TEST PROJECT ON COOPERATION IN EXECUTION OF VARIOUS MARITIME FUNCTIONALITIES AT SUB-REGIONAL OR SEA-BASIN LEVEL IN THE FIELD OF INTEGRATED MARITIME SURVEILLANCE (CoopP)

Final Report of Work Package 5:

Specifications of Common Data Formats and Semantics

Final Version

January 31, 2014

Co-Financed under European Integrated Maritime Policy

Table of Contents

Table of Figures

1 Executive Summary

Composed by 35 experts from 11 different member states, from several authorities and different user communities, and supported by the Joint Research Centre (JRC), the Directorate-General for Informatics (DIGIT) and 4 external experts, WP5 specified the data model and services necessary to support the development of an initial operational capability for the cross-border and cross-sectorial automatic exchange of maritime surveillance information, among European Union (EU) and European Economic Area (EEA) public authorities and agencies, in the context of the Common Information Sharing Environment (CISE). All the work has been developed during 12 months, within and between 5 face to face meetings and also within countless virtual meetings, all this supported by the collaborative tools provided by the JRC.

The data model is composed by 18 data entities, 7 main and 11 complimentary; defined in natural language and also specified in Unified Modelling Language (UML), Web Ontology Language (OWL), and XML –Extensible Mark-up Language– Schema Definitions (XSD). This model represents over 50% of the information needs identified by the Technical Advisory Group (TAG) for the development of the CISE, and over 64% of its definitions are reusing existing definitions from 34 related standards, systems and initiatives.

Fifteen services and five messaging patterns have been defined by WP5 which, when combined, form the framework to enable the use cases defined by WP2 and further enhanced by WP5. This framework, known as service model, is defined in UML and formalized in XSD. The patterns, as well as some of the services, have been formalized in Web Services Definition Language (WSDL).

Both the data model and the services have been validated and verified, first by the experts within WP5 and afterwards by two independent experts, encompassing the business and technical perspectives, whose conclusions and recommendations are provided as annexes to this report. The data entities are fully used by the use cases defined by WP2 and the services fully enable the use cases defined by WP2 and further enhanced by WP5.

WP5 has therefore developed the fundamental and preliminary pieces for the technical and semantic interoperability within the CISE, the common data model and the services, and the next steps towards the CISE should encompass the specification of a technical reference architecture aligned with the architectural visions and principles of CISE, as well as the definition of an appropriate governance structure. Both steps should be built following an iterative and incremental approach, so that risk is minimized and the necessary experimentation is conducted before bringing this capability into production.

Co-Financed under European Integrated Maritime Policy

2 Background

2.1 Assigned Tasks and Their Expected Outputs

Work Package 5 was assigned with following tasks:

- For each identified information service, define supporting data sets, common data formats and semantics (multilingualism needs to be carefully considered when designing the data formats and semantics). This includes:
- Identification of existing standards to be used - agreement on the standards if several may apply, using for example the 'Common Assessment Method for Standards and Specifications methodology';
- Detailed description of data elements (sufficient to produce a schema), making best use of available standards;
- List of code lists or enumerations

The abovementioned tasks were expected to be accomplished with following outputs:

Expected Outputs	**Output reached: yes/no; reference chapter/annex for results discussion**
List and complete description of data elements for each information service. The data model shall be made available in HTML, in XML and in RDF/RDFS formats to all user communities executing maritime functionalities in all EU/EEA Member States, irrespective of their participation to the present project, to the future FP7 project and to the industry. To allow for wider possible reuse, the beneficiary will make the data model available through the internet in a widely accessible form and may not implement more stringent condition for access/download and reuse than those specified in the ISA Open Metadata license.[1]	Yes. See chapter 3 and related annexes.

[1] https://joinup.ec.europa.eu/category/licence/isa-open-metadata-licence-v10

2.2 Objectives

The Cooperation Project was expected to reach following objectives:

Objective 1: To define and agree on a selection of use cases with related information services and attached access rights (WP 2 and WP 4)

Objective 2: To define common data formats and semantics (WP 5)

Objective 3: To contribute to the cost-benefit analysis of Integrated Maritime Surveillance (WP 3)

From these objectives, Objective 2 falls under the responsibility of WP 5. A table of predefined Output and Result Indicators is presented in Annex II with a summary of delivered outputs and results.

3 Data model

When faced with the purpose of WP5, to define common data formats and semantics (i.e., a common data model[2]), immediately three questions came to our mind: Why should such data model be common, and why should we define a new one, when at the EU and MS levels already exist so many initiatives regarding cross-sectorial and cross-border information sharing, not mentioning the widely existing and accepted standards on this matter produced by renowned organizations? Why cannot we simply take and use one of them?

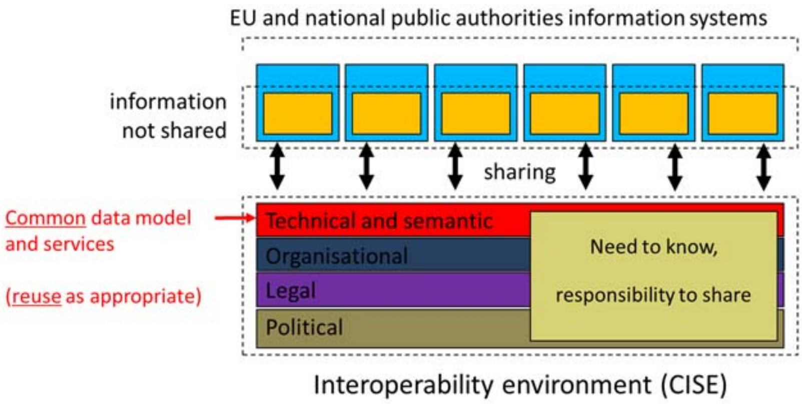

Figure 1: CISE as an interoperability environment

Let us focus for a moment on the role of the EU Common Information Sharing Environment (CISE) – an interoperability environment (Figure 1). Presently there is still plenty of information laying in the existing EU and national public authorities' information systems that is not shared to those needing it to perform their activities, thus hampering their efficiency and effectiveness and consequently Blue Growth.

To overcome this situation, it is of the outmost importance that their interoperability is enhanced at all levels, legal, organisational, technical and semantic, as per the European Interoperability Framework (EIF). This is

Core concepts

LEGAL Interop. Aligned legislation so that exchanged data is accorded proper legal weight.

ORGANISATIONAL Interop. Coordinated processes in which different organizations achieve a previously agreed mutually and beneficial goal

SEMANTIC Interop. Precise meaning of exchanged information which is preserved and understood by all parties

TECHNICAL Interop. Planning of technical issues involved in linking computer systems and services

[2] Although a data model is quite different from an information model, as per the IETF RFC3444, for simplicity reasons we shall use, throughout this report, the term data model to designate simultaneously the information model (in UML) and the data models (in XSD and OWL) defined by WP5.

where the CISE steps in, as an interoperability environment encompassing all the levels, based on the principles of the 'need to know' and the 'responsibility to share', to increase the interoperability maturity of all those involved and, consequently, the information shared among them. CISE is not the system of the existing systems. It shall lay among them, to enhance them.

To enable information sharing across computer systems we have to develop the appropriate services and, assuming that the data models of the existing systems are not all the same, to transform the data, back and forth, between the data models involved, while preserving the meaning of the information being exchanged.

