I-SEAMORE

D2.7 – MONITORING AND EVALUATION/VALIDATION FRAMEWORK

WP2 – I-SEAMORE Continuous SELP Landscape Assessment & Procedures Definition

Integrated surveillance ecosystem for European Authorities responsible for Maritime Operations leveraged by reliable and enhanced aerial support

Co-funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement 101073911

D2.7 – ANALYSIS OF CITIZENS' AWARENESS AND ACCEPTANCE OF EU MARITIME SECURITY SYSTEM

Due date	29/02/2024
Submission date	29/02/2024
Deliverable lead	ISIG
Version	1.0
Nature of the deliverable	Report
Dissemination level	PU – Public
Authors	Riccardo Laterza (ISIG), Olivia Ferrari (ISIG), Paola Lorenzoni (ISIG), Marina Andeva (ISIG)
Reviewers	Marius Budileanu (TS) Ionut Morcan (RBP)
Abstract	The deliverable defines the overall methodological framework for the implementation of monitoring and evaluation/validation activities, within Task T2.4 - Monitoring and Evaluation/Validation Framework. The report sets the basis of a feedback mechanism, able to generate and validate knowledge by means of validation and monitoring activities engaging end-users (both internal and external to the consortium) and stakeholders, by adopting a participatory approach (i.e., by means of focus-group, interviews, table-top exercises, etc.). The report provides a preliminary definition of monitoring and evaluation theoretical framework, criteria, strategy and action plan. The final version of the report, presenting the achieved framework for monitoring and evaluation will be submitted at M30 (June 2025).

DOCUMENT REVISION HISTORY

Version	Date	Description of change	List of contributors
0.1	02/02/2024	Table of content	Riccardo Laterza (ISIG)
0.2	19/02/2024	First draft	Riccardo Laterza (ISIG), Olivia Ferrari (ISIG)
0.3	23/02/2024	Consolidated draft for internal review	Riccardo Laterza (ISIG), Olivia Ferrari (ISIG), Marina Andeva (ISIG), Paola Lorenzoni (ISIG)
0.4	28/02/2024	Contribution by AEAT	Miriam Fernández Suárez (AEAT)
0.5	23/02/2024	Internal review	Marius Budileanu (TS) Ionut Morcan (RBP)
1.0	29/02/2024	Final version for quality check and submission Submitted version	Riccardo Laterza (ISIG), Olivia Ferrari (ISIG), Paola Lorenzoni (ISIG)

DISCLAIMER

This project has received funding from the European Union's Horizon Europe research and innovation program under the Grant Agreement 101073911. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

COPYRIGHT NOTICE

EXECUTIVE SUMMARY

Deliverable 2.7 "Monitoring and evaluation/validation framework" (hereinafter D2.7 or Deliverable) illustrates all the activities undertaken and the results achieved under the objectives of Task T2.4 "Monitoring and Evaluation/Validation Framework" of Work Package 2 "Continuous SELP Landscape Assessment & Procedures Definition".

This Deliverable is the first version of the I-SEAMORE Monitoring and Evaluation/Validation (M&E) Framework and sets the basis of a feedback mechanism, able to generate and validate knowledge by means of validation and monitoring activities engaging end-users (both internal and external to the consortium) and stakeholders, by adopting a participatory approach (i.e., by means of focus-groups, interviews, table-top exercises, etc.).

The deliverable is structured as follows:

- **Section 2**: provides an overview of the objectives covered by the Deliverable, including relevant relations established with other I-SEAMORE Tasks.
- **Section 3**: presents the methodology defined for the purposes of I-SEAMORE M&E activities, i.e., the theoretical framework of reference and its operational transposition.
- **Section 4**: outlines I-SEAMORE M&E Strategy, including the identification of indicators in line with the characteristics set in the methodology, and Data Collection and Data Evaluation processes. In this Section, tasks are assigned to I-SEAMORE project partners (i.e., I-SEAMORE M&E Task Force) and M&E preparation steps are outlined.
- **Section5**: concluding remarks and next steps in I-SEAMORE project activities.

This report will feed its outcomes to T6.1 and T7.3 for the preliminary and final evaluation activities foreseen with end-users and stakeholders. The final version of the report (D2.8), presenting the achieved framework for monitoring and evaluation will be submitted at M30 (June 2025).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

AB	Advisory Board
AI	Artificial Intelligence
DoA	Description of the Action
EO	Earth Observation
GA	Grant Agreement
KPI	Key Performance Indicator
M	Month
M&E	Monitoring and evaluation/validation
MOC	Maritime Operation Centre
P-FE	Principles-Focused Evaluation
RF	Radio Frequency
SOP	Standard Operating Procedures
SotA	State-of-the-art
T	Task
UAV	Unmanned aerial vehicle
UNDP	United Nations Development Programme
USV	Unmanned surface vehicle
UxV	Unmanned system
V&V	Verification & Validation
WP	Work Package

1 INTRODUCTION

1.1 Purpose of the Document

In its first version, Deliverable 2.7 presents the overall methodology for the implementation of monitoring and evaluation activities (M&E Activities) within I-SEAMORE project. In its second and final version, it will also provide an overview of the achieved framework for monitoring and evaluation activities.

The document provides for a tool for accompanying the development pathway of I-SEAMORE solutions during their whole design cycle, from drafting to piloting, so to assess their maturity level. It also provides for a consistent approach towards (internal and external) end-users' and stakeholders' engagement via a participatory approach.

1.2 Structure of the Document

The document is structured as follows:

- **Section 2** provides for a comprehensive overview of the objectives covered by the Deliverable, including relevant relations established with other I-SEAMORE Tasks.
- **Section 3** presents the methodology assembled for the purpose of I-SEAMORE M&E activities, i.e., the theoretical framework of reference and its operational transposition.
- **Section 4** defines the I-SEAMORE M&E Strategy, including the identification of indicators in line with the characteristics set in the methodology, and Data Collection and Data Evaluation processes. In this Section, tasks are assigned to I-SEAMORE project partners (i.e., I-SEAMORE M&E Task Force) and M&E preparation steps are outlined.
- Eventually, **Section 5** outlines concluding remarks and next steps in I-SEAMORE project activities.

Three Annexes are reported at the end of the document, namely: Annex A, providing a draft of the technical KPIs assessment questionnaire; Annex B, presenting the questionnaire addressed to end-users focusing on operational aspects linked to the adoption of the AI systems; Annex C, providing the Ethics questionnaire for testing, validation, and demonstration activities participants.

2 OBJECTIVES

2.1 Goals of M&E Structure

The purpose of M&E Structure is rooted in I-SEAMORE Grant Agreement. The M&E Structure covers all the four phases of the project (namely: Definition and Co-design Phase; Iterative Development Phase; Integration, Testing, Validation; Demonstration in realistic environment and end-users' final evaluation) contributing to two specific project objectives:

- O1: "Increase the level of cooperation between end-users, industrial and research players through the definition and implementation of co-design and co-creation methodologies and by involving external entities".
    - This objective is pursued by the M&E Structure by feeding the co-design and co-creation process with a monitoring and evaluation framework.
- O6: "Maximize the I-SEAMORE outreach, uptake and acceptance of results by end-users and stakeholders".
    - This objective is pursued by the M&E Structure via the final evaluation performed by the end-users on the project outcomes, which will result in the definition of a set of best practices and guidelines for future adoption of I-SEAMORE results.

Also, an effective M&E Structure should contribute to diminish some of the risks foreseen by the project, namely:

- Weak involvement of end-users: this risk can be mitigated by putting in place a co-creation and co-design approach and a dedicated framework for validation and evaluation activities. This should contribute to encouraging participation.
- Low level of precision of KPIs: the establishment of M&E solutions allows for an iterative approach for developing and updating KPIs throughout the project lifecycle.

2.2 Relations with other Tasks

I-SEAMORE Description of the Action (DoA) provides for relevant connections between the M&E activities and other Tasks foreseen in the project implementation. The following paragraphs briefly describe these connections, intended both as contribution coming from other Tasks (section 2.2.1), and contribution of M&E outputs/outcomes to other Tasks (section 2.2.2).

2.2.1 Contribution from other Tasks

The capability to correctly identify the M&E object of interest, as well as to choose the right means to guarantee the participation of end-users and stakeholders to the M&E process, and to guarantee a broader and open dialogue on the issues inquired through the M&E activities, relies on tasks previously implemented within I-SEAMORE project.

The first goal will be reached by leveraging on the outputs of T3.3 - Definition of Use Cases, Design of Operational Concepts and Elicitation of User Requirements & KPIs. Within this Task, potential use cases, design operational concepts, and the elicitation of stakeholder requirements and KPIs, are addressed. More specifically:

- Within T3.3.1, partners identified the potential use cases, that will be tested live in demonstration. The analysis following demonstration activities constitutes a relevant source of information for the M&E activities.

- Within T3.3.2, KPIs have been set, based on the outcome of the operational concepts endeavour. Benchmarks related with the KPIs constitute the base of the End-user validation process.
- T3.3.3 identified and prioritised users' requirements for border surveillance and operations in maritime environments. Verifying the achievement of these requirements constitutes a part of the M&E activities.

The second goal will be pursued by deploying the work done in T3.2 – Setup of Co-Design & Co-Creation Processes from Technological & Societal Perspectives. This Task aims to establish a framework and related methodology to enable open cooperation between technological organisations, maritime security practitioners, civil society representatives and other relevant stakeholders in co-designing and co-creating solutions for effective maritime and border security. Within this framework, participation of end-users and stakeholders will be leveraged also to implement the M&E process.

The third goal is encompassed within the activities foreseen by T8.2 - Stakeholders Engagement & Advisory Board Management. This Task involves the creation and mobilisation of a wide community of key stakeholders and experts within the Advisory Board (AB). The AB main aim is to facilitate a structured and informed dialogue among all stakeholders. In this sense, the GA mentions the necessity to establish a feedback loop so to:

- Map stakeholders' interests towards the development of I-SEAMORE solutions.
- Validate the results of the project and provide advice on how to improve the developed solutions, following the methodology defined in T3.2.

