding how to handle the system visualisation screen, and the operation of the keyword search, that was linked to particular system components.

- The Flexibility question involved the user's estimate that CS-AWARE was to be used by experienced users only. An affirmative answer to this question was scored negative on the scale for flexibility. A flexible system is a system that can be used by all types of users, experienced and less experienced. The low score here reflects the users' opinion that effective use of the system requires certain expertise, which is not necessary a negative outcome.

We feel safe to conclude that, in spite of (or because of) the additions to the system interface causing additional reflection and discussion, the score for Usability is above average. The scores on questionnaire 2 for usability are not very different from those of questionnaire 1, where usability was dealt with by one general question. The scores for that question were 0.6 for Rome and 0.8 for Larissa, for system administrators

VIII.3 Evaluation Level 3: Awareness

The awareness questionnaire 3 was administered to the same users as questionnaire 2, after the usability test (for Rome) or after sufficient experience with the system (in Larissa). This questionnaire does not inquire about the success or ease of certain activities during the usability simulation, but on user awareness of the different aspects of the phases of decision-making. Users are therefore asked about the aspects of this process that they considered and took into account, and the communicative actions that they undertook. Table 14 gives an overview of the scores on awareness of each of the phases.

Awareness, cycle 2	Rome	Larissa
N	3	4
Exp	2	5
S1: Provide Cybersecurity Awareness		
Perception	0.73	0.72
Comprehension	0.72	0.84
Projection	0.67	0.85
Decision	0.76	0.76
S2: Sharing		
Sharing	0.54	0.80
S3: Self-healing		
Self-Healing	0.72	0.63
Mean	0.65	0.74

Table 14: Outcomes for Questionnaire 3 (cycle 2), for each requirement, in KPI (0-1). Exp means experience of the user (1-5).

The relatively low score for sharing in Rome can be explained by the fact that users in Rome indicated they did not succeed in sharing, due to reasons explained under usability.

Figures 11a and 11b show the comparison in awareness scores from baseline, cycle 1 and cycle 2. Note, that the first two scores (baseline and cycle 1) were interpretations from qualitative information. As expected, all users in Rome indicate more awareness about threat perception and comprehension, which are strong assets of the CS-AWARE system. We see the same tendency for Larissa. Projection, or the awareness of the threat in the context of the system network, which is facilitated by visualisation in CS-AWARE, also increased. We interpret this as a realisation in Rome of the importance of this aspect. In Larissa, in a less complex system, awareness of the system is more obvious. There, we see a highly developed awareness overall, with the exception for self-healing. This aspect will receive more attention during cycle 3.

As a conclusion we can say there was a very satisfactory increase of awareness during cycle 2, especially in Larissa. This increase can be attributed to the many discussions about the system, and the time spent on working with the system, especially in Larissa. We will see if this proves sustainable in cycle 3, where the same instruments will be used.

VIII.4 Evaluation Level 4, Organisation in cycle 2

Questionnaire 4 was filled in by 12 managers from Rome and by 4 system administrators and their manager in Larissa. Table 15 shows the results.

We can see in table 15 the overall tendency of the managers in Rome to respond to the questionnaires with very positive ratings, especially for questions about the organisational level. Their experience with the CS-AWARE system may be low, but it can be supposed that there are many internal discussions in

Rome about CS-AWARE, at the organisational level. This may be less the case in Larissa, but we see that the expectations (and experience) of the system administrators in Larissa are high for this level. As a conclusion for the organisational level of evaluation we can say that KPIs are met, which is an excellent outcome with promising perspective for the actual success of CS-AWARE in the municipal context.

VIII.5 Evaluation Level 5, the business level in cycle 2

In spite of the organisation of a workshop in Barletta, for mayors and managers of the region, we could not administer questionnaire 5. This questionnaire was designed at a later stage, and at that moment the corona-virus hit Italy hard. Due to the COVID situation, in particular, local officials are reluctant to dedicate much, if any time, to new projects, no matter how important they might be. They are having enough problems with dealing with jobless, homeless, access to social services, etc. in their communities**.** Funding from the national government for now is being dedicated to restarting activities but the expectation is that like in earlier years the national government will reduce its allocation to local governments.
Instead, we provide here a brief report of the meeting, written by CS-Aware partner Cesviter, and their impression of chances for further marketing from our dissemination coordinator.

Report of meeting in Barletta: Lucia de Mari, a journalist in charge of conducting the presentation opened the meeting. She opened explaining the general information about the local interest in Cyber Security and in the CS-AWARE project. After that, she introduced the key speakers. First to talk was Marco Barone, the director of "Patto Territoriale nord Barese Ofantino" (Territorial pact nord Barese Ofantino), a public association that gathers both professionals and public officers together to build a stronger economic tie for that area of southern Italy. He explains why they are supporting us in this project (basically the idea of leading the market and pursuing innovation). After him, Fabio Roccuzzo, Focus Europe's director (an association that gathers small municipalities together to participate in European projects), explains its support to the CS-AWARE project and why it is considered the project-pioneer in the field of Cybersecurity awareness. After him, Sabino Mansi (Vice-president of the Engineers Order of Barletta, Andria and Trani) spoke. He gave his regards, also speaking for the President of the Order, and gave a statement on the scientific importance of the project. After him, we received the biggest endorsement by far: the next keynote talker was Cosimo Cannito, Barletta's mayor. After having stated his excitement for the project, he actually said that, after the presentation, he would have brought the CS-AWARE and its functionalities to a meeting with other mayors from the Province at a meeting with the Prefect (the official in charge of public safety in every city). After this roundtable of guests, we started the CS-AWARE presentations. First to talk was Manuel Leiva, generally introducing explaining the general information about CS-AWARE and the project partners. After him, Massimo Della Valentina spoke and introduced the concept of Active Cybersecurity, Awareness and the project's general functionalities.
After a small coffee break, John Forrester explained the dashboard and some screenshots of the software. Then, the Q&A followed. The presentation went smoothly, no problems or any sort of interruption. We gathered a lot of interest and a lot of Support letters to the project. It took approx. 2.30h of time, excluding the coffee break.

Informal evaluation: Following the various interactions via our dissemination channels with both companies and LPAs in various EU countries (UK, France, Netherlands, Italy, Greece) we got the following validation and feedback:

- overall, none of the LPAs or big companies in the above countries have any similar solution at the same level of complexity and offering as CS-AWARE
- all have various security tools, focusing on specific items such as networking monitoring.
- all the contacts we discussed with had various breaches, some that were not publicly made due to the fear of losing business and reputation
- the overall feedback was that such a tool would be welcome and very interesting for their security needs.

IX. Conclusions for Cycle 2

Achievements in cycle 2: Table 16 summarises the outcomes for cycle 2. Users from both teams indicated that the components of the deployment scenario, as established in November 2019, were still valid and no changes were necessary.

Organisation, cycle 2	Rome	Larissa
N	12	5
Exp	1.8	4.4
S10: Compliance with Internal Regulations		
Perception	0.80	0.76
S10: Service Delivery		
Comprehension	0.74	0.74
S1: Provide Cybersecurity Awareness (in organisation)		
Projection	0.79	0.72
Mean	0.78	0.74

Table 15: Outcomes for Questionnaire 4 (cycle 2), for each requirement, in KPI (0-1). Exp means experience of the user (1-5).

The specific monitoring patterns (for both Larissa and Rome) were instantiated, constructed from the input of the 3rd round of SDA workshops. The visualisation and information sharing protocols were refined and better integrated. New social media (general security warning) use case was instantiated at both locations. Through our usability testing we had good opportunities to test these additions and their impact on users.
The outcomes of the questionnaires overwhelmingly show that our intervention was well received at both locations. The most important difference between Rome and Larissa was the relatively high manager scores (and involvement) in Rome, and the relatively high scores from system administrators in Larissa. In addition, in Larissa, the system administrators spent considerable time in appropriating the system, therefore, their feedback is considered especially valid.
The evaluation at the business level through formal methods was seriously harmed by the corona-crisis. We were not able to obtain questionnaire feedback on already organised events, and upcoming events all had to be cancelled.
The main goal of cycle 2 was to look at information sharing. This has technical implications (what to share about the threat) as well as organisational ones (who is allowed to share what to whom, and what is our policy?). In addition, there were privacy issues: what sensitive personal information is involved in sharing? We achieved to put this on the management agenda of the two municipalities, and agreed about a protocol, allowing the CS-AWARE tool to store 'shared' information in a database, with details to be selected by the users. This means that sharing works well at the technical level.

Lessons learnt in cycle 2: The project work was affected by the second review, but the municipalities were not disturbed by this at all. However, the corona-crisis had an impact on the availability of people for further feedback and testing. Methodologically, we saw that the questionnaires were returned reasonably well, in spite of these drawbacks. We learnt that the lifespan of a questionnaire probably is more than three months, and not biweekly, as we originally intended. Users indicated
that they would not appreciate frequent consultation through questionnaires as nothing much would have changed within a cycle of three months. Whilst the outcomes of the questionnaires where encouraging, we learnt most from the usability testing, where we came to understand why our sharing of information protocol could pose problems to the users.

Implications for the next phase: As a general conclusion at the end of cycle 2, we have confidence that our approach is promising and has a good possibility for collecting further interest and potential new users, albeit at stages after the project. In cycle 3, we will come with some more results.

Deployment	**Rome**	**Larissa**
Technical achievements	**Instantiation of Roma and Larissa specific monitoring patterns finished. Refinement of system and dependency graph protocol finished. Internal Information sharing communication protocol between information sharing and visualization component refined. Social media monitoring (general security warning use case) instantiated for Rome and Larissa use cases.**	
Deployment Team	**13 participants**	**6 participants**
Objectives	**Efficient detection, better collaboration**	**Effective detection**
Foreseen Impact	**Improved Communication with service providers**	**Learning, reputation**
Desired	**Alerting and reporting**	**Reporting**
Evaluation	**Rome**	**Larissa**
1) Technical		
method	**Questionnaire 1**	
participants	**4 sysadmin, 10 managers, 6 service users**	**3 sysadmin, 1 service user**
Mean score	**0.71**	**0.81**
2) Usability		
method	**Cognitive walkthrough**	**Cognitive walkthrough**
participants	**3 (1 Data Center manager and 2 sysadmin)**	**3 (system admin)**
outcomes	**Observations and implemented changes**	**Observations and implemented changes**
processes	**Distributed**	**Collaborative**
method	**Questionnaire 2**	
participants	**2 sysadmin, 1 manager**	**4 sysadmin**
Mean score	**0.66**	**0.75**
3) Awareness		
method	**Questionnaire 3**	
participants	**2 sysadmin, 1 manager**	**4 (sysadmin)**
Mean score	**0.65**	**0.74**
4) Organisation		
method	**Questionnaire 4**	
participants	**12 (managers)**	**1 manager, 4 sysadmin**
Mean score	**0.78**	**0.74**
5) Business		
method	**Questionnaire 5**	
participants	**none**	**none**
scores	**NA**	**NA**
impression	**No similar tool exists, there is a clear need**	

Table 16: Summary of outcomes of cycle 2 for the two LPAs

X. DEPLOYMENT CYCLE 3

The main goals of the third and final cycle of piloting (table 17) were twofold. Firstly, to deploy the self-healing functionality at both municipalities, and second, to conduct final tests with the system.

Timing	M32	M33	M34
Deployment			
Technical	**System Development**		
Users	**Deployment team meetings**		
Evaluation			
Level 1: Technical			**Questionnaire 1**
Level 2: Usability			**Final Usability test Questionnaire 2**
Level 3: Awareness			**Questionnaire 3**
Level 4: Organisation			**Questionnaire 4**
Level 5: Business		**Contacting new users and sending Questionnaire 5**	

Table 17: Activities for deployment and evaluation during cycle 3 (April-June 2020)

Technical deployment: The technical developments in cycle 3 focused on the deployment of the last missing core feature of CS-AWARE – the self-healing feature - to the Rome and Larissa specific use case. This included a concrete instantiation of the self-healing policies for the Rome and Larissa specific monitoring patterns (described in D2.3 and instantiated in pilot cycle 2, as described in Section VII of this document). The general self-healing policies for each pattern have been described in D2.3 as well. The work in this cycle focused on instantiating the corresponding technical commands that would allow the policy to be implemented in specific cases.
Since it is seen as a bad practice to implement self-healing in the live systems of the Municipalities, a virtualised server environment was set up to act as the receiver and target of self-healing actions, instead of using the live systems of Rome and Larissa. It should be noted that no architectural and/or communication protocol changes to the CS-AWARE architecture were necessary at this point, and the architecture and protocol reported in deliverable D2.4 is still valid for all aspects relating to self-healing.

X.1 Evaluation Level 1: Technical

Questionnaire 1 was again administered to users in Larissa and Rome. Table 18 shows the outcomes for each municipality, separated for type of user and the requirements that were tested. It should be noted that the response was lower than the first time, due to questionnaire fatigue. Also, our contacts noted that 'once is enough', meaning that not much difference was to be expected.

The following observations seem relevant, where we should keep in mind that no firm conclusions should be based on these low response rate:

- Although the response rates were low, overall, for all requirements and for all target groups, scores were consistent and slightly higher than for cycle 2.
- Scores were much higher for requirements that scored relatively low during the previous round: sharing information and self-healing.

We take these outcomes as a confirmation that CS-AWARE was well received, in general, and that it technically performed according to the requirements. The improvements made as a result of user feedback (Annex 3.3) were appreciated as well. The users now appreciated and understood self-healing.

Q1 CY-CLE3	Sysadmin	Managers	Service users	Mean
Rome				
N	1 (4)	4 (10)	5 (6)	10 (20)
Exp	2	1.5	1.8	1.7
S1	0.7 (0.74)	0.89 (0.82)	0.81 (0.75)	0.83 (0.78)
S2	0.73 (0.57)	0.77 (0.72)	0.72 (0.61)	0.74 (0.66)
S3	0.6 (0.6)	0.75 (0.72)	0.75 (0.59)	0.74 (0.66)
S9	0.8 (0.6)	0.85 (0.78)	0.76 (0.73)	0.8 (0.73)
S10	0.6 (0.6)	0.87 (0.70)	0.73 (0.67)	0.77 (0,72)
Mean	0.69 (0.64)	0.82 (0.76)	0.76 (0.67)	0.78 (0.71)
Larissa				
N	3 (3)	1 (0)	2 (3)	6 (6)
Exp	5		2	4.67
S1	0.88 (0.82)	0.80	0.93 (0.9)	0.88 (0.86)
S2	0.76 (0.56)	0.80	0.93 (0.91)	0.89 (0.69)
S3	0.83 (0.8)	0.80	1.0 (0.87)	0.87 (0.84)
S9	0.93 (0.8)	1.0	0.9 (0.93)	0.93 (0.87)
S10	0.87 (0.8)	1.0	0.9 (0.87)	0.9 (0.83)
Mean	0.84 (0.74)	0.84	0.94 (0.9)	0.89 (0.81)

Table 18: Outcomes for Questionnaire 1 (cycle 2 between brackets), for each requirement, in KPI (0-1). Exp means experience of the user (1-5).