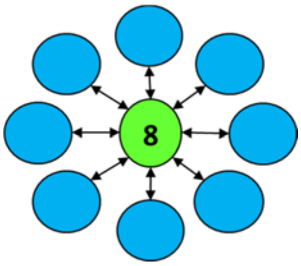

Figure 2: Common data model

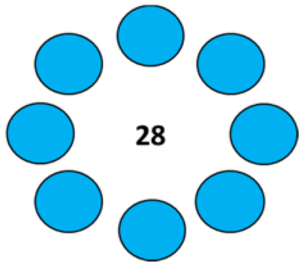

Figure 3: No common data model

If we use a common data model (Figure 2) in the services deployed in the existing systems with the purpose of implementing the CISE technical and semantic interoperability, the number of transformations that needs to be implemented is equal to the number of already existing data models. That is, we will only have to implement a transformation between each of the already existing data models and the common data model.

On the contrary, if we decide not to use a common data model (Figure 3), we need to implement transformations between each of the existing data models from the existing systems. Therefore the number of transformations that needs to be implemented equals $(n * (n-1)) / 2$, where n is the number of different existing data models.

This means that, for the same number of data models involved, the number of transformations that needs to be implemented is by far larger if we do not use a common data model. It also means that the number of transformations that needs to be implemented is exponential if we do not use a common data model, while using a common data model that number has a linear growth, proportional to the data models involved. Consequently, by using a common data model, we will have a far less complex and expensive solution both in terms of implementation and maintenance.

The reason to use a common data model is therefore the cost-effectiveness[3] it provides, when compared to the alternative solution.

The other questions we have to answer is why don't we simply use one of the existing systems or initiative's data models? Or even better, why don't we just use completely one of the existing systems to achieve the goals of CISE?

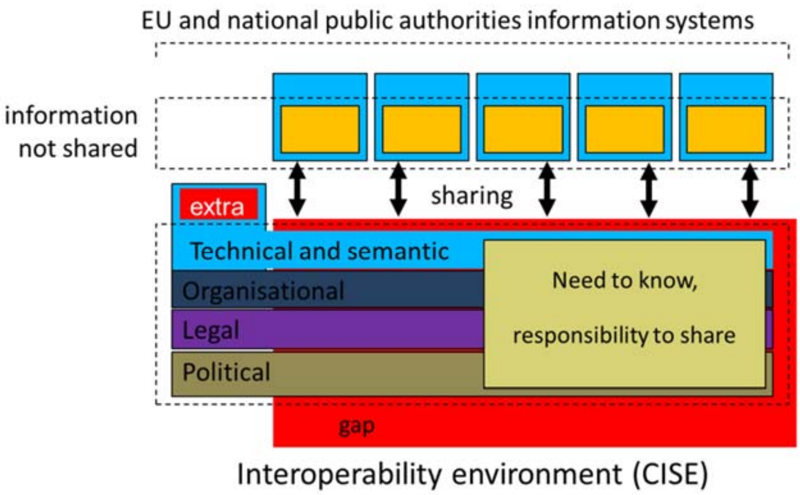

Figure 4: Differences between existing data models and a common one

The existing systems are operational systems which encompass many more components than the ones required to implement the technical and semantic interoperability level that CISE aims to. This means that by adopting one of the existing systems to implement the CISE, we would be using a far more complex solution than really needed.

Also the fact that existing systems are driven by operational requirements, and not by sector neutral technical and semantic interoperability requirements is important. This means that their data models and other components address primarily operational concerns, while a dedicated and sector neutral interoperability solution shall address primarily interoperability concerns. This also means that if a sectorial solution is used to implement interoperability, conflicts will certainly occur regarding the priorities to assign when operational and interoperability requirements are competing.

[3] Presently, throughout Europe, 360 different entities are performing coast guard functions. If each of them has its own system, the difference between one solution and the other, in terms of the transformations between data models that need to be implemented, is from 360 (common data model) to 64 620 (no common data model).

Additionally, the existing solutions have a sectorial purpose, meaning that their data models do not encompass the data requirements of other sectors and, consequently, if such a solution is adopted, the original data model and related components have to be enhanced to fulfil the gaps, thus implying additional investment and eventually raising legal issues regarding the owner of the system accessing data without the need to know.

Finally, if a sectorial solution is adopted to implement the technical and other interoperability levels aimed by the CISE, all the remainder sectors will have to adopt the data model used by that particular sector, thus creating a strong inter-dependence across sectors, meaning that when that sector decides to update or change their data model all the remainder sectors will have to change theirs accordingly. Last but not least, a governance issue can arise. If the interoperability data model and related components belong to a sector, how can other sectors change those components?

All in all, by adopting an existing data model or even a complete existing system to implement the technical and other interoperability levels aimed by the CISE (Figure 4) we would have more complexity, and most certainly legal and governance issues.

3.1 Methodology

During the project, WP5 was split into two subgroups and followed closely the Process and Methodology for developing CORE vocabularies[4] defined in the context of the Interoperability Solutions for European Public Administrations (ISA) Programme. This methodology has proven very useful throughout the project. Still, several challenges arose, and the ways found to deal with the major ones are further described in this section.

3.1.1 Selecting the core data entities

The first major challenge was to select the data entities which would be the building blocks of the data model. This implied knowing the data entities to consider and also developing a consensus around those that, for some reason, would be selected as core.

To understand which data entities should be considered, we had two options. The first one was to wait for the results of WP2, which encompassed the services definitions and, within them, the identification of the necessary data entities, as inputs and outputs of those services. The second option was to start working as soon as possible, before the end of WP2, by analysing the use cases and the data matrix developed by the Technical Advisory Group (TAG), then analysing the use cases developed by WP2 and, finally, analysing the services definitions, also made by WP2.

[4] https://joinup.ec.europa.eu/elibrary/document/isa-deliverable-process-and-methodology-developing-core-vocabularies

On one hand, the first option would, in theory, require less effort, as long as the results of WP2 would be precise enough and if WP5 would be effective and efficient enough to deliver the data specifications as required in such a short time frame. On the other hand, it was soon perceived that the tasks committed to WP5 encompassed several risks; high technical complexity, large amount of related data models and standards to analyse, very short time frame, experts with heterogeneous professional background, maturity and technical knowledge, as well as the language and geographical barriers, always present in such international projects.

To overcome these challenges, time was required; time for building the necessary processes and cohesion in WP5, and also to develop the inherent time consuming activities. Not only the second option allowed for this time, it also allowed for the development of the necessary cohesion and synergies between WPs, especially between WP2 and WP5; therefore, this was the chosen option.

WP5 started then by analysing the use cases and the data matrix developed by the TAG. This data matrix had over 400 possible data entities that could not be specified within this project. Therefore, it was necessary to narrow down the number of data entities.

As soon as the use cases and services were made available by WP2, some of these use cases have been selected for further enhancement which, in turn, originated an enhancement of the related services definition. By doing this, it was possible to be more precise regarding the identification of the data entities needed for the development of the CORE vocabularies and the services specifications. This task has been developed by an external expert who enhanced use cases 25b, 37 and 57 and also the related services definitions. The results are presented in annexes III and V, respectively.

After analysing the use cases developed by the TAG and those developed by WP2 and enhanced by the external expert, it was possible to reduce the number of possible data entities to over 150. Still, it was a number too huge to manage with the available resources and time. It then became evident that WP5 had to follow another approach.

Within the scope of the WP5, it was then agreed that a CORE data entity is such that it is encompassed in the broad description of the activity under analysis – maritime surveillance.

To arrive at such description, the activity of maritime surveillance has been described, by experts in the domain, by answering the following questions during a short brainstorming session:

- Why (is maritime surveillance performed)
- Who (performs maritime surveillance)
- What (is actually surveyed)

- How (is maritime surveillance performed)
- When (is maritime surveillance performed)
- Where (is maritime surveillance performed)

After arriving at a consensual description, the concepts used in it were considered as core data entities and a first version of the list of core data entities has been developed. Examples of such concepts include Illegal activity, Authority, Cargo, Period of time and Area.