2.2.2 Contribution to other Tasks

The indications contained in this Deliverable, as well as the outcomes of its implementation, will in turn contribute to shape the activities carried out in other I-SEAMORE Tasks. This particularly applies to validation activities performed both by end-users and stakeholders, and to the drafting of policy recommendations, as follows:

- End-users Validation Activities will be the focus of T6.1. The project foresees two cycles of validation activities (at M18 and M28), making use of the M&E framework presented in this Deliverable, in order to ensure that all components are ready to move to the final demonstrations (WP7) and meet the requirements and KPIs initially established and updated within T3.3.
- Stakeholders Final Evaluation Activities will be performed in T7.3. I-SEAMORE outcomes will be inquired from the perspective of involved stakeholders, using the M&E framework presented in this Deliverable. Beside the M&E activities, participants are also expected to outline a set of best practices and guidelines for future adoption and commercial use of I-SEAMORE Ecosystem for Maritime Surveillance operations.
- Policy recommendations based on lessons learned will be the object of T2.5. End-users' policy recommendations will particularly benefit from the participatory analysis of the results of M&E activities, as the analysis of the results of Demonstration activities in the light of KPIs and users' requirements will define the lessons learned from the implementation of I-SEAMORE solutions.

3 METHODOLOGY

The following section provides for both Theoretical and Operational Frameworks, which in turn allow to understand the rationale behind the implementation of I-SEAMORE M&E activities. More specifically, the Theoretical Framework contains an overview of relevant concepts related to the issue of monitoring and evaluation activities, including previous examples of theoretical approaches to M&E in the field of design and implementation of technological solutions, then focusing on the Realistic Model as the most suitable model for the purposes of the project. As the name suggests, the Operational Framework provides for an operational translation of the theoretical concepts outlined in the previous paragraph to the field of I-SEAMORE project, recalling Use Cases and User Requirements related to the foreseen technological solutions, and presenting Phases, Framework, and Criteria applicable to I-SEAMORE M&E process.

3.1 Theoretical Framework

The scrutiny and assessment of any design activity through M&E activities are crucial for analysing its content, implementation, and impact. This process allows for a comprehensive understanding of the design's quality, significance, and effectiveness. Additionally, monitoring and evaluation play a pivotal role in ensuring the legitimacy of the identified solutions. It involves considering potential success, scrutinizing actual results and accomplishments, and conducting analyses within the context of implementation, such as the target social groups, i.e., stakeholders, and the social and economic environment. Mere compliance with existing rules and administrative frameworks is insufficient; a thorough examination of the performance and outcomes of the developed solution is essential.

3.1.1 UNDP (United Nations Development Programme) Official Definition

The UNDP defines M&E aim as «to provide the main parties with timely information about the progress, or lack thereof, in the production of outputs and the achievement of outcomes. This serves as a basis for decision-making to improve the performance of the programme or project and to feed into the learning processes» (2003, p. 18).

M&E activities are considered pivotal when it comes to publicly funded interventions and public bodies to:

- Verify the rationale of a public intervention.
- Identify reproducible successes and/or failures not to be repeated.
- Be accountable to citizens.

These objectives, in turn, encompass three primary purposes:

- **Cognitive aim**, which involves the necessity to assess the effects of interventions on society as objectively as possible and comprehend the underlying principles of their functioning.
- **Normative aim**, which is focused on aiding policy-makers and all stakeholders in forming judgments regarding the value of these interventions.
- **Instrumental aim**, which involves the imperative to significantly contribute to the enhancement of interventions by raising stakeholders' awareness of the meaning, conditions, and consequences of their actions and decisions.

According to this framework, a M&E process allows:

- To evaluate the merit of a developed action, which involves deciding whether to initiate, persist, expand, or curtail the action at hand, while ensuring stakeholders are informed about the decision and the rationale behind it.

- To enhance the aforementioned activity, by applying changes to its implementation process informed by the M&E process.
- To expand knowledge about the operational mechanisms of the action, by gaining insights into the processes of social change; this also means to indirectly validate the application of theoretical concepts that boosted the design of the implemented solution.
- To boost the learning capacity of end-users and stakeholders, through reflections on the actions undertaken and the resulting consequences, more particularly those within the co-design activity.

3.1.2 The Principles-Focused Evaluation (P-FE)

According to the Principles-Focused Evaluation (P-FE) approach (Patton, 2018), the purpose of any evaluation activity, informing broader frameworks of understanding and action, can be identified deploying several parameters:

- **Summative evaluation**: making overall judgments of merit, worth, and significance, supporting project development decisions.
- **Formative evaluation**: identifying strengths and weaknesses and providing feedback to improve the system development plan.
- **Developmental evaluation**: supporting innovation and adaptation in control environments, such as large deployments and demonstrations.
- **Knowledge generation** to extrapolate lessons to inform new initiatives and evaluation adjustments.

Following this approach, any evaluation system needs to be structured according to evaluation types, evaluation approaches, and evaluation roles. In the Operational Framework proposed by this Deliverable, evaluation types correspond to M&E Criteria (see Par. 3.1.3.2), evaluation approaches correspond to M&E Data Collection Tools (see Section 4), and evaluation roles are defined as M&E Evaluation Procedures (see Section 4).

3.1.3 The Realistic Evaluation Model

In order to better transpose the abovementioned theoretical concepts related to the issue of setting up M&E activities in a more operational field, the reality which has to be scrutinised should be understood according to a defined evaluation model. Evaluation models are systematic, organic, and cohesive sets of reflections that identify and isolate the fundamental and constitutive elements of the subject of evaluation and the expected interaction among these elements.

The realistic model is one of the most deployed evaluation models. It posits that the success of an intervention, such as the implementation of a technological solution, is dependent on the interplay between the intervention's mechanisms and the specific context in which it unfolds. Mechanisms are not isolated entities, as their interaction with the surrounding environment evolves as the implementation process discloses.

According to this model, mechanisms refer to the underlying theoretical assumptions from various disciplines (psychology, sociology, pedagogy, etc.) that validate the logic of the developed solution. They serve as the rationale grounding the content and design of the solution. The context comprises the characteristics of the technology's design and implementation, stakeholders and end-users attitudes and experiences, and the environment in which the implementation of the solution occurs.

In this sense, every M&E process should be intended as context-specific. The evaluation task encompasses an analysis of both the context and the outcomes. Evaluation criteria play a crucial role in this process, providing a framework for asking pertinent questions about the context-mechanisms-outcome

relationships. For instance, questions may focus on whether the level of social acceptance of a developed technology depends more on the initial level of stakeholders' perceptions and beliefs, rather than on the implementation process itself.

Operationally, the Realistic Evaluation Model involves:

- Developing the evaluation design: ideally, this should be prepared before the start of the intervention, preferably integrated into the intervention planning.
- Identifying criteria: carefully and analytically identifying features that adequately describe the context dimensions and outcome components.
- Identifying indicators: identifying indicators that empirically represent the identified dimensions and components and allow for proceeding with the data collection.
- Identifying relevant empirical references (i.e., data sources).
- Establishing the data collection design: considering all identified indicators and data sources in the design process and allowing for data collection as an ex-ante, in-itinere (i.e., monitoring), and ex-post activity.

3.1.3.1 URSO Paradigm

In line with the principles of the Realistic Evaluation Model, the Council of Europe (2018) proposes a set of four principles which serves as a way to orient both the identification of consistent M&E criteria, and their correct implementation. In the light of this paradigm, any M&E system should respond to the following principles:

- Being **Useful**, i.e., providing concrete tools to practitioners to monitor and evaluate the implementation of foreseen solutions both in single Use Cases and at a project level.
- Being **Relevant**, i.e., strengthening end-users capacity in managing the M&E system and setting up a comprehensive M&E strategy to be implemented beyond the project obligations.
- Being **Sustainable**, i.e., ensuring both the possibility of being correctly implemented (*internal sustainability*) and introducing a feedback mechanism aimed at improving existing programmes and policies, thus adapting the solutions to the everchanging context and technical and societal challenges (*external sustainability*).
- Ensuring **Ownership** by the practitioners that may wish to implement it, guiding end-users step-by-step in its implementation, and allowing adaptation to the different contexts addressed by the project. Ultimately, the ownership principle foresees the empowerment of end-users, allowing them to apply the M&E scheme independently.

3.1.3.2 Realistic Evaluation Criteria

The ability to form a judgment through a Monitoring and Evaluation (M&E) process is enabled through the selection of various evaluation criteria. It is preferable to explicitly define these criteria prior to initiating the action, linking them with the establishment of a specific standard level for one or more goals. These chosen evaluation criteria serve as concrete guides for the M&E activity, facilitating the identification of processes to observe and indicators for analysis. Essentially, the criteria function as both theoretical and operational foundations necessary for evaluating a set of collected data, allowing for a comprehensive assessment of the intervention/adopted solution. Each criterion sheds light on specific aspects of the situation under investigation; this should not put in the background other, perhaps unexpected, features of the situation as it can be observed during the implementation of the foreseen solutions, which might be not fully covered by the selected indicators.