X.2A Evaluation Level 2: Usability, Qualitative

Figure 12: The opening screen for cycle 3 (example)

For the final test, we prepared a set of use cases for participants to resolve. As explained above, because we were testing self-healing, we used a test environment, and not the live system of the municipalities. We checked, in addition, that self-healing also worked in the target systems, so our usability test would be valid for the actual environment as well. We will discuss the outcomes for usability for the main screens of the console.
Figure 12 shows an example of the opening screen for the usability test. We asked users to resolve 4 different threats. Those that show a question mark in the 'state' column have self-healing options. The colours indicate the severity of a threat, in the order red-orange-yellow-blue, and green. Green threats are the outcome of a keyword search from relevant social media. The example screen and the user actions refer to 'threat perception', one of the affordances of CS-AWARE.

For the test, the following use cases were presented for participants to resolve, for each participant in a different order, where two of the cases involved self-healing:

1. Social media report: during the previous cycle, we implemented the option for the system to engage in search in relevant social media on the basis of user-defined keywords. We asked the participants to interpret and handle an outcome of a search, as well as to feed a new keyword.
2. Vulnerability pattern: a software vulnerability may put systems at risk
3. Suspicious behaviour: a system behaviour monitoring pattern is triggered that indicates suspicious behaviour
4. General Security warning: a possible Denial of Service attack

The opening screen (Figure 12) displayed all of these threats (and other) at once, and we asked users to focus on the 4 specific threats, one by one. For Larissa, the participants were 4 system administrators, for Rome the main participant was the database department manager, who assigned tasks to two system administrators, who also participated in the test. Both

Summary	**Usability study, Rome-Larissa, cycle 3 FINAL TEST**
use cases	**Social media report Vulnerability case Suspicious Behaviour General Security Warning**
components	**All components of the system are involved, we had specific interest in self-healing.**
Participants (Rome)	**2 system administrators (SUET), 1 manager (Head of the Data Center of Roma Capitale), 1 system administrator**
Participants (Larissa)	**4 system administrators, specialised in applications or in database services**
method	**Cognitive Walkthrough**
outcomes	**All participants passed this final test**

Table 19: Main features of usability tests in cycle 3

these situations were realistic. Table 19 summarises the usability final test.
All users immediately inspected in the table (on the right) the type of threat and the system component affected.
This clearly was acquired behaviour we did not find during our first testing, or not to be carried out as efficiently. Clicking in the table in the row of the threat (except when clicking in the 'where' column, see below) will take the participants to the threat description window (Figure 13).

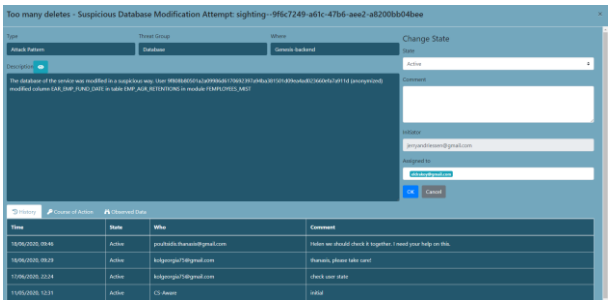


Figure 13: Threat description Window (example)

This window is the core screen from where to act. It describes the threat, at a general level, the observed data (technical details), the current course of action and the action that were already undertaken. Furthermore, it allows the participant to comment on a current state or action, to assign the threat to someone else (ticketing, as an outcome of cycle 1), and to mark a threat as resolved, when applicable.
For users, the comprehension of the threat and undertaking appropriate mitigation actions are done from this window. For example, a user can decide to accept the self-healing suggestion and apply it, as has been the case in the example in Figure 14. This can be done quickly, or with ample reflection, or through assigning it to one or more experts within the user group. Participants had no problems with

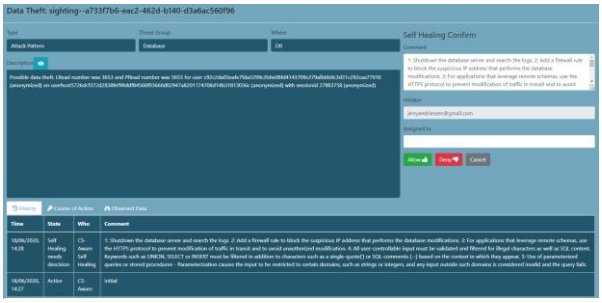


Figure 14: Threat Description window with self-healing suggestion (example)

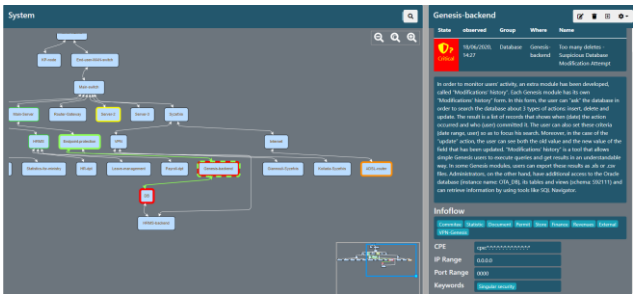


Figure 15: System visualisation (example)

these possibilities, their main reflections involved thinking about what to comment about their chosen action. All had learned (from the previous cycle) to inspect and interpret the details of the observed data. The participants in our usability test acted very effectively, especially by assigning any threat to another expert when this was seen as necessary.
When a self-healing suggestion is available, it is displayed in this screen (Figure 14).
While, during cycle 1, participants were hesitant to accept the suggestion, and information on the mitigation actions involved in self-healing was not as complete as in the current version, the final test showed more appreciation of self-healing and its implications. Users appreciated that this process was not automatic: there was the possibility to accept or reject the self-healing option, and the additional possibility to comment and decide if the threat had been resolved or if additional work was required. User agency in self-healing appeared to be an important asset. Participants were able to apply self-healing, often after consultation of the relevant experts.
Concerning projection of a threat, to study possible impact of a threat in the municipality network, CS-AWARE offers the network visualization, which (since cycle 2) immediately focuses on the part of the network implied by the threat (Figure 15).
Many things are shown in this screen. First, there is the network component that may be compromised, as well as the system nodes that are linked and possibly in danger. Also, the other threats are displayed. The threat that is in focus, and its main information can be seen in the right-side window. There (infoflow) it can be seen in what processes (pilot-scenarios, D2.2) the implied node is involved. Finally, we would like to mention the 'keywords' option, where users can insert keywords they thing are relevant to look for in social media. Participants all knew how to handle these keywords, as these were already investigated during cycle 2. It should be noted that not all participants inspected the system visualization. Threats can be resolved without looking at it, and this was sometimes what happened. This means that the description of a threat in the respective window was often sufficient for understanding and resolving a threat. But we can also conclude that handling of an emergency often took precedence over learning and deeper understanding. This was clearly due the nature of the usability test, but as well to the nature of their job description in general. Explicitly engaging into learning and increasing awareness takes place when there is no immediate threat and a user has time (allowed by their manager…) to 'play' with the system. We did not test that directly, nor did the usability and awareness questionnaires (2 and 3). This is an organizational matter as well: is development of awareness seen as incidental or is it an explicit feature of professional activity.
On the basis of the final usability test, we can however safely conclude that all users had mastered the system and were able to perform all essential activities efficiently and flawlessly. This clearly was not the case during cycle 1. The users have learned, and the system has improved.

Usability, CYCLE 3	Rome	Larissa
N	3	4
Experience	2	5
S9: Knowability		
Clarity	0.82 (0.67)	0.97 (0.85)
Consistency	0.73 (0.76)	0.98 (0.85)
Memorability	0.69 (0.69)	0.82 (0.72)
Helpfulness	0.8 (0.5)	0.70 (0.70)
S9: Operability		
Completeness	0.84 (0.78)	0.88 (0.73)
Precision	0.87 (0.78)	0.83 (0.72)
Internationalisation (S13)	0.8 (0.67)	0.85 (0.85)
Flexibility	0.53 (0.47)	0.45 (0.55)
Mean	0.76 (0.66)	0.81 (0.75)

Table 20: Outcomes for Questionnaire 2 (for cycle 2 between brackets), for each requirement, in KPI (0-1). Experience of the user (1:low -5:high).

X.2B Evaluation Level 2: Usability, Quantitative

Questionnaire 2 was filled in by the users in Rome and Larissa immediately after the usability test. Table 20 displays the outcomes. The following comments can be made:

- Overall, the scores are somewhat higher than for cycle 2, which we take as an indication that the improvements (Annex 3.5) were appreciated.
- Moreover, we can assume that users have advanced in their appropriation of the system, an

idea which is supported by the qualitative analysis of usability in the previous section.

- Two questions have mixed scores, one is about moving back and forth between screens, which can be taken in different ways. Users who often want to check their ideas often go back to a previous screen. However, this may also signify lack of experience and understanding. So, the question is ambiguous, but still received high scores (signifying infrequent movement between screens) in cycle 3. The other question is about the opinion of the users that CS-AWARE requires experience to be used, which received mixed scores. Here, as well, mixed interpretations are possible. This implies that our measure for flexibility (of the system: a component of Operability) is unreliable.

The usability sessions, as well as the reliable handling of all feedback from users, have contributed much to the appreciation and appropriation of the CS-AWARE tool by the users.

X.3 Evaluation Level 3: Awareness

Questionnaire 3 was filled in by the users in Rome and Larissa immediately after the usability test. Table 21 displays the outcomes. The following comments can be made:

- Overall, scores are somewhat higher than for cycle 2, which we take as an indication that the improvements (Annex 3.5) were appreciated, and of increased appropriation.
- Scores for awareness of the phases of decision-making were very high. This can be interpreted as high user agency for resolving threats, as a result of greater awareness of this process.
- The relatively new assets of the CS-AWARE approach, sharing information and self-healing, receive good scores. We interpret that as users indicating improved awareness of sharing information and self-healing.
- The high scores for comprehension indicate that the information that the CS-AWARE system provides to the users is very well received.

Participants of questionnaire 3 attributed very high scores to their awareness. This awareness pertains especially to the decision-making process when there is a threat.

Awareness, CYCLE 3	Rome	Larissa
N	3	4
Experience	2	5
S1: Provide Cybersecurity Awareness		
Perception	0.80 (0.73)	0.90 (0.72)
Comprehension	0.79 (0.72)	0.95 (0.84)
Projection	0.75 (0.67)	0.95 (0.85)
Decision	0.8 (0.76)	0.91 (0.76)
S2: Sharing		
Sharing	0.7 (0.54)	0.86 (0.80)
S3: Self-healing		
Self-Healing	0.71 (0.72)	0.89 (0.63)
Mean	0.72 (0.65)	0.89 (0.74)

Table 21: Outcomes for Questionnaire 3 (for cycle 2 between brackets), for each requirement, in KPI (0-1). Experience of the user (1: low -5: high).

X.4 Evaluation Level 4, Organisation

Questionnaire 4 was returned by 8 managers from Rome, and 1 from Larissa. Table 22 shows the outcomes. These are almost identical to the scores from cycle 2.

Because this level deals with managers, we asked the deployment teams to provide feedback on the 'look-and-feel' of the output table of CS-AWARE. This table is an excel file, that can be downloaded from the 'closed threats' page. This table has the overview of the threats that were closed over a period of one month. It contains the following slots: *closed at, first observed, group, Stix type, where, name, and description*. We got replies from both municipalities that these were OK. This means that, for now, the main desired artefact (deployment scenario, slot 3), from which many tables can be constructed, is satisfactory for the managers.

X.5 Evaluation Level 5, Business

The procedures outlined in section IV.2, set out by Cesviter and OTS have failed because of the coronavirus. Although we expected potential clients, at the municipal and regional levels to be open for our efforts, many local authorities are fully occupied with the impact of the corona pandemic, and do not have much attention left for anything else. Although questionnaire 5 was sent out by OTS to their network covering all municipalities in Greece, at the moment of finalisation of this Deliverable, there has been no single questionnaire returned.

Annex 2.4.5 has a list contacts of our Italian partner Cesviter, some of these were contacted during cycle 3. There also, no requests were returned.

Finally, through our partner Peracton, dissemination leader in CS-AWARE, some requests were sent to their contacts, with a null result.

Organisation, CYCLE 3	Rome	Larissa
N	8 (12)	1 (5)
Exp	1.8	1
S10: Compliance with Internal Regulations		
Perception	0.80 (0.80)	0.88 (0.76)
S10: Service Delivery		
Comprehension	0.76 (0.74)	0.84 (0.74)
S1: Provide Cybersecurity Awareness (in organisation)		
Projection	0.83 (0.79)	0.80 (0.72)
Mean	0.79 (0.78)	0.84 (0.74)

Table 22: Outcomes for Questionnaire 4 (cycle 2 scores between brackets), for each requirement, in KPI (0-1). Exp means experience of the user (1-5).

XI. CYCLE 3: CONCLUSIONS

Achievements in cycle 3: Table 23 provides an overview of the cycle 3 outcomes. Figure 20 provides an overview of all requirements and scores for cycle 3. As can be seen, no score is below 70%, and most scores are above 80%. In terms of these scores then, we can say that CS-AWARE has achieved all requirements in terms of KPIs.

The main achievement for the CS-AWARE system was the successful implementation of specific self-healing policies to the Rome and Larissa monitoring patterns.

Also, we can say that users successfully completed the four exercises that were part of the final usability test. This means that our users, and very probably also for users in other municipalities and professional contexts, with similar motivation and experience, can work with the CS-AWARE system to perceive, comprehend, project and mitigate cybersecurity threats. Moreover, we have shown that working with CS-AWARE increases awareness of cybersecurity, for individual users, and also within their organisation. This is the basis for realising the objectives that users formulated in the deployment scenario: efficient detection of cyberthreats, and better communication and collaboration between departments, between system administrators and managers, and between departments and citizen users of services. Better collaboration, learning, and increased reputation can be built on these achievements.

Lessons Learnt: The impact of the corona-virus during this cycle was twofold. In the first place, the number of questionnaires returned was about half compared to cycle 2. Of course, this could also be due to the impression by the pilots having to do the same thing twice, with no apparent changes. This explanation may hold, as the scores were the same or even higher than for the previous cycle of piloting. Especially self-healing and sharing information had higher scores. A final reason for lower return of questionnaire could be the very tight time frame. As we aimed for a final test, we wanted the questionnaires be filled in after the final test, which gave the respondents about a week for sending back the questionnaires, which might have been too short. Nevertheless, in spite of the low number of questionnaires returned, the scores were not affected.

The second impact of the corona virus seemed the impossibility to get time from potential new clients. Is was clear that the municipal level of government was most entangled in handling the corona-virus, which affected all professionals in those municipalities. As a result, we could not organise events, and could not persuade them to look at our questionnaires.

