ASSURING AUTONOMY
INTERNATIONAL PROGRAMME

DEMONSTRATOR PROJECT

FINAL REPORT

BOAUT
Boundaries Of AUTonomy

FEBRUARY 2022

Final Report of the BOAUT Project – The Practitioners' Perspective on Commercial MASS Operating in Territorial, Near-Coast Waters

This is the final report from the Boundaries of Autonomy (BOAUT) project, which was funded by Assuring Autonomy International Autonomy (AAIP) from 2021-02-01 to 2021-01-31.

The BOAUT project investigated whether it is possible to assure that, with the support of a remote operations centre (ROC), maritime autonomous surface vessels (MASS) can operate as safely as manned ships. MASS are designed to deviate from their planned missions when detecting hazardous situations. However, deviations based on erroneous information or system faults can still lead to collisions, loss of ships, etc. This can be harmful to humans, the environment and economic values.

This report details the project's results relating to the project's assurance objectives. The implications of the limitations of the project, such as access to particular configurations of vessels and data, are clarified and discussed. This includes discussing, based on the project partners' expertise, the possibility to generalise the results to other configurations of the context.

Published	2022
Authors	Fredrik Asplund, Pernilla Ulfvengren, Jonas Envall, Johan Holmqvist, Reine Larsson, Stein Magnus Undheim
Contact	fasplund@kth.se, +4687907405, Brinellvägen 83, 10044 Stockholm

1 CONTENTS

2 Executive Summary 5
3 Introduction 7
4 The Implications of ROC for Assuring the Safety of MASS 9
4.1 Limitations Based in Assurance Argumentation 9
4.2 Limitations Based in External Validity 10
5 Approach, Demonstrator, Scenarios and Validation Criteria 11
5.1 Demonstrator 12
5.2 Scenarios 16
5.3 Validation Using the Scenarios 21
6 Practitioners' Perspectives 23
6.1 Demonstrator Feedback 23
6.2 Feedback from the Trials Relevant to Safety Assurance 23
7 Limitations 29
8 The Safety Case 31
9 Summary 33

2 EXECUTIVE SUMMARY

This deliverable provides one possible definition of the role of Remote Operations Centres (ROCs) in relation to the safety assurance of Maritime Autonomous Surface Ships (MASS). In other words, how one can expect a ROC to contribute to the safety (assurance) of truly autonomous surface vessels. This role is defined as one of being responsible to act on hazardous situations which MASS do not recognize.

Tools, methods, and scenarios are then described, which allowed for probing this role definition through 2-to-3-hour long shadow trials organized with 8 maritime practitioners. Results from the trials indicate that the approach of defining the role of ROCs as one of handling unknown unknowns (from the perspective of the MASS) is reasonable. However, this approach implies a choice regarding who to recruit to ROCs, based on how they should (or whether they are even expected) to apply their previous experience when a hazardous situation involving unknown unknowns occurs. Regardless, this approach will rely on ROC support systems that can direct an operator's attention according to the distance to relevant threats; the definition and monitoring (through suitable infrastructure) of critical maritime areas; enabling MASS to receive feedback from operators (particularly from local VTS operators) on phenomena that are difficult, or rely on several information sources, to interpret; and providing and structuring the access of a wider set of stakeholders to making information digitally available.

Part of these concerns are already being addressed by researchers and practitioners, but we identified three topics that, if they received increased attention, would facilitate the introduction of maritime autonomy:

- The ways through which ROC operators could be unfairly blamed when automation breaks down in the maritime domain, as their influence on a situation is not meaningful.
- How to support AI reasoning with human insights instead of having to resort to taking full control of a MASS.
- How to include information from untrusted sources (such as the public) in common operational pictures in a dependable way.

3 INTRODUCTION

Although a traditionally conservative industrial domain, the shipping industry has seen a lot of changes throughout the last few decades. As an example, advanced automation to support navigation and situation(al) awareness on the bridge is already commonplace. More automation is expected in the near future, with several research projects looking at the introduction of Maritime autonomous surface vessels (MASS). The Swedish Transport Agency even expects to start in-service trials of autonomous road ferries in 2022.

However, many challenges still exist to making MASS the norm on the seas, such as:

- The maritime domain can put extreme requirements on equipment and infrastructure. There is a lack of sufficiently robust technology.
- Phased introduction of MASS might be necessary to avoid risk, requiring MASS to be crewed during specific phases and activities (e.g. in territorial waters, when docking, etc.). To comply to requirements, during this phase investment in and design of MASS technology might not be able to avoid costs for personnel.
- Maritime law and regulations are not made with autonomy in mind, but with captains, seamanship, etc. Even small changes require extensive analysis, to avoid making maritime stakeholders uncertain about their liability when utilizing MASS.
- The responsibilities held by (certain) crew members might need to be moved to other entities. This can require both training and recruitment of certain people, to ensure that the new responsibility associated with a role can be shouldered properly.

Remote operations centres (ROC) address many of these challenges. They are seen as an intrinsic part of increasing automation at sea to the point of introducing MASS. However, ROC do not necessarily have a role when surface vessels are truly autonomous. In fact, perhaps the most straight-forward scenario involving MASS is when they go from port to port without ever requiring human interaction beyond the setting of new routes, the routine maintenance and repair of equipment, etc. This scenario does not require a ROC.

Nevertheless, maritime practitioners often (at least) refer to ROC as important in ensuring that MASS can operate as safely as manned ships despite deviations based on erroneous information or system faults. Even if MASS are carefully designed to deviate from their planned missions when detecting hazardous situations, ROC are seen as an appropriate part of a defence in depth strategy in case these deviations still risk leading to collisions, loss of ships, etc.

This deliverable starts by defining what this perspective on MASS and ROC means in terms of safety assurance. In other words, how do we expect a ROC to contribute to the safety (assurance) of truly autonomous surface vessels. In Chapter 4 we then describe the methodological approach taken by BOAUT to probe practitioners regarding the role of a ROC when dealing with highly advanced, reliable MASS. This includes a description of the demonstrator and scenarios used during workshops conducted with practitioners. Chapter 5 describes the results and summarizes the outcomes from the workshops. Chapter 6 discusses the limitations of the results, and Chapter 7 provides a safety case pattern based on them.

4 THE IMPLICATIONS OF ROC FOR ASSURING THE SAFETY OF MASS

The BOAUT project investigated whether it is possible to assure that, with the support of a remote operations centre (ROC), maritime autonomous surface vessels (MASS) can operate as safely as manned ships. To clarify the meaning of this goal we describe the associated limitations agreed on within the project, while Chapter 6 describes the implications of these limitations for the project's results.