The following table (adapted from Council of Europe, 2018) provides a summary of some widely used evaluation criteria in the field of public policies, which have to be considered as possible general criteria for the development of criteria specifically applicable to the field of action of I-SEAMORE project:

TABLE 1: LIST OF PUBLIC POLICIES EVALUATION CRITERIA

Criteria	Criteria description in the *ex-ante* evaluation	Criteria description in the *ex-post* evaluation
Input	• Resources allocated for the intervention.	• Resources actually used to implement the intervention.
Output	• Number of programmed activities. • Number of hypothetical beneficiaries of the intervention.	• Number of implemented activities. • Number of effective beneficiaries of the intervention.
Results	• Expected results for the beneficiaries of the intervention. • Expectation of change in the context in which the intervention is implemented.	• Results obtained for the beneficiaries of the intervention. • Change produced in the context in which the intervention is implemented.
Efficiency	• Hypothetical ratio between costs and output. • Hypothetical ratio between costs and outcomes.	• Final ratio between costs and outputs. • Final ratio between costs and outcomes.
Efficacy	• Foreseen capacity of the intervention to produce results.	• Results produced by the intervention.
Impact	• Overall amount of foreseen – positive, negative, intended and unintended – effects in the context in which the intervention is implemented.	• Overall amount of – positive, negative, intended and unintended – effects produced in the context in which the intervention is implemented.
External consistency	• Logical consistency between the intervention goals and the goals of the policies which are supposed to be improved.	• Logical consistency between the goals reached through the intervention and the goals of the policies which were supposed to be improved.
Internal consistency	• Logical consistency among the programme foreseen activities.	• Logical consistency among the programme implemented activities.
Adequacy	• Foreseen level of coverage of the stakeholders' needs/interests.	• Reached level of coverage of the stakeholders' needs/interests.
Compliance	• Expected capability of the intervention to satisfy the beneficiaries expectations.	• Capability of the intervention to satisfy the beneficiaries expectations.
Reliability	• Dependability of the responsible for the implementation of the intervention.	• Proved dependability of the responsible for the implementation of the intervention.

Criteria	Criteria description in the *ex-ante* evaluation	Criteria description in the *ex-post* evaluation
Sinergy	• Potential added value which can be produced through the coordination among different sectors/actors.	• Added value produced through the coordination among different sectors/actors.
Process	• interactions among context, input and output which are expected to cause the expected outcomes.	• Occurred interactions among context, input and output which caused the effective outcomes.
Transferability	• Replicability of the intervention in other contexts.	• Replicability of the intervention in other contexts.

3.2 Operational Framework

In this Section, the general indications provided in the previous dissertation are translated into the field of I-SEAMORE project so to define the Operational Framework of the project M&E activities. After presenting some general considerations regarding its overall architecture, the I-SEAMORE M&E Operational Framework takes into account three complementary fields where to perform adapt the I-SEAMORE M&E approach:

- Technical M&E
- Societal M&E
- Ethical M&E

3.2.1 I-SEAMORE M&E General Architecture

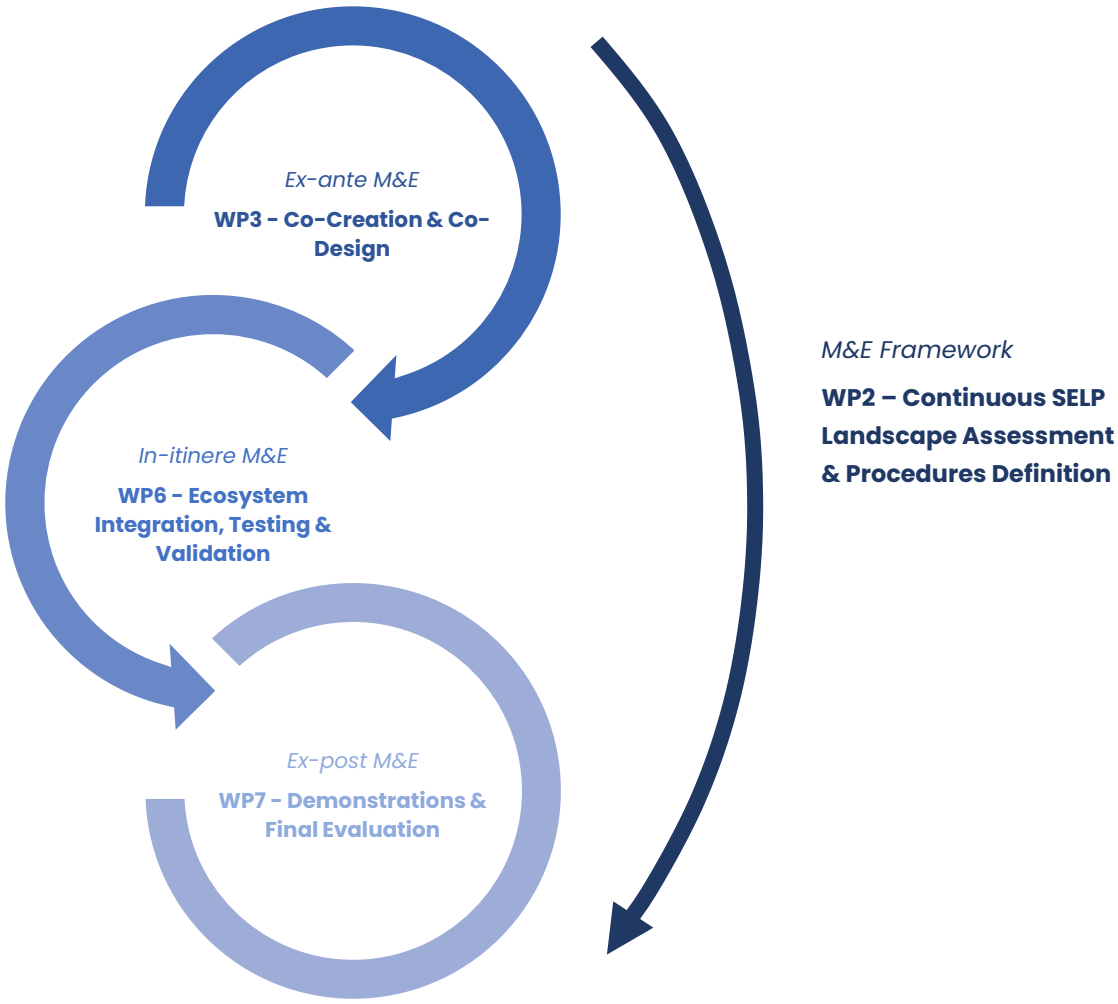
I-SEAMORE M&E process should be seen as a set of activities which accompanies the project throughout all its development phases, as follows:

- **Ex-ante activities**: during the design phase, forecast studies can be conducted to anticipate the potential impacts and outcomes resulting from the pilot action implementation. Ex-ante activities also involve assessing the existing situation and evaluating the current state of the art – for instance, in terms of technologies already available on the market, or in terms of regulatory provisions, before the piloting begins. This check can be extended towards expert analysis and validation of specific technological features that can be assessed prior the beginning of Testing/Demonstration activities, so to verify the compliance of the actual developed solution with technical minimum requirements stated before the implementation.
- **In itinere activities**: throughout the implementation phase, ongoing monitoring is carried out to compare the initial results with the original objectives of the pilot action. These activities should be organised to be effectively carried out during Testing activities, to continuously or discretely monitor features related to the behaviour of implemented technologies as the scenario in which the Testing is held evolves. This monitoring also allows for the assessment of unintended consequences, both positive and negative, with the primary aim of providing feedback to those executing the action and the stakeholders involved.
- **Ex-post activities**: when the piloting activities are concluded, an assessment can be conducted based on the identification and quantification of outputs and outcomes. This assessment involves analysing the pilot action according to predetermined criteria, for instance benchmarks, or more simply improvement in respect to the SotA. The data collected during this phase facilitate a

comparison with the initial data, enabling an evaluation of the overall results (i.e., outputs) and how the context has changed as a result of the pilot action (i.e., outcomes of the Testing/Demonstration activity). The latter can also be interpreted as a forecasting exercise regarding the formulation of scenarios where the piloted technology is operationally applied on a broader scale. As the Demonstration ends and all in-itinere data is collected, post-demonstration assessment might also involve the collection of additional data, deriving from further analysis and assessment activities conducted together with end-users who directly experienced, or even managed, the testing activities. Quantitatively speaking, data can be collected through dedicated monitoring devices. Qualitatively speaking, end-users' contribution can be enhanced through specific user-oriented demonstration sessions. Other means of assessment are the administration of surveys or interviews. In a longer temporal horizon, the results of the Demonstration activities can be used as a basis for further critical review by a broader range of actors, end-users, and stakeholders. This can be usually done through dedicated meeting such as workshops, focus groups etc. Common themes, weak points, areas of excellence and strength, and perspectives for further improvement, can be identified through this kind of activities, which are also of interest for the formulation of recommendations.

Within the framework of I-SEAMORE project, different WPs contribute to cover all the phases of a complete M&E process. The following figure represents this combination:

FIGURE 1: I-SEAMORE M&E GENERAL ARCHITECTURE



3.2.1.1 Definition of I-SEAMORE Use Cases

Use Cases for I-SEAMORE technological solutions are presented in Section 2 of D3.3 – Operational Concepts, KPIs and User Requirements. This paragraph briefly presents the main contents of that chapter for the purpose of the definition of I-SEAMORE M&E Operational Framework.

The Use Cases identified within the project are related to two complex and multifaceted issues that affect the European territory: irregular migration and drug smuggling. Both challenges require enhanced collaboration among authorities, improvements in data sharing and data elaboration, especially in relation to maritime surveillance. In this sense, the setting up of practical scenarios would help in evaluating the efficacy and effectiveness of I-SEAMORE solutions vis-à-vis these challenges.

The following tables summarise the main aspects of the two issues mentioned above, as well as the response provided by I-SEAMORE, which will be the object of M&E activities so to validate the efficacy and effectiveness of foreseen solutions.