Deployment	**Rome**	**Larissa**
Technical achievements	**Instantiation of Roma and Larissa specific self-healing policies for monitoring patterns finished.**	
Depl. Team	**13 participants**	**6 participants**
Objectives	**Efficient detection, better collaboration**	**Effective detection**
Foreseen Impact	**Improved Communication with service providers**	**Learning, reputation**
Desired	**Alerting,reporting**	**Reporting**
Evaluation	**Rome**	**Larissa**
1) Technical		
method	**Questionnaire 1**	
participants	**1 sysadmin, 4 managers, 4 service users**	**3 sysadmin, 1 service user, 1 manager**
Mean score	**0.78**	**0.89**
2) Usability		
method	**Cognitive walkthrough**	**Cognitive walkthrough**
participants	**3 (1 Data Center manager and 2 sysadmin)**	**4 (system admin)**
outcomes	**Final test passed**	**Final test passed**
processes	**Distributed**	**Collaborative**
method	**Questionnaire 2**	
participants	**2 sysadmin, 1 manager**	**4 sysadmin**
Mean score	**0.76**	**0.81**
3) Awareness		
method	**Questionnaire 3**	
participants	**2 sysadmin, 1 manager**	**4 (sysadmin)**
Mean score	**0.72**	**0.89**
4) Organisation		
method	**Questionnaire 4**	
participants	**12 (managers)**	**1 manager**
Mean score	**0.79**	**0.84**
5) Business		
method	**Questionnaire 5**	
participants	**none**	**none**
scores	**NA**	**NA**

Table 23: Summary of outcomes of cycle 3 for the two LPAs

Implications: We think the main implication is that our work in cycle 3 confirms very clear benefits of the CS-AWARE system for the pilot municipalities. They now have a cybersecurity system that detects and explains cybersecurity threats, which makes them more aware of cybersecurity, their system and the state of their organisation. In the next section, which is the final chapter, we will elaborate somewhat more on this, trying to answer the main questions for deployment and evaluation, formulated in section 1 of this deliverable.

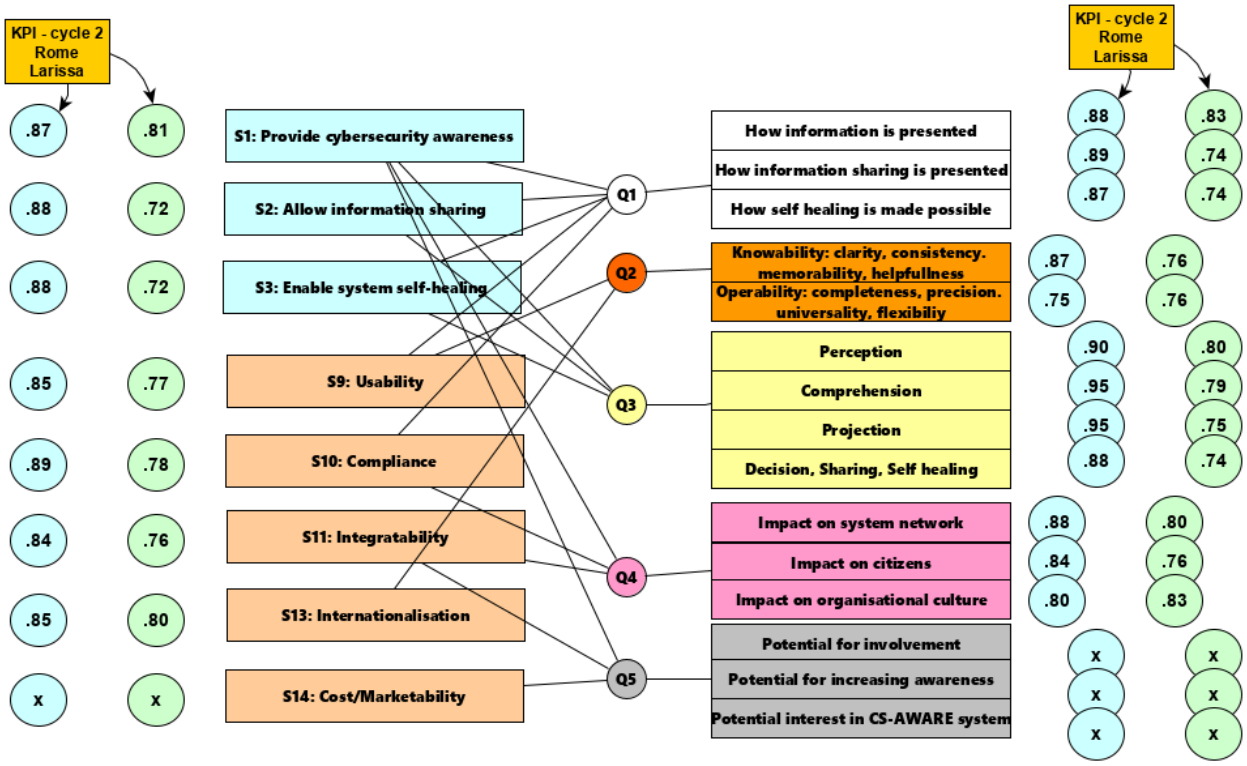


Figure 20: Cycle 3, requirements and KPIs

XII. FINAL CONCLUSIONS OF PILOTING

1) TO WHAT EXTENT IS THE TECHNICAL IMPLEMENTATION OF CS-AWARE EFFECTIVE?

From the technical perspective, piloting results have shown that the general framework as initially presented in deliverable D2.4 works as intended, and the technical functionality of each component is given in the generic case, as well as in the two pilot specific examples in Roma and Larissa. We have not identified any shortcomings or misconceptions of the technical framework that would necessitate a major revision of the implementation. This has been confirmed by the functional testing as reported in D4.2, as well as in the usability sessions with the users of the municipalities of Rome and Larissa. The positive outcomes have been confirmed by the evaluated results of questionnaires 1 and 2 in piloting rounds 2 and 3, with all KPI indicators showing a positive result. The analysis of the organizational and business level evaluation (represented by questionnaires 4 and 5) have not revealed any technologically relevant challenges that need to be addressed by the CS-AWARE framework.

Through the usability sessions with the users in Rome and Larissa we could confirm that the awareness created through CS-AWARE, and visualized to the user through the CS-AWARE user interface, provides insights into the security relevant system behaviour, and provides the user with information that is currently not available through any other technical means. A clear benefit and potential for efficient reaction to and mitigation of security issues has been identified. Those results have been confirmed through the results of questionnaire 3, with all relevant KPI indicators showing positive results.

One aspect that worked exceptionally well is the interaction between workshop participants during the socio-technical system analysis (according to the methodology specified in deliverable D2.5), which has proven to be an exceptionally positive element in the creation of individual and organizational cybersecurity awareness. Throughout the workshop the users gained a deep technical understanding of their own systems and their security relevance, which simplifies the interactions with the technical part of the CS-AWARE system, since incidents based on behaviour monitoring detected by CS-AWARE are instantly comprehended and understood by users that helped to define them. Furthermore, since this analysis leads to system monitoring that the actual users and administrators of the system actually care about, an increased motivation to address potential issues detected by the CS-AWARE system could be observed. This is an observation that we could only capture qualitatively through our design-based evaluation method, and is impossible to quantify through KPIs.

It was shown that the two advanced concepts implemented by the CS-AWARE interface, self-healing

and information sharing have been received well by the users and the potential for inclusion of those functionalities in the day-to-day work flow has been confirmed. It was observed for both functionalities that, since those concepts are relatively new and have not been available in any way before CS-AWARE in the two piloting municipalities, the time to get familiar with the relevant interfaces provided by CS-AWARE has been longer than for the other tested functionalities. This observation is confirmed by the analysis of results from questionnaire 1, with the KPIs for self-healing and information sharing initially scoring lower in round 2, but increasing in round 3. One aspect that cannot be quantified by KPIs but is a significant result relates to information sharing. It was brought forward by the municipalities that information sharing is not seen as a major technological challenge, and the functionality provided by CS-AWARE fulfills the technological need. However, sharing of security relevant information with parties outside the organization provides significant organizational and policy challenges that need to be solved first and are currently not addressed. It was acknowledged however that due to the reporting obligations of the NIS directive and the GDPR, municipalities are currently under pressure to get the relevant policies in place to address security relevant information sharing.

To conclude the analysis of the technical perspective of piloting results we would like to note that piloting has shown that the implemented CS-AWARE user interface concept is very well received by the users, and can be used to its full extent after a short time by new users due to its focus on simplicity and efficient work flow. The usability improvements that have been implemented based on feedback and observations during the usability tests - reported in Annex 3.1 (round 1), Annex 3.2 (round 2) and Annex 3.3 (round 3) - helped to further refine and significantly improve the workflow of the user interactions. Piloting has shown that the general framework works as intended and expected, and the technical functionality of each component is given in the general case, as well as in the two pilot specific examples in Roma and Larissa.

2) To what extent is the CS-AWARE system usable by expected target users?

From the perspective of usability, the users attributed very high scores to this aspect of the CS-AWARE system, and scores increased over time, which implies that they appreciated the improvements to the system that were made on the basis of their feedback and the outcomes of tests.
CS-AWARE is not a simple system to use. This is especially due to the complexity of information for cyberthreats, the difficulty of comprehending this information, understanding the role of compromised software or technology in the context of the municipal network, and making the correct decision for threat mitigation and protection of the network and the citizens. Such is the work of professionals, and they cannot make mistakes without consequences for the network, departments in the municipality, and, ultimately, the citizens.
We have deployed and tested the system in the context of two municipalities, with very different ways of working. In the municipality of Larissa, system administrators performed all the tasks, or consulted with their direct colleagues. In Rome, expertise was distributed, and various tasks where handled by different persons.
For each of these tasks, the system provides options. Users indicated that:

- The elements presented by the interface were clear, and their structure and functions were clear
- The elements were presented in a consistent manner
- The way the system works is easy to remember
- Working with the system does not require much help, with some time for learning it
- The presented information is complete, for making decisions
- The descriptions are clear and precise
- The system allows flexible use, users can go back and forth as they please, and retract most earlier steps or decisions
- The system presents the information in the language of the user, translations work well
- The options that the system presents are sufficient for acting

Moreover, the learning curve for using this complex system requires a couple of hours of practice, perhaps in the form of a guided session with an experienced coach, for example a system administrator who has already worked with the CS-AWARE system.
For new users understanding how to work with the system, we highly recommend usability sessions with new and beginning users. This could be part of new user sites co-designing and appropriating the system through socio-technical workshops.
To conclude, the CS-AWARE interface supports immediate detection of threats (depending on the cycle for scanning the municipality network), and supports comprehension of the nature of the threat, the part of the system network that is compromised and the other nodes that may be involved, and offers understandable self-healing suggestions when these are available.

3) To what extent is the awareness of users affected by discussing and using the system during deployment?

First, we note that the awareness of cybersecurity was already positively affected and enhanced as

soon as the project started, and most significantly for users participating in the SSM workshops. For the purpose of evaluation, we started interpreting awareness after the story workshops at the beginning of year 2, and again at the start of deployment at the beginning of year three.

We defined awareness as a concept with 6 components: *knowledge* of cybersecurity threats (1), the system network (2), the organisation (3), external cybersecurity-related organisations and communities (4) as well cybersecurity *agency*: knowing how to act in case of a threat (5), and acting when there is no threat (6).

Agency can be defined here as the possibility for actively contributing to cybersecurity. It alludes to the capacity of humans to distance themselves from their immediate surroundings and it implies recognition of the possibility to intervene in, and transform the meaning of situated activities (Mäkitalo, 2016)[xv]. With our qualitative approach (stories, deployment scenario), we could interpret organisational and communal awareness, as well as agency when there is no threat. With the usability exercises (also qualitative), we could investigate user agency when dealing with threats, and, to some extent, how users exploited the affordances of the CS-AWARE tool for comprehension, also in the context of their system network, and for sharing with external organisations. Finally, with the awareness questionnaire, we could check how users rated their use of the affordances of the CS-AWARE tool during the phases of decision-making in case of a threat. We assumed that higher ratings for use signify greater awareness.

Concerning knowledge about threats (1), we can say that instead of laborious sessions inspecting logs and internet sources, which happened before CS-AWARE, this is now handled by the CS-AWARE system, and in such a way, that information about threats is readily and immediately available for the user. This means that potentially, threat awareness is increased. Users agree with this statement, reflected by a score for requirement 1 (provide cybersecurity awareness) that is very good. We therefore conclude that knowledge of threats is increased with every new threat, especially when users are sharing their experience with others, during or after threat resolution.

The knowledge by users of their municipal system network (2) was already greatly fostered during the SSM workshops, taking place before deployment. We have evidence that the awareness ratings for system visualisation and projection of threats has gone up during deployment, and can therefore conclude that also system network awareness is addressed in a more than sufficient manner.

The organisational level of awareness is addressed below, through evaluation question 4. The level of external organisations, communities, or other departments involved in cybersecurity (awareness component 4), who all would have an interest in sharing information about cyberincidents, as it stands now, is merely a technical asset. The concept of sharing information involves knowing what to share, and with whom, for what reason. Our participants were aware of the need for sharing information in general, but there was no installed policy for this, neither was there explicit mention of some collaboration with other communities or departments in this respect. Both municipalities were aware that this issue needs follow-up.

We understand most about the awareness component of user agency for handling cyberthreats (5). When we say that CS-AWARE is an expert system, we crucially refer to the process of dealing with threats. This process was different in the two pilot municipalities. In Larissa, the system administration department handled all cyberthreats. This is a small department, with experts who can work together to resolve a threat. The way they currently work does not have to fundamentally change with CS-AWARE. In Rome, expertise on all aspects of the extended network of nodes and services, is highly distributed. Handling cybersecurity threats requires a central expert who delegates different tasks to different experts, who are responsible for their particular system or service. More often than not, handling a cyberthreat will involve more than two system administrators, who work in different departments. For the communication that this process requires, Rome already had a ticketing system in place. CS-AWARE is made to work in both contexts, so a ticketing system was implemented as well These differences have implications for awareness and on how expertise is distributed between users.

We distinguished four main phases in the process of dealing with threats, and users indicated (in the questionnaires) good awareness of all of these phases. From the usability tests, we learned a number of specific things for awareness:

- Concerning perception, the opening screen of CS-AWARE provides immediate awareness for those involved in monitoring.
- Concerning comprehension, we noted that users attend to the main characteristics of a threat (type, date, system component involved), but not always to all details (detailed description, system information and history). This may have good reasons, linked to a user's expertise, and the need for immediate resolution may require efficient handover. However, we observed that threat comprehension has more attention from those who are responsible for all aspects of

threat resolution, than from those who deal with some part of that process.

- Concerning projection, the same applies as for comprehension. While some users study the network visualisation extensively, others do not look at it, and focus on their own 'section' of the network. It should be noted that the actual repair, by inspecting log files of the affected system component, still takes place 'outside' of CS-AWARE, except in the case of self-healing. We highly recommend training for new users to focus on projection of threats through system visualisation with CS-AWARE.
- Concerning decision-making, we noted that handover of threat mitigation to other users was the rule rather than an exception. Also, we noted that most users have the habit of checking if their decision was implemented correctly (e.g. threat now listed in resolved threats, or handover now included in current threats). It was clear that this already was part of their normal routines, and now made explicit (and recorded) through CS-AWARE.

