European Maritime Safety Agency

SPECIFIC MASS CONCEPTS & RISK EVALUATION TECHNIQUE PROPOSED FOR TESTING THE RBAT

REPORT 3

Version: 2021-1343, Rev. 0

Date: 06/05/2021

EMSA

Table of contents

EXECUTIVE SUMMARY

Introduction

EMSA has contracted DNV to perform a functional study for developing a Risk-Based Assessment Tool (RBAT) for maritime autonomous surface ships (MASS). As outlined in DNV's proposal (DNV GL, 2020b) and EMSA's Tender Specifications (EMSA, 2020), the RBAT study consist of three parts:

- Part 1: Develop a framework for a generic MASS risk assessment tool
- Part 2: Test the risk assessment tool on specific cases and develop software tool prototype
- Part 3: Re-iterate testing on more complex cases and finalize the software tool

The study is currently halfway into Part 2 which includes the following scope of work:

a) Identify and select specific MASS concepts and sub-functions for testing RBAT

b) Develop a risk evaluation technique appropriate to be applied to MASS concepts

c) Perform a gap analysis of RBAT and further develop the framework

d) Develop test cases for the identified MASS concepts and sub-functions, and test RBAT

e) Based on the results from the test cases, update RBAT and develop a first version of a functional software prototype

f) Draft and submit reports elaborating the tasks above, also by explaining the identified gaps, the changes made and providing appropriate graphic material.

Activities a) to c) are documented in this report, namely the first report of Part 2 and the third report of the RBAT study.

Activities d) and e) will be documented in a separate report, namely the second report of Part 2 and the fourth report of the RBAT study.

Identify and select MASS concepts and sub-functions

A set of MASS concepts and sub-functions were identified and selected to be developed for testing the RBAT framework and methodology. The approach for how to do this consisted of the following four steps:

- Define selection criteria
- Identification and screening
- Assessing relevant MASS concepts and sub-functions
- Present proposed solution for EMSA and flag state representatives

The selection criteria were defined to ensure that the selected concepts are those with the (expected) broadest commercial usage, and that they allow RBAT to be tested using a variety of (expected) inputs.

Concepts and sub-functions were identified and screened through interviews with subject matter experts internally in DNV, as well as external key industry actors and information available online. A total of 47 different projects related to autonomous/remotely operated vessels and technologies were identified. A screening revealed that the following five concept categories where the most frequent:

- Suppliers (of assisted solutions and autonomy/automation related products) – 11;23%
- Un-crewed surface vessels (USVs) – 8;17% (out of scope)

- Small passenger ferries – 7;15%
- Short sea cargo – 6;13%
- Research – 5;11%

An assessment against the selection criteria concluded that short-sea cargo vessels, small passenger ferries, and autonomy/automation related products were most suitable for being brought forward as test cases. A RO-PAX was chosen to represent autonomy/automation related products.

The concepts and sub-functions which will be developed into test cases are presented in sub-chapter 2.6, Table 2-2. This includes details about vessel characteristics, manning and other roles involved, operating areas, mission phases, operations, traffic density, and type of supervision.

Although USVs are commercially strong, they were considered out of scope due to the USVs relatively small size and simpler configuration compared to conventional vessels being regulated by administrations and class societies. Concepts developed as part of research activities are diverse and it was difficult to see a trend in what would be brought forward commercially.

Develop a risk evaluation technique appropriate to be applied to MASS concepts

A technique suitable for evaluating risks associated with MASS concepts has been developed. Many of the risks identified for MASS are expected to be control related, with the failure causes stemming from software. Because the probability of such risks is inherently difficult to predict it was decided to abandon the classical definition of risk as a function of probability and consequence. Other industries, such as automotive, have faced similar challenges as part of their work with safety assurance. The proposed approach draws on such experiences but tries to adopt them in such a way that they meet the specific needs in the maritime industry, and do not deviate from established industry practices and frameworks.

So instead of directly applying the classical definition of risk (probability*consequence), RBAT evaluates risk as a combined function of:

- How severe is the worst-case outcome from an undesired event?
- How effective are the concept's mitigations to prevent losses?

Risk acceptance criteria (RAC) have been proposed to allow demonstrations of risks being made as low as reasonably practicable (ALARP). These have been compared and calibrated against other RAC commonly found in other safety standards, also those applied by other industries.

Perform a gap analysis of RBAT and further develop the framework

A gap analysis was performed to assess if there are any gaps between the RBAT framework developed in Part 1 and what is required to address the MASS concepts and sub-functions selected as test cases.

A standard format was used for the gap analysis. This included identifying any gaps in the current framework ("current state"), describing how the framework should look like ("desired future state"), and proposing recommendations for how to close the gaps.

The analysis revealed sixteen gaps. Some gaps could be closed as part of the activities documented in this report (marked "Implemented"), while the remaining will be addressed as part of developing the test cases (marked "Planned"). A complete list of gaps and recommendations is provided in sub-chapter 4.3, Table 4-1.

DEFINITIONS

Terms	Definitions
abnormal situation	A disturbance in the normal operation which can potentially result in an accident.
accident	An unintended event involving fatality, injury, ship loss or damage, other property loss or damage, or environmental damage (IMO, 2018).
accident category	A designation of accidents reported in statistical tables according to their nature, e.g., fire, collision, grounding, etc. (IMO, 2018).
accident scenario	A sequence of events from the initiating event to one of the final stages (IMO, 2018).
agent	Human or machine (computer) responsible for performing or supervising control actions.
annunciated failure	An annunciated failure condition is one which fails 'actively', i.e., in such a manner as to inform crew of the failure, either by virtue of indicators or via vessel behaviour obviously attributable to it (adapted from Kritzinger, 2016).
anticipated event	Events which do not force the system outside the safe operating envelope (SOE), and which can be handled while also maintaining normal operations.
automation	The execution by a 'machine' *agent* (usually a computer) of a function that was previously carried out by a human" (Parasuraman & Riley, 1997).
autonomy	"Technology operates alone". See sub-chapter 3.3.1 in Report 1oo2 for Part 1 of RBAT (DNV GL, 2020a).
causal factors	The minimum combination of causes required to initiate the unsafe condition/mode. May comprise of a single initiating cause, a combination of multiple causes, or initiating causes in the presence of other enabling events.
common cause failures	Failures of multiple items, which would otherwise be considered independent of one another resulting from a single cause (IEC, 2018).
ConOps	Document describing the characteristics of a proposed system from the viewpoint of an individual who will use that system (ISO/IEC/IEEE 15288:2015).
context	External and internal environment in which the organization seeks to achieve its objectives (ISO, 2009).
control	Purposeful action on or in a process to meet specified objectives (IEC, 2013).
control function	Control actions performed by humans or machines for the accomplishment of a functional goal (adapted from IEC, 2000).
control action	Acquisition of information, analysis of information, decision-making, or implementation of physical actions performed as part of a control function.