4.1 LIMITATIONS BASED IN ASSURANCE ARGUMENTATION

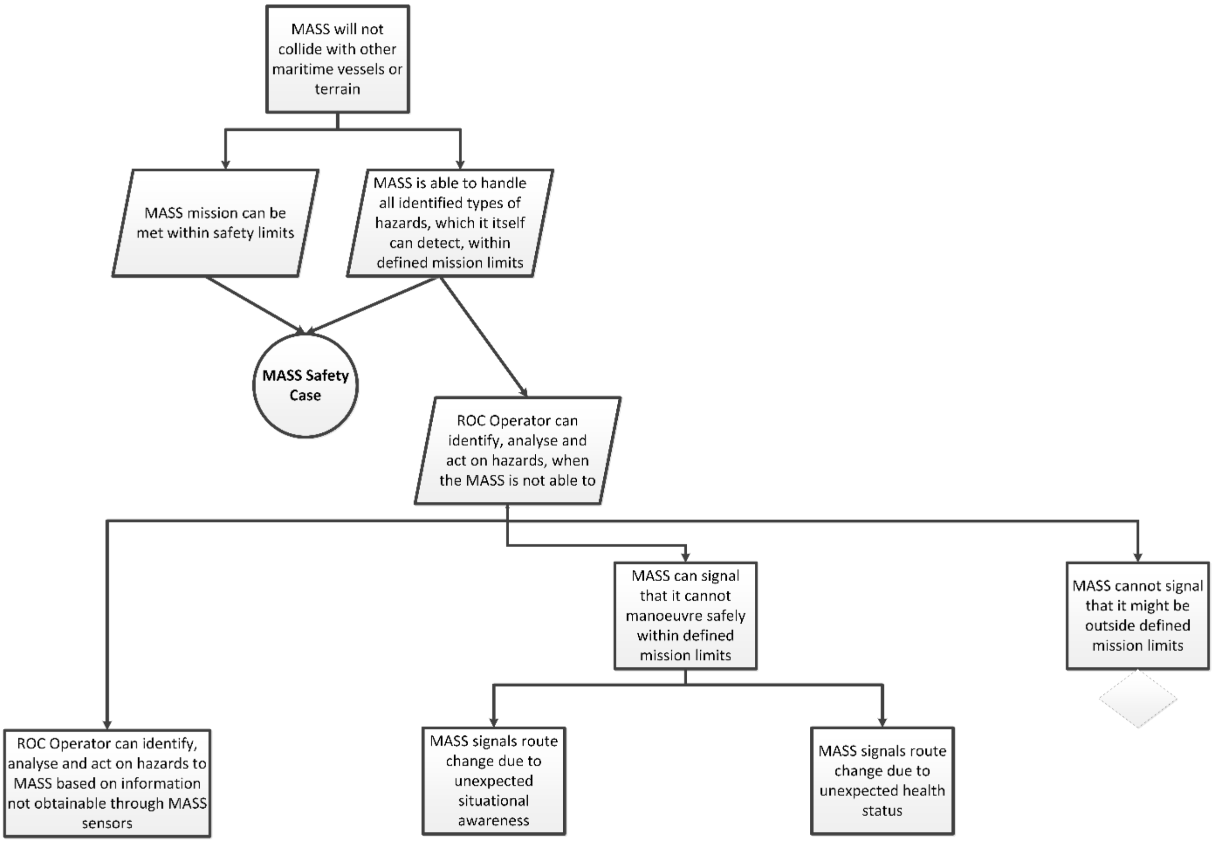

MASS will not collide with other maritime vessels or terrain
MASS mission can be met within safety limits
MASS is able to handle all identified types of hazards, which it itself can detect, within defined mission limits
MASS Safety Case
ROC Operator can identify, analyse and act on hazards, when the MASS is not able to
MASS can signal that it cannot manoeuvre safely within defined mission limits
MASS cannot signal that it might be outside defined mission limits
ROC Operator can identify, analyse and act on hazards to MASS based on information not obtainable through MASS sensors
MASS signals route change due to unexpected situational awareness
MASS signals route change due to unexpected health status

Figure 1 Overview - Targeted Assurance

As shown in Figure 1 BOAUT started by limiting the safety goal of the targeted assurance as that of MASS not colliding with other maritime vessels or terrain. There are other safety goals that might be considered, such as capsizing due to bad weather. However, for a ROC operator the primary task is to "remove obstacles", and this safety goal is thus the primary one.

BOAUT then rests on the assumption that a MASS will not be deployed if its owners are not confident that they can detect and handle all identified types of hazards within specified safety limits. These safety limits will be reflected in all MASS mission limits. This assumption is based on input from maritime stakeholders and is the defining difference between the (highest) level of automation and lower levels. This confidence might be misplaced, but it will define the relationship between ROC and MASS. In other words, ROC operators will be present to deal with unknown unknowns (from the perspective of MASS) – identifying, analysing and acting on hazardous deviations from the norm when a MASS is not able to *as these deviations are outside of what is specified in the MASS own safety case*.

This means ROC operators will not provide safety assurance for all deviations – a MASS will be correct to carry out some deviations, possibly even some deviations due to unknown unknowns. MASS will for instance *routinely deviate from their plans* to avoid other maritime vessels, but within specified mission limits. MASS can also deviate silently from mission limits, which, as long as ROC operators are not exclusively assigned to specific MASS, puts different requirements on ROC operators from situations when they signal their deviation. Finally, the deviation can also be the failure to act (omission) when circumstances suggest action is required.

This suggests that a ROC operator can provide safety assurance in three ways:

1. By acting on hazardous situations when the MASS cannot signal that it is outside of its mission limits.
2. By acting on hazardous situations when the MASS can signal that it is outside of its mission limits.
3. By identifying, analysing and acting on hazards to the MASS based on information that is not obtainable through MASS sensors.

All of these topics are relevant. However, (1) is more related to automated monitoring than ROC operator monitoring. ROC operators can of course spend time trying to analyse the behaviour of MASS in an ad hoc manner, but the expectation cannot be that they should consistently identify hazardous behaviour in complex behaviour when pre-defined alarms cannot. (1) is thus excluded by BOAUT.

4.2 Limitations Based in External Validity

The maritime operating context is highly variable, including large differences to phenomena such as depth, tides, sea ice, communication degradations, nation state zone restrictions, seasonal changes, wind, current, surface waves, and much more. To carry out an investigation with reasonable external validity BOAUT has chosen limitations that allow a discussion in Chapter 5. These define that the:

- MASS considered will be *near-coast*, i.e. in the territorial sea. This should open up the possibility for enabling more data-intensive functionality.
- Investigation will focus on *monitoring*, i.e. the phase before a ROC operator decides to take (remote) control over a MASS.
- Considered MASS will be unmanned and fully automated, i.e. not involving levels of autonomy in which some functions are automated while some are not.

5 APPROACH, DEMONSTRATOR, SCENARIOS AND VALIDATION CRITERIA

E-OCVM is a framework for the validation of traffic management originally developed for the avionics domain and based on systems engineering processes for concept development and validation. It allows for the validation of operational concepts from early phases of development to full implementation. Developed for situations when several independent R&D organisations have to collaborate, it provides consistency by providing a coherent approach and comparability across validation activities. Validating an assurance argument is difficult, since the effectiveness of the concepts in a real situation might be unacceptable due to hidden or unexpected behaviour by the participants. E-OCVM suggests the use of gaming and shadow mode trials in these circumstances. It is often fruitful to combine such trials with interviews to verify the researchers' interpretation of stakeholders' actions.

2-to-3-hour long BOAUT trials were thus organized with 8 maritime practitioners, in which they were each subjected to scripted events in an environment consisting of an early prototype of our technical concepts. This *demonstrator* is described in Subsection 4.1. During the trials the practitioners were continuously encouraged to speak about their impressions of what happened in the demonstrator, and elaborate on related thoughts on MASS.

During trial development, we talked extensively about sceneries, scenarios, scenes, mission limits, situations, local situational awareness (system/operator), and global situational awareness (system/operator)[1]. The two subgoals remaining from Section 3.1 were instantiated during the trials through scenarios, which challenged the practitioners by:

1. A MASS deviating from its planned route due to:
    a. Malfunctioning internal systems. (Type 1)
    b. An unexpected local situational awareness. (Type 2)
2. A global awareness that suggests that a MASS *should* deviate from its planned route. (Type 3)

As these subgoals might have to be fulfilled concurrently, combinations of the scenario types must also be considered. Similarly, as several hazards of the same type might occur in sequence, more complex scenarios must be considered. The scenarios defined by BOAUT are provided in Subsection 4.2, according to Table 1.

[1] *Scenery* is the background in where a *scenario* is taking place. Scenery is static and invariant over a scenario lifetime. A *scene* is a temporal snapshot comprising the scenery and the dynamical entities acting in the scenery. A scenario is a sequence of such snapshots making up the time interval between a start time and an end time. A MASS *mission limits* can thus be defined as the route it intends to take as it navigates a scenario, planned beforehand but subject to change if safety requires it. Such a route change will constitute a clear signal that a MASS is not able to manoeuvre safely within mission limits. A *situation* is related to an entity's comprehension of the scene and the scenario, past events and predicted, or assumed, future events. A situation is therefore an ego perspective and the internal projection and representation of scenario at timepoints. A MASS *local* situational awareness is thus its comprehension of its surroundings (the situation) through its own sensors. A ROC operator monitoring the MASS will have a personal *global* situational awareness, which includes comprehension of a situation from many different sources beside the MASS at hand.