As a conclusion for agency in handling threats, as a component of cybersecurity awareness, we can say that CS-AWARE greatly facilitates user agency, making detection and mitigation more efficient and effective, with the additional asset of better comprehension and projection of threats. The extent to which comprehension and projection abilities of users increases, depends on the extent to which users pay attention to this information. The very positive outcome of the questionnaire seems to be relative to the roles and actions of users during threat mitigation.

Finally, how about user agency when there is *no* immediate threat (6)? We know that system administrators in Larissa have been 'playing' with CS-AWARE when there were no threats to resolve. Of course, this is an important activity for gaining experience. System Administrators in Larissa stated their ambition for learning in the deployment scenario. They had an interest in an improved reputation for their department, as a consequence of improved services. This could lead to considerations about weaknesses in the network components and development of new services for citizens. On the other hand, in Rome, ambitions were formulated at the management level, in terms of more and more effective interactions between departments in the context of cybersecurity. Although the managers in Rome were very positive about possible organisational impact (see the next section), it remains to be seen how these expectations will be realised.

Concluding, through deployment of the CS-AWARE system, we can say that cybersecurity awareness in both pilots has been greatly increased, at the level of threat detection and mitigation, and, to a somewhat lesser extent, to understanding and learning about threats, also in the context of the system network. Further work is expected for the exploitation of increased organisational awareness in Rome, and the elaboration of sharing threat information with relevant external agencies and authorities.

4) To what extent does using the CS-AWARE system have impact on cybersecurity awareness at the organisational level?

As has been previously set out, system requirement 1 (S.1) of the CS-Aware System is to raise and maintain awareness of cybersecurity both at an operational and organisational level. During cycles 1, 2 & 3 the users completed questionnaires that covered a range of topics, each of which included potential aspects of awareness. The responses to these questionnaires are summarised in the preceding sections of this report. The responses we received to questionnaire 3 in cycles 2 & 3 enabled us to gauge the growth of the respondents' cybersecurity awareness as the project progressed, the baseline having been established by interpreting the outcomes of the story telling workshops. In terms of the organisational effects of the CS-Aware System, questionnaire 4 established the extent to which the CS-Aware System impacted upon organisational culture and security awareness.

It is clear from the responses to questionnaires 3 & 4 that the CS-Aware System has, in the opinion of the respondents, significantly improved both their cybersecurity awareness and that of the organisation(s) in which work. However, and despite positive responses concerning organisational impact, we are concerned about the issue of maintaining cybersecurity awareness at the organisational level of an enterprise over a period of time.

We might surmise that as those staffs that were initially involved in the SSM workshops and in the consequent adoption of a CS-Aware System move onto to other roles – possibly in other organisations, the state of "organisational awareness" of cybersecurity in an organisation may begin to decay. This is of course both a problem of knowledge management and of organisational memory – the issue of how an organisation maintains its knowledge base in despite the inevitability of staff turnover.

In our view this can be avoided in three ways: Firstly, by instituting an annual SSM workshop involving personnel from both systems admin and senior management to review and if necessary, re-tune and update their organisation's CS-Aware System – ideally as part of an IT audit in compliance with ISO

20701. Secondly, ensuring that a formal management reporting line is established that provides at least quarterly reports concerning the effective operation of the CS-Aware System and outputs from it to senior management at board level of the organisation. Thirdly, by using the CS-Aware System console to maintain a record of mission-critical security incidents and of how and the extent to which the consequences of these incidents were mitigated.

5) TO WHAT EXTENT CAN OTHER MUNICIPALITIES BE INVOLVED IN OUR APPROACH?

It is clear from our experience of working closely with the two pilot cities Larissa and Rome, that the CS-Aware approach is broadly applicable to a wide range of LPA's and that therefore, other municipalities can be involved in our approach.

All LPA's within the EU operate against the background of EU-wide legislation and within very similar operational frameworks and constraints. For example, the GDPR in particular, places clearly defined legal responsibilities on LPA's to ensure the security of personal and sensitive data and information. Under the terms of the GDPR, non-compliance can cost organisations up to 4% of their global annual turnover – or up to 20 million Euros. Such fine for a public sector organisation would be devastating for the provision of public services.

In recent times LPA's and public sector organisations in general, have been the victims of targeted cybersecurity attacks. For example, in 2018, attacks were launched at public sector organisations such as the Bank of Spain, the UK Parliament, the, the German military. In 2019 the Spanish Ministry of Defence and UK local Government networks were also the targets of attack. In the UK LPA's experienced an average of "800 cyber attacks every hour with more than 263 million incidents in the first 6 months of 2019 alone. [2]

It follows that all LPA's within the EU are, of necessity and with support from ENISA, actively seeking to protect their IT infrastructures and applications from the rapidly growing risk of attack, while at the same time often experiencing the consequences of economic austerity.

Following the various interactions via our dissemination channels with both companies and LPAs in various EU countries (UK, France, Netherlands, Italy, Greece) we got the following validation and feedback:

- overall, none of the LPAs or big companies in the above countries have any similar solution at the same level of complexity and offering as CS-AWARE
- all have various security tools, focusing on specific items such as networking monitoring.
- all the contacts we discussed with had various breaches, some that were not publicly made due to the fear of losing business and reputation
- the overall feedback was that such a tool would be welcome and very interesting for their security needs.

The CS-Aware system has enabled the two pilot cities in the project to understand their IT systems and their vulnerabilities in a way unparalleled hitherto. It has also enabled them to monitor their key IT operations. The project consortium intends to provide the EU public sector organisations with a cost-effective cybersecurity solution and make the CS-Aware system available to LPA's across the EU.

[2] https://www.openaccessgovernment.org/cyber-security-in-the-public-sector/78477/

REFERENCES

[i] Schønheyder, J. F., & Nordby, K. (2018). The use and evolution of design methods in professional design practice. *Design Studies*, *58*, 36–62. https://doi.org/10.1016/j.destud.2018.04.001

[ii] Strauss, A. (1988). The articulation of project work: An organizational process. *The Sociological Quarterly*, *29*(2), 163–178.

[iii] Hogan, M., Ojo, A., Harney, O., Ruijer, E., Meijer, A., Andriessen, J., ... & Groff, J. (2017). Governance, Transparency and the Collaborative Design of Open Data Collaboration Platforms: Understanding Barriers, Options, and Needs. In: A. Ojo, & J. Millard (Eds.). *Government 3.0–Next Generation Government Technology Infrastructure and Services: Roadmaps, Enabling Technologies & Challenges* (Vol. 32). (pp. 299-332), Springer.

[iv] Berliner, D. C. (2002). Comment: Educational Research: The Hardest Science of All. *Educational Researcher*, *31*(8), 18–20. https://doi.org/10.3102/0013189X031008018

[v] The Design-Based Research Collective. (2003). Design-Based Research: An Emerging Paradigm for Educational Inquiry. *Educational Researcher*, *32*(1), 5–8. https://doi.org/10.3102/0013189X032001005

[vi] Paavola, S., Lakkala, M., Muukkonen, H., Kosonen, K., & Karlgren, K. (2011). The roles and uses of design principles for developing the trialogical approach on learning. *Research in Learning Technology*, *19*(3). https://doi.org/10.3402/rlt.v19i3.17112

[vii] Christensen, K., & West, R. E. (2017). The development of design-based research. *Foundations of Learning and Instructional Design Technology*.

[viii] Engeström, Y. (2011). From design experiments to formative interventions. *Theory & Psychology*, *21*(5), 598–628. https://doi.org/10.1177/0959354311419252

[ix] Allen, I. E., & Seaman, C. A. (2007). Likert scales and data analyses. *Quality progress*, *40*(7), 64-65

[x] Hibshi, H., Breaux, T. D., Riaz, M., & Williams, L. (2016). A grounded analysis of experts' decision-making during security assessments. *Journal of Cybersecurity*, *2*(2), 147–163. https://doi.org/10.1093/cybsec/tyw010

[xi] Anttila, J., & Knowledgist, V. (2006). General Managerial Tools for business-integrated information security management. Paper, paper at IPICS Winter School of the University of Oulu. http://www.qualityintegration.biz/InformSecPDCA.html

[xii] Mahatody, T., Sagar, M., & Kolski, C. (2010). State of the Art on the Cognitive Walkthrough Method, Its Variants and Evolutions. *International Journal of Human-Computer Interaction*, *26*, 741–785. https://doi.org/10.1080/10447311003781409

[xiii] Kurtz, C. (2014). *Working with Stories in Your Community or Organization: Participatory Narrative Inquiry*. Third Edition. New York: Kurtz-Fernhout Publishing

[xiv] Andriessen, J. & Baker, M. (2020). *On Collaboration: personal, educational and societal arenas*. Leiden/Boston: Brill/Sense Publishers.

[xv] Mäkitalo, Å. (2016). On the notion of agency in studies of interaction and learning. *Learning, Culture and Social Interaction*, *10*, 64–67. https://doi.org/10.1016/j.lcsi.2016.07.003

D5.1
ANNEXES

Grant Agreement number:	740723
Project acronym:	CS-AWARE
Project title:	A cybersecurity situational awareness and information sharing solution for local public administrations based on advanced big data analysis
Principal author:	Jerry Andriessen, Wise & Munro, jerry@wisemunro.eu
Co-author(s)	Thomas Schaberreiter, Kim Gammelgaard, Chris Wills
Document version:	1.0

Table of Contents

1 Annex 1: Deployment scenarios Rome & Larissa, November 2019

1.1 Deployment scenario Larissa

We organised a workshop at the municipality of Larissa on Thursday, October 10th, 2019. The people attended are mentioned in the 'Larissa Deployment Team' section. Together we discussed each slot of the deployment template. For some slots (local objectives, artefacts, behaviour, and impact), the participants discussed their ideas first in their own language, and then, in a plenary session, wrote their ideas on a flip over chart. Pictures of these charts were taken and further analysed. The pictures were shared in our joint (confluence) repository. Discussions between the local participants took on the average about half an hour, and plenary presentation and discussion for the respective slots required a similar amount of time. The atmosphere was constructive and awareness of slots resulted in the final planning table (table 2 below) that we prepared to be accepted by all without much discussion needed.

1. **Larissa Deployment Team: t**he role of the deployment team is to design, prepare and monitor execution of the deployment scenarios in conformance with the requirements. The Larissa deployment team includes the following roles and people:
   - Owner: The stakeholder who is supposed to locally exploit the outcomes of the scenario, or of several scenarios, who initiates the scenario and has a great interest in successful outcomes. The scenario owner is Kostoulas Aristotelis, manager of the system department of Larissa Municipality.
   - Representatives of the users, that is the system department: Poultsidis Thanasis, Kolovou Georgia, Drakou Heleni and Topalidis Christos. The first three people are supposed to work with CS-Aware on a daily basis.
   - Internal service users, employees of the municipality: Karioti Dimitra, Basdeki Dimitra.
   - Technical support: technical coordinator of CS-Aware (Thomas Schaberreiter), implementation coordinator (Kim Gammelgaard), data manager (Stefania Tola)
   - Exploitation partner, OTS (did not attend the workshop): Apostolopoulos George
   - Evaluators/moderators: Wise & Munro (Jerry Andriessen), Caris (Chris Wills).

The people above will together constitute the Deployment team. The team (or some members of it) will meet every two weeks to discuss on-going issues.

2. **Objectives (user requirements)**

Objectives will serve as performance indicators for evaluation. The general CS-AWARE requirements of deployment (see page 8/9 above) were presented to the participants before the session. These general objectives were taken up locally for the formulation of expectations, for each type of local stakeholder (see deployment team, above). This leads to additional indicators for evaluation, which will be taken up in the awareness and organisation questionnaires (see section 3 on Evaluation)

The Larissa **system administrators**, the actual users of the technology, formulated the following expectations:

1. To be informed on time about cybersecurity threats
2. No false alarms from the system
3. The system provides system administrators up-to-date information about solutions, that comes from official and trusted sites
4. Room for system administrator judgements, the system should not solve issues on its own

The objectives that the system administrators formulated for deployment are derived from their own existing situation. In this situation, that we analysed and reported in D2.2 by eliciting stories, we described a system department focused on solving individual user's issues, trying to avoid the blame for these issues as much as possible, by reminding the user about the rules and regulations, and by doing their job as well as possible. Such work is facilitated by timely alerts, less false alarms, as is often the case. The other feature of these objectives is that it will take time before these administrators develop trust in the CS-AWARE solution. In fact, all four expectations can be seen as prerequisites for developing this trust. Collecting feedback on their development of trust will therefore be included in the awareness level of the evaluation plan.

For the **manager (**Kostoulas Aristotelis**)** the following expectations were formulated:

5. No service down time
6. No extra costs or resources needed
7. No reputation damages

The objective from the manager's perspective is that deploying the system does not generate a disproportionate burden for the organisation, in terms of time, money, and other resources. In the evaluation at the organisational level we will include this expectation. It should be noted that the role of the manager in dealing with the presence of the CS-AWARE solution in daily practice, is still unclear. Development of increased awareness most probably implies handling more information about cybersecurity. On the other hand, we may expect increased awareness to evolve into different objectives.

For **the internal service user,** the following expectations were formulated:

8. No additional burden for their work
9. Feel safe and protected
10. Be clearly and concisely informed on time
11. Not being watched

The objective from the internal users' perspective is to be able to work without any interruptions caused by security issues. This includes being informed on time, without additional control or regulations. These expectations may evolve with increased awareness.

3. Desired Artefacts

Artefacts are the tangible output (reports, logs, all other information) of the CS-AWARE solution that allow stakeholders to see and reflect on what has happened in their system and network. For example, the manager can read a weekly summary table of threats and how these were resolved. These artefacts can serve an important role as mediators of awareness of cybersecurity within the organisation.

The **system administrators** desired the following artefacts:

- Ranking of threats by frequency (monthly)

- List of nodes in the system that were affected (weekly)
- Report of information shared by other LPA's (monthly)

From **the manager's** perspective, the following artefacts (to be sent by email) were envisioned:

- Weekly summary of threats (detected, resolved, ignored, still active)
- Report of services affected (weekly)

The **service user** could be informed by a weekly report of threat sources.

Most of these artefacts can be generated by editing the tabulated output that already is provided. The deployment team will further discuss the generation and use of these artefacts for deployment cycles 2 and 3.

4. Desired Behaviour

Desired behaviour is the behaviour of users that would be beneficial for achieving their objectives. Formulation of this behaviour would enable monitoring and developing awareness of what is desirable during deployment. Simply said: users should learn to see what activities are beneficial for handling the CS-AWARE system.

The **system administrators** identified the following behaviour:

- To assign monitoring roles on a daily basis (who will do it and what will be done?)
- To actively learn from the system and to getting informed about cybersecurity
- Collaboratively discussing solutions

These activities will be monitored for evaluation. They are excellent manifestations of growing awareness of cybersecurity and their development of trust in the CS-AWARE solution.