TABLE 2: USE CASE NO. 1 – IRREGULAR MIGRATION

IRREGULAR MIGRATION	
Issue in a nutshell	Irregular migration across EU borders has surged by 64% in 2022[1], indicating a growing concern and the need for proactive measures to detect and mitigate unauthorized border crossings, exploited by organized crime groups. Enhanced collaboration among States is imperative to address this challenge effectively.
I-SEAMORE Objective	Detection of unauthorised border crossings and enhancement of timely response by operators, by leveraging advanced technologies, data analysis, and improved cooperation mechanisms.
I-SEAMORE Use Case Scenario	The use case focuses on detecting irregular migration near the maritime borders of Spain and/or Portugal. It involves a scenario where a migrant vessel attempts to evade detection by not reporting AIS signals, exploiting foggy conditions and low sunlight. Advanced technologies like radar, thermal imaging, data analytics, and pattern recognition algorithms are employed to enhance border authorities' capabilities in identifying and intercepting such vessels. Specialised sensors and situational awareness tools are crucial in optimising detection in challenging maritime environments.
I-SEAMORE Tested Solution	The PUD UAV 'ONE 150' drone detects RF signals indicating the presence of humans in a suspicious maritime corridor, and I-SEAMORE's services confirm the vessel as carrying migrants, triggering an alert. As the MOC performs a comprehensive assessment of the situation, the Mission Commander directs the UAV to lower altitude for monitoring, combining the activity with EO data analysis and coordinates interception efforts with relevant authorities. Authorities intercept the vessel, while the UAV and ECA's USV 'INSPECTOR MK2' provide ongoing support and gather valuable intelligence.
Excellence	Advanced technologies ensure effective detection and response to irregular migration, safeguarding both migrants and communities while upholding border control integrity.

TABLE 3: USE CASE NO. 2 - SMUGGLING OF DRUGS

SMUGGLING OF DRUGS	
Issue in a nutshell	The European Union is actively combating illicit drug trafficking by sea, aiming to enhance criminal intelligence and coordinated police action in international waters.
I-SEAMORE Objective	Fusion of heterogeneous vessel information in order to detect anomalies, risks, and threats associated with the smuggling of illicit goods through persistent surveillance methods. This will address the main challenges nowadays faced by end-users, namely: a) obtaining early-warning for suspicious vessels b) enhancing cooperation and coordination in the utilisation of UxVs c) improving the technology system installed on UxVs, especially those related with potentiated operability in nocturnal and low visibility conditions.

[1] https://www.frontex.europa.eu/media-centre/news/news-release/eu-s-external-borders-in-2022-number-of-irregular-border-crossings-highest-since-2016-YsAZ29

SMUGGLING OF DRUGS	
I-SEAMORE Use Case Scenario	The scenario involves the smuggling of narcotics from Morocco to Europe in the waters of southern Iberian Peninsula, focusing on nighttime illicit activities with small, fast boats interacting with merchant vessels.
I-SEAMORE Tested Solution	The I-SEAMORE solution responds to suspicious behaviour by initiating coordinated actions. It analyses AIS data to identify a merchant vessel deviating from its route, triggering a suspicious activity alert for the Mission Commander. UAVs are deployed to survey the area and detect a smaller boat approaching the vessel. Additional UAVs and USVs collaborate to monitor the event, providing data for identifying illicit cargo. Manned means are activated for interdiction, with Portuguese authorities pursuing suspects, seamlessly transferring control to Spanish authorities if needed.
Excellence	This cooperative approach ensures an effective response, enabling uninterrupted monitoring and tracking across jurisdictional boundaries.

3.2.1.2 Technical and Societal Aspects of I-SEAMORE M&E Process

The following table builds on the M&E Phases outlined in the previous section, by adding specific indications on tools and procedures that can be enabled both in relation to technical and societal/ethical aspects related to the I-SEAMORE technologies. The table provides a comprehensive framework of the available instruments which can contribute in developing a thorough M&E approach:

TABLE 4: TECHNICAL AND SOCIETAL/ETHICAL ASPECTS OF M&E PROCESS

	Technical aspects	Societal/Ethical aspects
Ex-ante (WP3)	• Evaluation of the current state of the art (i.e. available technologies, constraints, legal framework etc.) – *review factsheet* • Definition of benchmarks • Expert analysis and validation of specific technological features prior the beginning of Demonstration activities - *checkboxes*	• Forecast studies to anticipate the potential impacts and outcomes – *workshop activities with end-users and/or stakeholders* • Preliminary assessment of the perception of foreseen technologies by end-users and stakeholders - *survey*
In-itinere (WP6)	• Ongoing technological monitoring – data collection through monitoring devices, observation • Assessment of unintended consequences – critical observation, review factsheet	• Identification of potential societal and ethical issues – *survey,* workshops/interviews with stakeholders and end-users • Assessment of unintended consequences – *critical observation, review factsheet*

	Technical aspects	Societal/Ethical aspects
Ex-post (WP7)	• Identification and quantification of outputs and outcomes – *survey and checkboxes* • Forecasting exercise on potential impacts of broader application of the technology – *workshop* • Formulation of recommendations - *report*	• Collection of additional data through ex-post assessment activities targeting end-users and stakeholders – *survey, workshop* • Further critical review of the obtained results – *workshop* • Formulation of recommendations - *report*

3.2.2 I-SEAMORE Technical M&E Operational Framework

3.2.2.1 I-SEAMORE User Requirements

User Requirements are presented in Section 5.5 of D3.3 - Operational Concepts, KPIs and User Requirements. This paragraph briefly presents the main contents of that chapter for the purpose of the definition of I-SEAMORE M&E Operational Framework.

Within D3.3, User Requirements are grouped into the areas mentioned below. All User Requirements were prioritised (as high, medium, and low priority) as a result of a workshop conducted with the participation of end-users and other participants.

FIGURE 2: USERS' REQUIREMENTS

Id	Title	Description	Source	Assigned Priority
I-SEAMORE.01	AIS Near Real-time Data Analysis	System shall create alerts without human intervention analysing AIS information.	DoA, Use cases, output Co-creation design	H
I-SEAMORE.02	Object detection based on sensor data	The system may detect small boats using sensor data (e.g., Radar Sensors, Optical Sensors, Satellite Sensors, Acoustic Sensors).	DoA, Use cases, output Co-creation design	L
I-SEAMORE.03	Anomaly detection	The system shall detect abnormal behaviour (e.g., change of speed, change of course, entering a prohibited area, rendezvous) without human intervention (in near Real-time).	DoA, Use cases, output Co-creation design	H
I-SEAMORE.04	Integration of multiple data sources	The system shall perform integrate data from different sources (AIS, sensor data, images from UV and Satellite information) to allow a better detection process.	DoA, Use cases, output Co-creation design	H
I-SEAMORE.05	Automation of routine tasks	The system should generate alerts using pre-defined rules identified by end-uses.	Use cases, output Co-creation design	M

Id	Title	Description	Source	Assigned Priority
I-SEAMORE.06	Enhance situational awareness	The system shall contribute to maritime picture.	DoA, Use cases, output Co-creation design	H
I-SEAMORE.07	Processing large amount of data	The system shall provide near real-time analysis of large amounts of data generated by maritime surveillance systems.	DoA, Use cases, output Co-creation design	H
I-SEAMORE.08	Vessel monitor	The system should perform automatic identification and tracking of vessels.	DoA, Use cases	M
I-SEAMORE.09	Real-time Data Sharing	The system shall provide connectivity to CISE data stream.	DoA, Use cases	M
I-SEAMORE.10	Service Configurability	The I-SEAMORE system should configurable and deployment at end-user Operational Centres.	DoA, Use cases	M
I-SEAMORE.11	Annotations tool	The system should provide an annotation tool to store data for the debriefing tool.	Workshop	M
I-SEAMORE.12	Visualization tool	The I-SEAMORE system should provide a visualization tool to be used command centre.	Workshop	M
UaV.01	I-SEAMORE UAV Module	I-SEAMORE shall include UAV operation module.	DoA, Use cases, output Co-creation design	H
UaV.02	Search and rescue operations	I-SEAMORE shall include UAV operation. UAV quickly search large ocean areas for survivors or debris, reducing response times and increasing the likelihood of a successful rescue.	DoA, Use cases, output Co-creation design	H
UaV.03	Inspection and maintenance	I-SEAMORE Drones shall perform inspection maritime infrastructure, such as, offshore platforms, and ships, reducing the need for human operators to work in hazardous environments.	DoA, Use cases, output Co-creation design	H
UaV.04	Object Detection for Monitoring of illegal activities	I-SEAMORE Drones shall have the ability to detect objects (until 4 meters long) in the maritime environment. This is used to monitor illegal activities in the maritime environment, such as smuggling, illegal fishing, and piracy, providing near real-time intelligence that can support rapid response and enforcement operations.	DoA, Use cases, output Co-creation design	H

Id	Title	Description	Source	Assigned Priority
UaV.05	Multi-asset mission	The system shall have the ability to support multi-asset mission planning and execution (e.g., UAVs, USVs).	DoA, Use cases	N/A
UaV.06	Infrared Vision	The system should include Unmanned Surface Vehicles (USVs) with infrared vision for nocturnal detection is designed to enhance the USVs' ability to detect and track objects or activities during low light conditions, such as at night or in low light environments. This infrared vision should allow the USV to identify and track objects that may not be visible using conventional cameras.	DoA, Use cases	H
UaV.07	Augmented SIGINT Capabilities	The system shall provide UAVs with augmented SIGINT capabilities (e.g. receivers and antennas capable of intercepting a wide range of frequencies, Onboard Processing, Secure Data Transmission and Resilience to Electronic Warfare).	DoA, Use cases	H
UaV.08	UAVs with multiple PTZ	The system shall provide UAVs with multiple PTZ for augmented maritime monitoring.	DoA, Use cases	H
UaV.09	Debriefing	I-SEAMORE shall have the ability to analyse the data collected during surveillance operations.	DoA, Use cases, Workshop	H
UaV.10	Increase data range transmission	I-SEAMORE Drones should act as a relay to increase data range transmission in maritime surveillance operations.	Workshop	M
SatData.01	I-SEAMORE Copernicus satellite module	The I-SEAMORE system shall integrate Copernicus satellite data for maritime surveillance.	DoA, Use cases, output Co-creation design	H
SatData.02	I-SEAMORE Copernicus satellite data store	The I-SEAMORE system shall store Copernicus satellite data.	DoA, Use cases, output Co-creation design	H
SatData.03	I-SEAMORE Copernicus satellite data analytics	The I-SEAMORE system shall process Copernicus satellite data to create knowledge.	DoA, Use cases, output Co-creation design	H
SatData.04	Web Interface for Copernicus satellite data	I-SEAMORE should provide a web interface or the integration with the existing C2C system.	DoA, Use cases, output Co-creation design	M

3.2.2.2 Technical Key Performance Indicators (KPIs) – T.3.3

T3.3 defined a broad range of KPIs aimed at measuring the success and effectiveness of the I-SEAMORE solution in comparison to existing state-of-the-art (SotA) approaches. KPIs are here re-organised as grouped within KPI dimensions, and they are associated with means of verification which comply with the characteristics of the available tools and instruments presented above.