Table 1 Scenarios

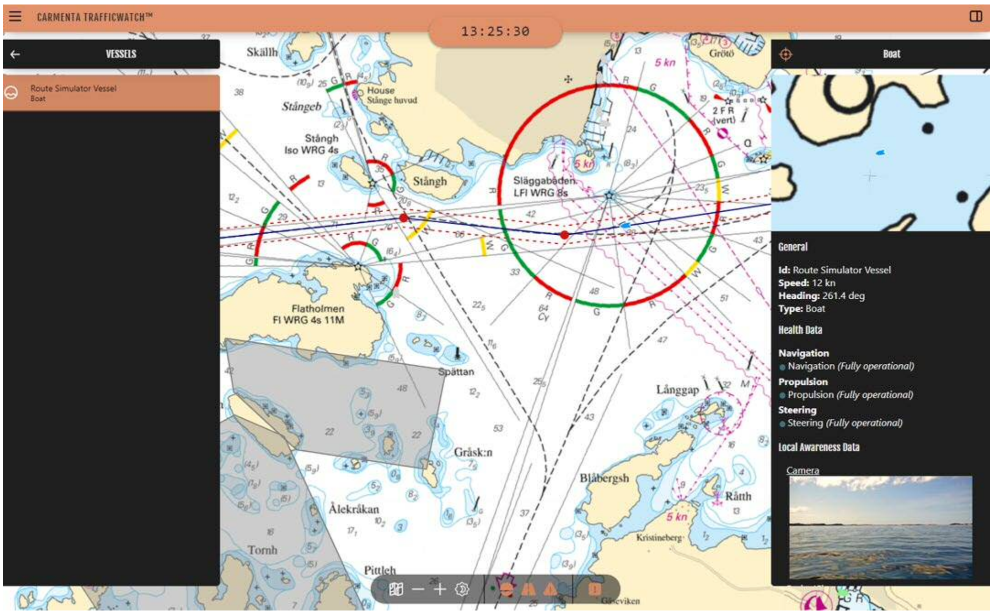
	Type 1	Type 2	Type 3
Type 1	Simple: Section 4.2.1 Complex: Section 4.2.2	Section 4.2.3	Section 4.2.4
Type 2	Section 4.2.3	Simple and Complex: Section 4.2.5	Section 4.2.6
Type 3	Section 4.2.4	Section 4.2.6	Simple: Section 4.2.7, Section 4.2.8 Complex: 4.2.9, 4.2.10

5.1 DEMONSTRATOR

The Carmenta TrafficWatch platform was configured and extended to meet the demands specified by the project parameters and scenarios. Carmenta TrafficWatch was used for supervising vessels and surroundings to generate warnings of upcoming dangers or anomalies, as well as supervising the completion of planned routes, thus providing an operator with situation awareness.

5.1.1 Introduction

From an operator's point of view, Carmenta TrafficWatch displays a large map where vessels, routes and obstacles are shown. The routes are drawn on top of sea charts and there are optional maps available such as satellite image and bathymetry layer.



Route Simulator Vessel
Boat

General

Id: Route Simulator Vessel
Speed: 12 kn
Heading: 261.4 deg
Type: Boat

Health Data

Navigation
- Navigation *(Fully operational)*

Propulsion
- Propulsion *(Fully operational)*

Steering
- Steering *(Fully operational)*

Local Awareness Data

Camera

Figure 2 Carmenta TrafficWatch User Interface

Vessels are listed in different panels depending on their type. Simulated MASS vessels send their data directly to Carmenta TrafficWatch, including extended sensor data. AIS vessel positions are received automatically from transceivers on actual ships.

In addition to standard parameters such as speed and heading, the simulated MASS can also send sensor data to Carmenta TrafficWatch. The sensor data may contain system health data (status indicators for navigation, propulsion, steering etc.) and local awareness data (camera image, radar image etc.). This provides the operator with critical information about on-board systems, e.g. that a failure in the propulsion system has occurred in the MASS.

Carmenta TrafficWatch analyzes data in real-time and displays warnings or alerts to the operator. An example is when a plotted route is intersecting a hazard area, indicating that the MASS and/or operator would need to act.

When all these different sources of information were combined, Carmenta TrafficWatch provided great situation awareness for the operator. The scenarios contained different combinations of routes, sensor data and obstacles to examine and make observations on how operators will respond to situations in a wide range of difficulty levels. The result was a clear picture on how different types of situations may affect the safety of an autonomous vessel.

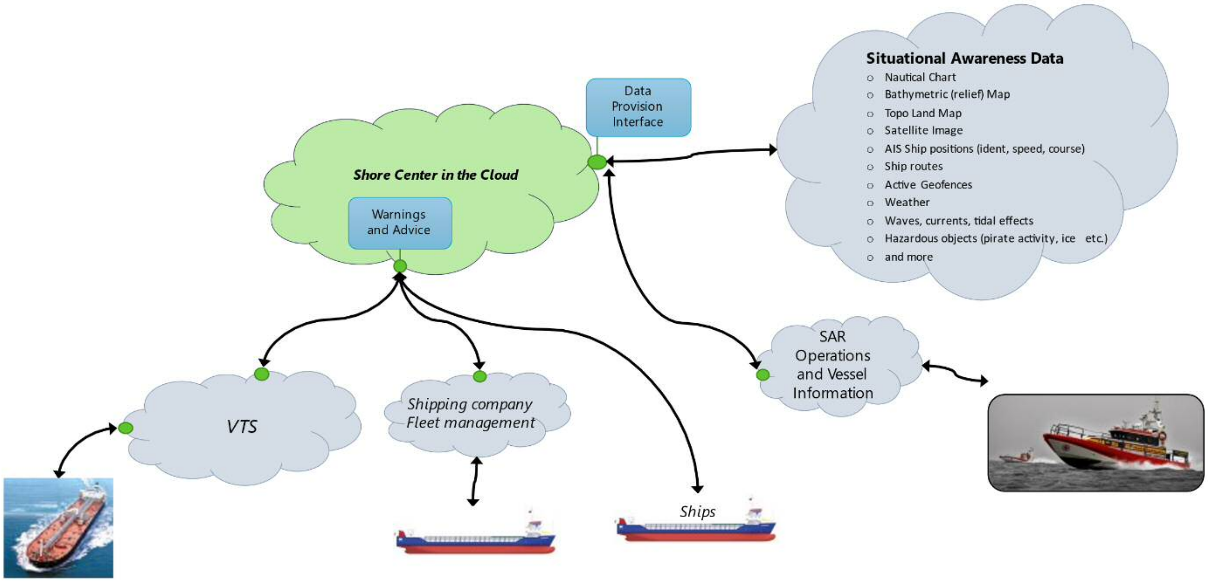

Situational Awareness Data
Nautical Chart
Bathymetric (relief) Map
Topo Land Map
Satellite Image
AIS Ship positions (ident, speed, course)
Ship routes
Active Geofences
Weather
Waves, currents, tidal effects
Hazardous objects (pirate activity, ice etc.)
and more
Data Provision Interface
Shore Center in the Cloud
Warnings and Advice
SAR Operations and Vessel Information
VTS
Shipping company Fleet management
Ships

Figure 3 Example Applications for Situational Awareness Platform

5.1.2 System overview

Carmenta TrafficWatch is a modular cloud-based solution. It consists of specialized software modules which can be combined to quickly provide a wide range of functionality. The core function is to supervise and control vessel operations by collecting and analyzing data about the surrounding environment and traffic situation.

The ROC is using a web-based client connected to Carmenta TrafficWatch, visualizing the objects and environment to provide operators with full situational awareness.