The **manager** is expected to read the reports, which is expected to lead to more trust in the work of his system administrators. As for the **service users**, it is expected they will develop more discipline in following rules and regulations. The current way of working is that individual users contact the IT-department in case of issues. We can track changes in the frequency and nature of such interactions.

5. Participants

The main participants in Larissa that we will monitor during the phases of deployment and evaluation have been identified as the three system administrators who are members of the deployment team. Their current knowledge and experience are highly relevant and sufficiently adapted to using the CS-AWARE solution. The service users will be members of the municipality, and more particularly (for evaluation) the two service users that are part of the deployment team.

6. Impact

Impact of deployment concerns what we expect to change when deployment objectives have been realised. Impact can include several cycles of scenarios, not just one, in other words, a long-term perspective. Thinking about desired impact allows us to understand 'how far we are' – and what can be realistically achieved. It may be required to implement additional activities to realise the desired impact.

The local stakeholders in the deployment team expect the following impact:

1. Prevention of attacks will be improved
2. Cybersecurity issues will be resolved faster and easier
3. More investment in training personnel or equipment
4. More trust by the general public in the municipality services
5. Increased status of Larissa municipality among other LPA's

These perspectives will be taken up as benchmarks for evaluation at the organisational level.

7. **Evaluation**

Evaluation specifies and assesses the criteria by which we can claim that the CS-Aware solution was successfully deployed. Because we engage in formative evaluation (evaluation during and not only after deployment) we have specified 5 levels of evaluation, and 3 cycles of assessment. Evaluation will be discussed in section 3 of the deliverable.

8. **Support**

Support during the first cycle of deployment for Larissa will be technical, as the CS-AWARE team is responsible for implementation of the CS-AWARE system within the local context. Furthermore, the users will receive a short manual of the operations of the console of the system, and their feedback will be discussed within the deployment team.

9. **Preparation and Planning**

Table 2 on the next page shows the planning for deployment in Larissa. Please note that planning for cycle 2 and 3 may change as a result of ongoing activity.
After the deployment team has been established in October 2019, they will have online meetings every two weeks, for discussing current issues, planning and feedback. System implementation will continue until mid-November. During the first cycle, formal usability and system testing will take place. After the system has been implemented, we will collect feedback on using the system on a biweekly basis. This may lead to technical work in the periods that follow. The nature of the feedback and the technical work will be documented and reported.

1.2 Deployment scenario in Rome

We organised a workshop at the municipality of Rome on Thursday, October 24th, 2019. The people attended are mentioned in the 'Rome Deployment Team' section below. Together we discussed each slot of the deployment template. For some slots (local objectives, artefacts, behaviour, and impact), the participants discussed their ideas first in their own language, and then, in a plenary session, wrote their ideas on a flip over chart. Because of the number of participants, they were divided into three teams, each presenting a chart with their ideas. Pictures of these charts were taken and further analysed. The pictures are shared in our joint (confluence) repository. Discussions between the local participants took on the average about half an hour, and plenary presentation and discussion for the respective slots required a similar amount of time. The atmosphere was constructive and awareness of slots resulted in the final planning table (table 3 below) that we prepared to be accepted by all without much discussion needed.

1. Rome Deployment Team

The role of the deployment team is to design, prepare and monitor execution of the deployment scenarios in conformance with the requirements
The Rome deployment team includes the following roles and people:

- Owner: The stakeholder who is supposed to locally exploit the outcomes of the scenario, or of several scenarios, who initiates the scenario and has a great interest in successful outcomes. The scenario owner is Arianna Bertollini, Rome Capitale, manager.
- Representatives of the users, that is the system department: Omar Parente, Andrea Quatrini, and Raffaele Conforte.
- There are several other managers involved in the deployment team: Claudio Guido Ferilli, Massimo Ferrarelli, Ivan Bernabucci.
- External service providers, from SUET: Antonio La Malfa, Angelina Marchio and Valerio Voci
- Internal service users, employees of the municipality: Valentina Modesti, Stefano Vallocchia
- Technical support: technical coordinator of CS-Aware (Thomas Schaberreiter), implementation coordinator (Kim Gammelgaard), data manager (Stefania Tola)
- Exploitation partner, Cesviter : John Forrester, Manolo Leiva, Massimo della Valentina
- Evaluators/moderators: Wise & Munro (Jerry Andriessen), Caris (Chris Wills).

The people above will together constitute the Deployment team. The team (or some members of it) will meet every two weeks to discuss ongoing issues.

2. **Objectives (user requirements)**

Objectives will serve as performance indicators for evaluation. The general CS-AWARE requirements of deployment (see page 8/9 above) were presented to the participants before the session. These general objectives were taken up locally for the formulation of expectations, for each type of local stakeholder (see deployment team, above). This leads to additional indicators for evaluation, which will be taken up in the awareness and organisation questionnaires (see section 3 on Evaluation).

The **system administrators**, the actual users of the CS-AWARE technology, formulated the following expectations, formulated as benefits:

- Easy identification, and classification of the threat
- Therefore, reduction of team work time for understanding a problem
- Suggestions for solutions and prevention of (future) damage

The objective for the system administrator seems to become more effective by receiving the appropriate information on time. Appropriate information is information that works for them, to better understand what is going on and to find the right solution. In other words, their objective is increased awareness.

For the **manager** the following expectations were formulated:

- The possibility to support a more structured and collaborative relation between system administrators and stakeholders of the service (SUET) in dealing with cybersecurity
- To be better informed about threats and mitigation
- Improved quality of the service
- More satisfied citizens

The objectives from the manager's perspective are improved internal processes, because of the possibility for collaborative responsibility, and also about results and outcomes for service users and citizens. The collaboration has already started with the presence of SUET during all workshops of the CS-AWARE project.

For **the internal service user,** the following objectives were formulated:

- Increased quality of work (business continuity, less complaints of final users) and more confidence in data integrity
- Better internal support to be able to explain issues to final users

For the **external users** (service providers and final users) the following objectives were formulated:

- Improved quality of work (time-savings, efficiency)
- Increased service reliability for citizens
- Personal data protection
- Increased trust in LPA

For both internal and external users, the objectives are related to a better experience of working with the services, because there are less interruptions, and where there are interruptions, they are better explained.
Overall, we see a clear desire for improving the quality of work. This ambition will be taken up in our evaluation, at the organisational level.

3. Desired Artefacts

Artefacts are the tangible output (reports, logs, all other information) of the CS-AWARE solution that allow stakeholders to see and reflect on what has happened in their system network. These artefacts can serve an important role as mediators of awareness of cybersecurity within the organisation.

The **system administrators** desired the following artefacts:

- Push notifications (trouble ticket) through mobile device (or email) in case of attack

- Real-time information (on the dashboard) about the **number** of critical warnings, average **time** of resolutions, and a **classification** and **count** of attacks

These desired artefacts might signify that system administrators do not expect to regularly monitor the CS-AWARE console. Instead, they expect to be alerted only when this is necessary, and furthermore to always have an overview of attacks available. This is related to the complexity of the situation in Rome, where people with different roles and expertise are dealing with various technology-related activities. The system administrator role is therefore not a fixed assignment for a small number of designated individuals.

From the **manager's** perspective, the following artefacts (to be sent by email) were envisioned:

- Weekly incident reports (severity, priority, root causes, solution, time frame)
- Monthly trend reports

From the **internal** service user's perspective, the following artefacts were suggested:

- The possibility for the SUET platform to share and devlop security guidelines for the prevention of attacks
- Notifications by SUET about the nature of the problem, expected time for solution, and suggestion for workaround and help desk

For the **external user,** in this case the citizen's perspective, being notified about issues *before* they are solved was considered *un*desirable.

These suggestions about desirable artefacts will be taken up in the deployment team meetings, in order to decide and prepare the artefacts that will be used in the second and third deployment cycle.

4. Desired Behaviour

Desired behaviour is the behaviour of users that would be beneficial for achieving their objectives. Formulation of this behaviour would enable monitoring and develop awareness of what is desirable during deployment.

The **system administrators** identified the following behaviour:

- Use of the tool on a daily basis (a system administrator will be assigned this role)
- Interpretation of the history of past problems and solutions
- Regular communication with the rest of the technical team (including service providers) and internal users

The **manager** should engage in the following activities:

- All managers should improve their knowledge about security (training)
- Define formal objectives for security improvement (MBO)
- Proactive approach towards internal users, and to senior management

The **internal service user** is supposed to:

- Follow the guidelines and acquire information about an issue
- Apply this in communication with final users

These activities will be monitored for evaluation. They are excellent manifestations of growing awareness of cybersecurity and the transformation of rules and norms and changing working practices within and between departments.

5. Participants

To be discussed in the deployment team:

- The names of the users of the CS-AWARE tool
- The names of the managers and service users who participate in further feedback collection

6. Impact

Impact of deployment is what do we expect to change when deployment objectives have been realised. Impact can include several cycles of scenarios, not just one, in other words, a long-term perspective. Thinking about desired impact allows us to understand 'how far we are' – what can be realistically achieved. It may be required to implement additional activities to realise the desired impact.
The local stakeholders in the deployment team expect the following impact (H: High, M: Medium, L: Low):

	Sysadmin	Manager	Internal user	External User
Numbers	H	M	L	L
Awareness	H	H	M	M
Transparency	H	H	H	L
Roles and Responsibilities	H	H	M	M
Policies	H	H	H	L
Money	H	H	H	H

The categories in the first column refer to:

- Numbers: what is the impact of more attacks detected, and faster repair?
- Awareness: what could be the impact of more awareness of cybercrime?
- Transparency: what is the impact on transparency of services?
- Roles and Responsibilities: what is the impact of changing roles and responsibilities of system administrators and of the managers?
- Policies: what is the impact of greater cybersecurity awareness on policies?
- Money: what is the impact on saving money by greater efficiency of services?

Most impact is expected internally, that is at the level of system administration and management.
These perspectives will be taken up as benchmarks for evaluation at the organisational level.

7. Evaluation

Evaluation will be discussed elsewhere.

8. Support

Support during the first cycle of deployment for Rome will be technical, as the CS-AWARE team is responsible for implementation of the CS-AWARE system within the local context.

Furthermore, the users will receive a short manual of the operations of the console of the system, and their feedback will be discussed within the deployment team.
Formalising the support will be discussed in the deployment team.

9. Preparation and Planning

Table 3 shows the planning for deployment. Please note that planning for cycle 2 and 3 may change as a result of ongoing activity and feedback.
After the deployment team has been established in October 2019, they will have online meetings every two weeks, for discussing current issues, planning and feedback. System implementation will continue until end-November. After implementation, formal usability and system testing will take place. Also, we will collect feedback on using the system on a biweekly basis. This may lead to technical work in the periods that follow. The nature of the feedback and the technical work will be documented and reported.
Concerning the local objectives and expectations (item 2 of the deployment scenario), we will design a template for on-going feedback by the users (system administrators) to be completed biweekly, and discussed within the deployment team. This feedback will also include using the artefacts, as provided by the CS-aware team, and further activities required in the local team for reporting and communication. All feedback will be reported, including the subsequent activity.
Concerning evaluation at the levels 3-5, we will engage in planning and preparing materials and procedures during November and December 2019.

2 Annex 2: CS-Aware Requirements, KPI's and questionnaires

On reviewer request, a consolidated list of CS-AWARE requirements is added as an Annex to this deliverable. This list represents an updated version of the tables presented in Section 1 of deliverable D2.2 and follows the three categories of requirements defined in D2.2: the requirements for selection of external information sources (Section 1.1 of D2.2), the requirements for LPA specific analysis and information sources (Section 1.2 of D2.2) and the CS-AWARE system requirements (Section 1.3 of D2.2). A detailed description of the context of those requirements as well as the methodology used for requirements selection and the selection procedures can be found in the respective Sections of D2.2. The requirements I8 and I9 have been added at a later stage and are detailed in Section 2.1 of this document. Also, we have added the requirements for evaluation. All those requirements are linked to the main requirements, numbered S1-S14.

2.1 Requirements for External Information Sources

#	Requirement	Functional	Non-functional	End user viewpoint
E1	Overview of all potential data sources that can be collected to enhance cybersecurity in the context of a dynamic environment that requires constant re-evaluation and integration of new information		X	
E2	Information sources that can be collected dynamically		X	
E3	Information sources that facilitate collaborative and cooperative cybersecurity		X	
E4	Information sources that are community driven (non-commercial operators)		X	
E5	Information sources that follow common cybersecurity exchange standards		X	
E6	Information sources that are especially applicable to LPA cybersecurity context, based on initial risk analysis performed in CS-AWARE deliverable D2.1		X	
E7	Periodisation and selection based on these Requirements	X		

2.2 LPA specific analysis requirements

#	Requirement	Functional	Non-functional	End user viewpoint
I1	Identification of the critical assets (socio-technical) in the LPA system	X		X
I2	Identification of the critical dependencies between assets in the LPA system	X		X
I3	Identification of critical LPA services and service processes	X		X
I4	Identification of information flows of critical service processes through assets and dependencies	X		X
I5	Identification of monitoring points able to determine cybersecurity state related to information flows of critical service processes	X		X
I6	Improve the system understanding of LPA personal		X	X
I7	Increase cybersecurity awareness for LPA personal with respect to LPA systems		X	X
I8	Determine normal and abnormal behaviour related to critical service processes and information flows	X		X
I9	Interface analysis results with CS-AWARE technology solution for continuous monitoring, awareness and self-healing	X		

2.3 CS-AWARE system requirements

#	Requirement	Functional	Non-functional	End user viewpoint
Technical System Requirements				
S1	**Provide cybersecurity awareness**	X		X
S2	**Allow information sharing**	X		X
S3	**Enable system self-healing**	X		X
S4	Enable data collection from internal LPA and external cyber security information sources	X		
S5	Allow preprocessing to bring data into an unified format	X		
S6	Enable data analysis by setting external and internal data into context	X		
S7	Ensure international usability of the system by providing multiple languages	X		
S8	Identifying relevant internal and external sources	X		
General System Requirements				
S9	Usability (The usability of the CS-AWARE system, as determined by the end users)		X	X
S10	Compliance (Compliance to LPA regulations, policies and procedures)		X	X
S11	Integratability (Integratability of CS-AWARE system into LPA work flows)		X	
S12	Open Source (How much of the CS-AWARE components can be open sourced and how much is kept proprietary)		X	
S13	Internationalization (Integration into different cultural and language contexts)		X	
S14	Cost/Marketability of CS-AWARE solution		X	

2.4 Requirements for Evaluation (by users)

2.4.1 Evaluation Level 1: Technical

Evaluation at this level addresses the extent to which users assess the technology of CS-AWARE as sound (i.e.: it works as it is supposed to).

Method: A short questionnaire is constructed. The questionnaire will be answered at the pilot sites at the end of March, and another time at the end of deployment in June. The questionnaire items allow users to indicate the degree to which they agree with a statement (Likert Scale 1-5). We have set a baseline at KPI for every question at 60% appreciation, as this reflects the accepted level of satisfaction. We expect the score to remain stable or to go up over time. Example question: I like the way the information is presented on the opening screen

Users: All members of the deployment teams at the pilot sites: sysadmin, service providers, managers, and service users. We foresee 10 users in Larissa, and 20 in Rome.