Afterwards, the +150 data entities found in previous activities were reviewed with the purpose of finding core data entities missed in the developed description of maritime surveillance. The vast majority of the analysed data entities were considered either complimentary, synonyms or abstractions of the data entities initially considered as core. A residual number of new core data entities emerged and some of the data entities from the first version of the list were renamed or merged. The list of core data entities was updated accordingly, resulting on its second version.

Additionally, the services definitions developed by WP2 and enhanced by the external expert, which included the identification of input and output data, were reviewed with the purpose of finding core data entities missed in the developed description of maritime surveillance. Again, a residual number of new core data entities emerged and some of the data entities from the second version of the list were renamed or merged. The list of core data entities was updated accordingly, resulting on its third and final version.

By combining both approaches it was possible to arrive at a shorter list of over 20 core data entities, which was refined along the project and finally contained 18.

3.1.2 Selecting the data definitions to reuse

After having identified the core data entities, WP5 focused in understanding and modelling the main attributes and associations relevant for the use cases and services described by WP2. Afterwards, WP5 searched into the related standards, systems and initiatives specifications available for equivalent and additional attributes and associations of interest, in order to enrich the data model. Finally, WP5 continued to refine the data models, also with the support of external experts hired for this purpose.

Considering the huge amount of existing systems, initiatives and standards, related to the maritime domain, it soon became obvious that there was a very high reuse potential in the existing solutions throughout the maritime community, which could facilitate the work of WP5, but also that a proportional amount of resources and time was necessary to analyse them properly and decide on their usefulness, and that so much diversity would lead to difficulty on deciding which specifications to reuse.

First, a list of related systems, initiatives and standards of interest was compiled. Then WP5 identified points of contact for most of these systems, initiatives and standards and, with their support, obtained access to most of the inherent data specifications. Some of the data specifications from these systems, initiatives and standards were not analysed, mostly because either their data specifications were not available or because there was not enough time and resources to do it.

To easily identify what is being reused by the present data model, the fully qualified name of the data elements was used, so that the reader can easily understand the provenance of such data element. There are also several situations where although a direct reuse was not performed, the existing data models inspired the present proposal.

At the end of the day, what we expected to achieve by reusing or defining data inspired by existing solutions, was to reduce the amount of effort while designing the data model and to increase its understandability and usability. However, sometimes reuse conflicts with the objectives of usefulness and extensibility. Furthermore, many times conflicting or at least alternative data definitions that could be reused do exist. Therefore, the proposed model is a careful attempt to try to achieve all the objectives, considering as much as possible the existing reality, in an efficient and effective manner, and could not simply be a cut and paste of the data definitions of the existing systems, initiatives and standards into a single data model.

Whenever we found equivalent data definitions that could be reused, we had to select one of them. In these situations we decided that it was more important to foster the interoperability among the existing systems at the EU level relevant for maritime surveillance, rather than fostering the interoperability of CISE with other systems or even more general standards. Therefore we often chose primarily data definitions from existing systems at the EU level, in spite in most cases they are not standards or even existing standards could have been adopted instead. By taking such measures we intend to enable the integration of the existing systems relevant to maritime surveillance at the EU level to be faster and less expensive.

It was a constant concern that the data models developed should be flexible enough to easily accommodate change and to facilitate the transformations from and into the legacy systems' data models, since the information model developed is aiming to support the interoperability among the existing systems in the CISE communities. To achieve this objective, it was decided to keep the model at an adequate level of abstraction: not too specific, to facilitate transformations, and not too generic, to keep the focus and understandability. In this context, the way to increase the semantic robustness of a more generic data model, while keeping it extensible, is through the definition of constraints (invariants) in OCL.

After defining the data model in UML[5], it was transformed into their XSD and OWL counterparts by an external expert hired for this specific purpose. During this process, the OCL constraints were also included in the XSD and OWL formalizations as "annotations for use".

3.2 Results

The data model developed addresses the following general requirements:

1. Shall be useful, in the sense that the structure and data definitions it encompasses enables the use cases also defined in the scope of the COOP project

2. Shall be understandable, meaning that it has to be easily understood by those involved in future system integration activities

3. Shall be usable, in the sense that the effort it imposes to the systems' integration is acceptable

4. Shall be extensible, so that future evolutions at the use cases and services can be easily followed

The data model developed is presented in a series of core vocabularies specifications documents which are presented in annex IV. Each of these documents describes a part of the data model, also using examples, focused on one of the data entities defined. Additionally, each core vocabulary specification has its corresponding formalization in XSD and in OWL.

In the core vocabularies there is an UML representation of the data models, a standard widely acknowledged and used by the software engineering community. The data model has been defined using classes, attributes, associations and enumerations, which broad definitions and examples (and much more) can be found here[6] for a better understanding.

Despite the fact that several formalizations of the data model are provided, the one which shall be used as reference for the services implementation is the one provided in XSD, since all of them have slight variations in order to better adapt to the specific language being used, and the specification in UML is less formal than the remainder.

The data model developed is simple, sufficient and flexible. It comprises several special features to accommodate crosscutting concerns such as auditing, security and data reliability and validity. It comprises 18 data entities with 271 data attributes, representing over 50% of

[5] http://www.uml.org
[6] http://www.agilemodeling.com/artifacts/classDiagram.htm

the data requirements expressed in the TAG data matrix. Over 65% of the data definitions in the data model are imported from over 34 different existing related data models.

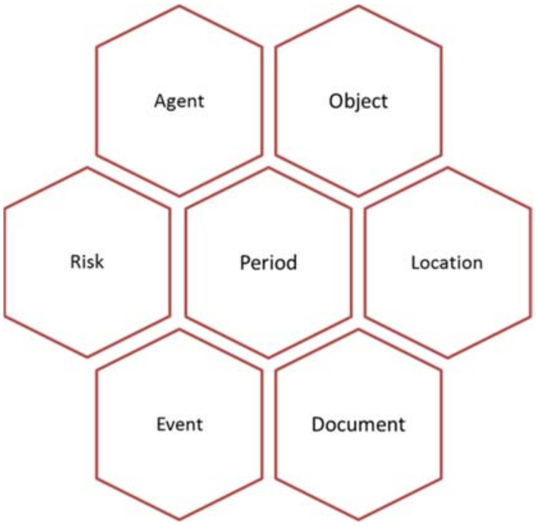

Figure 5: Data model essential data entities

The data model comprises 7 data entities (Figure 5) which are essential to share information regarding maritime surveillance. These are:

- **Agent** - an agent is an operative entity that plays a role in any Event; owns, handles or operates Objects, such as Cargo or Assets; and creates and exploits Documents

- **Object** - Holds information about physical entities from the maritime domain like vehicles (vessels, aircrafts and land vehicles) and cargo packages

- **Risk** - Risk is the possibility that something bad or unpleasant may happen. There are also several types of risk and severity, as well as different risk levels. Risks can be further described by documents and can involve events, places, objects persons and organizations

- **Period** - Defines a time interval which can be expressed in several different ways

- **Location** - Location is the spatial dimension of maritime surveillance. A location is a specification of a physical location on earth. It can be an address, a point, an area, or a port. Events occur in locations, organizations and objects are in locations which can be further described by documents and can be associated with risks

- **Event** - Something that occurs in a certain location, during a particular interval of time, especially one of some importance (movements, anomalies, incidents and operational actions). An event can involve persons, organizations, objects, can imply risks and can be further described or explained by existing documentation. Additionally, persons and organizations can play specific roles in events.

- **Document** - Documents are digital representations of paper documents. These can be presented as images, pdf files or even more structured documents, which format can also be shared among partners for their automatic interpretation, even if their specific data model is not part of this data model. Documents are used to further describe every aspect of interest for maritime surveillance. As such, many different kinds of documents can be produced by organizations and shared.