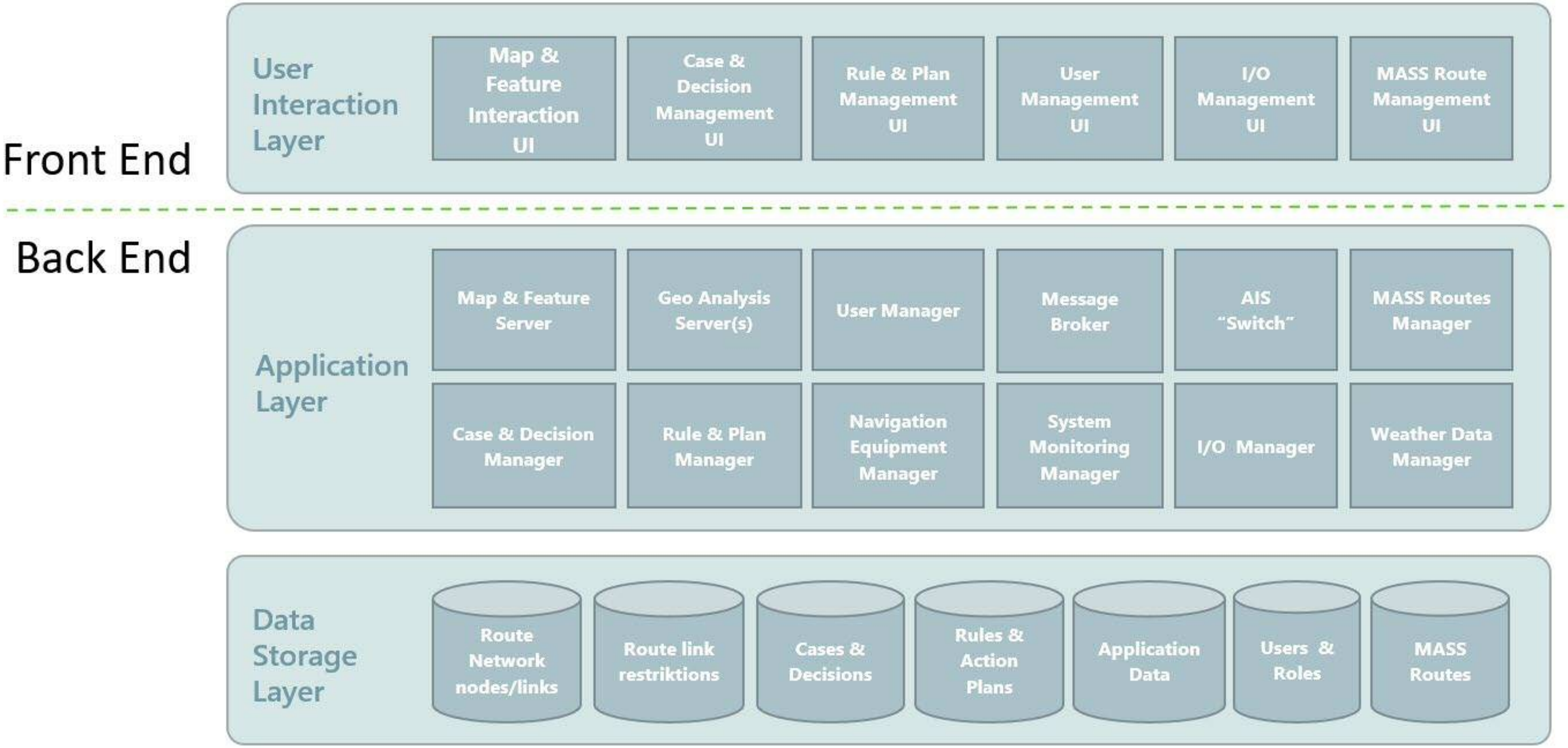


Figure 4 Example of modular architecture in Carmenta TrafficWatch

In the BOAUT project, there was two main categories of input data to Carmenta TrafficWatch:

- MASS send information regarding position, route, sensor health data and local awareness data. This data was transmitted according to a specific protobuf format, using the MQTT messaging protocol. Routes used RTZ, a standardized format for route plan exchange.
- Several external sources send different types of data, depending on purpose. Examples are sea chart data used as background maps and AIS position data used for visualizing vessels.

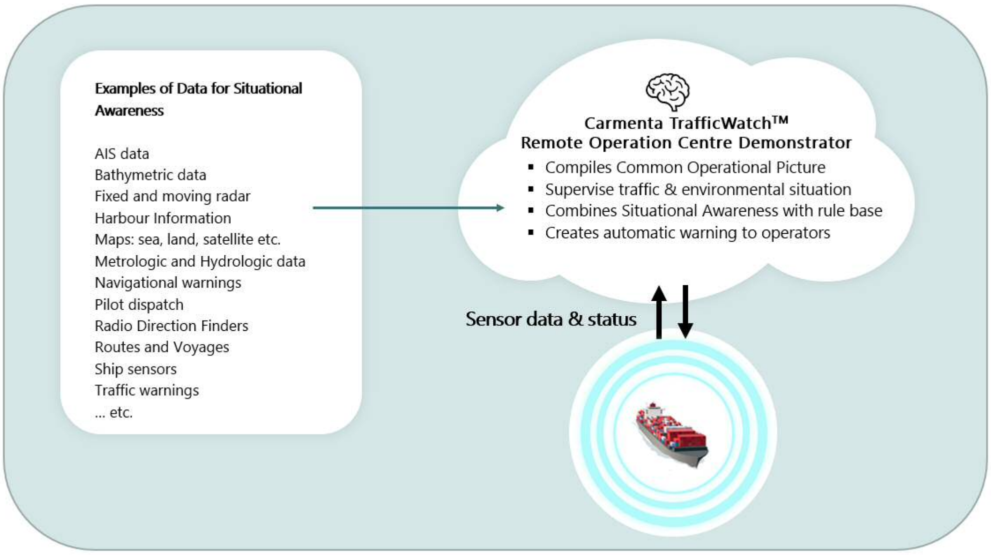

Examples of Data for Situational Awareness
AIS data
Bathymetric data
Fixed and moving radar
Harbour Information
Maps: sea, land, satellite etc.
Metrologic and Hydrologic data
Navigational warnings
Pilot dispatch
Radio Direction Finders
Routes and Voyages
Ship sensors
Traffic warnings
... etc.
Carmenta TrafficWatch™
Remote Operation Centre Demonstrator
▪ Compiles Common Operational Picture
▪ Supervise traffic & environmental situation
▪ Combines Situational Awareness with rule base
▪ Creates automatic warning to operators
Sensor data & status

Figure 5 Examples of system input to Carmenta TrafficWatch

5.1.3 Additional Health View

A very simple system health system was used to prompt practitioners to start talking about possible faults in MASS and how they would like to diagnose and handle them. One example of this system is shown in Figure 6.

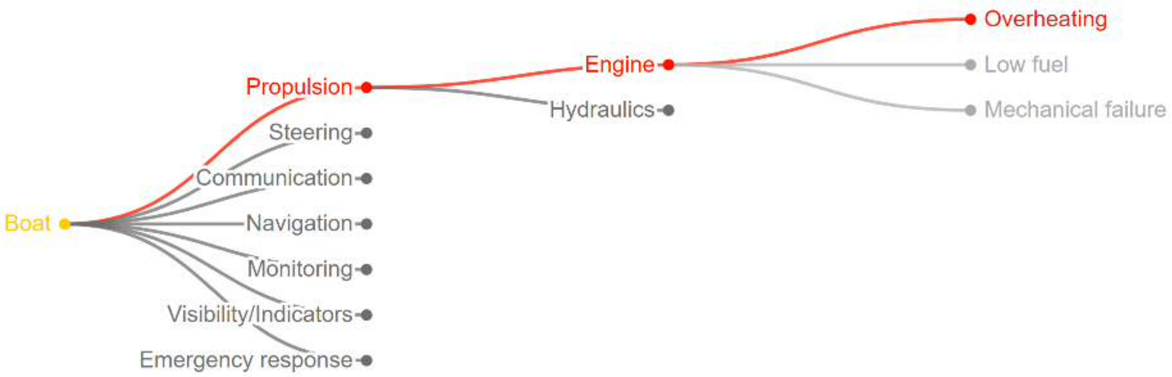


Figure 6 The Health System

5.2 SCENARIOS

This subsection contains the scenarios defined to challenge operators. Where relevant we highlight subtle configurations we designed to spark discussion. These were not critical to the investigation, but are included as they could be of interest to the reader.

5.2.1 Scenario 1 – MASS Breakdown (Type 1)

1) MASS sends initial route to ROC.
2) MASS sends continuous position updates which can be seen by operators in real time.
3) MASS has a system health breakdown (variable breakdown). MASS updates state and sees risk on route, prompting route recalculation[2].
    a. A health data message is sent to ROC.
    b. A route update message is sent to ROC. (Including reason for change.) This leads to the ROC operator being notified.
4) ROC lookup action to take for malfunction, which is to increase the safety margins. ROC infrastructure (also) detects that the increased safety margins and a hazard overlap and generates a notification.
    a. *Variability*, as in different types of breakdowns affecting MASS capability differently: Simple breakdown with complete loss of functionality, and Complex breakdown where capability is reduced rather than completely lost.
5) Operators see notifications. (Prompted by vessel and self-identified.)
6) ROC updates vessel information:
    a. Route change
    b. Capability reduction
7) Operators can look at notifications, MASS information, etc. (Local and global awareness.)