Technical (addresses requirements S1, S2, S3, S9, S10		
CS-AWARE system requirements	**Items Questionnaire 1**	**KPI**
S1: Provide cybersecurity awareness	1. User satisfaction with how information is presented in dartboard overview 2. User satisfaction with how information is presented in top threats and threats overview 3. User satisfaction with how information is presented in detailed threat view 4. User satisfaction of how information is presented in system overview	% Sum of ratings/20
S2: Allow information sharing	5. User satisfaction with how information is presented in information sharing overview 6. User satisfaction with how information is presented in detailed information sharing description 7. Clarity of technical implementation information sharing process	% Sum of ratings/15
S3: Enable system self-healing	8. User satisfaction with self-healing information provided via threat overview 9. Clarity of technical implementation of self healing process	% Sum of ratings/10
S9: Usability (The usability of the CS-AWARE system, as determined by the end users)	10. User satisfaction with usability of the interface, menus and options	% Sum of ratings/5
S10: Compliance (Compliance to LPA regulations, policies and procedures)	11. Estimated level of compliance of CS-AWARE technology with LPA regulations, policies and procedures	% Sum of ratings/5

Other technical requirement evaluation, not based on questionnaires

Some CS-AWARE system requirements (in particular the functional requirements, except S9-S11) are also evaluated based on technical criteria, that do not involve users. The functional requirements mentioned are evaluated by requiring passing the functional tests defined and reported in CS-AWARE deliverable D4.3.

Technical (addresses requirements S1, S2, S3, S4, S5, S6, S7, S8, and S12)		
CS-AWARE system requirements	**Requirements**	**KPI**
S1: Provide cybersecurity awareness	System is able to present selected information	Passes functional tests T6.X Passes functional tests TI.X
S2: Allow information sharing	System is able to share information with external parties	Passes functional tests T7.X
S3: Enable system self-healing	System is able to provide self-healing for selected cases	Passes functional tests T8.X
S4 Enable data collection from internal LPA and external cyber security information sources	System is able to collect data from selected information sources	Passes functional tests T2.X
S5 Allow preprocessing to bring data into an unified format	System is able to pre-process data to the defined unified format	Passes functional tests T3.X
S6 Enable data analysis by setting external and internal data into context	System is able to set information in context	Passes functional tests T4.X
S7 Ensure international usability of the system by providing multiple languages	System is able to translate information to selected languages	Passes functional tests T5.X
S8 Identifying relevant internal and external sources	Select ten most relevant external information sources	Information source scores is <= 10 according to scoring system defined in D2.2 (Section 4)
S8 Identifying relevant internal and external sources	Identify the system and dependency information (assets, dependencies, processes information flows, log files, monitoring patterns) in each analysed organization.	Consent reached in SSM workshops within the analysed organization
S12: Open Source	Ensure open-source availability of the core CS-AWARE system	60% or more of CS-AWARE system code base is open source

QUESTIONNAIRE 1 Satisfaction with the CS-AWARE system Please indicate your agreement to the statements below by ticking the appropriate box	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Don't Know/ Not applicable
1 I am happy with how information is presented in dartboard overview	☐	☐	☐	☐	☐	☐
2 I am unhappy with how information is presented in top threats and threats overview	☐	☐	☐	☐	☐	☐
3 I like how information is presented in detailed threat view	☐	☐	☐	☐	☐	☐
4 I do not like how information is presented in system overview	☐	☐	☐	☐	☐	☐
5 I am happy with how information is presented in information sharing overview	☐	☐	☐	☐	☐	☐
6 I am happy with how information is presented in detailed information sharing description	☐	☐	☐	☐	☐	☐
7 I think the technical implementation information sharing process is clear	☐	☐	☐	☐	☐	☐
8 I am unhappy with the self-healing information that is provided via threat overview	☐	☐	☐	☐	☐	☐
9 The technical implementation of the information sharing process is unclear	☐	☐	☐	☐	☐	☐
10 I like the usability of the interface, menus and options	☐	☐	☐	☐	☐	☐
11 I think the CS-AWARE technology complies with our regulations, policies and procedures	☐	☐	☐	☐	☐	☐
12 My role in my organisation (tick more if applicable)	System Administrator ☐	Manager ☐	Service Provider ☐	Internal User ☐	Other ☐	Citizen ☐
13 My experience with CS-AWARE (tick more if applicable)	Workshops ☐	Demo ☐	Active use (more than 10 hours) ☐	Active use (between 2-10 hours) ☐	Incidental use (less than 2hours) ☐	None ☐

☐

2.4.2 Evaluation Level 2: Usability

Here, we address the usability of the interface. Users are the intended users of the CS-AWARE system, 3 in Larissa. For Rome, the situation is more complicated, expertise is more distributed, and resolving a threat involves several participants.
Method: Usability is evaluated by hands-on usability tests (users actually working with the system, thinking aloud) **as well as** by post-hoc questionnaire on usability. The questionnaire for usability has a Likert scale, and is always administered after a session where the users are requested to resolve several exercises involving the resolution of threats. We have set a baseline at KPI for every question at 60% appreciation. We expect the score to go up over time in the pilots, but with new additions to the system, this may be different. We intend to have one usability session in cycle 2, and two more in cycle 3, and a final session at the end of deployment.

The main **requirements** for usability that we test are **Knowability** (user can understand, learn and remember how to use the system) and **Operability** (the system provides the user with the necessary functionalities). Both are split up into more detailed requirements[1].

Usability (addresses requirements S9 and S13)			
CS-AWARE system requirements	**Usability Requirements**	**Items Questionnaire 2**	**KPI**
S9: Usability	**U1: Knowability - Clarity** (elements are clear, structure is clear, function is clear)	1. The elements (options, colours, categories) that were used on the different screens were clearly presented 2. The different elements on the screens were easy to find 3. It was clear to me what the functions of the elements were	% Sum of ratings/15
S9: Usability	**U2: Knowability - Consistency** (elements are consistent, structure is consistent, function is consistent)	4. There was no contradictory information on different screens 5. The different screens have a consistent structure 6. The same elements on different screens always mean the same thing	% Sum of ratings/15
S9: Usability	**U3: Knowability-Memorability** (elements are remembered, structure is remembered, function is remembered)	7. The information presented about the threats is easy to remember 8. The information presented about the threats is well structured 9. I often go back to a previous screen to check some information	% Sum of ratings/15
S9: Usability	**U4: Knowability - Helpfulness** (documentation, assistance)	10. I want more help in using the interface	%Score/5

[1] Alonso-Ríos, D., Vázquez-García, A., Mosqueira-Rey, E., & Moret-Bonillo, V. (2009). Usability: A Critical Analysis and a Taxonomy. *International Journal of Human-Computer Interaction*, *26*(1), 53–74. https://doi.org/10.1080/10447310903025552

S9: Usability	**U5: Operability-Completeness** (Information is complete, User knows what to do)	11. The information that the system provides is complete 12. The information presented about the threats helps me in deciding what to do 13. The information about that the system provides is reliable	% Sum of ratings/15
S9: Usability	**U6: Operability-Precision**	14. I have enough possibilities to find out what I need for understanding a threat 15. I have enough possibilities to find out what I need for resolving a threat 16. I would like to have more information about threats	% Sum of ratings/15
S9: Usability and S13: Internationalisation	**U7: Operability-Universality** (language)	17. The information about threats is clearly written	%Score/5
S9: Usability	**U8: Operability-Flexibility** (Controllability, Adaptiveness)	18. I would like to have more options for acting	%Score/5

QUESTIONNAIRE 2

Usability of the CS-AWARE system

Please indicate your agreement to the statements below by ticking the appropriate box

		Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Clarity						
1	The elements (options, colours, categories) that were used on the different screens were clearly presented	☐	☐	☐	☐	☐
2	The different elements on the screens were easy to find	☐	☐	☐	☐	☐
3	It was clear to me what the functions of the elements were	☐	☐	☐	☐	☐
Consistency						
4	There was no contradictory information on different screens	☐	☐	☐	☐	☐
5	The different screens have a consistent structure	☐	☐	☐	☐	☐
6	The same elements on different screens always mean the same thing	☐	☐	☐	☐	☐
Memorability						
7	The information presented about the threats is easy to remember	☐	☐	☐	☐	☐
8	The information presented about the threats is well structured	☐	☐	☐	☐	☐
9	I often go back to a previous screen to check some information	☐	☐	☐	☐	☐
Helpfulness						
10	I need more help in using the interface	☐	☐	☐	☐	☐
Completeness						
11	The information that the system provides is complete	☐	☐	☐	☐	☐
12	The information presented about the threats helps me in deciding what to do	☐	☐	☐	☐	☐
13	The information about that the system provides is reliable	☐	☐	☐	☐	☐
Precision						
14	The system provides good information to understand a threat	☐	☐	☐	☐	☐
15	The system allows me to do what is needed to resolve a threat	☐	☐	☐	☐	☐
16	There is not enough information about threats	☐	☐	☐	☐	☐
Universality and Flexibility						
17	The language used on the various screens is very clear and helpful	☐	☐	☐	☐	☐
18	This system can only be used by experts who know their system	☐	☐	☐	☐	☐

		System Administrator	Manager	Service Provider	Internal User	Citizen	Other
0	My role in my organisation (tick more if applicable)	☐	☐	☐	☐	☐	☐

		Workshops	Demo	Active use (more than 10 hours)	Active use (between 2-10 hours)	Incidental use (less than 2hours)	Discussion
0	My experience with CS-AWARE (tick more if applicable)	☐	☐	☐	☐	☐	☐

2.4.3 Evaluation Level 3: Awareness

Here, we address the awareness outcomes after using the system. Users are the intended users of the CS-AWARE system, 3 in Larissa. For Rome, the situation is more complicated, expertise is more distributed, and resolving a threat involves several participants.

Method: The questionnaire for awareness is always administered after the usability session where the users are requested to resolve several exercises involving the resolution of threats. The questionnaire also exploits the benefits of a Likert scale. We have set a baseline at KPI for every question at 60% appreciation. We expect the score to go up over time in the pilots, but with new additions to the system, this may be different. We intend to have one usability session in cycle 2, and two more in cycle 3, and a final session at the end of deployment.

The **specific requirements for awareness** that we test are: **Perception** (user perceives a threat), **Comprehension** (User understands the cures and explanations), **Projection** (user foresees consequences of actions), **Decision** (User makes a decision)[2]. To those, we add awareness of sharing and self-healing. The requirements for awareness link to the general requirements for cybersecurity awareness (S1), sharing (S2), and self-healing (S3).

Awareness (addresses requirements S1, S2, S3)			
CS-AWARE system requirements	**Awareness Requirements**	**Items Questionnaire 3**	**KPI (> 60%)**
S1 Provide cybersecurity awareness	**A1: Perception** (what elements did the user look at?)	(Dartboard screen) 1. The user perceives a threat immediately 2. The user perceives the possible impact of a threat 3. The user visits all elements of the table	% Sum of ratings/15
S1 Provide cybersecurity awareness	**A2: Comprehension** (interpretation of cues, explanations)	(Threats table) 4. The user understands the state of a threat 5. The user understands when a threat was first observed 6. The user understands to whom the threat was assigned 7. The user understands the group of the threat 8. The user understands the location in the network of the threat 9. The user understands the name of the threat (Threat Description) 10. The user reads the description of the threat 11. The user understands the description of the treat	% Sum of ratings/60
S1 Provide cybersecurity awareness	**A3: Projection** (foreseen consequences of interpretation)	(System Visualisation) 12. The user looks at the system visualisation	% Sum of ratings/15

[2] Hibshi, H., Breaux, T. D., Riaz, M., & Williams, L. (2016). A grounded analysis of experts' decision-making during security assessments. *Journal of Cybersecurity*, *2*(2), 147–163. https://doi.org/10.1093/cybsec/tyw010

		13. The user reads the information about the component 14. The user can understand what node of the system is affected 15. The user can understand what other nodes are in danger ('Threats' table) 16. The user can understand what services are in danger 17. The user knows when to look for more information (History table) 18. The user knows when to inspect the history table	
S1 Provide cybersecurity awareness	**A4: Decision** (What is the best thing to do?)	('Threat action' screen) 19. The user (considers) communicating with another team 20. The user (considers) communicating with colleagues 21. The user (considers) communicating with the manager 22. The user knows how to change a state 23. The user knows how to describe his activity 24. The user knows if a threat is resolved (Resolved threats screen) 25. The user closes a threat 26. The user verifies that a threat is resolved	% Sum of ratings/40
S1 Provide cybersecurity awareness **And** **S2 Allow information sharing**	**A5: Sharing** **(**What is shared?)	('Sharing' screen) 27. The user indicates a need for sharing with cybersecurity authorities 28. The user indicates understanding about sharing with cybersecurity authorities 29. The user understands the policies for sharing with cybersecurity authorities 30. The user considers what to share with cybersecurity authorities	% Sum of ratings/20
S1 Provide cybersecurity awareness **And** **S3 Enable system self-healing**	**A6: Self-Healing** (Is self-healing applied?)	('Self-healing needs decision' action screen) 31. The user looks for the self-healing option 32. The user understands if the self-healing option is available or not 33. The user understands the description of the self-healing option 34. The user understands the consequences of applying self-healing 35. The user makes a decision about self-healing 36. The user applies a self-healing option	% Sum of ratings/30

Users need experience with the CS-AWARE system, meaning, having worked with it, before they can answer this questionnaire.

Qualitative approach to awareness

Awareness is an important aspect of our qualitative approach to deployment. The deployment scenarios that were constructed at both sites in cycle 1, will be revisited at the end of cycle 2

and at the end of the project. The approach is also relevant for the organisational level 4. Some outcomes of the scenario have been integrated in the various questionnaires. This will be described in more detail in D5.1.
The **baseline** for awareness will be constructed by interpreting the outcomes from user stories and deployment scenarios workshops in November 2019. For example, for **perception**, we know that the users only perceive the threat when an employee comes to them with an issue (not immediate), they have to research what caused the issue (no immediate diagnosis) and what kind of threat is causing the issue. This will give no points for immediate perception. For comprehension, some baseline understanding may be assumed, once the threat has been diagnosed, but not all details will be known. We expect a score of 1 point, maximum. For projection, the measures will be adequate, but probably too rigorous, even at the level of not allowing certain applications to be used. For decision, some interactions and communication has been implemented, but not always, and not in every case. Similar conjectures apply to sharing and self-healing. We will look into this more precisely, and we expect differences between the two municipalities, for this base-level.