Terms	Definitions
direct cause	Events which singly, or in few numbers, can cause an accident (and severe losses) if they occur in the presence of a hazard.
enabling event	Occurrence of a failure or presence of a hazard which contributes to escalating an unsafe condition/mode into an accident.
failure	Loss of the ability of an item to perform the required (specified) function within the limits set for its intended use. This occurs when the margin (to failure) is negative (DNV, 2021b).
failure cause	Set of circumstances that leads to failure (IEC, 2018).
failure condition	A condition with an effect on the vessel and its occupants (if present), both direct and consequential, caused or contributed to by one or more failures, considering relevant adverse operation or environmental conditions (SAE, 1996).
failure effect	A description of the operation of a system or an item as the result of a failure; i.e., the consequence(s) a failure mode has on the operation, function or status of a system or an item (SAE, 1996).
failure frequency	The number of failures expressed in failures per unit of time (calendar or operational).
failure mechanism	Process that leads to failure (IEC, 2018). The process may be physical, chemical, logical, psychological or a combination thereof.
failure mode	The observed way in which the failure (of an item) occurs (adapted from SAE, 1996 and DNV, 2021b).
function	Specific purpose or objective to be accomplished, that can be specified or described without reference to the physical means of achieving it (IEC, 2020). In RBAT functions refer to how systems perform to successfully accomplish operations. Sub-functions are offspring (sub-goals) of higher-level, parent function.
functional allocation/ assignment	Distribution of functions between human and machine (ISO, 2000). Functional allocation can also be referred to functional assignment (IEC, 2000).
functional analysis	The examination of the functional goals of a system with respect to available manpower, technology, and other resources, to provide the basis for determining how the function may be assigned and executed (IEC, 2009).
functional goal	The performance objectives that shall be satisfied to achieve a higher-level corresponding function (adapted from IEC, 2009).

Terms	Definitions
functional hazard analysis	A systematic, comprehensive examination of functions to identify and classify failure conditions according to their severity (SAE, 1996).
function tree	Hierarchical breakdown of high-level key functions into a set of sub-functions.
hazard	A potential to threaten human life, health, property or the environment (IMO, 2018). For the purpose of RBAT, this is interpreted as the source of harm which, unless managed, has the potential to cause accidents involving harm or losses. In terms of *safety*, a hazard therefore often refers to conditions, situations, or states in which various sources of energy, biological or chemical agents are present.
hierarchical goal structure	Relationship between a goal and sub-goals structured in a hierarchical order (adapted from IEC, 2009).
human-automation interaction	The way a human is affected by, controls and receives information from automation while performing a task (Sheridan & Parasuraman, 2006).
human error	Discrepancy between the human action taken or omitted, and that intended or required to achieve a task goal (adapted from IEC, 2018).
incident	Occurrence of any event, other than an accident, that is associated with a ship or its required infrastructure and affects or could affect its safety.
independent mitigation layer	Measures preventing unsafe conditions or modes from resulting in losses and which are independent of the causal factors which initiates the event.
initiating event	The first of a sequence of events leading to a hazardous situation or accident (IMO, 2018).
internal mitigation layer	A control function's internal capacity to withstand or self-recover from a failure so that normal operations are not disrupted and can continue safely.
item	Subject being considered (IEC, 2018).
key function	High level functional goal shared by a set of control functions. Navigation, manoeuvring, and communication are examples of key functions. In RBAT, key functions are the highest level of functions in the Function Tree.
loss	A loss involves something of value to stakeholders. Losses may include a loss of human life or human injury, property damage, environmental pollution, loss of mission, loss of reputation, loss or leak of sensitive information, or any other loss that is unacceptable to the stakeholders (Leveson & Thomas, 2018).
minimum risk condition	A temporary as-safe-as-possible state that the ship enters when it experiences situations which, if continued, involves operating outside the safe operating envelope.

Terms	Definitions
mission	The commercial, political (e.g., defence) or public intentions which have contributed to and justifies the vessel concept development and operation.
mission model	Hierarchical breakdown of a vessel mission into a set of mission phases and operations.
mission phase	Subdivisions of the mission typically characterized by a recognizable shift in where the vessel is located in terms of geographical surroundings, or the start and end of one or more operations.
mitigation layer	See “Independent mitigation layer” and “Internal mitigation layer”. Measures capable of moving the ship into a minimum risk condition, or recover it from a degraded state and back to normal operations, are both examples of mitigation layers.
node	In RBAT a node is one operation for a mission phase under which a set of control functions and actions a grouped together for analysis.
operations	Activities performed as part of a mission phase in order to achieve the mission goal. Sub-operations are offspring (sub-goals) of higher level, parent operations.
operational goals	The ultimate purposes of a vessel (adapted from IEC, 2009). In RBAT operational goals are explained in terms of the mission, mission phases and operations.
performance	The performance of a technology is its ability to provide its specified functions (DNV, 2021b). These functions contribute to safety/reliability as well as the output or value generated by the system, equipment, or component when in operation.
performance margin	The difference between the achieved performance and the specified performance requirement (DNV, 2021b).
performance shaping factors	Human, workplace, or other contextual factors which have a significant effect on an operator’s or crew of operator’s performance.
process	Set of interrelated or interacting activities that transforms inputs into outputs (IEC, 2018)
reliability	The ability of an item to perform a required function under given conditions for a given time interval or at a specified condition (DNV, 2021b). In quantitative terms, it is one (1) minus the failure probability.
recovery actions	Actions taken to recover the system from a degraded, failed or unsafe state and back to a state which allow normal and safe operations to be continued.

Terms	Definitions
redundancy (of a system)	Having multiple capabilities for performing the same function, typically in parallel (DNV, 2021b). Alternatively, Provision of more than one means for performing a function (IEC, 2018).
risk control measure	A means of controlling a single element of risk (IMO, 2018). This may refer to […] measures taken to reduce the risks to the operation of the system, and to the health and safety of personnel associated with it or in its vicinity by (DNV, 2021b): — reduction in the probability of failure — mitigation of the consequences of failure *Guidance note:* The usual order of preference of risk control measures is: a) inherent safety b) prevention c) detection d) control e) mitigation f) emergency response.
risk control options	A combination of risk control measures (IMO, 2018).
safe operating envelope (SOE)	Conditions, both internal and external, in which a system can safely execute its normal and planned operations.
scenario	Possible sequence of specified conditions under which the system, item or process functions are performed (IEC, 2018). See also “accident scenario”.
severity	Relative ranking of potential or actual consequences of a failure or a fault (IEC, 2018).
situational awareness	Situational awareness or situation awareness (SA) is the perception of environmental elements and events with respect to time or space, the comprehension of their meaning, and the projection of their future status (Endsley 1995).
supervision	A role with an explicit responsibility to monitor system performance and detect abnormalities so that the desired outcome can be achieved through implementation of corrective responses.

Terms	Definitions
system	Combination of interacting elements organized to achieve one or more stated purposes, i.e. goals (IEC, 2018).
task	A set of [control] actions taken by humans to enable functions and perform operations. A task may involve interactions with several different functions, but also with humans. Task goals is the same as *operations*.
undetected/ unannunciated failures	An unannunciated failure is potentially a latent or passive failure condition, or one that is misleading. A failure is latent until it is made known to the crew or maintenance personnel (adapted from Kritzinger, 2017).
unsafe condition/ mode	Incident where a system is operating outside its normal (and safe) operating envelope due to degraded performance (e.g., failures) or exceeded capabilities which, if left unmitigated, has the potential to directly cause an accident.
worst-case outcomes	The most severe foreseeable outcome of an unsafe condition/mode when assuming there is no mitigation. In RBAT, worst-case outcomes assume the contextual presence of a *hazard*. For example, loss of steering (an unsafe condition) close to shore (a hazard) results in a grounding (a worst-case outcome).