5.2.2 Scenario 2 – Multi-Malfunctions in MASS subsystems (Type 1 Complex)

1) MASS sends initial route to ROC.
2) MASS sends continuous position updates which can be seen by operators in real time.
3) MASS has a system health breakdown (engine malfunction). MASS updates state and sees risk on route, prompting route recalculation.
    a. A health data message is sent to ROC.

[2] In this scenario the new route left a sheltered, but narrow, passage and extended into open waters. This can be seen as positive as there are more space for drifting and less risk for other traffic, but also more exposure (for an already faulty MASS) to e.g. wind, waves, etc.

    b. A route update message is sent to ROC. (Including reason for change.) This leads to the ROC operator being notified.

4) ROC lookups actions to take for engine malfunction, which is decided to increase the safety margins. ROC infrastructure (also) detects that the increased safety margins and a hazard overlap and generates a notification.
5) MASS has a system health breakdown (AIS malfunction). (Prompts no route change.)
    a. A health data message is sent to ROC.
6) MASS has a system health breakdown (Radar malfunction). MASS updates state and sees risk in current speed, prompting route recalculation.
    a. A health data message is sent to ROC.
    b. A route update message is sent to ROC. (Including reason for change.) This leads to the ROC operator being notified.
7) Operators see incidents and notifications. (Prompted by vessel and self-identified.)
8) ROC updates vessel information:
    a. Route change
    b. Capability reduction
9) Operators can look at notifications, MASS information, etc. (Local and global awareness.)

5.2.3 Scenario 3 – Restricted Area on Route, Seen by Local Awareness (Type 1 and Type 2 Combined)

1) MASS sends initial route to TW.
2) MASS sends continuous position updates which can be seen by operators in real time
3) MASS sends notification of route change, including reason (which is restricted area on route).
    a. *Variability*, as the restricted area could be due to:
        i. Different local weather phenomena identifiable by MASS (e.g. fog).
        ii. An old map (so, a health message probably sent much earlier, but which was not critical at that point in time.)
4) Operator gets notification that MASS has changed course due to restricted area one MASS route.
5) Operator investigates notification (global awareness), not seeing the restricted area.
6) Operator investigates MASS (local awareness).

5.2.4 Scenario 4 – MASS Breakdown (Type 1 and 3 Combined)

1) MASS sends initial route to ROC.

2) MASS sends continuous position updates which can be seen by operators in real time.
3) MASS has a system health breakdown (variable breakdown).
    a. A health data message is sent to TW.
    b. *Variability*, as in different types of breakdowns affecting capability differently (incident can thus be different, but here we use a grounding incident as an example): Simple breakdown which complete loss of functionality, and Complex breakdown where capability is reduced rather than completely lost.
4) ROC lookup action to take for malfunction, which is to increase the safety margins. ROC infrastructure (also) detects that the increased safety margins and a hazard overlap and generates a notification.
    a. *Variability*, as in:
        i. That he restricted area could be due to different global weather phenomena not updated on MASS.
        ii. The MASS could have an old map (so, a health message probably sent much earlier, but which was not critical at that point in time.)
5) Operators see incidents and notifications. (Prompted by vessel and self-identified.)
6) ROC updates vessel information:
    a. Capability reduction
7) Operators can look at notifications, MASS information, etc. (Local and global awareness.)

5.2.5 Scenario 5 – Other Ship on Collision Course with MASS (Type 2, and Type 2 Complex due to Variability)

1) MASS sends initial route to ROC.
2) MASS sends continuous position updates which can be seen by operators in real time.
3) MASS sends notification of route change, including reason (which is ship on collision course).
4) Operator gets notification that MASS has changed course due to other ship on collision course with MASS.
5) Operator investigates notification (global awareness), not seeing the ship.
6) Operator investigates MASS (local awareness).
    a. *Variability* in reason for collision risk due to different ships seen on radar and camera, i.e., ferry[3], super-tanker, kayaks, speedboat, slow cruiser and fishing boat.

[3] In this particular case the ferry was supposed to yield to the MASS, but it is not uncommon for certain types of commercial traffic to break the maritime regulations in this fashion. This can be expected by the MASS, but also go against “hard-programmed rules”.

b. *Information via radio* from e.g. fishing boat.

5.2.6 Scenario 6 – New route from MASS crosses restricted area (Type 2 and 3 Combined)

1) MASS sends initial route to ROC.
2) MASS sends continuous position updates which can be seen by operators in real time.
3) MASS notices object in route. MASS sends notification of route change, including reason (which is ship on collision course).
    a. *Variability* in different ships detected, i.e. ferry, super-tanker, kayaks, speedboat, slow cruiser, and fishing boat.
4) ROC gets notification that MASS has changed course due to other ship on collision course with MASS.
5) ROC gets notification that MASS route interferes with restricted areas.
6) Operator investigates notification (global awareness).
    a. *Variability* in different restricted areas, i.e. dangerous, fog, restricted, etc.
7) Operator check whether MASS sees same indication, which it does not (local situational awareness.)
8) ROC checks route and sees that it interferes with restricted area. (Global situational awareness.)

5.2.7 Scenario 7 – Other Ship on Collision Course with MASS (Type 3)

1) MASS sends initial route to ROC.
2) MASS sends continuous position updates which can be seen by operators in real time.
3) Operator gets notification that ship is on collision course with MASS.
4) Operator investigates notification. (Global situational awareness).
5) Operator check whether MASS sees same indication, which it does not. (Local situational awareness).

5.2.8 Scenario 8 – Restricted Area on Route, Seen by Global Awareness (Type 3)

1) MASS sends initial Route to ROC
2) MASS sends continuous position updates which can be seen by operators in real time

3) ROC gets alert from external system that there is a temporary restriction area[4] along the MASS route.
4) Operator investigates notification. (Global situational awareness).
5) Operator checks whether MASS sees same indication, which it does not. (Local situational awareness.)

5.2.9 Scenario 9 – Several Other Ships on Collision Course with MASS (Type 3 Complex)

1) MASS sends initial route to ROC.
2) MASS sends continuous position updates which can be seen by operators in real time.
3) ROC backend discover that the route crosses an area with other vessels (it receives these objects from AIS).
4) Operator gets multiple notifications that ships are on collision course with MASS.
5) Operator investigates notification (global awareness).
   a. *Variability* in reason for collision risk:
      i. Planned route on wrong side of fairway.
      ii. Area of fog.
      iii. Restricted area due to drifting, burning ship on route.
6) Operator check whether MASS sees same indication, which it does not. (Local situational awareness.)

5.2.10 Scenario 10 – Multi-Notifications from MASS (Type 3 Complex)

1) Several MASS sends initial Route to ROC.
2) Several MASS sends continuous position updates which can be seen by operators in real time.
3) Weather change prompts increase of risk contours across large area with several MASS.
4) Several MASS sends notification of route change, including reason (which is ships at risk of grounding).
5) One MASS does not, but operator gets notification that this MASS is at risk of grounding.
6) Operator investigates notifications (global awareness).
   a. *Variability* in different restricted areas, i.e. dangerous, fog, restricted, etc.
7) Operator checks whether MASS (1) sees same indication, which it does not.

[4] In this scenario we used a diving activity with time limits. The start of the activity might be after the MASS has passed or might be a "soft" boundary, which is a distinction a MASS might not be willing to consider or able to act on – but an ROC operator can take advantage of.

5.3 VALIDATION USING THE SCENARIOS

BOAUT was not aiming for the final validation of a particular technical concept, but rather to exploring challenges to assurance using a demonstrator – to challenge operators during trials to provide feedback relevant to future validation. Using the Concept Lifecycle Model of E-OCVM, BOAUT was thus a feasibility phase project with a special focus on assurance as a transversal case. For each of the identified scenarios validation activities must thus define an acceptable operator outcome related to the technical concept at hand. This meant that, at least initially, human and technology integration, operating procedures and communications requirements were analysed and tested, with a focus on performance, operability and acceptability of operational aspects. The defined high level validation criteria are shown in Figure 7.