Currently, the KPIs envisioned as a result of T3.3 activities encompass a comprehensive set of metrics to evaluate the effectiveness, efficiency, and user satisfaction in relation to I-SEAMORE solutions. KPIs are here re-elaborated and organised so to be in line with URSO principles as quoted in Par. 3.1.3.1.

Foreseen KPIs cover the following KPI dimensions:

- Efficacy/Efficiency: encompasses aspects like efficacy, efficiency, and responsiveness, emphasizing quality data gathering, integration of diverse data sources, timely threat detection, and optimisation of resource allocation.
- Operational Capability, Range and Coverage: focuses on the range, coverage, and endurance of unmanned assets.
- Ecosystem Capabilities: address the development of the environment of dedicated services and tools.
- Interoperability: evaluates the seamless integration between different assets and data sources, promoting effective collaboration and information sharing.
- Safety and Security: ensures the system's ability to monitor and report unsafe situations.
- Users' Satisfaction: gathers feedback from end users and stakeholders to assess their overall satisfaction with the I-SEAMORE solution.

The following table presents the identified KPIs, clustered per above mentioned KPI dimension and related means of verification:

TABLE 5: I-SEAMORE TECHNICAL KEY PERFORMANCE INDICATORS

KPI Dimension	KPI	Description	Means of verification
Efficacy/ Efficiency	Quality and efficiency of data collection	This KPI involves integrating data from a variety of sensors, including radar systems, AIS, GPS, cameras, and communication devices operating on digital and analogic radio frequencies. These sensors are crucial for providing comprehensive situational awareness, especially in detecting small boats at night and mapping their routes.	Data collection through: • EO/IR Systems • Radar Systems • LIDAR Systems • GRF Systems Post-Testing/Demonstration survey and/or checkboxes
	Rate of efficiently performed missions	This KPI evaluates the system's efficiency by quantifying the number of missions that have been successfully executed. It considers the validity of the collected data, taking into account various factors that	Data collection through testing observation Post-Testing/Demonstration survey and/or checkboxes

KPI Dimension	KPI	Description	Means of verification
		could impact mission performance, including recorder failure, disk data corruption, communication failure, electric power failure, or sensor failure.	
	Data Fusion Performance	This KPI assesses the performance of data fusion algorithms and techniques in effectively integrating diverse data from satellites, UAVs, and USVs. The objective is to enhance situational understanding and facilitate decision-making processes. By evaluating the performance of data fusion, the project can ensure the seamless integration of information from multiple sources, leading to improved situational awareness and more informed decision-making.	Data collection through testing observation Post-Testing/Demonstration survey and/or checkboxes
	False Positive Rate	This KPI measures the frequency of false alarms or false positive detections produced by the system. The objective is to minimize unnecessary alerts and enhance operational efficiency.	Data collection through testing observation Post-Testing/Demonstration survey and/or checkboxes
	Detection Accuracy	This KPI measures the accuracy of the I-SEAMORE in detecting and identifying suspicious vessel behaviour or illicit activities, such as irregular migration or drug smuggling. It encompasses the detection of various indicators, including changes in speed, course, entry into prohibited areas, and rendezvous activities.	Data collection through monitoring of the detection accuracy in two critical scenarios: • Detection of Small Ships without AIS • Detection of Suspicious Behaviour in Merchant Vessels with AIS Post-Testing/Demonstration survey and/or checkboxes
	Response Time	This KPI measures the system's ability to detect suspicious behaviour and promptly respond to potential threats or suspicious events. It encompasses:	Data collection through assessment of response time (i.e., data transmission and processing targeting the MOC), alternatively via: • Real-time data collection and analysis

KPI Dimension	KPI	Description	Means of verification
		• the duration between the actual occurrence of an event and its first detection by the system. • the time taken by the system to identify and alert the Maritime Operation Centre (MOC) about these incidents.	• Timely communication protocols Post-Testing/Demonstration survey and/or checkboxes
	Resource Utilisation	This KPI assesses the efficiency of resource allocation and utilization, with a focus on optimizing the deployment of unmanned assets to maximize operational effectiveness while minimizing costs and resource wastage. It evaluates how effectively resources are allocated and utilized to achieve the desired outcomes, ensuring that the right assets are deployed in the right place at the right time.	Post-Testing/Demonstration Assessment through survey and/or focus group
Operational Capability, Range and Coverage	Altitude Range and Coverage	This parameter describes the altitudes at which each drone can operate and the corresponding coverage it can provide. By understanding the operational altitude capabilities, the project can determine the areas that can be effectively monitored and surveilled by each drone.	Data collection through assessment of technical characteristics during testing Post-Testing/Demonstration Assessment through survey and/or focus group
	Endurance	Endurance refers to the duration for which a drone can operate before its energy or resources are depleted. Assessing the endurance capabilities of each drone is crucial for assigning missions and understanding any limitations that may impact the duration of its operations. This information is essential for effective mission planning and resource allocation.	Data collection through assessment of technical characteristics during Testing/Demonstration Post-Testing/Demonstration Assessment through survey and/or focus group

KPI Dimension	KPI	Description	Means of verification
	Geographical Coverage	The combination of altitude range, endurance, and the technical capacity of sensors determines the geographical coverage of each drone. By evaluating these parameters, the project can determine the areas that can be effectively monitored and surveilled by each drone, ensuring comprehensive coverage of the target areas.	Data collection through assessment of technical characteristics during Testing/Demonstration Post-Testing/Demonstration Assessment through survey and/or focus group
	Power Consumption Improvement	The I-SEAMORE system aims to achieve a minimum of 15% improvement in power consumption for drones compared to previous technologies. Additionally, reducing the acoustic signature of the drones is crucial for minimizing noise disturbances and maintaining covert operations. This factor contributes to the optimization of operational efficiency and the overall effectiveness of the system.	Data collection through assessment of improved features during Testing/Demonstration Post-Testing/Demonstration Assessment through survey and/or focus group
	Acoustic Signature	Reducing the acoustic signature of the drones is crucial for minimizing noise disturbances and maintaining covert operations. This factor contributes to the optimization of operational efficiency and the overall effectiveness of the system.	Data collection through assessment of improved features during Testing/Demonstration Post-Testing/Demonstration Assessment through survey and/or focus group
	Processing Capability	Each drone must have its own processing capability to handle the data collected during missions. This ensures efficient data processing and analysis onboard the drone, reducing the dependency on external processing resources and facilitating real-time decision-making.	Data collection through assessment of improved features during Testing/Demonstration Post-Testing/Demonstration Assessment through survey and/or focus group

KPI Dimension	KPI	Description	Means of verification
	Drone Location and Navigation	Location and navigation is essential to have a high level of certainty regarding the location of drones at all times.	Data collection through verification, prior to the Testing/Demonstration, that at least two drones within the system utilize the European Geostationary Navigation Overlay Service (EGNOS)
	Flights in adverse weather conditions	It is necessary to know the technical characteristics related to the meteorological conditions that drones could withstand. It is important to evaluate and know the minimum meteorological conditions that allow for a truly operational flight. Weather conditions affect the operation of the equipment and consequently the operation.	Assessment in scenario. Collect meteorological data on the day of use case and every time the system is tested.
Ecosystem Capabilities	Visual Analytics Tools	The project aims to develop and utilize advanced visual analytics tools that enable effective data exploration, visualization, and analysis. These tools enhance the understanding of complex data sets and support decision-making processes by providing intuitive visual representations and actionable insights.	Post-Testing/Demonstration Assessment through survey and/or focus group
	Modernised C4 Platform	The Command, Control, Communications, Computers, and Intelligence (C4I) platform utilized in the project will be modernized to meet the specific operational requirements of I-SEAMORE. This platform serves as the central hub for information management, communication, and coordination among the system components and operators, ensuring efficient and seamless operation.	Post-Testing/Demonstration Assessment through survey and/or focus group
	Dedicated UxV's Mission Planner	To optimize mission planning for unmanned assets (UxV), a dedicated mission planner will be developed. This tool will enable operators to plan and	Post-Testing/Demonstration Assessment through survey and/or focus group

KPI Dimension	KPI	Description	Means of verification
		schedule missions, define waypoints, and set specific parameters for each UxV, considering factors such as mission objectives, operational constraints, and available resources.	
	Mission Debrief Module	The mission debrief module provides a post-mission analysis and evaluation capability. It allows operators to review mission data, analyse performance metrics, identify areas for improvement, and extract valuable insights for future mission planning and optimization.	Post-Testing/Demonstration Assessment through survey and/or focus group
	Dedicated Modeling and Simulation (M&S)	The project will utilize dedicated modelling and simulation tools to simulate and evaluate different operational scenarios. This allows for testing and refining system components, algorithms, and operational procedures in a virtual environment before their deployment in real-world situations. Modelling and simulation enable the identification of potential challenges, performance optimizations, and the development of robust operational strategies.	Post-Testing/Demonstration Assessment through survey and/or focus group
Interoperability	External systems feeding	A set of customized services and tools, fed by at least three external systems, will ensure the successful interoperability of the system.	Post-Testing/Demonstration Assessment through survey and/or focus group
	Interoperability and seamless integration	This KPI assesses the degree of interoperability and seamless integration between various unmanned assets and data sources, facilitating effective collaboration and information sharing among stakeholders.	Post-Testing/Demonstration Assessment through survey and/or focus group