QUESTIONNAIRE 3 My awareness from using the CS-AWARE system Please indicate your agreement to the statements below by ticking the appropriate box		Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Don't Know/ Not applicable
Perception: the opening screen							
1	I was warned about a threat immediately	☐	☐	☐	☐	☐	☐
2	The possible impact of a threat was almost immediately clear to me	☐	☐	☐	☐	☐	☐
3	I inspected all elements of the table of threats on the opening screen before moving on	☐	☐	☐	☐	☐	☐
Comprehension: the threats table							
4	I could see how dangerous a threat is	☐	☐	☐	☐	☐	☐
5	I know when the threat was first observed by the system	☐	☐	☐	☐	☐	☐
6	I have understood if a threat has been assigned to someone else	☐	☐	☐	☐	☐	☐
7	I have found what group the threat belongs to	☐	☐	☐	☐	☐	☐
8	I have found what location in the network is in danger by the threat	☐	☐	☐	☐	☐	☐
9	The system tells me the name of the threat	☐	☐	☐	☐	☐	☐
Comprehension: Threat Description							
10	I have read the long description of the threat	☐	☐	☐	☐	☐	☐
11	I have undertood the information about the threat	☐	☐	☐	☐	☐	☐
Comprehension: Network Visualisation							
12	I have studied the local network visualisation	☐	☐	☐	☐	☐	☐
13	I have understood what node of the network was affected by the threat	☐	☐	☐	☐	☐	☐
14	I have read the information about the component affected by the threat	☐	☐	☐	☐	☐	☐

15	I have undersood what other nodes were in danger	☐	☐	☐	☐	☐	☐
Projection: Threats Table							
16	I have understood what municipality services were in danger	☐	☐	☐	☐	☐	☐
17	I tried to find information about how to resolve the threat (on the CS-AWARE system or elsewhere)	☐	☐	☐	☐	☐	☐
18	I have explored the table with the history of mitigation actions taken so far	☐	☐	☐	☐	☐	☐
Decision: The Threat action screen							
19	I contacted my colleagues of my department	☐	☐	☐	☐	☐	☐
20	I contacted colleagues at another department or service team	☐	☐	☐	☐	☐	☐
21	I contacted the manager of my department	☐	☐	☐	☐	☐	☐
22	I changed the state of the threat	☐	☐	☐	☐	☐	☐
23	I described what I did before changing the state	☐	☐	☐	☐	☐	☐
24	I understood that the threat was resolved	☐	☐	☐	☐	☐	☐
25	I have closed the threat	☐	☐	☐	☐	☐	☐
26	I verified that the threat was resolved	☐	☐	☐	☐	☐	☐
Sharing							
27	I thought about what to share about the threat with cybersecurity authorities	☐	☐	☐	☐	☐	☐
28	I understand why information about the threat should be shared with cybersecurity authorities	☐	☐	☐	☐	☐	☐
29	I am aware of our policies on sharing information about threats	☐	☐	☐	☐	☐	☐
30	Crucial information about the threat was shared with cybersecurity authorities	☐	☐	☐	☐	☐	☐

Self-Healing

31	I looked for a self-healing option	☐	☐	☐	☐	☐	☐
32	I know where to find the self-healing option	☐	☐	☐	☐	☐	☐
33	I understood the description of the self-healing option	☐	☐	☐	☐	☐	☐
34	I understood the possible impact of self-healing	☐	☐	☐	☐	☐	☐
35	I made a decision about applying or not applying self-healing	☐	☐	☐	☐	☐	☐
36	I succesfully applied self-healing	☐	☐	☐	☐	☐	☐

		System Administrator	Manager	Service Provider	Internal User	Other	Citizen
37	My role in my organisation (tick more if applicable)	☐	☐	☐	☐	☐	☐

		Workshops	Demo	Active use (more than 10 hours)	Active use (between 2-10 hours)	Incidental use (less than 2hours)	None
38	My experience with CS-AWARE (tick more if applicable)	☐	☐	☐	☐	☐	☐

2.4.4 Evaluation Level 4: Organisation

To what extent is the system used in the organisation, and are the managers aware of this use, and its impact? Foreseen participants are the managers within the organisation, for the system department, but also those of the departments involved in the pilot scenarios that have been selected for CS-AWARE. In addition, their managers will be addressed as well. Total number of participants: for Larissa: currently unknown; for Rome: 12.

Method:
We have developed a short questionnaire, to be administered at the end of March, and at the end of deployment. This allows us to monitor developments in organisational awareness. Baseline KPI will be the first responses that will be provided. We expect scores to go up over time. The requirements for organisational awareness link to the general requirement of the CS-AWARE system: to provide cybersecurity awareness (S1), as well as to the system requirements set for compliance (S10) and integratability (S11). We distinguish awareness of impact on the municipality network, on services for citizens, and on organisational culture.[3]

Organisational Awareness (addresses requirements S1, S10, S11)			
CS-AWARE system requirements	**Organisational Awareness Requirements**	**Items Questionnaire 4**	**KPI's (>.6)**
S10 Compliance and S11 Integratability	**O1 Impact on system**	-Securing citizens' personal information and data -Securing our organisation's sensitive information and data -Complying with GDPR -Systems are more secure and resilient	% Sum of ratings/20
S1 Provide cybersecurity awareness and S11 Integratability	**O2 Impact on Citizens**	More effective delivery of services	%Score/5
S1 Provide cybersecurity awareness and S10 Compliance and S11 Integratability	**O3 Impact on Culture**	-Senior management is more mindful about security issues and their system -Positive impact on general organisational culture -Positive Impact on security awareness	% Sum of ratings/15

[3] Anttila, J., & Knowledgist, V. (2006). General Managerial Tools for business-integrated information security management. Paper, paper at IPICS Winter School of the University of Oulu. http://www.qualityintegration.biz/InformSecPDCA.html

QUESTIONNAIRE 4 Effects the CS-AWARE system at the business level of the organisation Please indicate your agreement to the statements below by ticking the appropriate box		Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Don't Know/ Not applicable
1	I think that CS-Aware has enabled my organisation to further demonstrate due diligence in relation to securing our citizens' personal information and data	☐	☐	☐	☐	☐	☐
2	I think that CS-Aware has enabled my organisation to demonstrate due diligence in relation to us securing our organisation's sensitive information and data	☐	☐	☐	☐	☐	☐
3	I think that CS-Aware has assisted my organisation to comply with the reponsibilities placed upon us by the GDPR	☐	☐	☐	☐	☐	☐
4	I think that CS-Aware has made our systems more secure and resilient	☐	☐	☐	☐	☐	☐
5	I don't think that CS-Aware has affected my organisation's ability to comply with the reponsibilities placed upon us by the GDPR	☐	☐	☐	☐	☐	☐
6	I think that CS-Aware has enabled our organisation to improve the delivery of services to our citizens because our systems are now more secure and resilient	☐	☐	☐	☐	☐	☐
7	I don't think that CS-Aware has enabled our organisation to improve the delivery of services to our citizens	☐	☐	☐	☐	☐	☐
8	I don't think that CS-Aware has made the senior management of my organisation any more mindful of systems and data security issues than they were before we used CS-Aware	☐	☐	☐	☐	☐	☐
9	I think that CS-Aware has made the senior management of my organisation more mindful of systems and data security issues	☐	☐	☐	☐	☐	☐
10	I don't think that CS-Aware will affect the organisational culture in my enterprise	☐	☐	☐	☐	☐	☐
11	I think that CS-Aware will make a positive contribution towards improving organisational culture relating to security in my organisation	☐	☐	☐	☐	☐	☐
12	My role in my organisation (tick more if applicable)	System Administrator ☐	Manager ☐	Service Provider ☐	Internal User ☐	Other ☐	Citizen ☐
13	My experience with CS-AWARE (tick more if applicable)	Workshops ☐	Demo ☐	Active use (more than 10 hours) ☐	Active use (between 2-10 hours) ☐	Incidental use (less than 2hours) ☐	None ☐

2.4.5 Evaluation Level 5: Business

Here, we address the interested users and potential early adopters. These are either participants to workshops organised by Cesviter and OTS, or new users that we send the questionnaire. We are interested in decision makers, mayors and managers of departments. From these people, we want to know the degree to which the CS-AWARE system addresses their cybersecurity needs and organisational structure.
Three main business requirements are addressed:
1. Potential for organisational Involvement with cybersecurity: Policy ownership (It is important if someone in the top policy group (that is, the city council) is specifically charged with overseeing cybersecurity), operational oversight (Someone with authority on the administrative staff should be charged with managing cybersecurity activities). **2. Potential for increasing cybersecurity awareness**: Degree of involvement in the municipality for cybersecurity issues - is it talked about? Is awareness an issue? **3. Potential for administrative and policy interest in a system that provides awareness, self-healing and sharing of information**. These business requirements link to the system requirements of Marketability (S14) and Integratability (S11) and to the general awareness requirement (S1).

Business (addresses requirements S1, S11 and S14)			
CS-AWARE system requirements	**Business level Requirements**	**Items Questionnaire 5**	**KPI's (>.6)**
S14 Marketability and S11 Integratability	**B1: Potential for organisational involvement**	-Ownership of cybersecurity operations -Oversight of cybersecurity operations -Cybersecurity is discussed	% Sum of ratings/15
S1 Awareness	**B2: Potential for increasing Awareness**	-Importance of cybersecurity awareness -Cybersecurity only is a management issue -News of incidents has increased awareness -Cybersecurity is recurrent issue at meetings	% Sum of ratings/20
S14 Marketability	**B3: Potential interest for CS-AWARE system**	-Self-healing is a desirable technical asset -Self-healing is important policy - CS-information sharing is a desirable technical asset -CS- information sharing is important policy - CS-awareness is a desirable asset -CS-awareness is an important policy -Open Source is an important system requirement -CS-AWARE will all these features is particularly interesting	% Sum of ratings/40

QUESTIONNAIRE for Level 5 : Effect of the CS-AWARE system on the management levels of the municipality Please complete the questions below by ticking the appropriate box		Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Don't know/ Not applicable
1	In my municipality, one individual is charged with overall responsibility for ensuring the smooth running of our municipality's Cybersecurity operations	☐	☐	☐	☐	☐	☐
2	In my municipality, an elected member of the City Council is has responsibility for the oversight of Cybersecurity operations	☐	☐	☐	☐	☐	☐
3	The topic of Cybersecurity is rarely discussed at the management level of my municipality.	☐	☐	☐	☐	☐	☐
4	In the last 12 months, the rate of growth of Cybersecurity awareness has increased at both the administrative and policy-making levels in my municipality	☐	☐	☐	☐	☐	☐
5	In my municipality, Cybersecurity is only discussed at a management level when a Cybersecurity issue has arisen	☐	☐	☐	☐	☐	☐
6	News of incidents has enhanced the Cybersecurity culture within the administrative and policy areas of my municipality	☐	☐	☐	☐	☐	☐
7	Cybersecurity is a permanent item on the agenda of meetings at the policy level In my municipality.	☐	☐	☐	☐	☐	☐
8	A Cybersecurity system incorporating a self-healing facility, would be of particular interest to the management team at the administrative level of my municipality	☐	☐	☐	☐	☐	☐
9	A Cybersecurity system incorporating a self-healing facility, would be of particular interest to the management team at the policy level of my municipality	☐	☐	☐	☐	☐	☐
10	A Cybersecurity system incorporating an information-sharing facility, would be of particular interest to the management team at the administrative level of my municipality	☐	☐	☐	☐	☐	☐
11	A Cybersecurity system incorporating a Cybersecurity - awareness facility, would be of particular interest to the management team at the administrative level of my municipality	☐	☐	☐	☐	☐	☐
12	A Cybersecurity system incorporating an information-sharing facility, would be of particular interest to the management team at the policy level of my municipality	☐	☐	☐	☐	☐	☐
13	A Cybersecurity system incorporating a Cybersecurity - awareness facility, would be of particular interest to the management team at the policy level of my municipality	☐	☐	☐	☐	☐	☐
14	Considering the variety of advantages and flexibility offered by open source systems, a cybersecurity system incorporating open source elements would be of particular interest to the management team at the administrative and policy level of my municipality.	☐	☐	☐	☐	☐	☐
15	In general, a Cybersecurity system incorporating the elements described above of Cybersecurity - awareness, information sharing, and self-healing would be of particular interest to the management team at the policy level of my municipality.	☐	☐	☐	☐	☐	☐
16	My role in my organisation (tick more if applicable)	System Administrator ☐	Manager ☐	Service Provider ☐	Internal User ☐	Policy Maker ☐	Citizen ☐
17	My experience with CS-AWARE (tick more if applicable)	Workshops ☐	Demo ☐	Active use (more than 10 hours) ☐	Active use (between 2-10 hours) ☐	Incidental use (less than 2hours) ☐	None ☐

Selection of Commercial Contacts made by Cesviter Consulting in 2019 and 2020[1]

Region	Organization	Function	Areas	Consultation[2]
Umbria	Municipality of Urbino	Mayor's Office	Political staff	Once every 3 days in last month
Lazio	Municipality of Fondi	Mayor's Office	Political staff	Once every 3 days in last month
Lazio	Municipality of Tivoli	Mayor's Office	Political staff	Once every 3 days in last month
Lazio	Municipality of Cerveteri	Mayor's Office	Political staff	Sometimes in last month
Lazio	Municipality of Ladispoli	City Council	Technical Staff	
Lazio	Municipality of Civitavecchia	Mayor's Office	For smart city, innovation and digital city, European funding, etc.	
Sicilia	Municipality of Ribera	City Council	Technical Staff	Sometimes in last month
Sicilia	Municipality of Caltanisetta	City Council	Technical Staff	Sometimes in last month
Sicilia	Municipality of San Michele di Ganzaria	Mayor's Office		
Sicilia	Municipality of Catania	City Council	Sector covering Innovation and technical staff	
Abruzzo	Municipality of Assisi	Mayor's Office		
Abruzzo	Municipality of L'Aquila	City Council	Sector covering Innovation and technical staff	
Sardegna	Province of Cagliari	Provincial Offices		
Sardegna	Municipality of Cagliari	Mayor		

[1] Since the political and policy making areas are typically the most important areas, we have been explored the possibilities of the CS-AWARE project primarily with people in policy making positions.

[2] While there were other municipalities and agencies contacted in the North and in Sicily, we have included, for now, only those that seem the most promising. Unfortunately, we will have to wait until the end of the current emergency situation caused by the Chronovirus to be able to evaluate better the situation.