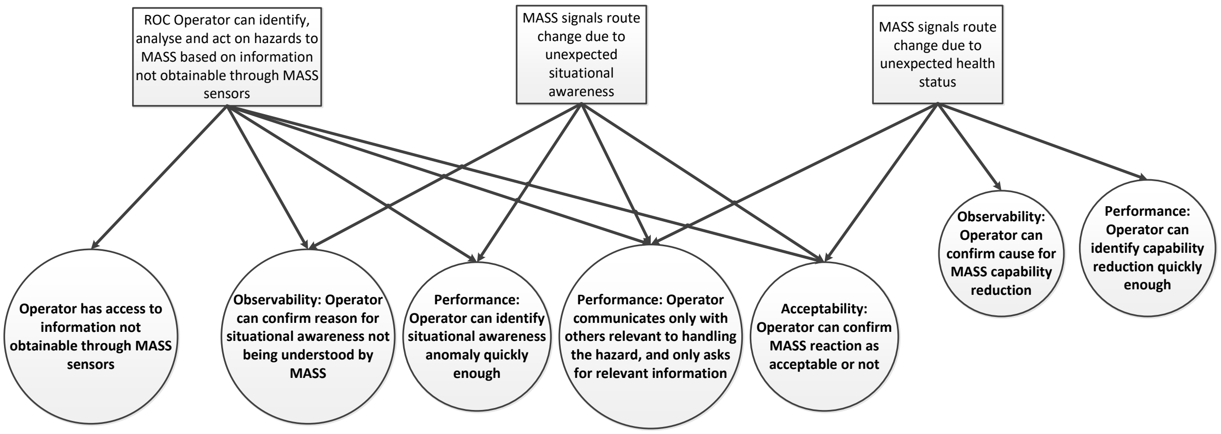

ROC Operator can identify, analyse and act on hazards to MASS based on information not obtainable through MASS sensors
MASS signals route change due to unexpected situational awareness
MASS signals route change due to unexpected health status
Operator has access to information not obtainable through MASS sensors
Observability: Operator can confirm reason for situational awareness not being understood by MASS
Performance: Operator can identify situational awareness anomaly quickly enough
Performance: Operator communicates only with others relevant to handling the hazard, and only asks for relevant information
Acceptability: Operator can confirm MASS reaction as acceptable or not
Observability: Operator can confirm cause for MASS capability reduction
Performance: Operator can identify capability reduction quickly enough

Figure 7 The Evidence Sought

This meant that the BOAUT trials included questions both before and after the scenarios that probed these validation criteria.

6 PRACTITIONERS' PERSPECTIVES

This chapter provides feedback on the demonstrator and trials described in Chapter 4 within the limitations described in Chapter 3.

6.1 DEMONSTRATOR FEEDBACK

Carmenta TrafficWatch was very useful during the workshops and the application was met with positive feedback from the interviewees. It was very valuable to observe reactions to events and challenges occurring in the system, rather than discussing "abstract" situations. However, the demonstrator was primarily provided as something concrete for the practitioners to talk "around". Nevertheless, some observations that arose are:

- As the scenarios involved anomalous situations triggered by alarms, many of the practitioners discussed the demonstrator's ability (or lack thereof) to provide predictions and direct the attention of the user. This included the likely location of and shared routes between surface ships, highlighting uncertainty, the risk of collision and associated alternatives for action. However, it also included small configurations, such as showing more of what was in front of a MASS when centering on it then what is behind it.
- Restriction areas were not very useful as static descriptions of an area. Interviewees wanted more information, such as who had reported it, time restrictions, who could be contacted (if relevant), the details of the restriction (exactly what was forbidden rather than what activity was taking place), etc. Related findings are provided in Subsection 5.2.5.
- The interviewees sought the ability to keep incidents active until deciding to remove them, thus returning to a "clean interface" if (and only if) all necessary actions have been taken.
- Several interviewees mentioned the need to the demonstrator GUI to be quickly reconfigurable to match the needs of each operator. There is significant variation in operators' interaction with the GUI. (As operator's might change places during their work.)

Obviously, a clear and intuitive user interface is important to quicky assess the situation and get correct information to act upon. Warnings and alerts need to be finely tuned in terms of appearance and danger level, to avoid sensory overload for the operator. The mission role of the operator is also very important. The technical environment for a ROC could differ depending on whether the operator has a general supervising role, remote operations role or a more specialized fleet role with deep access and knowledge of on-board systems. A few of these observations can thus be tied more directly to the feedback from the trials.

6.2 FEEDBACK FROM THE TRIALS RELEVANT TO SAFETY ASSURANCE

MASS travelling in territorial, near-coast waters will traverse a challenging environment with a large risk of accidents. Mixed traffic; vessels controlled by humans with little knowledge of maritime law and regulations; and the current, frequent reliance on informal communication are just a few examples of the potential causes of escalated risk levels raised by the interviewees.

Even if MASS was extremely reliable, it is highly likely that they will be put in situations not foreseen by their designers. Interviewees often came back to the need for ROC to support MASS facing unknown unknowns without putting the blame of an accident on the ROC operator. The latter was a real issue, as many of the situations foreseen as problematic involved little or no possibility for the ROC operator to meaningfully affect the outcome. Some of this was related to short time spans, such as a MASS relinquishing control to a ROC operator only a few seconds before a collision. However, there were also less straight-forward issues, such as the likely need to increase the role of Vessel Traffic Services (VTS) – which currently only provide advice to shipping (at least in some countries, including Sweden and Denmark).

Finding 1: Assumption of the role of ROCs as handling unknown unknowns is verified.

This means that the operator with the best chance of intervening successfully might not be the ROC operator acting on behalf of the shipping company owning the MASS. Other ROC operators, such as VTS operators, might not have the necessary authority and tools to act when MASS face unknown unknowns. The following subsections provide an analysis grounded in that challenge. As the details of how MASS and ROC will interact depends on future negotiations between flag states the implications of the analysis are by necessity broad. However, they are detailed enough to allow for the definition of a safety case in the next chapter.

6.2.1 Autonomy is a Change of Perspective

One interviewee described the trials as trying to envision a science fiction world with little time to prepare. Despite this, the interviews validated the logic described in Chapter 3. No interviewee believed that MASS would be allowed to operate if they were not reliable and able to handle all situations in their operating environment with an almost zero risk of causing an accident. Naturally, incidents outside of what had been imagined by MASS designers could still occur, and ROC operators would then be the last line of defence against accidents. The scenarios described in Chapter 4 were not seen as extreme, but rather as quite forgiving and good for discussing events that would give a ROC operator a meaningful influence.

That said, most interviewees quickly forgot that they were dealing with surface ships that were supposed to be capable of advanced recognition of entities in its environment and reasoning about their implications. On the one hand, the maritime experience of the interviewees was what allowed the interviewees to see hazards and reason about what the MASS should do. On the other hand, the same experience meant that mundane instead of exceptional explanations for the situation were sought. This does not have to be wrong, even the mundane can sometimes be dangerous. However, the interviewees were fully aware of the fact that an autonomous vessel, monitored on demand by a remote operator, cannot trigger an alarm for every other vessel that it passes. Despite this they frequently treated near-collisions as an ordinary meeting of ships, without seeking further explanations for a MASS course change.

Finding 2: The experience of operators can be both beneficial and detrimental to their ability to successfully intervene when faced with an unknown unknown (from the perspective of a MASS).

Similarly, it took time before most of the interviewees sought to identify and understand any internal malfunctions in the MASS. Whether this should be required will be discussed in

Subsection 5.2.3, but we here note that it did not come naturally to interviewees to diagnose faults remotely – even if they had experience from working with technical diagnosis, maintenance and repair onboard ships during their career.

6.2.2 A ROC Operator to Confront or Avoid Danger?

ROC operator ability to identify the extraordinary despite previous experience can be related to how interviewees often called for an explanation by the MASS for its decisions: What in a situation was seen as extraordinary? Why was a particular route chosen instead of another? How uncertain was the MASS about possible alternatives? When dealing with unknown unknowns, the interviewees sought a transparent or explainable AI. However, this had different implications depending on what the interviewee thought was the main approach to ensuring safety.