1 INTRODUCTION

1.1 Background

EMSA has contracted DNV to perform a functional study for developing a Risk-Based Assessment Tool (RBAT) for maritime autonomous surface ships (MASS). As outlined in DNV's proposal (DNV GL, 2020b) and EMSA's Tender Specifications (EMSA, 2020), the RBAT study consist of three parts:

- Part 1: Develop a framework for a generic MASS risk assessment tool
- Part 2: Test the risk assessment tool on specific cases and develop software tool prototype
- Part 3: Re-iterate testing on more complex cases and finalize the software tool

1.2 Objective

The objective of this report is to finalize and document a first version of RBAT and present an outline of (fictive) MASS concepts to be developed as case studies for testing RBAT.

1.3 Scope of work

The study is currently halfway into Part 2 which includes the following scope of work:

a) Identify and select specific MASS concepts and sub-functions for testing RBAT

b) Develop a risk evaluation technique appropriate to be applied to MASS concepts

c) Perform a gap analysis of RBAT and further develop the framework

d) Develop test cases for the identified MASS concepts and sub-functions, and test RBAT

e) Based on the results from the test cases, update RBAT and develop a first version of a functional software prototype

f) Draft and submit a final report elaborating the tasks above, also by explaining the identified gaps, the changes made and providing appropriate graphic material.

Activities a) to c) are documented in this report, namely the first report of Part 2 and the third report of the RBAT study.

Activities d) and e) will be documented in a separate report, namely the second report of Part 2 and the fourth report of the RBAT study.

1.4 Updates to the framework described in report 2/2 for Part 1

Updates to the framework described in report 2/2 for Part 1 (DNV, 2021a) have been documented as part of the gap analysis in presented in this report's Chapter 5.

2 IDENTIFY AND SELECT MASS CONCEPTS AND SUB-FUNCTIONS

2.1 Purpose

The purpose of this activity is to identify MASS concepts and a set of selected sub-functions[1] which provides suitable input for testing the RBAT framework and methodology.

2.2 Approach

The concepts and sub-functions have been identified and selected following four steps:

1. *Define selection criteria*. The criteria for selecting the test cases were defined with the purpose of achieving two goals. One is (as per EMSA's tender specification /2/) to select the concepts that are currently being planned or already implemented by the industry with the broadest commercial usage. The criteria's second goal is (as per DNV's proposal) to ensure that the selected cases allow RBAT to be sufficiently tested using a representative variety of expected input.
2. *Identification and screening*. MASS concepts and sub-functions were identified by interviewing experts internal in DNV as well as representatives from key industry actors. The interviews were used to obtain an overview of ongoing projects, but also to discuss challenges and motivations for these types of projects. Input from the interviews were further supplemented with information found on the internet and other available sources.
3. *Assessing relevant MASS concepts and sub-functions*. The compilation of MASS concepts and sub-functions identified as part of the screening were then evaluated against the criteria developed to guide the selection.
4. *Present and decide on MASS concepts and sub-functions*. A meeting with EMSA and representatives from EU flag states took place on the 15th of October. The goal of the meeting was reaching a final decision about which MASS concepts and sub-functions should be further developed into suitable input for testing RBAT.

2.3 Defining selection criteria

In DNV's proposal the rationale for autonomous concepts defined by the MUNIN-project (MUNIN, 2015) was suggested to serve as selection criteria for predicting broadest commercial usage. These are as follows:

- Economical sustainability
- Social sustainability
- Legal barriers (and opportunities)
- Ecological sustainability
- Environmental sustainability
- Environmental impact
- Feasibility of technical solutions

The initial thought was to further specify the criteria and make them more measurable, so that a scoring system could be developed which allowed a direct comparison between the different concepts.

It was early on decided to abandon this approach due to how the variation and uncertainties inherent in the available input data would not produce results which accurately enough reflected the scores. Furthermore, initial talks with in-house DNV experts gave the impression that the rationale behind realization of MASS concepts cannot be measured and ranked based on a set of single criteria. Instead, they result from complex interactions and synergies from a wide range of factors of different magnitude and importance, on a case-by-

[1] The term "sub-function" here refers to a lower-level function of what is referred to a "key function" in the RBAT function tree, i.e. the highest function level.

case basis. Thus, it was decided to use the proposed criteria as prompts and discussion topics during the interviews with industry representatives. The goal of identifying concepts with the broadest commercial usage was achieved through evaluating the following aspects:

- Feedback from the interviewees about commercial and technical considerations/ enablers
- The number of concepts identified as belonging to the same category
- How far the various concepts have come in the development stage

As part of evaluating which concepts have the broadest commercial potential, aspects such as operational flexibility (e.g., offering several types of services) and scalability of a fleet was taken into consideration. The number of projects currently being worked on under each concept-category gives an indication of which concepts the market has the most interest in.

To ensure that RBAT is tested based on a complete and representative set of cases, it was opted to not base the selection of concepts solely on the abovementioned criteria. In principle, this could result in selecting three concepts which would be too similar for RBAT to be tested against a representative variety of expected input. It was therefore decided that the combination of concepts and sub-functions should also be chosen based on the opportunity to cover the following system characteristics:

- An unmanned vessel
- A vessel with reduced manning
- A vessel with passengers onboard
- A vessel transporting cargo

Furthermore, the selection of sub-functions should include features such as:

- Sub-function(s) performed in an iterative manner, e.g., navigation w/ collision avoidance
    - One case where the sub-functions are remotely controlled
    - One case where the sub-function is remotely monitored (supervised)
    - One case where the sub-functions are unsupervised
- Sub-function(s) performed in a sequential manner, e.g., cargo handling
- Sub-function(s) which has a continuously demand/presence, e.g., integrated monitoring and control
- Sub-function(s) required as a response to an abnormal and potentially unsafe event

2.4 Identification and screening

The main source of data about MASS concepts was obtained from interviews which was further supplemented with information gathered online.

2.4.1 Interviews

Table 2-1 lists the people who have been interviewed as part of the screening activity. Based on the outputs, DNV is of the impression that the interviewees represented a good mix of various insights, ranging from the specific concepts they were involved in, to technological trends and commercial foresights. Most of the people interviewed had extensive networks with various types of stakeholders, which gave the screening an extra reach beyond the organizations they represented.

Table 2-1: Overview of interviewees

Name	Company	Role
Ørnulf Jan Rødseth	Sintef	Senior Researcher
Päivi Haikkola	One Sea	Naval Architect
Pia Meling	Massterly	VP Sales & Marketing
Tom Eystø	Massterly	CEO
Øyvind Smogeli	Zeabus	CTO
Henrik Stray	Zeabus	COO
Christian Cabos	Wärtsilä	Director Product Development Smart Vessel
Per Marius Berrefjord	DNV	Senior Vice President and Business Development Leader
Are Jørgensen	DNV	Senior Principal Engineer
Øystein Engelhardtsen	DNV	Senior Researcher
Arnstein Eknes	DNV	Business/Segment Director – Special Ships
Tom Arne Pedersen	DNV	Principal Researcher

Each interview lasted approximately 1 to 1,5 hours and were loosely structured around the following key questions:

- Introduction
    - Introduction of meeting participants
    - Brief about the RBAT project
    - The purpose and structure of the interview
- Which concepts or products do you know of?
- Which 3-4 of the mentioned concepts and products do you predict will achieve the broadest commercial usage in the coming years? And why?
    - What is the business case rationale? (Use selection criteria as prompts/discussion topics)
    - What are the technological, commercial, or political enablers driving the development?
    - What are the main challenges and obstacles?
- Any tips about who else to talk to?
- Any tips about information available online?