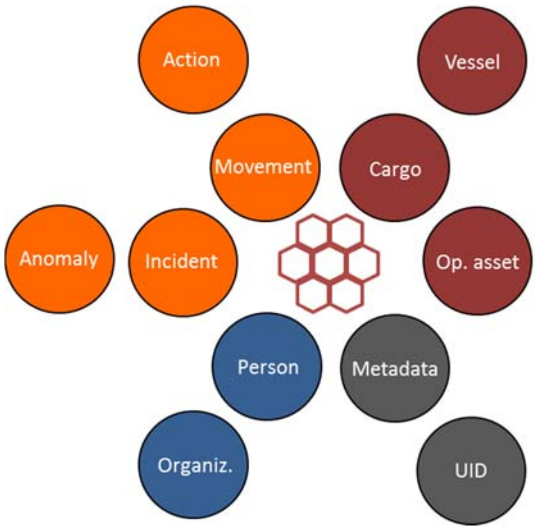

Figure 6: Data model complimentary data entities

Additionally, 11 other data entities (Figure 6) have been defined, to increase the overall expressiveness of the data model and to support the special features early mentioned. These are:

- **Action** - An action (or an operational action) is a specific type of event that is conducted by organizations or persons and that may follow an incident or an anomaly, or is performed to prevent these. Actions can also imply and/or be the consequence of movements

- **Movement** - A movement is also a specific kind of event, therefore it is related in the same manner to documents, risks, places, organizations. A movement can therefore represent the movement of an object, a person, from one place to the other, within a specific Period, further described by documents

- **Anomaly** - An anomaly is something unusual that has, is occurring or is expected to occur, that can lead to an incident

- **Incident** - An incident is an unexpected and usually unpleasant thing that happens. There are several types of incidents and, as a specific kind of event, they are likewise related to documents, risks, places, organizations, objects and persons exactly like events

- **Vessel** - A vessel is a special kind of an object and, as such, it is also related in the same manner to locations, risks, documents, and events. As a special kind of craft, a vessel also has passengers and crewmembers

- **Cargo** – A set of goods transported by a ship between two ports

- **Operational Asset** - An Operational Asset is an Object (in particular means of observation or transportation, but also including associated sensors, means of communication and means of intervention such as deterrence or neutralization of threats, firefighting, pollution containment etc.) enabling operational Actions (most often at sea or on sea shores) of the Agents mandated by public Organizations in charge of Maritime Surveillance

- **Person** - Persons are humans of interest for maritime surveillance. They can play roles in organizations or objects such as vessels or operational assets. Persons can be further described by documents and can be associated to risks

- **Organization** - Organizations are entities that involve persons, can own objects, be involved in risks and/or events, are usually in certain locations and can be further described by documents

- **Metadata** - Provides information about the properties of the data communicated through the system, excluding the content of the data

- **Unique identifier** – A unique identifier is a piece of information present in any data object being shared, randomly generated by partners at a specific point in time, so that each and every data object being shared can be referred to and used without ambiguity.

3.3 Verification and validation

As important as defining and formalizing the data model was to ensure that:

- it was expressive enough
- it was technically correct

For this matter WP5 counted with the support of two external experts which followed closely the data modelling activities and decisions and produced independent assessments.

One of these assessments is presented in annex VII and comprises a review, from a business perspective, of the core vocabularies specifications developed. Within this review several recommendations are produced, most of which have already been incorporated in the final version of the core vocabularies.

Another assessment made is part of the enhancement of the use cases developed by WP2, presented in annex III. During this exercise, the external expert verified that 100% of the data entities defined are used by the use cases enhanced.

The last assessment performed in this scope is presented in annex VIII and comprises a review, from a technical perspective, of the core vocabularies specifications developed. Within this review, many recommendations have been produced, most of which have already been incorporated too, in the final version of the core vocabularies. It is worth mentioning that although there is room for improvement, no major issues have been reported.

Additionally, as soon as the data model was ready to be presented to an external audience, even though not yet at its final version, a meeting with the Technical Advisory Group (TAG) was arranged, so that it could be validated. The outcome of this meeting was very fruitful, resulting in a series of contributions and recommendations which were fully adopted and oriented the evolution of the data model towards its final version.

4 Services

4.1 Methodology

To enable the technical and semantic interoperability among existing systems, in this particular case those related to maritime surveillance, a common data model is a very important piece; however, there are other pieces as important, such as the integration services.

Core concepts

MESSAGING PATTERN Template for a solution to a specific set of service usage requirements

ENTITY SERVICE Manage access to (business) operational entities. Provide the information needed to implement the tasks

The rationale behind developing a set of services which is used by all the partners consistently, therefore designated as common, is pretty much the same used for developing a common data model. In a nutshell, it is a much less complex and expensive approach, as already discussed.

Therefore, WP5 embraced the task of specifying such services and started by analysing the services definitions proposed by WP2, which were developed based on the technique proposed by WP5 and encompassed a set of messaging patterns aiming the fulfilment of several information exchange requirements, regarding the ways information can be requested and provided.

Upon analysing the list of services proposed by WP2, it was found that further refining and specifying such a list, up to the level required to start their implementation, would be a huge task, in terms of the time required to perform it, which this project could not afford. Therefore, and following the recommendations coming from the meeting with the TAG, the decision was to focus the work only on the entity services, which by themselves would already bring a great deal of added value to the results of the project and would be enough to develop an information sharing initial operational capability. With this in mind, WP5 has developed a catalogue of services.

However, having a set of messaging patterns and a set of entity services was not enough. It was also required to define how the services and the patterns should be used altogether, for which a service model has been developed.

Core concepts

SERVICE Modular unit of business (operational) functionality that is made available through a service contract

INTEGRATION Connection to and exposure of existing applications and/or data as services

INTEGRATION SERVICES Expose existing application and data as services

The final step was then the formalization of the messaging patterns and some of the services in WSDL. The services selected were those necessary to fulfil the information exchange requirements inherent to some[7] of the use cases developed by WP2 and enhanced by WP5.

To develop the information model and to specify the patterns and the services, WP5 was supported by external experts.

4.2 Results

4.2.1 Messaging patterns

The five messaging patterns developed by WP5 and explained below describe the different ways by which the different services can interact to enable the exchange of the desired information between the computer systems involved, and their usage depends on being technically implemented by the specific service, and also on the given operational context.

Pull: the consumer knows the exact provider and asks for the information which is immediately made available (synchronous)

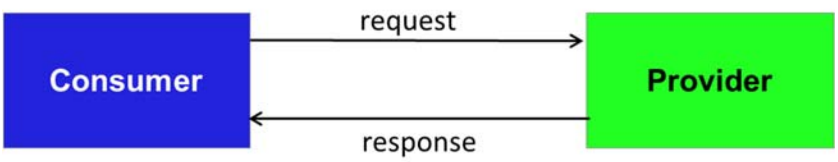

Figure 7: Pull messaging pattern

Pull delayed: the consumer knows the exact provider and asks for the information which is made available only if and when possible (asynchronous)

Figure 8: Pull delayed messaging pattern

[7] Use Case 57+70

Broadcast pull: the consumer does not know the exact provider and asks for the information to all the possible providers. The information is made available only if and when possible (asynchronous). Several responses may occur.