KPI Dimension	KPI	Description	Means of verification
	Ground Control Station (GCS) connection	It is essential to ensure that the drones are equipped with a Ground Control Station (GCS) capable of communicating with the I-SEAMORE system.	Post-Testing/Demonstration Assessment through survey and/or focus group
Safety and Security	Corrupt or faulty data rate	Rate of corrupt or faulty data due to sensor(s) failure onboard, which may compromise the accuracy and reliability of collected data.	Data collection through monitoring during Testing/Demonstration Post-Testing/Demonstration Assessment through survey and/or focus group
	Drone failure rate	Rate of drone failure that prevents the completion of the mission, requiring immediate attention to mitigate any potential risks.	Data collection through monitoring during Testing/Demonstration Post-Testing/Demonstration Assessment through survey and/or focus group
	System failures rate	Rate of system failures, such as issues with the measurement algorithm, detection logic, or software configuration, which can impact the system's overall performance and effectiveness.	Data collection through monitoring during Testing/Demonstration Post-Testing/Demonstration Assessment through survey and/or focus group
	Undesirable outcomes	Undesirable outcomes resulting from operator actions, necessitating corrective actions to ensure the integrity of the mission and the system.	Data collection through monitoring during Testing/Demonstration Post-Testing/Demonstration Assessment through survey and/or focus group
Users' Satisfaction	Satisfaction related to usability	Evaluation of the ease of use and user-friendliness of the I-SEAMORE solution. This KPI assesses factors such as intuitive interfaces, clear instructions, and efficient workflows to ensure that users can effectively navigate and utilize the system.	Post-Testing/Demonstration Assessment through survey and/or focus group
	Satisfaction related to reliability	Examination of the system's reliability in terms of its performance and stability. This KPI assesses the frequency of system downtime, responsiveness, and the occurrence of any technical issues or glitches that may impact user satisfaction.	Post-Testing/Demonstration Assessment through survey and/or focus group

KPI Dimension	KPI	Description	Means of verification
	Satisfaction related to overall Performance	Analysis of the overall performance of the I-SEAMORE solution in meeting the specific operational needs of end users and stakeholders. This KPI evaluates whether the solution effectively addresses their requirements and provides the expected outcomes.	Post-Testing/Demonstration Assessment through survey and/or focus group

The above presented KPIs are operationally translated into potential questions for I-SEAMORE solution in-itinere and ex-post assessment in Annex A.

3.2.3 I-SEAMORE Social and Ethical M&E Operational Framework

3.2.3.1 M&E on SELP Concerns

Within I-SEAMORE, the monitoring and evaluation of ethical and societal aspects is ensured by the adoption of the principles and tools presented in D2.1 – Periodic Report on SELP Concerns.

Monitoring and assessment tools to be adopted by all project partners are listed below (for a more detailed description of each tool please refer to D2.1):

FIGURE 3: SELP MONITORING AND ASSESSMENT TOOLS

The results of the assessment and monitoring activities linked to the adoption of I-SEAMORE ethical principles and requirements will be reported in:

- Deliverable "Periodic report on SELP concerns" (D2.3) at M15, March 2024.
- Deliverable "Periodic report on SELP concerns" (D2.5) at M30, June 2025.

3.2.3.2 M&E on Societal Aspects

The assessment of societal aspects will be covered by the activities performed in Task 2.2 (and described in D2.2 – Analysis of citizens' awareness and acceptance of EU maritime security system), namely:

- **I-SEAMORE workshops and in-depth interviews**: I-SEAMORE workshops and in-depth interviews are aimed to provide a common platform and opportunities for structured debate for project

partners and relevant stakeholders to increase participation/engagement of the latter on specific policy topic, thus ensuring that aspects like protection of fundamental rights and freedoms of citizens, or other social and ethical issues, are taken into consideration along the whole project. The first workshop will be organized in 2024.

- **I-SEAMORE citizens' survey**: I-SEAMORE citizens' survey is intended to investigate aspects linked to citizens' acceptance and perception of EU maritime security systems. A draft version of the survey is provided in Deliverable D2.2 and the final result will be reported in D2.4 (M24).

Within the framework of the present deliverable, two additional data collection tools, to be adopted within pilots and demonstration activities, are proposed:

- **End-users' questionnaire on operational aspects linked to the adoption of AI systems** (Annex B): this questionnaire has been developed after the preliminary assessment on AI (by adopting the Assessment List for Trustworthy Artificial Intelligence -ALTAI), performed in February 2024. The questionnaire is aimed at collecting information from end-users on operational aspects (and potential issues) linked to the adoption of AI systems.
- **Ethics questionnaire for testing, validation, and demonstration activities participants** (Annex C): questionnaire aimed at assessing compliance with I-SEAMORE ethical principles within the framework project activities.

4 I-SEAMORE M&E STRATEGY AND ACTION PLANS

The present section first recalls the actions related to M&E, testing and demonstration foreseen by the Grant Agreement. It then outlines the main steps composing I-SEAMORE M&E process and their relationship with the above-mentioned actions. Every step should be validated and operationally defined by the responsible project partner. This will lead to the drafting of the I-SEAMORE M&E Action Plan in relation to the identified Round of Integrated Test, or to the Final Demonstration.

4.1 Overview of I-SEAMORE Tasks Related with Testing and Demonstrations Activities

The following table presents all the M&E, Testing and Demonstration activities mentioned in I-SEAMORE Grant Agreement:

TABLE 6: I-SEAMORE TESTING AND DEMONSTRATION ACTIVITIES

What	How	Who	When
Model based State-of-the-art (T3.1, D3.1)	Model based SotA resulting from the activities performed under T3.1 regarding concepts and capabilities for Maritime Surveillance.	INOV	M6 – June 2023
Definition of Use Cases, Design of Operational Concepts and Elicitation of User Requirements & KPIs (T3.3, D3.3)	Definition of the potential use cases, design operational concepts, and the elicitation of stakeholder requirements and KPIs.	MPT with the support from INOV all end-users and technical partners	M6 – June 2023
1st Release of M&E Framework (T2.4, D2.7_v1)	First version including the overall methodology for the implementation of M&E activities.	ISIG	M14 – February 2024
1st and 2nd round of Preliminary Testing Activities (T6.1)	These rounds will target the preliminary testing of functionalities from standalone components as well as some aspects of the developed interoperability layer (WP4, WP5).	MPT with strong support from all partners involved in WP4-WP6 activities	M12-13 – December 2023-January 2024 M17-18 – May-June 2024
1st round of end-users Validation Activities (T6.1)	End-users are involved in the validation activities in order to ensure that all components are ready to move to the final demonstrations and meet the requirements and KPIs.	MPT with strong support from all partners involved in WP4-WP6 activities	M18 – June 2024
Definition of Use Cases, Design of Operational Concepts and Elicitation of User Requirements & KPIs (T3.3, D3.5)	Definition of the potential use cases, design operational concepts, and the elicitation of stakeholder requirements and KPIs.	MPT with the support from INOV all end-users and technical partners	M24 – December 2024

What	How	Who	When
3rd and 4th round of Preliminary Testing Activities (T6.1)	These rounds will target the integration results (T6.2, T6.3) and the preliminary deployment of the Ecosystem (T6.4)	MPT with strong support from all partners involved in WP4-WP6 activities	M22-23 – October-November 2024 M27-28 – March-April 2024
2nd round of end-users Validation Activities (T6.1)	End-users are involved in the validation activities in order to ensure that all components are ready to move to the final demonstrations and meet the requirements and KPIs.	MPT with strong support from all partners involved in WP4-WP6 activities	M28 – April 2024
Technical Verification & Validation (V&V) (T6.5, D6.3)	Multi-step verification and validation approach leading to a qualified system for the trials in WP7.	TNO	M21-M28 – September 2024-April 2025
Definition of I-SEAMORE Demonstrations (T7.1)	Final definition of the demonstrations taking place in T7.2.	MPT, with the support of INOV and WP6 partners	M24-M29 – December 2024-May 2025
I-SEAMORE Demonstrations in realistic operational scenarios (T7.2, D7.1)	Final deployment of the I-SEAMORE Ecosystem at the MPT's OEC and monitoring of the KPIs and metrics previously defined.	MPT with strong collaboration of all partners	M26-M30 – February 2025-June 2025
Stakeholders Final Evaluation Activities (T7.3, D7.2)	Final evaluation of the I-SEAMORE outcomes from the end-users perspective.	RBP with high support of MPT, AEAT, UKBF, Atos, TNO, INOV, ISIG, and INI	M28-M30 – April-June 2025
Final Release of M&E Framework (D2.7_v2, T2.4)	Final version covering the achieved framework for monitoring and evaluation.	ISIG	M30 – June 2025
SOP and Policy Recommendations (D2.6)	Set of guidelines and recommendations based on the lessons learned.	RBP	M30 – June 2025

4.2 M&E Strategic Preliminary Actions

4.2.1 I-SEAMORE M&E Task Force

As diffusely recalled in this Deliverable, I-SEAMORE project architecture provides for many connections among different Work Packages and Tasks, especially in the field of M&E processes. In order to effectively and correctly implement all the foreseen M&E activities, then, a strong and close cooperation among PPs is required. More particularly, some Members of the I-SEAMORE Consortium play a pivotal role in guaranteeing that guidelines and directives outlined in this Deliverable properly apply to all stages of M&E activities:

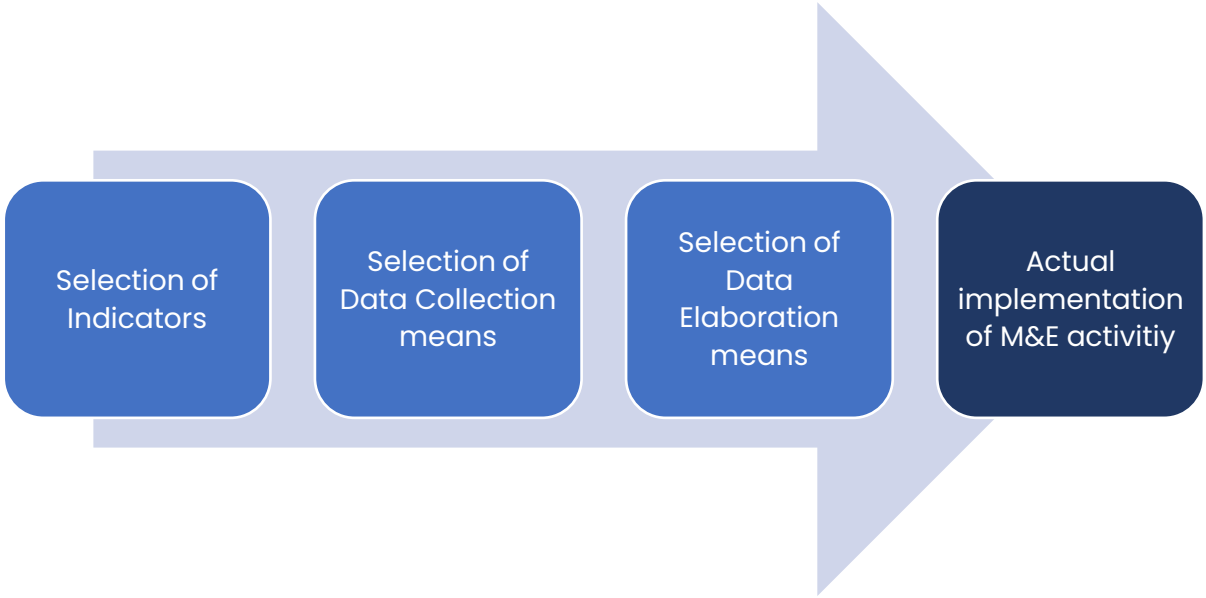
- EVIDEN, Project Coordinator and WP1 and WP5 Leader, guarantees both the compliance of the I-SEAMORE M&E Strategy with the overall project objectives and is responsible for the correct implementation of I-SEAMORE Platform & Services.
- ISIG, WP2 Leader, is responsible for the Work Package which contains the coordination core of M&E/V Framework, thus ensuring the overall consistency of M&E activities and their compliance with project's expectations.

- INOV, WP3 Leader, coordinates the co-creation and co-design phases of I-SEAMORE project. Within this effort, both the definition of the SotA, and the outlining of Operational Concepts, KPIs, and User Requirements are of particular relevance to the M&E Activities.
- EXAIL, WP4 Leader, is in charge of implementing the technological solutions related with UxVs, Satellites, and Payloads.
- TNO, WP6 Leader, coordinates the I-SEAMORE Ecosystem Integration, as well as the whole phase of Testing & Validation of developed solutions, including end-users' assessment and technical V&V Activities.
- MPT, WP7 Leader, covers the final Demonstrations and Evaluation of I-SEAMORE solutions.

4.2.2 I-SEAMORE M&E Preparation Steps

T2.4 aims at strategically filling the gaps between the M&E general framework and the above-mentioned M&E activities foreseen by the Grant Agreement, thus allowing to effectively operationalise the M&E process outlined by the I-SEAMORE project. The crucial, preliminary steps to effectively implement I-SEAMORE M&E activities are resembled in the following figure:

FIGURE 4: M&E PREPARATION STEPS



More specifically:

- Selection of Indicators refers to the necessity of identifying the most appropriate Indicators to be deployed for the operationalisation of relevant KPIs in every testing/demonstration activity foreseen by the project. The selection of the best available pool of indicators should maximise the capability of gathering and interpreting a sufficient amount of data to perform a consistent evaluation process.
- Selection of Data Collection means refers to the necessity of identifying the Data Collection instruments which suit the most with the foreseen testing/demonstration modalities, with the object of testing/demonstration, and with the scenario of operation. This includes:
    - Setting, when relevant and applicable, the technical means of collection, manipulation, and gathering of relevant data.
    - Setting, when relevant and applicable, the participants to the activity and its ex-post assessment, being them PPs, external experts, end-users, and/or stakeholders.

- Selection of Data Elaboration means refers to the necessity of Identifying contexts, timings, and instruments of Data Elaboration, in relation to the selected Indicators and Data Collection means, as well as with to the targeted participants to the Data Elaboration activities. This includes, for instance, the selection of the proper tools to gather feedback from the participants, as:
    - Checkboxes to easily assess the compliance of the tested/demonstrated technology against benchmarks and thresholds deriving from KPIs selection and translation into indicators.
    - Questionnaires so to gather feedback, mainly in the form of quantitative evaluation, in relation to the tested/demonstrated features, so to gather relevant insights to be then adapted and/or combined in order to fit into the selected indicators.
    - Interactive activities, such as focus groups and workshops, so to gather feedback, mainly in the form of qualitative appraisal, in relation to the tested/demonstrated features, so to gather relevant insights to be then adapted and/or combined in order to fit into the selected indicators.

All the above-mentioned steps should be performed taking into account the specificities of every testing/demonstration/evaluation activity and should be finalised prior their actual implementation.

The following table provides for a proposal for the combination of the M&E preliminary strategic steps with the foreseen testing/demonstration activities, as well as with the final evaluation activities:

TABLE 7: M&E PRELIMINARY STRATEGIC STEPS

<table>
<tr><th rowspan="2">Activity</th><th rowspan="2">Starting date</th><th colspan="3">Preliminary Steps</th></tr>
<tr><th>Selection of Indicators</th><th>Selection of Data Collection means</th><th>Selection of Data Elaboration means</th></tr>
<tr><td>1st and 2nd Round of Preliminary Testing Activities, and associated end-users Validation</td><td>M12[2] - December 2023</td><td colspan="3">M16 –April 2024
I-SEAMORE M&E Task Force</td></tr>
<tr><td>3rd and 4th Round of Preliminary Testing Activities, and associated end-users Validation</td><td>M22 – October 2024</td><td rowspan="2">M19 – July 2024
I-SEAMORE M&E Task Force</td><td colspan="2" rowspan="2">M20 – August 2024
I-SEAMORE M&E Task Force</td></tr>
<tr><td>Technical Verification and Validation</td><td>M21 – September 2024</td></tr>
<tr><td>Demonstrations in Realistic Operational Scenarios</td><td>M26 – February 2025</td><td>M24 – December 2024</td><td>M25 – January 2025</td><td>M26 – February 2025</td></tr>
</table>

[2] Please note that this activity started before the submission of the first release of the present Deliverable.

Activity	Starting date	Preliminary Steps		
		Selection of Indicators	Selection of Data Collection means	Selection of Data Elaboration means
Stakeholders Final Evaluation Activities	M28 – April 2025	I-SEAMORE M&E Task Force + RBP, AEAT, UKBF, INI	I-SEAMORE M&E Task Force + RBP, AEAT, UKBF, INI	I-SEAMORE M&E Task Force + RBP, AEAT, UKBF, INI

The operational steps foreseen for the implementation of the I-SEAMORE M&E Framework within the testing, demonstration, and validation activities foreseen by the project, are:

- Preliminary steps linked to the 1st and 2nd Round of Preliminary Testing Activities, and associated end-users Validation: M16 – April 2024; I-SEAMORE M&E Task Force.
- Preliminary steps linked to the 3rd and 4th Round of Preliminary Testing Activities, and associated end-users Validation, as well as Technical Verification and Validation:
    - Selection of Indicators: M19 – July 2024; I-SEAMORE M&E Task Force.
    - Selection of Data Collection and Data Elaboration means: M 20 – August 2024; I-SEAMORE M&E Task Force.
- Preliminary steps linked to the Demonstrations in realistic operational scenarios, as well as Stakeholders Final Evaluation activities:
    - Selection of Indicators: M24 – December 2024; I-SEAMORE M&E Task Force + RBP, AEAT, UKBF, INI
    - Selection of Data Collection means: M25 – January 2025; I-SEAMORE M&E Task Force + RBP, AEAT, UKBF, INI
    - Selection of Data Elaboration means: M26 – February 2025; I-SEAMORE M&E Task Force + RBP, AEAT, UKBF, INI

5 CONCLUSIONS

This Deliverable is the first version of the I-SEAMORE Monitoring and Evaluation/Validation (M&E) Framework and intends to:

- Set the basis of a feedback mechanism, able to generate and validate knowledge by means of validation and monitoring activities that engage end-users and stakeholders, by adopting a participatory approach.
- Provide an overall methodological framework for M&E activities that will be implemented within I-SEAMORE.
- Define interrelations between I-SEAMORE tasks and activities covering monitoring, evaluation, and validation aspects.
- Define the operational framework of I-SEAMORE M&E activities, covering three main dimensions (i.e., technical, ethical, societal).
- Provide a strategy (including a time plan and tasks) for M&E activities foreseen in the second and third year of project implementation.

Deliverable 2.7 ultimately lays the groundwork for forthcoming I-SEAMORE M&E activities, in synergy with WP3, where KPIs and end-users' requirements.

This report will feed its outcomes to T6.1 and T7.3 for the preliminary and final evaluation activities foreseen with end-users and stakeholders.

The next steps of the M&E framework are the following:

- I-SEAMORE M&E 1st Action Plan: 1st-2nd Round of Preliminary Testing Activities and Associated End-Users Validation
- I-SEAMORE M&E 2nd Action Plan: 3rd-4th Round of Preliminary Testing Activities, Associated End-Users Validation, and Technical Verification and Validation
- I-SEAMORE M&E 3rd Action Plan: Demonstration in Realistic Operational Scenarios and Stakeholders Final Evaluation Activities

The main results of Task 2.4 will be reported in D2.8 (due in M30, June 2025), which will present the achieved framework for monitoring and evaluation.

REFERENCES

Council of Europe (2018). *Monitoring and Evaluation Methodology Toolkit.* Available at: https://rm.coe.int/methodology-for-monitoring-and-evaluation-of-training-programmes-for-p/16808ace55

Patton, M. Q. (2018). *Principles-Focused Evaluation: The GUIDE*. New York: The Guilford Press.