2.4.2 Identified concepts

The screening identified 47 different projects related to autonomous vessels and technology. A complete list of all the concepts can be found in Appendix A. Figure 1 shows the distribution of different concept across generic categories. It includes a wide variety of projects ranging from those currently being designed and constructed, to those still being developed on a conceptual stage.

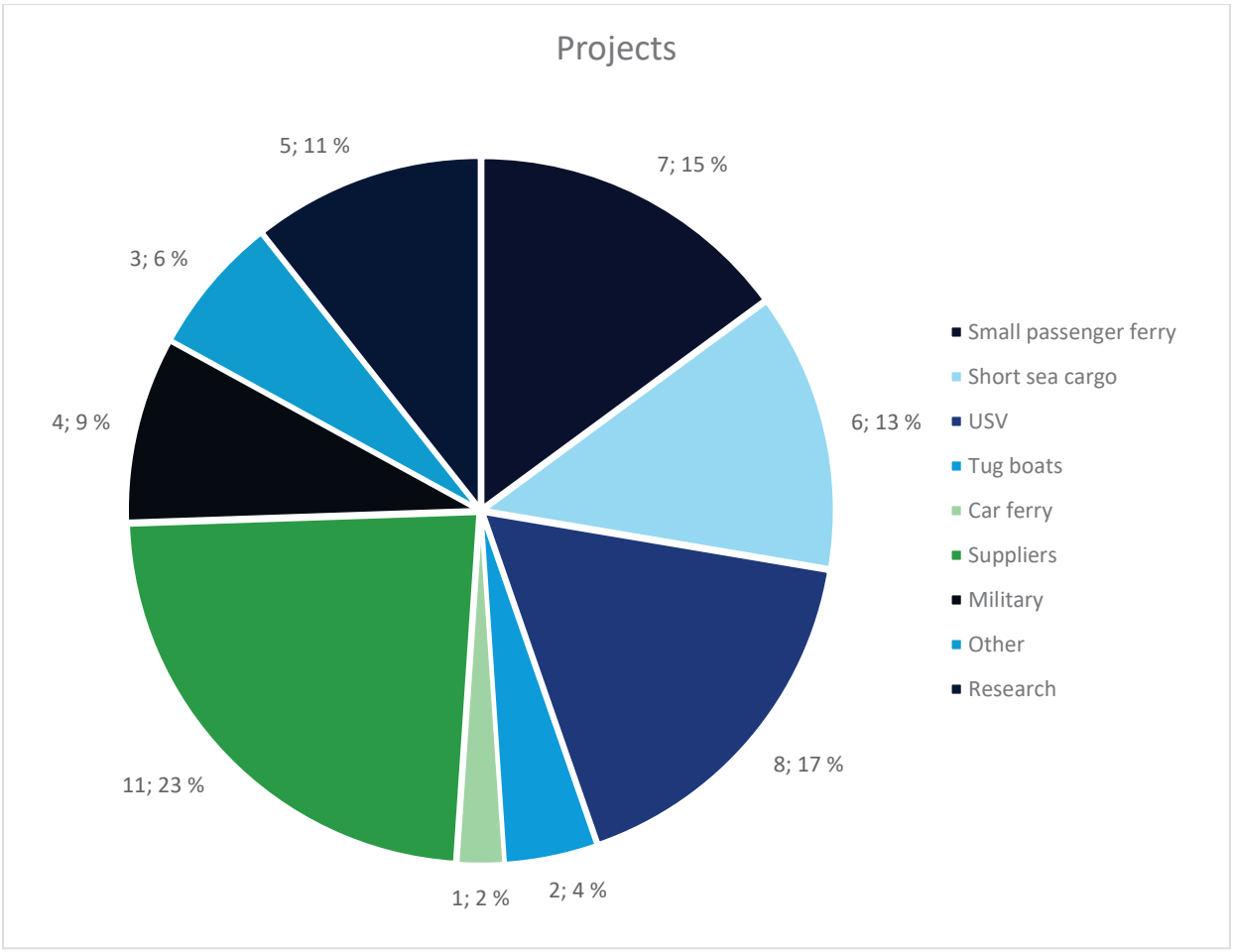


Figure 1: Distribution of concept types identified as part of the screening

As can be seen, when defining a lower cut-off limit of 10%, the five main categories are:

- Suppliers (of assisted solutions and autonomy/automation related products) – 11;23%
- Un-crewed surface vessels (USVs) – 8;17% (out of scope)
- Small passenger ferries – 7;15%
- Short sea cargo – 6;13%
- Research – 5;11%

Due to the USVs relatively small size and simpler configuration compared to conventional vessels being regulated by administrations and class societies, these were considered out of scope for testing RBAT. While they may have been good in testing the concept of multi-vessel fleets, this can be covered by other vessel types such as the small passenger ferries. Based on the number of projects in development, the three top

candidates for further assessment were autonomy/automation related products (functionality), small passenger ferries and short sea cargo vessels.

2.5 Assessing relevant MASS concepts and sub-functions

Data gathered through the interviews suggest that the drivers behind current MASS-related projects tend to be specific in terms of what they intend to achieve. In general, and as already indicated by others (MUNIN, 2015), the goals tend to revolve around increased flexibility, scalability, improved safety, and operational performance, as well as environmental benefits.

Several of the interviewees argued that concepts which include vessels that sail close to shore, with access to clean energy, will be some of the first concepts to be realised. Electrically (battery) powered vessels with reliable machinery which require less maintenance during operations are desirable for vessels with no or reduced manning.

Operations that include repeated crossings and simple routes are easier to automate than those which require more complex navigation. They are also easier to remotely monitor which again allows increased scalability. The possibility to increase a fleet size relative to the required number of remote-control centres and/or vessel operators was also emphasized as essential for being able to develop a sustainable and good business case for several of the concept types.

The feedback also suggested that specific technological solutions (automated/autonomous functionality) can be equally important as the overall vessel and fleet concept, and that future developments may be driven forward equally much by technology suppliers, as by ship owners and/or stakeholders of infrastructures and supply chains. Again, scalability of the technology and products is important when considering the commercial potential.

Although being the most frequently used “buzzwords”, autonomy/automation and remote operations (reducing operational expenditures by enabling reduced manning levels) may not be the only major mechanism driving the current developments. Improved supply chain logistics and fuel efficiency based on smarter data utilization may prove to become equally important. Many of the projects also try to solve a bigger challenge instead of just for example transporting goods from one port to another. A big part of the technological transformation we are seeing comes from the desire to take more control over the whole value chain and thus solving logistical challenges. As for smarter operations, an example can be made by how it is not necessary for a fishing vessel to have enough fuel for a trip around the globe. It’s better to have just the right amount of energy for the exact trip. This may be applicable for a lot of vessels around the world.

The overall assessment of MASS concepts supported the notion to pursue autonomy/automation related products, small passenger ferries, and short sea cargo vessels as candidates for testing RBAT. The following sub-sections summarize the assessment done of each concept. The section structure follows the main questions asked during the interviews.