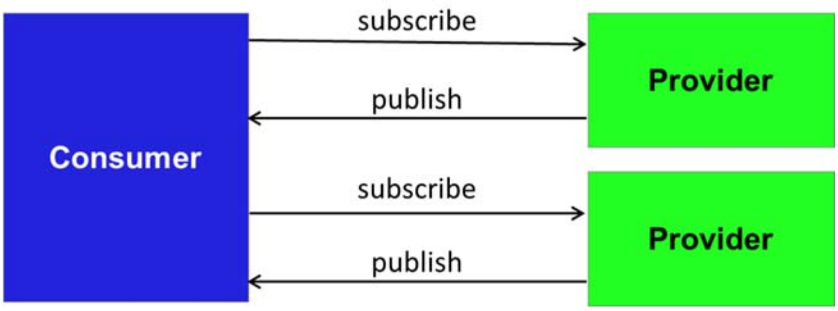

Figure 9: Broadcast pull messaging pattern

Push: the provider knows the exact consumer and sends it to it regardless of it not having previously requested such information (synchronous)

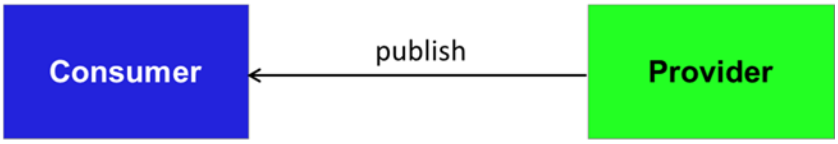

Figure 10: Push messaging pattern

Broadcast push: the provider does not know who is willing to consume, therefore it broadcasts the information to all possible consumers (synchronous)

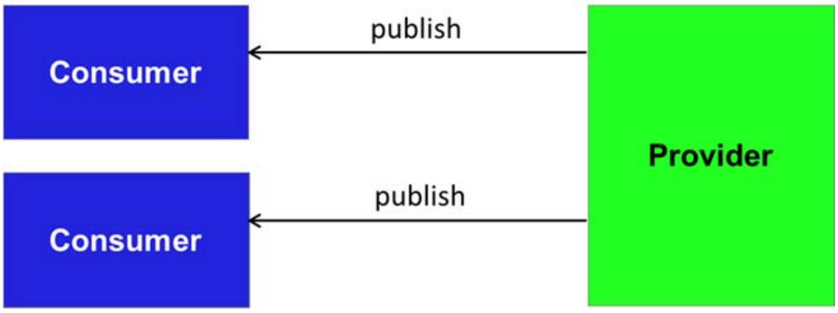

Figure 11: Broadcast push messaging pattern

4.2.2 Services catalogue

The services catalogue is, in broad terms, a list of the services that should be developed in order to support the information exchange activities required by the use cases and services developed by WP2 and further enhanced by WP5.

As agreed during the meeting with the TAG, considering the amount of services specified by WP2 and the available resources and time for WP5 to accomplish its tasks, it would not be feasible to specify all the services necessary. Since it was found that developing only entity services would be a good starting point, WP5 decided then to take a bottom-up approach and, focused on the core data entities defined earlier, proposed the entity services presented in the following table, which are defined by their name, their input and output parameters, the patterns which would make sense to use with it, and its brief description.

Name	Input	Output	Pattern	Description
Agent services				
PersonService	Period, Location, Object, Event, Risk, Document	Person with associated entities	Pull, pull delayed, broadcast pull	Delivers information on Persons related to defined Location, Object, Event, Risk, Document and/or Period
OrganizationService	Period, Location, Object, Event, Risk, Document	Organization with associated entities	Pull, pull delayed, broadcast pull	Delivers information on Organizations related to defined Location, Object, Event, Risk, Document and/or Period
Event services				
ActionService	Period, Location, Object, Agent, Risk, Document	Action with associated entities	all	Delivers information on Actions related to defined Location, Object, Agent, Risk, Document and/or Period (can also be request for action)
AnomalyService	Period, Location, Object, Agent, Risk, Document	Anomaly with associated entities	all	Delivers information on Anomalies related to defined Location, Object, Agent, Risk, Document and/or Period
IncidentService	Period, Location, Object, Agent, Risk, Document	Incident with associated entities	all	Delivers information on Incidents related to defined Location, Object, Agent, Risk, Document and/or Period
MovementService	Period, Location, Object, Agent, Risk, Document	Movement with associated entities	all	Delivers information on Movements related to defined Location, Object, Agent, Risk, Document and/or Period

Name	Input	Output	Pattern	Description
Object services				
VesselService	Period, Location, Agent, Cargo, Event, Risk	Vessel with associated entities	Pull, pull delayed, broadcast pull	Delivers information on Vessels related to defined Location, Agent, Cargo, Event, Risk and/or Period
CargoService	Period, Location, Agent, Vessel, Event, Risk	Cargo with associated entities	Pull, pull delayed, broadcast pull	Delivers information on Cargo related to defined Location, Agent, Vessel, Event, Risk and/or Period
Location services				
PersonLocationService	Period, Person	Location	Pull, pull delayed, broadcast pull	Delivers the Location(s) of Persons within given Period
OrganizationLocationService	Period, Organization	Location	Pull, pull delayed, broadcast pull	Delivers the Location(s) of Organizations within given Period
VesselLocationService	Period, Vessel	Location	Pull, pull delayed, broadcast pull	Delivers the Location(s) of Vessels within given Period
CargoLocationService	Period, Cargo	Location	Pull, pull delayed, broadcast pull	Delivers the Location(s) of Cargo within given Period
Other services				
RiskService	Period, Location, Object, Agent, Event	Risk with associated entities	all	Delivers the Risk related to Location, Object, Agent, Event and/or Period
DocumentService	Period, Location, Object, Agent, Event, Risk	Document	Pull, pull delayed, broadcast pull	Delivers the Document related to defined Location, Object, Agent, Event, Risk and/or Period
OperationalAssetService	Period, Location, OperationalAsset	OperationalAsset	Pull, pull delayed, broadcast pull	Delivers information on available OperationalAssets (with required capabilities) in given Location and Period

During the definition of the services catalogue, WP5 has considered that Period, Metadata and UniqueIdentifier were not worth having dedicated entity services, since these are essentially data entities to support technical concerns of the data model.

Additionally, the following requirements shall apply to all services:

1. In all services, it shall be possible to use any of the attributes of the data entities identified as input parameters.

2. In all services, it shall be possible to specify the data entities and attributes that the consumer is willing to receive as output.

3. In case of all pull services, there is the need for an optional parameter (e.g. "ResponseDeadline") so that requests can be discarded when the information requested is no longer useful.

4.2.3 Service model

The services catalogue defines a list of business services which are necessary to support the activities defined in the use cases. To ensure consistency, reusability and extensibility of the CISE technologies, it is useful to define a model for these services which is independent of their content and function in the context of a specific activity or use case. This allows parties to easily develop new business services which are automatically compatible and interoperable with existing ones, and which can be easily integrated into the CISE network.

A service model has therefore been devised which is independent of the specific activities of the use cases and is only loosely coupled to the data model which has been developed by WP5. It is based around the notions of communication patterns and data templates, promoting reuse and extensibility. The service model is characterised as follows:

The 5 messaging patterns can be viewed on 3 different levels or “layers”:

- *Business* – the conceptual transaction between businesses
- *Messaging* – the layer in which the transaction is realised through a sequence of specific messages
- *Transport*[8] – the layer which defines how individual messages are transmitted on a technical level

The transport layer defines 5 “fundamental” low-level operations for transferring messages between systems:

- *RequestObjectFromService*
- *RequestNotificationFromService*
- *CallbackNotification*
- *PushCISEMessage*
- *BroadcastCISEMessage*

Each operation takes an input and returns an output. The inputs and outputs of an operation are called messaging objects. Different messaging objects are passed back and forth in order to conduct the desired information transaction. The same transaction with different messaging

[8] The transport layer defines how the messages are transmitted whilst the messaging layer defines which service takes place

objects will have a different outcome. Across all the patterns, there are 7 pairs of messaging objects in total, each pair consisting of:

- A RequestObject (sent as part of the service operation invocation)
- A ResponseObject (received synchronously in response to a service operation invocation)

The 14 messaging objects are shown bellow (Figure 12).