Sardegna	Municipality of Sant'Andrea Frius	Mayor		
Sardegna	Municipality of Quartu Sant'Elena	Mayor		
Sardegna	Municipality of Sassari	Mayor's Office e City Council	Sector covering Innovation and technical staff	
Sardegna	Municipality of Escalaplano	Mayor's Office and Technical staff		
Sardegna	Municipality of Alghero	City Council	Sector covering Innovation and technical staff	
Sardegna	Council of Local Autonomies of Sardinia	Provincial Council		
Sardegna	Municipality of Silius	Mayor		
Sardegna	Municipality of Collinas	City Council	Sector covering Innovation and technical staff	
Puglia	Municipality of Andria	Mayor		
Puglia	Municipality of Barletta	Mayor		
Puglia	Municipality of Trani	City Council	Sector covering Innovation and technical staff	
Puglia	Unione Comuni del Puglia Nord	Union Offices		
Puglia	Municipality of Spinazzola	Mayor		
Puglia	Chamber of Commerce of Province Barletta, Andria and Trani	Director		
Puglia	Artisan confederation of Barletta, Andria and Trani	Director		
Puglia	Engineers Order of Barletta, Andria and Trani	Director		
Puglia	Agency for the occupation and development of the Nord Barese Ofantino area	Director		

Puglia	Mediterranean Network Foggia.	Director		
Puglia	BAT (Province of Barletta-Andria-Trani)	Provincial Offices		
Sicily[3]	San Michele di Ganzaria	Mayor		
Sicily	Mirabella Imbacari	Mayor		
Sicily	Licodia Eubea	Mayor		
Sicily	Mazzarone San Cono	Mayor		
Sicily	San Cono	Mayor		

[3] Unfortunately, due to a lack of time and travel funds we were unable to stage a workshop and presentation for these and other municipalities in the Province of Catania. The ChronoVirus emergency made even a limited event in practicable.

3 Annex 3: Technical deployment tables

3.1 Updated system and dependency graph protocol

This Section shows a generic example of the system and dependency graph protocol format to illustrate the extensions done during the second round of piloting. The major modifications were the addition of two additional objects of the type "pattern" and "x_logfile", and the relevant parameters "logfile_ref" and "pattern_ref" in the asset object type "node" to be able to associate logfiles and patterns to specific assets. The "x_logfile" object is able to model information about logfiles and individual parameters. The "pattern" object can model patterns by defining pattern parameters that are based on one or more log file parameters, and specifiying the ranges for which those parameters are considered within or outside the norm that should trigger the pattern.


```
{
  "type": "region",
  "id": "region--bbb4faf4-5490-4b43-88b0-981e4161485e",
  "name": "Generic example",
  "created": "2019-07-17T11:48:47.856Z",
  "modified": "2019-08-01T11:17:09.783Z",
  "version": "1.0",
  "objects": [
    {
      "type": "node",
      "id": "node--0b14e229-b5ff-4fef-896a-ee8891c321c7",
      "name": "Asset 1",
      "created": "2019-07-24T12:21:04.000Z",
      "modified": "2019-07-24T12:21:04.000Z",
      "description": "Asset description",
      "source": [
        "node--333d7548-b35f-4e97-8b72-3b5bc1358934",
        "node--8141673a-5cd3-417f-a36d-3fb3b7fd8393",
        "node--8e8c8ae4-59a9-4cc3-b51d-6b479143fe9e",
        "node--d7eb5902-7dc2-4a33-ac20-1639de0305f5"
      ],
      "x_infoflow": [
      ],
      "x_cpe_list": [
        "cpe_<SoftwareName>:*:*:*:*:*:*:*:*:*:*:*:*:*"
      ],
      "pattern_ref": [
        "pattern--544946f9-bf08-472d-be2e-c51593267b47"
      ],
      "logfile_ref": [
        "x_logfile--8c207347-456c-4a6c-8bca-b340e27267d4"
      ],
```

```
    "x_ip_range": "0.0.0.0",
    "x_port_range": "0000",
    "x_categories": []
  },
  {
        "type": "pattern",
        "id": "pattern--544946f9-bf08-472d-be2e-c51593267b47",
        "name": "Pattern name",
        "pattern_type": "Behaviour monitoring",
        "created": "2020-04-17T11:48:47.856Z",
        "modified": "2020-04-17T11:48:47.856Z",
        "version": "1.0",
        "description": "Description of the pattern.",
        "x_paramlist": [
                {
                        "type": "x_parameter",
                        "id": "x_parameter--b2654f51-5800-4f81-8cf2-b496edbaa513",
                        "name": "Pattern Parameter 1",
                        "context": "",
                        "object_refs": [
                                {
                                        "x_logfile--8c207347-456c-4a6c-8bca-b340e27267d4":
[
                                                "x_parameter--b2815614-f637-425c-affa-
c8d0d223bd84",
                                                "x_parameter--869177e9-29f8-49d5-a4f6-
e640cd983d1e"
                                        ]
                                }
                        ],
                        "x_equation": "<Text>",
                        "resource_level": "<Number>",
                        "x_range_normal": "<Text>",
                        "x_max_range": "<Text>"
                }
        ]
  },
  {
        "type": "x_logfile",
        "id": "x_logfile--8c207347-456c-4a6c-8bca-b340e27267d4",
        "version": "<xml>",
        "description": "General description of the log file",
        "name":"Log file name",
        "value": "<Text>",
        "x_paramlist": [
                {
                        "type": "x_parameter",
```

```
                "name": "parameter1",
                "description": "Parameter description.",
                "id": "x_parameter--b2815614-f637-425c-affa-c8d0d223bd84"
            },
            {
                "type": "x_parameter",
                "name": "Parameter 2",
                "description": "Parameter description",
                "id": "x_parameter--869177e9-29f8-49d5-a4f6-e640cd983d1e"
            }
        ]
    }
  ]
}
```


3.2 Updated protocol for communication between information sharing and visualization

The protocol for exchanging information regarding InformationShare has been modified to allow for modification of data to share. The header section is more or less unmodified from previous version. A data section has been added to allow for sharing and modification of data.
Each chunk of data can be marked to be editable and deletable from InformationShare when send to Visualisation. In Visualisation the user can decide if the piece of information should the shared or mark is as “isDeleted”:”true” to not share the information. Likewise, the user can edit the information before sharing it. The piece of information is then returned in modified form and marked with “isChanged”:”true”.
InformationShare updates the information before sharing it with the cybersecurity community.

A sample record with just two pieces of data. The format is open to allow for any number of data pieces.


```
{
        "type": "InformationShare",
        "id": "InformationShare--d718e05d-9ced-4b15-bd19-69cfc4b7d969",
        "bundleID": "bundle--bc1855bf-1ffc-4057-a238-be2f2369d664",
        "state": "SharingAccepted",
        "summary": "bundle--bc1855bf-1ffc-4057-a238-be2f2369d664 - Too many deletes - Suspicious Database Modification Attempt. ...",
        "created": "2020-06-25T13:07:34.000Z",
        "modified": "2020-06-25T13:07:34.000Z",
        "data": [
            {
                "id": "attack-pattern--da242c6d-35b7-40f9-a031-ff325079e5e0",
                "name": "Too many deletes - Suspicious Database Modification Attempt",
                "editable": true,
                "deletable": true,
```

```
                "isChanged": false,
                "isDeleted": false,
                "information": "The database of the service was ..."
            },
            {
                "id": "-",
                "name": "timestamp",
                "editable": false,
                "deletable": true,
                "isChanged": false,
                "isDeleted": false,
                "information": "2019-02-21T00:00:08"
            }
        ]
}
```


3.3 Observations and resulting system changes from piloting cycle 1

#	View	Description	Enhancement initiated
1	All	Checkout button with name of logged in person - to be able to log out of authentication system: better security and better awareness	Larissa User Testing, Autumn 2019
2	Threat details	Improved User interface graphics in the Threat details window.	Internal testing, Summer 2019
3	Threat details	Observed data added to Thread details window - more information for the system administrator	Roma Capitale User testing, Late 2019
4	Threat details	Assignment/ticketing system with email- for bigger organisations like Roma Capitale	Roma Capitale User testing, Late 2019
5	Threat overview	Improved visualisation when no threats shown: better awareness	Larissa User Testing, Autumn 2019
6	Threat overview	More consistent colour scheme in threat overview: Better awareness /usability	Internal testing, summer 2019
7	Threat views	Instant translation of descriptions	Larissa User Testing, Autumn 2019
8	Threat Views	Added filtering by Threat Group, Assigned person and location	Internal testing, Late 2019
9	Threats	Adding the ID of threats to overview	Larissa User Testing, Autumn 2019
10	User management	Added live filtering of users	Roma Capitale User testing, Late 2019

11	System Overview	Added details of system components (e.g. CPE, IP-address, infoflow)	Internal testing, Summer 2019
12	Closed Threats	Added Closed Threats Excel export functionality for management reporting	Internal testing, September 2019
13	System Overview	Infoflow can now be used for single flow view, e.g. Finance, showing all nodes with infoflow “Finance"	Internal testing, Summer 2019
14	System Overview	Improved Zooming in System Overview	Larissa User Testing, Autumn 2019
15	System Overview	Search for nodes in System Overview	Internal testing, Late 2019
16	System Overview	Preparation for individual shapes for easier overview in System Overview	Larissa User Testing, Autumn 2019
17	Information Sharing	Enhanced Information Sharing with LPA-specific filtering, for removal before sharing	Internal testing, December 2019
18	About	System version in the About section	Internal testing, December 2019

3.4 Observations and resulting system changes from piloting cycle 2

When	View	Description	Necessary change	Implemented in cycle 2
Pilot testing round 2 Usability test Roma, 1 April	Information sharing	Information sharing concept is not understood adequately by users	Potential: Place introductory message on top of page	Message included in pop-over on InformationShare page.
Pilot testing round 2 Usability test Roma, 1 April	Threat detail	The individual log file parameters are not understood by the users	Primary cause: anonymisation	Use of non-anonymous data in on-premise Rome deployment
Pilot testing round 2 Usability test Roma, 1 April	Threat view	Link between threats in Threat overview and System Overview missing	Location in Threat Overview and Overview could link directly to System Overview	Links are added in threat tables where column.
Pilot testing round 2 Usability test Roma, 1 April	System Overview	System Overview zoomed in from start. - for some users this should be zoomed out	Zoom behaviour	Changed zoom behaviour
Pilot testing round 2 Usability test Roma, 1 April	System Overview	Performance sometimes lacking at startup and when multiple persons simultaneously logged in. No consistent measurable problem.	Code review	Code optimized for more users
Pilot testing round 2	Analysis, visualisation	Attribution of social media is not possible (which	This information is already available in the	**(SoMe)** Social media reports may need different

Usability test Roma, 1 April		source does it come from, which user/ group postet it)	STIX created by watcher/analysis, but is not utilized in the visual concept.	approach depending on user input and workflow Will not be done in this version of visualisation.
Pilot testing round 2 Usability test Roma, 1 April	System Overview	Filter/search functionality is not easily discovered by users. This is especially apparent in the system overview, which may require significant scrolling/zooming in order to reach specific assets. The already implemented search functionality can significantly reduce the time to access specific assets.	Make search functionality more easily discoverable in the visual concept.	Change in Zoom behaviour changes the need.
Pilot testing round 2 Usability test Roma, 1 April	Visualisation	It is not easily apparent (especially in the thread details view) which keyword triggered the social media message.	The visual concept does not include presenting of trigger keywords.	No change in this version of visualisation, as it would affect performance more than help usability.
Pilot testing round 2 Usability test Roma, 1 April	Closed threats	Better information in xls-export	Location should not be ID, but the name of such.	Node id mapped to node name.
Pilot testing	InformationShare	1) Link information	Threat ID (Sighting-ID)	1) ID is not easily identifiable. No

round 2 - Deployment team comments Larissa, March 12		sharing messages to threat message is not obvious 2) e.g. a reminder (and jump) to share information when closing threat may help better identifying	should be passed on as ID in info sharing message for identification in list of Threats. Different workflows hinders automatic reminder of sharing.	implementation expected. 2) Suggested change does not work for all workflows, hence not implemented
Pilot testing round 2 - Deployment team comments Larissa, March 12	Closed Threats	Nodes in threat excel report are not clear text	Node ID instead of clear text is exported by mistake - minor correction in code required	Node id mapped to node name – same as found by Rome. Fixed.
Pilot testing round 2 - Deployment team comments Larissa, March 12		An option for mass resolve of threats by type or by selection is useful, for example resolve of all “report” type threats.	Mass handling of threats needs to be implemented (low priority) Affected components: • Visualization	See **(SoMe)** Social media reports may need different approach depending on user input and workflow
Pilot testing round 2 - Deployment team comments Larissa, March 12		For “report” type threats there should be further options such as “read” or “informed” for state	A feed reader like management system (including sharing options) for threats needs to be implemented (low priority) Affected components: • Visualization	See **(SoMe)**
Pilot testing round 2 -		“reports” stated as “resolved” or “ignored” should	Handling of reports needs to be discussed. Still no	**(SoMe)** Social media reports may need different

Deployment team comments Larissa, March 12		be in a different tab, not in the “closed threats”	clear opinion if "reports" (aka. social media information) should be handled exactly like threats. (Low priority, since current visual concept - full integration, but with dedicated label "report" and colour "green" is the concept we fixed and no clear objection to the concept could be observed) Affected components: • Visualization	approach depending on user input and workflow
Pilot testing round 2 - Deployment team comments Larissa, March 12		The way that objects are created in “information sharing” tab is not yet understandable. How can information regarding closed threats be shared with others? In addition, how can closed threat information be received from others, for example Roma Capitale?	Similar as comment 1 in the for Rome usability above. Same solution applies.	Message included in pop-over on InformationShare page.

3.5 Observations and resulting system changes from piloting cycle 3

When	Where	Observation	Necessary change	Implemented in or after cycle 3
Pilot testing round 3 Usability test Roma, 16th of June	Visualisation	A self-healing option currently allows to assign a threat to a specific person, yet the threat cannot be passed without accepting/denying the self-healing action.	Potential: Remove the ability to pass the threat to assess/discuss the self-healing option first Potential: Allow to assign the threat to other persons first to assess/discuss the threat, without the need to accept/deny the self-healing option first	Change in Progress
Pilot testing round 3 Usability test Roma, 16th of June	Visualisation	It is only possible to assign the threat to one person at once. It is not possible to assign to multiple persons at the same time. Real world usage scenarios in Rome have shown that a threat/ticket is usually assigned to multiple people at the same time to assess/check/resolve the issue in specific parts of the system.	Provide a workflow that allows to assign a threat to multiple persons/groups at once.	Demands restructuring of workflow depending on size and organisation of Municipality. To be considered (not in scope)

Pilot testing round 3 Usability test Roma, 16th of June	Analysis (+Visualisation and Self-healing)	In the "observed data" details tab, the origin of the data entries is not clear to the user. The suggestion was to highlight from which asset/appliance and which log file the individual data entries come from.	Provide better origin attribution in the "observed data" tab.	Logfile information may be enhanced. (not in scope)
Pilot testing round 3 Usability test Larissa, 18th of June	Visualisation	No direct link possible from detailed threats view to system overview. Users would like to have this (has been observed in Rome tests as well)	Provide a link from the asset name in the detailed threats view to the system graph.	Change in Progress
Pilot testing round 3 Usability test Larissa, 18th of June	Visualisation	The users would prefer a self-healing workflow that keeps the threat open after applying self-healing, in order to check if the system is functioning correctly and provide a closing comment.	Do not automatically close threat after self-healing action is applied. – similar to comment from Rome.	Workflow Change in Progress
Pilot testing round 3 Usability test Larissa, 18th of June	Self-Healing / Visualisation	The technical description (command(s) to be performed) of the self-healing action is not shown. It would be good to have in addition to human readable description.	Provide the technical description in the "course of action" tab.	Change in Progress