2.5.1 Short sea cargo vessels

2.5.1.1 What is the business case rationale?

This concept may not become the largest in terms of numbers but is predicted to have broad commercial impact in terms of being developed in full scale to accommodate the specific needs of larger companies, especially those owning the transported cargo. It can sail in territorial waters and consequently be under national laws and regulations. Much like the small passenger ferries, the ongoing projects involve using such vessels for relatively short distances and following a fixed route. This allows for concepts being electric with zero emission supported by a dedicated infrastructure, which is considered as commercially beneficial. Not having to conduct changes of crew shifts onboard, an unmanned vessel also has the possibility of

transporting goods more often and with a higher frequency compared to manned ships. Most of these concepts will be part of larger logistical chains and can replace/reduce use of trucks and transport on the road, making it an environmentally friendly alternative for transport.

2.5.1.2 What are the technological, commercial, or political enablers?

Increased flexibility, safety and more environmentally friendly solutions. Cargo owners can get control over the whole value chain and to a greater extent utilize otherwise expensive assets (e.g., more accurate loading/unloading schedule for trucks, resulting in less waiting time).

2.5.1.3 What are the main challenges and obstacles?

The key challenges to solve are navigation and efficient and reliable cargo handling systems. Another important aspect with unmanned vessels is solving the challenges with moving personnel from ship to shore, as part of a transitional phase. How this should be managed and how the new type of tasks will be dealt with is yet to be seen. The concepts also require a certain fleet size to be controlled from the remote-control centres, to justify the costs of having a remote instead of onboard vessel operator.

2.5.2 Small passenger ferry

2.5.2.1 What is the business case rationale?

This concept is socio-economical beneficial due to relative low-cost and high benefit of increasing transport efficiency. It also gives flexibility and a new tool for city planners, enabling new means of transport and opening new areas. By using the waterway, these ferries can help reduce traffic on the roads, and due to short and fixed routes, be fully electric, making it an environmentally friendly alternative.

A fully autonomous unmanned passenger ferry also allows for on-demand transport and full availability for passengers. An interesting perspective with this concept is the fleet perspective. Operating on short distance routes, at low speed, and with relatively simple navigational requirements, the vessels can potentially be made with a high degree of automation. A human supervisor on shore may therefore be able to supervise a large number of ferries, only interfering in case of an alarm. In case the ferries can be safely operated unmanned, crew costs related to this concept can be kept relatively low.

Another aspect related to the business case rationale, is that the technology in the concept is scalable, making it possible for stakeholders/developers to expand the business to other geographical areas.

2.5.2.2 What are the technological, commercial, or political enablers?

There is potentially a big market as there are many canal-cities, especially in Europe. This is a concept which will sail short, fixed routes, while being in shallow and sheltered waters, thus, the small autonomous passenger ferry is a suitable concept of testing and developing solutions for automated navigation. The concept can also help reduce traffic on roads and bridges as well as improve collective transport.

2.5.2.3 What are the main challenges and obstacles?

Passenger safety and legislation is the main challenge. Today, safety personnel are required onboard for passenger vessels. The success of small passenger ferries may (to a larger extent than other MASS concepts) rely on being allowed to operate crewless to keep the operational costs at a viable level, and so exemption from or changes to the existing rule set is likely to be needed.

2.5.3 Autonomous functions (assisted solutions)

2.5.3.1 What is the business case rationale?

Assisted solutions could potentially help reduce fuel and crew costs, and improve overall operation, making the vessels more reliable, flexible, and safer. By introducing increased automation for specific functions, like auto-crossing, auto-docking, cargo handling, or certain engineering functions, the technology can be gradually introduced and tested without having to significantly challenge the current regulations. The effects

are expected to be improved operator support, more efficient use of resources (fuel, battery power, winds/currents/waves), and reduced workload for the operators.

2.5.3.2 What are the technological, commercial, or political enablers?

The advent of technologies enabling real-time transfer of high-quality data, which makes it possible to monitor specific performances of several vessels in a fleet, are a big driver for autonomous/automated and remote solutions. A remote-control centre can help relieve certain tasks onboard and serve as experts for various purposes, such as planning and troubleshooting. The fleet perspective is also an interesting aspect and involves that vessels can be in preparedness for each other and thereby increasing safety.

2.5.3.3 What are the main challenges and obstacles?

Much of the technology is not yet fully developed and is still subject to research.

2.6 Proposed MASS concepts and sub-functions

Both EMSA's tender specification and DNV's proposals state that "concepts" and "sub-functions" shall be the starting point selecting and developing test cases. However, since these documents were authored, the RBAT framework has matured significantly, and in a way which influences this process. Specifically, Part 1 of the RBAT project developed a "Mission Model" and "Function Tree" which together forms a structure for how use of automation and remote control is assessed. The proposed test cases are therefore outlined according to the RBAT methodology, by also describing additional characteristics such as roles present in remote control centres, vessel manning, energy sources for propulsion, fleet size, geographical operating areas, traffic density, and a breakdown of the vessel mission into phases, operations, and functions.

The proposed MASS concepts are to be developed as RBAT test cases are described in Table 2-2. Note that a RO-Pax ferry has been included as a third Concept C, despite how there was only one case identified as being under development (see Figure 1 and Appendix A). This is because it is used to represent a vessel type which is likely to be adopting emerging technologies for assisted solutions, which was one of the three proposed concept categories.

Table 2-2: Proposed MASS concepts to be developed as RBAT test cases

	Concept A	Concept B	Concept C
Vessel type	**Short sea cargo**	**Small passenger ferries**	**RO-Pax Ferry**
Length	**80 meters**	**15 meters**	**120 meters**
Remote control centre	- Vessel operator - Chief engineer	Vessel operator	Chief engineer
Vessel manning	Unmanned	Unmanned, but with passengers	Reduced manning - Bridge crew - Deck crew
Other roles	Dock operator	Emergency preparedness org.	--
Energy source	Battery	Battery	Battery/diesel (hybrid)
Fleet size	3 sister vessels	10 sister vessels	5 sister vessels
Area of operation	Partly enclosed and open waters	Enclosed/sheltered waters	Short route in narrow waters (fjord)
	Concept-function combination #1		
Mission phase	Arrival in port	Transit to location	Arrival in port
Traffic density	Medium	High	Low
Operation	Perform harbour manoeuvring	Navigate through enclosed/sheltered waters	Perform docking
Function i	***Perform navigation:*** *- Observe surroundings* *- Avoid collision and grounding* ***Perform manoeuvring:*** *- Provide steering* *- Provide acceleration/deceleration* ***Perform mooring:*** *- Prepare mooring line* *- Deploy mooring line* *- Fix/secure mooring line to quay*	***Perform navigation****:* *- Perform voyage planning* *- Observe surroundings* *- Follow planned route* ***Perform manoeuvring:*** *- Provide steering* *- Provide acceleration/deceleration*	***Perform navigation:*** *- Determine vessels position & relative distance* ***Perform manoeuvring****:* *- Provide steering* *- Provide acceleration/deceleration* *- Maintain position* ***Embark/disembark crew & passengers:*** *- Operate ramp*
Supervision	Active supervision	Passive supervision	Active supervision