- No interviewee thought that a ROC operator would be able to handle all situations involving a MASS encountering an unknown unknown (from its perspective), but on one side of the continuum the scenarios were handled as time critical go/no-go decisions. MASS explanations were required to be short and additional information only involve the most obviously safety-relevant factors in the environment. As an example, interviewees could request (only) the location, predicted trajectories, and safety depth of the MASS and anything else moving in its vicinity. An explanation was required to allow a quick decision on whether the MASS decision-making should be overridden.
- On the other side of the continuum, interviewees required explanations to be presented in a way that allows the operator to understand any trade-offs involved in the MASS' decision-making, not just a simplified cause and effect. As an example, interviewees could request to know why a specific route was chosen in favour of a few others, making a distinction between high waves and strong wind when bad weather was signalled in a specific area.

Part of this was related to the role envisioned by the interviewees, as either a representative of the MASS' shipping company or a VTS operator. Typically, the former would be linked to situations closer in time to an accident. However, interviewees expressing the latter approach still acknowledged that quick emergency action could be required even by VTS operators, when there is no time to transfer responsibility to someone else.

The need for different types of explanations of MASS reasoning can thus be tied to the same operator, even if one can envision a nominal process in which e.g. VTS operators order or advice a shipping company to start monitoring a specific MASS more closely.

Finding 3: Regardless of a ROC operator's role, ROC support systems must be able to direct the operator's attention according to the distance to relevant threats.

6.2.3 An Expert, or an Operator Relying on Experts?

The same differences in approach also had implications for the reliance on others. Specifically, whether the ROC operator should be an expert in steering surface ships and/or MASS systems, or simply an operator that filters anomalies.

- If time critical go/no-go decisions are the focus, then the operator might be an expert in steering surface ships, but should not be an expert in MASS systems. Any lengthy fault diagnosis will be detrimental to the situational awareness of the ROC operator and might be hazardous in itself. This diagnosis should instead be shifted to a maritime engineer, who can provide suggestions on how to act or reconfigure e.g. radar systems for better situational awareness.

- Otherwise, there is a case to be made for the ROC operator being an expert in both surface ship steering and MASS systems, or in neither.
    - A ROC operator who is an expert in both would be able to reason about uncertainty in the MASS decision-making, the underlying reasons for its actions due to its design, and whether a fault is likely to increase and when it will aggravate problems. Such a ROC operator might be better at supporting other traffic or ROC operators in other roles. This is further discussed in Subsection 5.2.5.
    - According to some interviewees you can become reasonably competent in handling all types of maritime vessels. An operator who is then not an expert in the steering a specific MASS, or its internal systems, might ask the kind of "stupid questions" that someone with more experience might not – but really should, to identify a problem related to a unknown unknown. Such an operator could then be better suited for simply filtering anomalies, to identify those that other experts might need to assess for risk.

Finding 4: Finding 2 implies a choice of who to recruit for positions at ROCs, which has further implications for necessary support systems.

Regardless of the implications due to differences in the approach, there was also a case made for why differences in expertise should be completely taken out of the equation. According to this perspective there should be detailed guidance created for what to do when specific faults or circumstances occur. As an example, a fault in the steering might mean a MASS cannot enter certain narrow passages. This is not an uncommon approach and often depends on the demands of a particular industry. Less requirements on expertise when controlling a system leads to the policy to strictly follow guidance when risk increase, and vice versa. In this case the argument was mostly made from a legal perspective, i.e. based on that it would protect ROC operators from the liability should an accident occur. In light of the purpose of the ROC operator in regard to MASS being to handle unknown unknows, it seems doubtful that this is a good direction to take. However, we note that lower levels of automation might actually encourage the development of such strict guidance.

6.2.4 Critical Areas

Regardless of the role of ROC operators, the interviewees stressed the need to understand the area through which a MASS moves to help ensure its safety. This puts restrictions on ROC operators that control MASS on a global scale. The traffic patterns and behaviour of other vessels can differ greatly between countries, which poses a risk when a ROC operator intervenes in an unknown context. It is highly likely that this will maintain the need for VTS operators, which can support ROC operators that follow MASS along a long geographical route. This is further discussed in Subsection 5.2.5.

More importantly, these concerns suggest that critical areas of higher risk need to be defined within larger VTS areas. As an example, this can be a narrow strait, where ships should not meet. To ensure safety what happens in relation to it should be monitored to keep track of traffic levels, weather changes, and other environmental factors. This can involve the need to install special infrastructure, not only to monitor the area but to enhance MASS capabilities[5]. It can also require the extension of current VTS areas, as far from all territorial waters are currently covered by these.

Finding 5: Critical areas needs to be identified and equipped with suitable infrastructure.

With such areas, ROC operators would better know which MASS to focus on at any given moment. They would also be less prone to infer the behaviour of one MASS based on other ships, even if they are affected by the same local phenomena within the larger VTS area (such as a change in weather). The behaviour of specific MASS would then be weighed against the risks of critical areas, and not based on the actions of other MASS that might be different in a critical way.

6.2.5 A Common Situational Awareness and New Stakeholders

Interviewees stressed that special circumstances are often *broadcast* locally to let those in the area know of and adapt appropriately. As implied in previous subsections, this suggests that it would be suitable to staff a VTS operator role with those who are experts in MASS systems and safety-relevant factors in the local environment. Information is, even today, often shared by VTS operators based on what might be of interest to operations, rather than what is officially required by the role. Similarly, VTS operators often talk to different stakeholders to see whether they can change their plans to allow the overall traffic system to work more smoothly. However, this can also be solved through VTS operators contacting experts at the shipping company owning MASS, when necessary to change or ensure their behaviour.

More interestingly, it raises the question on how to make MASS aware of special circumstances. As an example, if a MASS is passing close to a group of kayaks, then a ROC operator might spot the odd kayak about to turn back into the path of the MASS. This can be handled through communication between different ROC operators, but also through the ability to highlight potential problems to the MASS' reasoning directly.

Finding 6: ROC operators (particularly VTS operators) must be able to provide cues to MASS reasoning.

Furthermore, such information about special circumstances might come from new types of stakeholders or put new requirements on old stakeholders to share information in a digital format. As an example of the former, there might be a need to elicit information from the public to get it in a timely manner; and as an example of the latter, it might be quite complex to share information about a boat race that allows MASS to avoid the area within a suitable time frame.

Finding 7: A wider set of stakeholders must be able to provide information digitally to the maritime traffic system.

[5] Interviewees assumed that MASS would be able to enter modes in which they are less efficient, but are able to navigate much more exactly.

7 LIMITATIONS

The identified issues with supporting MASS through ROC were the primary one's to today's operators. That they thought their role would be to handle unknown unknows (from the perspective of the MASS), is reasonably clear.

Woods and Christoffersen[6] describe a moving target for system development because of the changing nature of demands, pressures, and resources within fields of practice. The non-static nature of a field of practice makes it difficult to foresee the effects of the envisioned system in the actual future context, a problem they call the envisioned world problem.

A complete safety case involving ROC and MASS cannot be defined as which stakeholder will fill which role, and how these roles will interact, is not yet defined. On the one hand, this was positive as the interviewed practitioners were quick to reflect quite broadly on nautical problems, exactly because there are no clear rules for MASS yet. An odd course or weather update was not ignored as impossible regardless of whether it had been designed into the scenario intentionally or by mistake. On the other hand, this was especially negative as the interviewees did not easily extend the set of entities and phenomena that populate their current working environment.

Limitation 1: Study not suitable for identifying completely new entities and phenomena.

The *purpose* of the envisioned MASS is thus the most obvious limitation to the conclusions of this study. Commercial shipping is monitored and operate within a (more or less clear) system for cooperation in today's maritime shipping ecosystem. Leisure vessels might operate outside established systems for cooperation and with designs implying very different operational envelopes. If such a MASS experiences a fault, will there even be anyone at hand that can diagnose it at hand? The results in Chapter 5 are only relevant if such MASS will operate under other regulations that either puts them on par with commercial shipping or forces them to limit their operations accordingly.

Limitation 2: Non-commercial shipping might not come to operate under regulations that enforce collaboration with ROCs.

This study was also carried out in Scandinavia. The interviewees made it clear that other *geographical locations* will have entirely different traffic patterns, boundaries for what is seen as acceptable behaviour, and other requirements on e.g. Vessel Traffic Service operations and personnel. Although several of the interviewees had experience from working across the globe, we cannot rule out that other contexts make the envisioned solutions impossible.

Limitation 3: Interviewees might have underrated the (un)importance of geographical location.