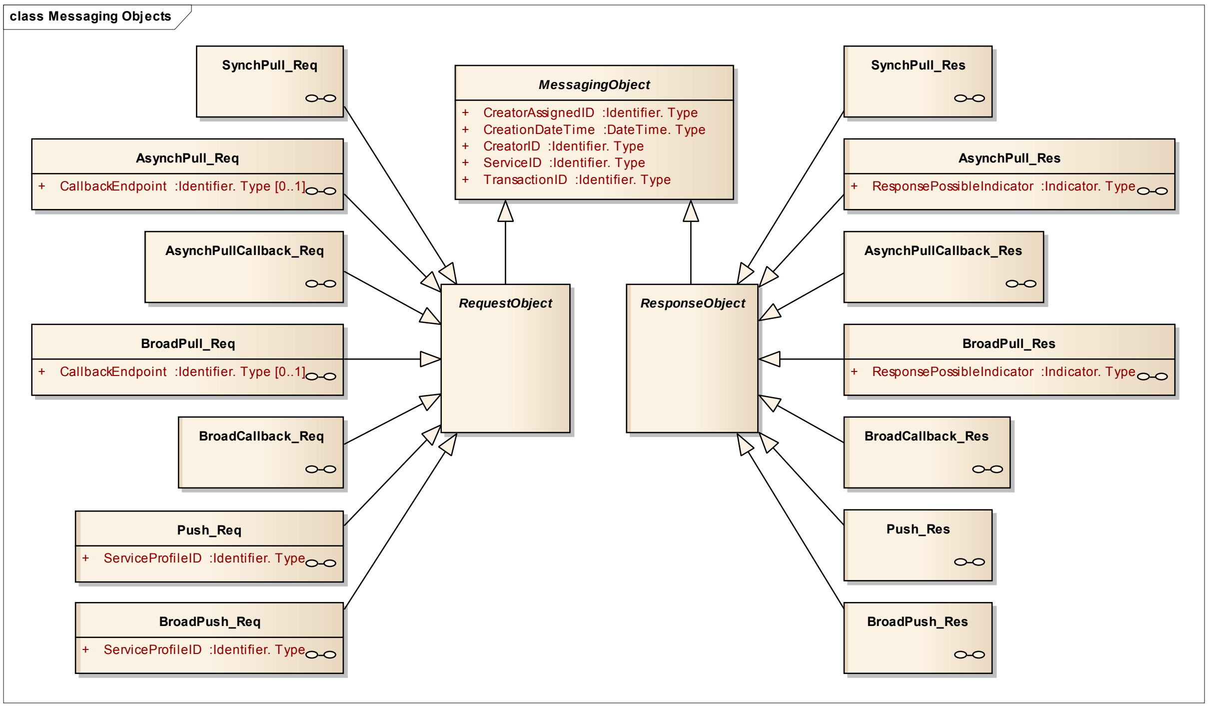

Figure 12: Messaging Objects

The correspondence between the different patterns, operations and messaging objects is mapped out in Table 1. The "Direction" column specifics the direction in which the operation is invoked: P = Provider, C = Consumer. For example, C → P indicates that the Consumer invokes the operation on the Provider. Numerical suffixes (P1, C2, etc.) are used to show the multiplicity of communication in certain patterns.

Table 1: Service Model Mapping

Pattern Name	Service Operation/Call	Direction	Request Object	Response Object
PULL	RequestObjectFromService	C → P	SynchPull_Req	SynchPull_Res
PULL	RequestNotificationFromService	C → P	AsynchPull_Req	AsynchPull_Res

Pattern Name	Service Operation/Call	Direction	Request Object	Response Object
DELAYED	CallbackNotification	P → C	AsynchPullCallback_Req	AsynchPullCallback_Res
BROADCAST PULL	RequestNotificationFromService	C → P1	BroadPull_Req	BroadPull_Res
	RequestNotificationFromService	C → P2	BroadPull_Req	BroadPull_Res
	CallbackNotification	P1 → C	BroadCallback_Req	BroadCallback_Res
	CallbackNotification	P2 → C	BroadCallback_Req	BroadCallback_Res
PUSH	PushCISEMessage	P → C	Push_Req	Push_Res
BROADCAST PUSH	BroadcastCISEMessage	P → C1	BroadPush_Req	BroadPush_Res
	BroadcastCISEMessage	P → C2	BroadPush_Req	BroadPush_Res

The Services Catalogue defines a number of business services for maritime surveillance information exchange. A Service has the concept of an Input and an Output which contain information in the form of Core Entities as specified by the CISE core vocabulary data models.

Each Service can follow one or more of the messaging patterns. Each Pattern is defined by a specific sequence of Operations and corresponding Request and Response Objects. Both types of Messaging Object – Request and Response – may contain information in the form of Core Entities, thus conducting the process of information exchange.

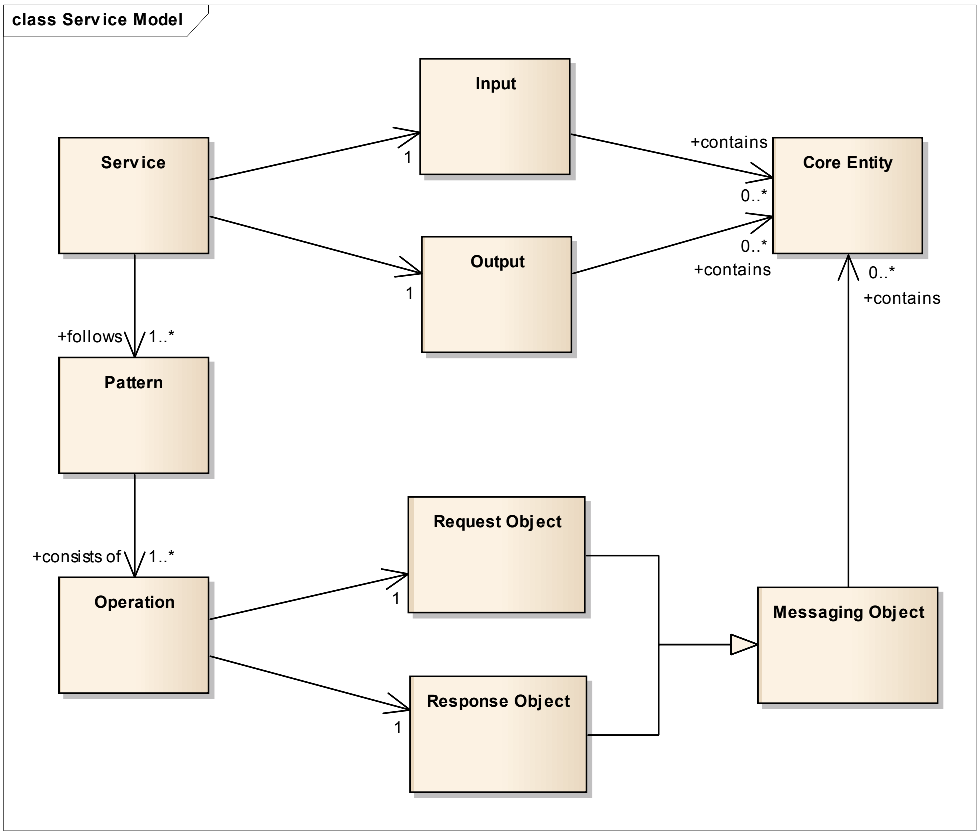

Figure 13: Service Model

The service model is described above (Figure 13), showing how business Services are realised through the exchange of Messaging Objects using a combination of Operations defined in the transport layer. The Messaging Objects act as containers for carrying Core Entities between providers and consumers to enact a specific business Service. The business Service is thus decoupled from the technical implementation of the communication infrastructure. The Input and Output will be the same, regardless of the pattern used, but the messaging objects will be different.

4.3 Verification and validation

To ensure that all the outputs of this phase were adequate and technically correct, WP5 was supported by two external experts.