	Concept A	Concept B	Concept C
Vessel type	**Short sea cargo**	**Small passenger ferries**	**RO-Pax Ferry**
	Concept-function combination #2		
Mission phase	Transit to location	Transit to location	Activities in port
Traffic density	Medium	High	NA
Operation	Navigate through enclosed waters	Navigate through enclosed/sheltered waters	Re-plenish consumables
Function ii	***Perform collision and grounding avoidance:*** *- Detect vessels/objects* *- Classify vessels/objects* *- Observe vessels/objects movements (heading and speed)* *- Determine vessels/objects position and relative distance* *- Determine CPA/TCPA for vessels/objects* *- Implement collision and grounding avoidance strategy*	***Perform collision and grounding avoidance:*** *- Detect vessels/objects* *- Classify vessels/objects* *- Observe vessels/objects movements (heading and speed)* *- Determine vessels/objects position and relative distance* *- Determine CPA/TCPA for vessels/objects* *- Implement collision and grounding avoidance strategy*	**Perform manoeuvring:** *- Maintain position* **Provide electrical power:** *- Charge/receive electrical power from shore*
Supervision	Active supervision	Passive supervision	Passive supervision
	Concept-function combination #3		
Mission phase	Activities in port	Transit to location	Depart from port
Traffic density	NA	High	Medium
Operation	Perform loading/unloading	Navigate through enclosed/sheltered waters	Perform harbour manoeuvring
Function iii	***Handle and monitor cargo:*** *- Plan & prepare cargo handling* *- Un-secure cargo* *- Unload cargo* ***Perform ballasting & trim:*** *- Calculate and verify trim & stability* *-Operate ballast pumps*	***Maintain communication:*** *- Communication/data link between RCC and ferry*	***Embark/disembark crew & passengers:*** *- Operate ramp* ***Provide electrical power:*** *- Generate power* *- Distribute electrical power*

	Concept A	Concept B	Concept C
Vessel type	**Short sea cargo**	**Small passenger ferries**	**RO-Pax Ferry**
	Maintain communication: *- Communication between vessel and dock operator*		
Supervision	Active supervision	Passive supervision	Passive supervision
	Concept-function combination #4		
Mission phase	Transit to location	Emergency response in Transit	Transit to location
Traffic density	Medium	High	Medium
Operation	Handle loss of communication link	Perform evacuation (fire)	Handle blackout (back-up power available)
Function iv	***Maintain communication (data, voice/sound, visual signalling):*** *- Use AIS and light signals to notify other ships* *- Notify VTS* ***Perform manoeuvring:*** *-Maintain position until communication is established (MRC)*	***Provide means for evacuation***: *- Mitigate fire* *- Guide passengers* *- Evacuate vessel* ***Perform navigation:*** *Observe surroundings* ***Perform manoeuvring:*** *- Provide steering and speed adjustments to move away from objects or approach quay/land (MRC)* *- Call for assistance (MRC)* *- Drop anchor*	***Maintain communication (data, voice/sound, visual signalling):*** *- Notify authorities* *- Notify other ships* *- Call for tug assistance* *- Use AIS and light signals to notify other ships* ***Integrated monitoring and control:*** *-Restart of system* ***Perform anchoring:*** *-Emergency release of anchor (MRC)*
Supervision	No supervision	Active supervision	Active supervision

3 DEVELOPMENT OF A RISK EVALUATION TECHNIQUE FOR RBAT

3.1 Updated scope of work for Activity 2b)

One of the activities originally requested by EMSA (2020) was to assess ALARP limits that would be appropriate to apply to MASS. Furthermore, it was stated that the Formal Safety Assessment (FSA) Guidelines (MSC-MEPC.2/Circ.12/Rev.2) should be used to develop an appropriate method of evaluating that the risk is within the ALARP limits. EMSA also expected that the contractor (i.e., DNV) develops risk levels that may be used for the assessment of the risk implied by the intended MASS concepts.

During the development of RBAT, as well as drawing on experience from research and other commercial projects, it became increasingly evident that a traditional approach to risk evaluation, such as that described in the FSA guideline, has limitations when being applied to complex and software intensive systems (such as MASS). Reference is made to a meeting with EMSA on the 7th of October 2021, where the objective was to discuss various options for how to perform risk evaluation of software related events using RBAT. Much drawing on the experiences from other industries, the main conclusion from this meeting was that it is not recommended to apply a probabilistic (quantitative) approach as part of evaluating risk associated with unwanted events which are software-related (e.g., failures). This is due to a complex nature of software-related failures and associated uncertainty related to the likelihood of such failures.

Software-related failures may be introduced at a system design requirements level e.g., through overlooked dependencies among the technical, operational, human, and organisational components of systems, specifications that are based on inadequate understanding of physical processes and the environment a system is operated in (e.g., being outside of operational envelope), or unexpected inputs for which no specific response has been specified.

Software-related failures may also be introduced at software design and implementation level, e.g., through unforeseen dependencies in the internal dataflow, or unsafe use of the programming language. In addition, failures that are caused by degraded but still operational hardware may in some cases manifest themselves as software-related failures.

Therefore, a system may fail to meet expectations in a substantial number of ways, and thus a practice of assigning likelihood to software-related failures come with an inherent uncertainty which makes confident decision-making difficult.

Due to the complexity of initiating events in which software-related failures are a significant contributor, and the uncertainty related to its likelihood, it is therefore proposed to direct the attention of RBAT towards event *mitigation*. Consequently, a qualitative approach is proposed to be developed, where the risk is evaluated as a function of how severe the potential consequences of an unwanted event/unsafe condition or mode are, combined with how effective mitigation measures are at recovering the system to a as safe-as-possible state or preventing losses.

As action from the meeting held on the 7th of October, DNV was tasked with proposing a new way for how best to implement Activity 2b) of the RBAT project. The following approach was subsequently reviewed and accepted by EMSA:

DNV shall develop a (qualitative) technique suitable for evaluating the risk associated with scenarios identified when using the RBAT. Risk shall be evaluated as a function of consequences of the unwanted event and mitigating measures implemented to reduce consequences/recover a system to a safe-as-possible-state. Two separate indexes shall be developed; one which defines levels of severity of consequence (e.g., from no effect to catastrophic), and another which defines levels of mitigation effectiveness (e.g., from fully recovered to a safe state to no recovery). Severity is understood as a degree

of impact on safety (e.g., human safety and system degradation leading to an accident), while mitigation refers to how successful a response is at reducing consequences of unwanted event by preventing the impact or losses. The two indexes shall be combined to form a risk matrix which can be used to define levels of risk by considering degree of severity against the degree of mitigation effectiveness. The risk matrix shall include criteria for what is considered acceptable and unacceptable risk levels.

3.2 Rationale for the proposed risk evaluation approach

The proposed approach is to a large extent based on what is already done in Maritime today. However, since the existing maritime safety philosophy for control systems relies heavily on operators being present onboard and a part of the safety loop, some changes are required when autonomous control functions are introduced. Thus, inspiration from other industries have been utilised in the proposed approach.

To provide a rationale for the proposal this section contains a high-level description of management of control-related risk in other industries, a similar high-level description regarding management of control risk in maritime, and finally a proposed high-level risk assessment scheme for MASS.

3.2.1 Management of control-related risks in other industries

Table 3-1 below shows a typical example of a risk matrix, that also contains an example of risk acceptance criteria for combinations of Severity and Probability of dangerous failure per year.