[6] Dekker, S.W., and D.D. Woods. 1999. "To Intervene or not to Intervene: The Dilemma of Management by Exception." Cognition, Technology and Work 1: 86–96. doi: 10.1007/s101110050035.

8 THE SAFETY CASE

Figure 8 returns to the limitations set out in Chapter 3, reworking them into a safety case. We refer to the associated MASS safety case as a solution (circle), as it is a cornerstone for the complete safety case. It could possibly have been referred to as an assumption or as part of the context. Strategies (rhombuses) are then laid out, defining how to approach the use of ROC operators to identify, analyse and act on hazards when a MASS is not able to. Strategies are related to solutions (rectangles), which define what the system must be able to do.

This safety case is not complete, in the sense that it is a blueprint for how to implement a safe maritime ecosystem involving MASS supported by ROC. To arrive at a full safety case pattern we would for instance have to define which stakeholders will operate in this ecosystem, their obligations and their relationships. This was not our intention and – in our opinion – not the most urgent endeavour. Exactly which stakeholder will be allocated which obligation will not change the fundamental problems that will have to be solved to allow MASS in territorial waters. Several of the solutions are thus capped by undeveloped element decorators (hollow diamonds), meaning that the goals associated with these solutions are yet to be developed.

Some of these undeveloped goals and strategies are already topics being discussed in the autonomy and maritime communities, such as:

- AI transparency and explainability
- How to appropriately direct the attention of operators through graphical interfaces.
- Maritime infrastructure for vessel traffic services

However, other topics have still not seen much interest, signalling related challenges:

- It is well known that operators are unfairly blamed when automation breaks down, but not what this implies in the maritime domain when introducing ROC as a support to MASS.
- How to support AI reasoning with human insights is insufficiently investigated, implying that the only recourse a human operator might have is to take full control of a MASS.
- The concept of a common operational picture is a well investigated topic, but not how to include information from untrusted sources (such as the public) in a dependable way.

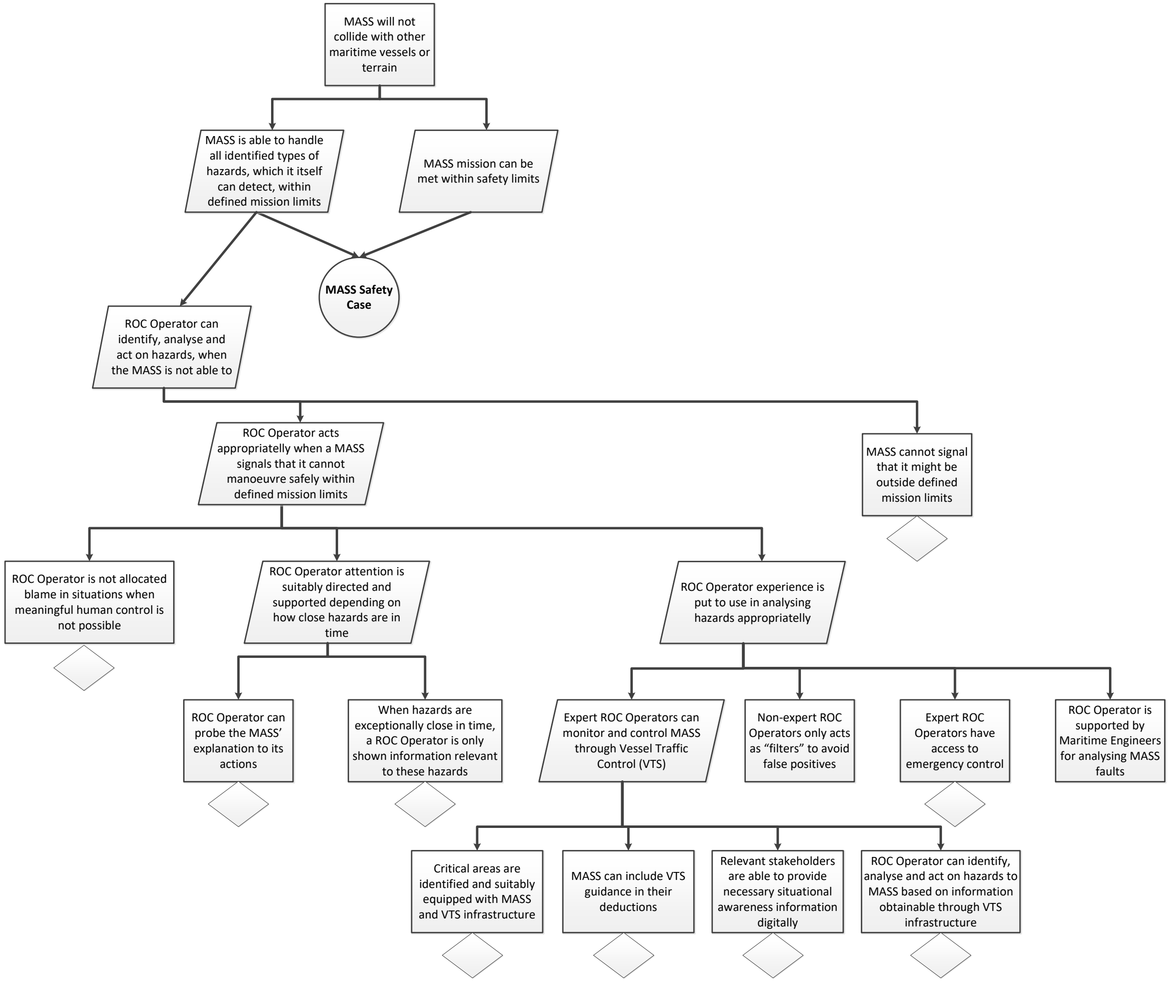

MASS will not collide with other maritime vessels or terrain
MASS is able to handle all identified types of hazards, which it itself can detect, within defined mission limits
MASS mission can be met within safety limits
MASS Safety Case
ROC Operator can identify, analyse and act on hazards, when the MASS is not able to
ROC Operator acts appropriately when a MASS signals that it cannot manoeuvre safely within defined mission limits
MASS cannot signal that it might be outside defined mission limits
ROC Operator is not allocated blame in situations when meaningful human control is not possible
ROC Operator attention is suitably directed and supported depending on how close hazards are in time
ROC Operator experience is put to use in analysing hazards appropriately
ROC Operator can probe the MASS' explanation to its actions
When hazards are exceptionally close in time, a ROC Operator is only shown information relevant to these hazards
Expert ROC Operators can monitor and control MASS through Vessel Traffic Control (VTS)
Non-expert ROC Operators only acts as "filters" to avoid false positives
Expert ROC Operators have access to emergency control
ROC Operator is supported by Maritime Engineers for analysing MASS faults
Critical areas are identified and suitably equipped with MASS and VTS infrastructure
MASS can include VTS guidance in their deductions
Relevant stakeholders are able to provide necessary situational awareness information digitally
ROC Operator can identify, analyse and act on hazards to MASS based on information obtainable through VTS infrastructure

Figure 8 A Safety Case

9 SUMMARY

There are several ways to define the role of ROCs in relation to MASS in a future traffic system. BOAUT defined the role as one of being responsible to act on hazardous situations which MASS do not recognize. This role was then investigated through workshops with practitioners.

Results indicate that the approach of defining the role of ROCs as one of handling unknown unknowns (from the perspective of the MASS) is reasonable. However, this approach implies a choice regarding who to recruit to ROCs, based on how they should (or whether they are even expected) to apply their previous experience when a hazardous situation involving unknown unknowns occurs. Regardless, this approach will rely on ROC support systems that can direct an operator’s attention according to the distance to relevant threats; the definition and monitoring (through suitable infrastructure) of critical maritime areas; enabling MASS to receive feedback from operators (particularly from local VTS operators) on phenomena that are difficult, or rely on several information sources, to interpret; and providing and structuring the access of a wider set of stakeholders to making information digitally available.

Part of these concerns are already being addressed by researchers and practitioners, but we identified three topics that, if they received increased attention, would facilitate the introduction of maritime autonomy:

- The ways through which ROC operators could be unfairly blamed when automation breaks down in the maritime domain, as their influence on a situation is not meaningful.
- How to support AI reasoning with human insights instead of having to resort to taking full control of a MASS.
- How to include information from untrusted sources (such as the public) in common operational pictures in a dependable way.

ASSURING
AUTONOMY
INTERNATIONAL PROGRAMME