While one of the experts evaluated if the integration services were adequate to suffice the information exchange requirements inherent to the use cases developed by WP2 and enhanced by WP5, the other expert evaluated if the services catalogue, the patterns and the services specifications in WSDL were technically correct and in-line with the information model.

The results of these evaluations were therefore the services review from a business and a technical perspective, which reports are presented respectively in annexes VII and VIII.

Several recommendations have been produced in the context of such reviews, and some of them have already been implemented in the last version of the products being delivered. There is still, of course, room for improvement.

In broad terms, these reviews confirm that the services are fit for purpose and technically correct.

5 Conclusions

The objectives of WP5 have been achieved and exceeded. Not only the common data formats and semantics have been specified as required, but also a set of fifteen services and five messaging patterns have been defined and formalized using the appropriate standards and reusing as much as possible from related data models. These are the building blocks of the technical interoperability framework necessary to support the information exchange activities required by the use cases and services developed by WP2.

The preliminary work developed by the TAG and the ISA process and methodology for developing CORE vocabularies have been extremely useful during the activities of WP5.

The decision to start working as soon as possible in the data models, with the inputs available, namely those from the TAG, while waiting for the results of WP2 and merging them along the way, has turned out to be a good one. It would have been impossible to deliver the necessary products in such a short time frame, if the original schedule has been followed. This also allowed for the development of synergies and cohesion among WPs, namely between WP2 and WP5, which have been fruitful for both.

The core data entities selected and specified by WP5 are very relevant for maritime surveillance information exchange. This has been confirmed by their wide usage in the use cases and services defined by WP2, as well as in related data models, although sometimes with different designations, and also by the TAG.

The data model and services developed by WP5 are technically sound as confirmed in the review performed by the external expert, although there are some issues that need to be addressed in future work.

The involvement of external experts in WP5, namely from the JRC, DIGIT and the TAG was a success factor for the activities of WP5. Without the wide experience and knowledge they carried into the project it simply would not have been possible to achieve these results, given the overwhelming amount of topics involved and the high technical complexity the activities of this WP encompassed.

Hiring external experts to strengthen the WP skills regarding specific technologies required by the project has proven to be a good decision. By combining the experience and knowledge of the experts from the member states regarding systems and data models in the maritime domain with external specific technical skills from the industry, allows the dissemination of the state-of-the-art initiatives in the maritime domain to the industry and the familiarization of the experts from the public authorities involved with state-of-the-art technologies and best practices. It also allowed for achieving good and technologically cutting-edge products in a small period of time.

6 Recommendations

According to the independent external reviews, although the outputs of WP5 are technically sound, there is still room for improvement and concrete recommendations on how to proceed. Therefore, we recommend a Project follow-on activity to improve the present Project Deliverable following those recommendations.

Additionally, any specification has to be implemented so that it can be further enhanced in order to be adequate, therefore the outputs of WP5 should be put into practice as soon as possible.

CISE is an interoperability environment and the results delivered by WP5 are only a part of its technical and semantic level. Although there is already plenty of information regarding the architecture of CISE, it is now necessary to fill the gap between that work and the one being delivered by WP5, namely by defining a technical reference architecture for the CISE gateways and nodes, foreseen in the CISE architecture hybrid vision, which encompasses the services and data model herein specified.

More precisely, the present work has already anticipated the importance of the "Meta-Data" entity to embed data security features, and WSDL service specifications[9] have been developed for 3 components of the eventual CISE information exchange infrastructure: a Service Registry, a Timestamp-Authentication-Authorization Broker (TAAB) and the CISE Gateway. The key elements are there, but what remains missing is essentially the "information exchange management mechanism" safeguarding the legal restrictions of accessing (legal clearance, need to know) and possibly using (legal restrictions of purpose) some of the CISE data or services. Addressed by the WP4 activity, the definition of Access Rights is only available now at the end of the present phase to be considered and implemented by future work, also preferably encompassing the User Registry and Authentication functionalities.

The CISE has a long way to go, and most of its important components are as complex as innovative. Therefore, an iterative and incremental approach is recommended, so that an evolutionary solution can be built, learning from experience and minimizing the risk of failure. An "Agile" development scheme[10] should be considered, with short cycles of

9 **Cf OUT 5.2.6** section **5**

10 **Agile software development** is a group of software development methods based on iterative and incremental development, where requirements and solutions evolve through collaboration between self-organizing, cross-functional teams. It promotes adaptive planning, evolutionary development and delivery, a time-boxed iterative approach, and encourages rapid and flexible response to change. It is a conceptual framework that promotes foreseen tight interactions throughout the development cycle. (Wikipedia)

components development immediately followed by field validation, with several iterations until completion.

Last but not least, any solution needs to be governed, and the CISE is not an exception. We recommend to start a discussion about the governance of CISE as soon as possible and also to implement it in an iterative and incremental way.

7 Acknowledgements

All the work done by the members of WP5 was highly supported by several entities, to whom we hereby express our most sincere gratitude. In particular, we would like to emphasize our appreciation to the following:

To DG-MARE and its team, for providing the vision, direction and guidelines of the work to be developed, thus ensuring its alignment with the CISE framework and roadmap, so that our results could become a positive contribution towards a wider and longer effort in the field of European Integrated Maritime Surveillance, for which we are very proud of.

To the Finnish Border Guard and the project management team, for their pragmatism and for providing us the necessary support and flexibility, indispensable to deliver the expected results, while simultaneously overseeing all project's activities to ensure their synergies and coherence.

To other WP's leaders and members, for their support and collaboration in several activities. For their insights on our work, as our primary customers, but also for their availability for listening to our insights on their work. All WPs were pretty much interconnected, although some more than other, and their sound interaction was very fruitful, needed and appreciated.

To JRC and DIGIT's teams and external experts, for complementing and polishing our work with their priceless experience and knowledge, and for their unlimited availability and unconditional support.

Last but not least, to the Portuguese Directorate-General for the Maritime Policy (DGPM), the European Union Joint Research Center (JRC), the Finnish Transport Agency (FTA), the Baltic Marine Environment Protection Commission (HELCOM), the Spanish Guardia Civil, the European Union Satellite Center (EUSC) and the Romanian Border Guard, which kindly hosted our meetings, always exceeding our expectations and giving their best to support an important part of this endeavor, where most of the work was refined and consolidated, and also where the indispensable cohesion and team spirit was built.

Annex I – List of Acronyms and Abbreviations

AIS	Automatic Identification System
BLUEMASSMED	Pilot Project for the integration of Maritime Surveillance on the Mediterranean Area and its Atlantic Approaches
BMM	See BLUEMASSMED
BSMF	Baltic Sea Maritime Functionalities
CISE	Common Information Sharing Environment
CleanSeaNet	Near-real-time satellite-based oil spill and vessel monitoring service
CoopP	Cooperation Project Maritime Surveillance
CSDP	Common Security and Defence Policy
DG MARE	Directorate-General for Maritime Affairs and Fisheries
DGPM	Portuguese Directorate General for the Maritime Policy
DIGIT	Directorate-General for Informatics
EEA	European Economic Area
EMODnet	The European Marine Observation and Data Network
EMSA	European Maritime Safety Agency
ESA	European Space Agency
EU	European Union
EUROPOL	European Police Office
EUROSUR	European Border Surveillance System
EUSC	European Union Satellite Center
FP7	EU Seventh Framework Programme for research and technological development
FRONTEX	The European Agency for the Management of Operational Cooperation at the External Borders of the Member States of the European Union
FTA	Finnish Transport Agency
GMES	Global Monitoring for Environment and Security
HELCOM	Baltic Marine Environment Protection Commission
IA	Impact Assessment
IMDatE	Integrated Maritime Data Environment
INSPIRE	Infrastructure for Spatial Information in the European Community
ISA	Interoperability Solutions for European Public Administrations
JRC	Joint Research Centre
LRIT	Long-Range Identification and Tracking of ships system
MARSUNO	Pilot Project: Maritime Surveillance North
MSEsG	Member States Expert sub-Group
POV	Pre-Operational Validations
SafeSeaNet	Vessel traffic monitoring and information system

SAR	Search and Rescue
SEIS	Shared Environmental Information System
TAG	Technical Advisory Group
THETIS	Information system for the Port State Control inspection regime of ships
User Communities	Border control, maritime safety and security, fisheries control, customs, marine environment, general law enforcement and defence
VMS	Vessel Monitoring System
WP	Work Package

Maritime Surveillance

Annex II – Summary of Results and Outputs

The reaching of project objectives under the responsibility of WP 5 was measured with the following output and result indicators. In the following table is a summary of the delivered outputs and results with a reference to more specified results.