Table 3-1: Example of classical risk matrix

			Severity				
Probability of failure per year			Negligible	Minor	Significant	Severe	Catastrophic
Frequent	>=1		Medium	High	High	High	High
Probable	>=1/10	To <1	Low	Medium	High	High	High
Occasional	>=1/100	To <1/10	Low	Medium	Medium	High	High
Remote	>=1/1000	To <1/100	Low	Low	Medium	Medium	High
Very remote	>=1/10000	To <1/1000	Low	Low	Low	Medium	Medium
Improbable		<1/10000	Low	Low	Low	Low	Medium

A problem with using this matrix in decision-making is that quantitative estimation of failure probabilities typically is not possible for systematic and systemic failures (see Annex D for an explanation of these failure types). The way this is dealt with when it comes to control functions in other industries like automotive, aviation, railway, machinery, process, oil & gas etc. can be summarised as follows:

1. Other industries will allocate Safety Integrity Levels (SIL) (IEC, 2010), Development Assurance Levels (DAL) (SAE, 2010), Performance Levels (PL) (ISO, 2015) or similar to individual control functions. Each level reflects a certain range of probability for random hardware failure, e.g. “Remote” in Table 3-1 above, corresponds to what is required to claim SIL 2, DAL-C or PL=d performance. Demonstration of probability of dangerous random hardware failure will typically be performed through quantitative analysis utilising e.g. fault tree analysis.
2. Many people associate the integrity levels with quantitative reliability analysis as described in pt. 1 above. However, the main topic in the various functional safety standards is the vast potential for systematic failures in control functions, in particular in software (see also Appendix B). Consequently, the achievement of a specific integrity/assurance/performance level requires the demonstration of so-called systematic capabilities. See item 5 below for further details.

3. Control functions considered safety critical are often emergency shutdown functions that are separated from the main control functions used during normal operations. This is typical for oil & gas, the process industry, machinery etc. These kinds of safety functions are typically capable of mitigating both random hardware and systematic/systemic failures in the main control functions, and consequently the main control functions are not considered safety critical at all.

4. If an independent safety function needs to be equally advanced as the main control function to mitigate all forms of systematic/systemic faults, a separate emergency function is typically not implemented. In such cases the main control functions are considered safety critical and must be developed to a SIL or similar. Such functions are found e.g., in automotive, railway and aviation. They typical contain a large number of mechanisms for Fault Detection Isolation and Recovery (FDIR) aimed at managing both hardware and software related faults. However, such functions will typically not be fault tolerant to all types for systematic/systemic failures.

5. Requirements for systematic capabilities applies to processes and methodologies used in risk analysis, requirements specifications, system/software/hardware architecture, detailed design, implementation, various forms of verification and validation. The required level of rigour and methods applied in those overall processes varies with required integrity/assurance/performance level. The goal is that the processes applied shall be so strong that the probability of dangerous systematic failure is small compared to the required probability of dangerous random hardware failure. Since, the probability of systematic failures is very difficult to quantify, it is also difficult to substantiate the claim that the probability really is that low, in particular for complex functions. However, experience from various industries seems to indicate that the various schemes are working quite well. This is further elaborated in 6 below.

6. Regarding what is achieved in the various industries, the standards alone do not provide all knowledge needed to achieve the required integrity. As an example, it is much more challenging to demonstrating Automotive Safety Integrity Level C (ASILC), for an adaptive cruise control function in a car compared to demonstrating a SIL3 for a relatively simple separate emergency shutdown function used in oil & gas. More advanced methodologies will be needed in the automotive case, for example when it comes to types of risk analyses being applied. Thus, the organisations involved need to have competence, processes and methods in place that are fit for purpose. In practice, the safety community within each industry will reach a level on consensus of what is standard industry practice for different types of applications within that industry. Sometimes this goes beyond what is the minimum requirements in the standards. The safety communities within various industries are also influencing each other. This can be observed when new revisions of the various standards and guidelines are issued. What today is required at comparable integrity levels cross industries is in a high-level perspective similar, although there are quite a lot of variations when it comes to more detailed requirements.

7. When performing allocation of SIL, DAL or similar, subfunctions or individual components may be allocated a lower target level than the overall function. Typically, such downgrading requires some level of *independence* between the various subfunctions and components.

8. Some safety standards e.g., the ones for machinery and automotive allows for reduction of required level based on exposure rate. E.g., ISO 26262 (ISO, 2011) used in automotive, allows for targeted Automotive Safety Integrity Level (ASIL) to be reduced by one level if the relevant hazards are estimated to be present less than 10% of the average operational time for the car, and two levels if it can be argued that the relevant hazards are present less than 1% of the average operational time.

9. The requirements in the functional safety standards do apply not only to top-level applications, but to all software and hardware components within a safety critical system, e.g. sensors, CPU boards, operating systems, communication protocols, network stacks, compiler libraries, low-level software packages proving support for a specific CPU board etc. Such components are typically available commercially of the shelf, and there is a huge market for components that are certified or qualified for use in safety critical applications. As an example, Wind River which is a supplier of the VxWorks real time operating system extensively used in maritime, also provide versions qualified for use in safety critical systems in automotive, aviation, railway etc. Qualified versions of commercial off-the-shelf software components may often have a somewhat more limited functionality than the standard versions and may also only work with specific hardware components for which a qualified board support package has been developed.

10. For separate safety systems used to provide emergency stop, the level of standardisation is particularly high, since vendors of safety controllers, e.g., Siemens, ABB HIMA etc., also will provide certified function blocks that are used to build quite simple cause & effect logic without programming from scratch. For such functions, use of the so called “de-energize to safe/trip” principle allows for system specific safety applications realising quite simple cause & effect logic. This combination of a very strong standardised platform and simple system specific safety applications limits the room for systematic/systemic faults, and allows for a simpler verification and validation process, which relies heavily on testing. The IEC 61511 standard (IEC, 2016) used in oil & gas and in the process industry, anticipates that all components used has been qualified as SIL capable, so that this kind simplified process can be applied. If that is not the case, IEC 61511 disqualifies itself and requires IEC 61508 (IEC, 2010) to be applied.

11. Regarding probability of dangerous failure in a continuous or frequently used control function that has not been developed to a Safety Integrity Level or similar, the lowest claim regarding probability typically corresponds to Probable in the risk matrix above, see e.g., IEC 61511. The limit reflects the start of the SIL1 range for such functions, and to claim a lower probability, the control function would need to be developed to a SIL or similar.

12. When complying with the above-mentioned IEC standards, the highest risk reduction factor that can be claimed for an independent safety function that will be activated on-demand but has not been developed to a Safety Integrity level is 10. This means that such a function is assumed to work 9 out of 10 times and that probability of failure on demand is 0.1. The limit reflects the start of the SIL1 range for on-demand safety functions, and to claim a lower probability the control function would need to be developed to a SIL or similar

13. Many functional safety standards provide a route were control functions and/or component that was not originally developed to a SIL or similar can be approved based on field experience as “proven in use”. However, these approaches are challenging as it is difficult to know whether a software has been exposed to all relevant combinations and sequences of inputs, because software may have been updated during its service and because the required amount of high-quality field data may not be available.