UNDP (2003). *Handbook on Monitoring and Evaluating for Results*. UNDP Evaluation Office. Available at: https://www.oecd.org/derec/undp/35134974.pdf

ANNEX A –TECHNICAL KPIS ASSESSMENT QUESTIONNAIRE DRAFT

This Annex presents a draft of possible questions related to the KPIs presented in Par. 3.2.2.2. Questions and Data Collection means need to be validated/integrated following the workflow envisioned in the M&E Strategy and operationalized in the M&E Action Plans. Sections marked with [*] will be completed in the next releases of the Deliverable with the contribution of the project partners involved in the definition of I-SEAMORE M&E next steps.

Efficacy/Efficiency

Quality and efficiency of data collection

- Is data integration from Earth Observation (EO) Infrared IR System, Radar Systems, LIDAR Systems, and GRF Systems, correctly working?
- Is data integration effectively working in providing situational awareness?
- Is data integration effectively working in detecting small boats at night and mapping their routes?

Rate of efficiently performed missions

- Which is the % of missions that have been successfully executed during the Demonstration?
- Which is the % of missions that was not successfully executed due to recorder failure?
- Which is the % of missions that was not successfully executed due to disk data corruption?
- Which is the % of missions that was not successfully executed due to communication failure?
- Which is the % of missions that was not successfully executed due to electric power failure?
- Which is the % of missions that was not successfully executed due to sensor failure?

Data Fusion Performance

- [*]

False Positive Rate

- Which is the % of false alarms or false positive detections produced by the system during the Demonstration?

Detection Accuracy

- How do you assess the detection accuracy of the system in relation to Small Ships without AIS?
- How do you assess the detection accuracy of the system in relation to suspicious behaviour in Merchant Vessels with AIS?
- Does the system correctly detect and assess changes in speed as an indicator of suspicious activities?
- Does the system correctly detect and assess changes in course as an indicator of suspicious activities?

- Does the system correctly detect and assess entry into prohibited areas as an indicator of suspicious activities?
- Does the system correctly detect and assess rendezvous activities as an indicator of suspicious activities?

Response Time

- How long does, on average, the system take to detect the occurrence of a relevant event?
- Is the duration between the actual occurrence of an event and its first detection by the system acceptable?
- How long does, on average, the system take to alert the Marine Operation Centre (MOC) about these incidents?
- Is the duration between the first detection and the system alert to the MOC acceptable?

Resource Utilisation

- Are resource allocation and utilisation efficient in relation to the desired outcomes?
- Are assets deployed in the right place and at the right time during the system operations?
- Are unmanned assets deployed in an optimised manner during the system operations?

Operational Capability, Range and Coverage

Altitude Range and Coverage

- Which is the maximum altitude at which each drone can operate?
- Is the altitude sufficient to reach the expected coverage?

Endurance

- Which is the maximum duration of operativity for each drone?
- Is the duration of operativity sufficient to comply with the mission planning and resource allocation?

Geographical Coverage

- Were all targeted geographical areas covered by the drones during the Demonstration?

Power consumption improvement

- Was the benchmark of 15% improvement in power consumption for drones compared to previous technologies reached during the Demonstration?

Lower Acoustic Signature

- Was the acoustic signature of the drones reduced, in comparison with previous technologies?

Processing Capability

- Is internal processing capability of single drones adequately developed and deployed?
- Does internal processing capability effectively help in reducing the dependency on external processing resources?
- Does internal processing capability effectively help in facilitating real-time decision-making?

Drone Location and Navigation

- Do at least two drones within the system utilise the European Geostationary Navigation Overlay Service (EGNOS)?

Flights in adverse weather conditions

- [*]

Ecosystem Capabilities

Visual Analytics Tools

- Are Visual Analytics Tools effective for data exploration, visualisation, and analysis?
- Is visual representation intuitive?
- Are insights provided through the system effectively actionable?

Modernised C4 Platform

- Does the C4I platform meet the operational requirements of I-SEAMORE?

Dedicated UxV's Mission Planner

- Does the dedicated UxV's Mission Planner correctly allow to plan and schedule missions?
- Does the dedicated UxV's Mission Planner correctly allow to Define waypoints?
- Does the dedicated UxV's Mission Planner correctly allow to set specific parametres for each UxV, considering factors such as mission objectives, operational constraints, and available resources?

Mission Debrief Module

- Is the Mission Debrief Module in line with mission planning and optimisation needs?

Dedicated Modelling and Simulation (M&S)

- Is the M&S System in line with performance optimisation and operational strategies development needs?

Interoperability

External systems feeding

- Is the set of customised services and tools fed by at least three external sources?
- Does the system configuration ensure the successful interoperability of the system?

Interoperability and seamless integration

- Does the Interoperability setting of the instrument facilitate effective collaboration and information sharing among stakeholders?

Ground Control Station (GCS) connection

- Are drones equipped with a Ground Control Station (GCS) capable of communicating with the I-SEAMORE system?

Safety and Security

Corrupt or faulty data rate

- Which is the % of corrupt or faulty data due to sensor(s) failure onboard?

Drone failure rate

- Which is the % of drone failure that prevents the completion of the mission?

System failures rate

- Which is the % of system failure, such as issues with the measurement algorithm, detection logic, or software configuration?

Undesirable outcomes

- Was there any undesirable outcome, resulting from operator actions, which might have jeopardised the integrity of the mission and/or the system?

Users' satisfaction

Satisfaction related to usability

- How much would you rate the overall user-friendliness of the I-SEAMORE solution?
- Are interfaces enough intuitive?
- Are provided instructions enough clear?
- Are set workflows efficient?

Satisfaction related to reliability

- How much would you rate the overall reliability of the I-SEAMORE solution?
- Is the system well performing in terms of system downtime?
- Is the system well performing in terms of responsiveness?
- Did you experience any technical issues or glitches?

Satisfaction related to overall performance

- How much would you rate the overall performance of the I-SEAMORE solution?

ANNEX B – END-USERS' QUESTIONNAIRE ON OPERATIONAL ASPECTS LINKED TO THE ADOPTION OF AI SYSTEMS

1. Do you have any protocol /policy of work when it comes to the interaction with an AI system or AI components within a system?
    a. Yes
    b. No
    c. Do not know

2. How would you rate your internal protocols / policies / procedures of using AI in you daily work activities and in decision-making processes?
    a. Non-existent
    b. Completely inadequate
    c. Almost adequate
    d. Adequate
    e. Fully adequate

3. If you have any protocol / policy, does it include the question of over-reliance of the AI system, on in other terms procedures to avoid over-rely on the AI system or AI components within a system?
    a. Yes
    b. No
    c. Do not know

4. If you have any protocol / policy, does it tackle the question of human autonomy while using the AI system or AI component within a system?
    a. Yes
    b. No
    c. Do not know

    Examples include, protocol for actions to be taken while using an AI which includes double check by a human, monitoring the work of AI by a human in order to avoid any affect to the human autonomy and thus interfering with the decision-making process in any other unintended way.

5. Do you have any protocol/policy/procedures of work where interpretation of results of the use of AI system or AI components within a system is tackled?
    a. Yes
    b. No
    c. Do not know

6. Do you have any protocol / policy / procedures of work to safely abort an operation using AI system or AI components of a system when needed?
    a. Yes
    b. No
    c. Do not know

7. Do you have any detection and response mechanisms for undesirable adverse effects of the AI system or AI components within a system?
    a. Yes
    b. No
    c. Do not know

8. Do you have a process to report potential vulnerabilities, risks or biases in the AI system of AI component within a system you use?
    a. Yes
    b. No
    c. Do not know

ANNEX C – ETHICS QUESTIONNAIRE FOR TESTING, VALIDATION AND DEMONSTRATION ACTIVITIES PARTICIPANTS

ASSESSING ETHICS

for _____

(title of testing, validation and demonstration activities activity)

The questionnaire aims at investigating your personal experience before, during and after the _____ (*testing/validation/demonstration*).

		I-SEAMORE ETHICAL PRINCIPLE
1	Before the event I have received an informed consent in which: 1.1 Consistent, clear and general (needed) information on the event to be organized was given. [(1) Strongly disagree; (2) Disagree; (3) Neither agree nor disagree; (4) Agree; (5) Strongly agree]	INTEGRITY
	1.2 Activities foreseen and explained in the informed consent ensure honest, fair and respectful treatment of participants / volunteers: [(1) Strongly disagree; (2) Disagree; (3) Neither agree nor disagree; (4) Agree; (5) Strongly agree]	INTEGRITY
2	Was your participation to the event voluntary? Yes/No/I prefer not to answer	VOLUNTARY PARTICIPATION
3	The host institution and the project protected my physical social and psychological well-being during the event: [(1) Strongly disagree; (2) Disagree; (3) Neither agree nor disagree; (4) Agree; (5) Strongly agree]	PROTECTION OF THE VOLUNTEER PARTICIPANTS AND HONOURING TRUST
4	The host institution and the project as a whole took precaution measures in order to minimize any disturbance during the event:	ANTICIPATING HARMS

	[(1) Strongly disagree; (2) Disagree; (3) Neither agree nor disagree; (4) Agree; (5) Strongly agree]	
5	The participation in the event was a disturbing experience for me. [(1) Strongly disagree; (2) Disagree; (3) Neither agree nor disagree; (4) Agree; (5) Strongly agree]	AVOIDING UNDUE INTRUSION
6	My privacy and confidentiality have been ensured and respected before and during the event. [(1) Strongly disagree; (2) Disagree; (3) Neither agree nor disagree; (4) Agree; (5) Strongly agree]	CONFIDENTIALITY AND ANONYMITY
7	The host institution and the parties involved provided a clear definition and directions during the course of the event, and the activities performed were explained properly. [(1) Strongly disagree; (2) Disagree; (3) Neither agree nor disagree; (4) Agree; (5) Strongly agree]	REPRESENTATION AND MISUSE OF EXPERTISE