[**Outputs** are all the tangible results, milestones, and specific activities that were achieved, in order to complete the project. They directly result from the activities carried out in the project. They report on the main activities carried out during the project. They do not lead to a qualitative judgment on the project's outcomes.

In other words, it is not because the project organizes a high number of workshops that it will necessarily be successful. Output indicators are typically measured in physical units such as the number of meetings, seminars, site visits, conferences, participants, publications.

Results are direct and immediate effects resulting from the project and from the production of the outputs. They do not report on the 'what' but on 'why' the project is delivering the specific outputs. The organization of interregional events and meetings, the identification and dissemination of good practices, etc. are only means to an end. These activities are carried out in order to achieve specific effects that the result indicators should be able to assess and measure in quantified terms. Therefore, compared to the outputs, they imply a qualitative value. They also have to be measured in physical units such as the number of staff with increased capacity, the number of good practices successfully transferred or the number of policies improved.]

Output Indicators:	**Delivered Outputs**
Number and reports of meetings organized (working groups/sub-groups meetings, meetings between maritime authorities executing different maritime functions) and number of participants.	5 meetings were organized, hosted by different public authorities and agencies, in different member states, as follows: Meeting #1 – Hosted by the Directorate-General for the Maritime Policy (DGPM), this meeting was held in Lisbon from the 20th to the 21st of March 2013 (34 participants). Meeting #2 – Hosted by the JRC, this meeting was held in Ispra from th e29th to the 30th of May 2013 (21 participants). Meeting #3 – Hosted by FTA and HELCOM, this meeting took place in Helsinki on the 9/10 July 2013 (28 participants)

	Meeting #4 – Hosted by the Guardia Civil and EUSC, this meeting was held in Madrid from the 4th to the 5th of September 2013 (23 participants). Meeting #5 - Hosted by the Romanian Border Police, this meeting was held in Bucarest from the 13th to the 14th of November 2013 (25 participants). A presentation on the work package methodology and preliminary results has been conducted at the SEMIC13 international conference.
Results Indicators:	**Delivered Results:**
List of data sets defined in each information service. Target value corresponds to all data sets needed so that the information services are fully defined.	1 data model encompassing 18 core data entities which fulfill 100% of the data required by the use cases defined by WP2 and enhanced by WP5, and also 100% of the data required by the services defined by WP5
Information services (not required)	15 different services have been defined in natural language, to provide an initial technological support to the information exchanges required by the use cases developed by WP2. 5 different messaging patterns have been defined, based on the use cases developed by WP2, so that the information services developed could support the 'need to know' and the 'responsibility to share' paradigms and the information exchanges required by the use cases developed by WP2

Annex III – Enhanced use cases

This annex encompasses the following documents:

- CoopP WP5 expert support 258444 Deliverable OUT-5-1-1

Annex IV – Core vocabularies specifications

This annex encompasses the following documents:

- COREVocabulary_Action
- COREVocabulary_Agent
- COREVocabulary_Anomaly
- COREVocabulary_Cargo
- COREVocabulary_Document
- COREVocabulary_Event
- COREVocabulary_Incident
- COREVocabulary_Location
- COREVocabulary_Metadata
- COREVocabulary_Movement
- COREVocabulary_Object
- COREVocabulary_OperationalAsset
- COREVocabulary_Organization
- COREVocabulary_Period
- COREVocabulary_Person
- COREVocabulary_Risk
- COREVocabulary_UniqueIdentifier
- COREVocabulary_Vessel
- Usage examples
- CISECCTS
- COOPP Data Model (non printable)
- XSD (non printable)

- OWL (non-printable)

Annex V – Enhanced services definitions

This annex encompasses the following documents:

- Services definitions - UseCase25b FDM20131104
- Services definitions - UseCase37 FDM20131104
- Services definitions - UseCase57+70 FDM20131104

Annex VI – Services specifications

This annex encompasses the following documents:

- CISE SOA Patterns
- Choreography1-Pull_Synchronous
- Choreography2-Pull_direct_Asynchronous(delayed)
- Choreography3-Broadcast_Pull_Asynchronous
- Choreography4-PushBroadcast
- Choreography5-Push
- CISEGateway (non printable)
- CISEServiceRegistry (non printable)
- TAAB (non printable)

Annex VII – Business review

This annex encompasses the following documents:

- CoopP WP5 expert support 258444 Deliverable OUT-5-1-3
- CORE Vocabulary - Agent_final BG
- CORE Vocabulary - Document_final BG
- CORE Vocabulary - Location_final BG
- CORE Vocabulary - Metadata_final BG
- CORE Vocabulary - Movement_final BG
- CORE Vocabulary - OperationalAsset_final BG
- CORE Vocabulary - Organization_final BG
- CORE Vocabulary - Person_final BG
- CORE Vocabulary - Unique Identifier_final BG
- CoopP WP5 expert support 258444 Deliverable OUT-5-1-2

Annex VIII – Technical review and service model

This annex encompasses the following documents:

- Core Vocabularies Technical Review v1_(OUT5.2.4)_20131219
- CORE Vocabulary - Action_final_TC
- CORE Vocabulary - Agent_final_TC
- CORE Vocabulary - Anomaly_final_TC
- CORE Vocabulary - Cargo_TC
- CORE Vocabulary - Document_final_TC
- CORE Vocabulary - Event_final_TC
- CORE Vocabulary - Incident_final_TC
- CORE Vocabulary - Location_final_TC
- CORE Vocabulary - Metadata_final_TC
- CORE Vocabulary - Movement_final_TC
- CORE Vocabulary - Object_final_TC
- CORE Vocabulary - OperationalAsset_final_TC
- CORE Vocabulary - Organization_final_TC
- CORE Vocabulary - Period_final_TC
- CORE Vocabulary - Person_final_TC
- CORE Vocabulary - Risk_final_TC
- CORE Vocabulary - Unique Identifier_final_TC
- CORE Vocabulary - Vessel_final_TC
- SC technical review and service model
- Final Technical Review & Recommendations_(OUT5.2.6)_20140118

- CISE Service Model (non printable)
- Messaging Objects XSD (non printable)

Annex IX – Meeting minutes and list of participants

This annex encompasses the following documents:

- WP5 Meeting 1 – Minutes
- WP5 Meeting 1 - Participants 20MAR13
- WP5 Meeting 1 - Participants 21MAR13
- WP5 Meeting 2 – Minutes
- WP5 Meeting 2 – Participants
- WP5 Meeting 3 – Minutes
- WP5 Meeting 3 - Participants 09JUL13
- WP5 Meeting 3 - Participants 10JUL13
- WP5 Meeting 4 – Minutes
- WP5 Meeting 4 – Participants
- WP5 Meeting 5 - Minutes - State of play
- WP5 Meeting 5 – Participants
- WP5 Meeting 5 Agenda - 20131113