3.2.2 Management of control related risks in maritime

The maritime industry is relying on prescriptive rules (e.g., class) and does not typically apply functional safety standards. The overall approach and differences to other industries can be summarised as follows:

1. When it comes to mitigation of random hardware faults, the maritime industry relies on redundancy in the same way as other industries. Risk analyses, like failure mode effect and criticality analysis (FMECA), are performed to analyse threats to the redundancy concept. Such risks are evaluated

qualitatively only. It is not required to perform quantitative analysis in order to demonstrate that the probability of two dangerous random hardware failures, affecting both channels in a redundant system is below a specific target.

2. Regarding mitigation of systematic/systemic failures manifesting themselves in software, the maritime strategy has been to rely on a fallback chain activated automatically or by the operator. E.g., if autonomous operation of a DP system has been aborted due to systematic failure in software, the fall-back chain is as follows: “Independent Joystick” which means that the operator can manually control all thrusters; individual lever per thruster; and, local control levers in the engine rooms. DP represent a border case when it comes to applicability of the current maritime approach, since that fallback chain may not be effective for all types of operations where DP may be used.

3. When it comes to reducing the probability of systematic faults, there are relatively few mandatory requirements compared to other industries. Testing is required and very important when it comes to removal of systematic/systemic faults, but as described in Annex D, the number of possible input combinations and possible execution paths typically prevents exhaustive testing even when using a simulated environment. This means that testing typically can only demonstrate the presence of conditions that can lead to failures and not their absence. The lack of detailed requirements means that the various suppliers enjoy a considerable freedom regarding how much effort to put in and what methodologies to use in order to remove systematic and systemic failures.

4. The maritime industry does not require the use of hardware and software components that has been qualified for use in safety critical systems as described in 3.2.1, item 9 and 10. However, even if it is not formally required, such components are used onboard ships, typically in separate safety functions providing various types of emergency stops.

5. As briefly discussed in 3.2.1, item 6, there is a big difference in demonstrating integrity for a complex main control function, such as one used for automated navigation, and a relatively simple emergency shutdown function. In maritime the differences between such safety critical systems will be handled through prescriptive class rules.

6. Maritime does not have the same type of scheme for approval based on “proven in use“, as described in 3.2.1, item 13. However, as in other industries, the suppliers benefit from having their systems onboard many operational units. This typical provides a continuous stream of feedback that is used to improve the systems. E.g., suppliers of DP systems benefit from decades of feedback from a large number of different vessels and types of operations. This is one of the reasons that DP systems today are much more reliable than earlier, which again has led to DP being used in some very critical operations. It is important to consider that the same level of maturity cannot be expected from novel autonomous functions. It is also important to considered that suppliers sometimes make major changes to system architecture and/or hardware/software execution platform for existing types of system. Such new generations may introduce a new set of systematic faults, that was not present in the previous generation. Maintenance releases issued more frequently also carries the risk of unwanted side-effects.

3.3 Proposed approach for risk evaluation of MASS

The following sub-chapters present the two indexes which together form the risk matrix to be used in RBAT. Criteria for what is considered acceptable and unacceptable risk levels are also suggested.

3.3.1 Worst-case outcome severity index

An index for ranking the severity of worst-case outcomes is presented in Table 3-2 below. It adopts the four levels of severity described in the FSA guideline (IMO, 2018), with some slight changes to the choice of wording. “No effect” and “Negligible” has also been added as options to be used for events with no or less significant effects in terms of safety.

Table 3-2: Severity index for worst-case outcomes

Severity	Effects on human safety
No effect	No injuries
Negligible	Superficial injury
Minor	Single injury or multiple minor injures
Significant	Single serious or multiple injuries
Severe	Single fatality or multiple serious injuries
Catastrophic	Multiple fatalities (more than one)

3.3.2 Mitigation effectiveness index

Regarding initiating events that manifest themselves in control functions, it is not foreseen that the maritime industry will seek to reduce the likelihood of such failures by going in the direction of building high-integrity main control functions according to functional safety standards[2].

The occurrence of a failure where external intervention by human or another system is required, should be estimated as *Probable*, if a classical risk matrix of the type shown in Table 3-3 below is being used. This is in line with the classification that would have been made if a functional safety standard had been applied. Since control functions on a ship are complex, for which dangerous faults could occur in many ways, the top-level risks are aggregated. Thus, this assumption is not seen as very conservative.

A standard redundant control system used in maritime is considered to (stand-alone) have a moderate strength when it comes to mitigate internal failures *without* any operator intervention. Such a system is expected to be tolerant to single random hardware failures, and there may also be mitigating measures that can prevent losses from some type of systematic faults, e.g., in software, but there will typically be types of systematic faults which the system itself will not be able to mitigate if they were to occur.

Table 3-3: Classical risk matrix based on probability estimation

		Severity				
Probability of failure per year		Negligible	Minor	Significant	Severe	Catastrophic
Frequent	>= 1	Medium	High	High	High	High
Probable	>= 1/10 to < 1	Low	Medium	High	High	High
Occasional	>= 1/100 to < 1/10	Low	Medium	Medium	High	High
Remote	>=1/1000 to < 1/100	Low	Low	Medium	Medium	High
Very Remote	>=1/10000 to < 1/1000	Low	Low	Low	Medium	Medium
Improbable	< 1/10000	Low	Low	Low	Low	Medium

Since the maritime industry does not require risk mitigating control functions to have a high integrity, it is suggested that maximum risk reduction factor that can be claimed for such functions is 10. This

[2] In a similar way to what is done e.g., in automotive, aviation, and railway, ref. 3.2.1.

classification is also in line with what is done in functional safety standards in that the maximum risk reduction factor that can be claimed for a risk mitigating control function not developed according to such a standard is 10 (see also 3.2.1 item 12 in 3.2.1). A risk reduction factor of 10 corresponds to one probability class lower in the probability-based risk matrix (Table 3-3).

Drawing on these arguments it is therefore suggested to avoid the use of probability altogether and instead use a risk matrix that focus on available risk mitigation layers like one shown in Table 3-4.

For control systems the thinking behind the mitigation scale is as follows:

- For a control function that is not fully redundant, the effectiveness of internal risk mitigation is considered *Low*. There may be mitigation measures that can prevent losses from some types of random hardware failures, but the function being analyzed is not fully hardware fault tolerant nor fully tolerant to systematic/systemic faults.
- As discussed in the introduction to this section, a standard critical control system used in maritime is expected to be redundant. This implies that there at is least one *internal* mitigation layer that can prevent losses from various types of random hardware failures. There may also be mitigation measures that can prevent losses from some types of systematic faults, but for such systems there will typically be types of systematic/systemic faults that cannot be mitigated without external intervention. Thus, the effectiveness of the internal mitigations in the system should be classified as *Moderate*.
- An *independent* mitigation layer will increase the strength of the mitigating measures by one level. For example, an independent emergency function that can mitigate a control failure in a standard control system will raise the strength from Moderate to Medium. A further strengthening to High will require a second independent mitigation, and so on.

Table 3-4: Effectiveness of mitigation layers

Effectiveness	Description
Very high	At least three effective *independent* mitigation layers that for the assessed scenario can prevent losses regardless failure cause.
High	At least two effective *independent* mitigation layers that for the assessed scenario can prevent losses regardless failure cause.
Medium	At least one effective *independent* mitigation layer that for the assessed scenario can prevent losses regardless failure cause.
Moderate	At least one *internal* mitigation layer that can prevent losses from random *hardware* failures. The control function has additional capacities for self-recovery from other types of failures, however, for the assessed scenario these are not effective regardless failure cause.
Low	The control function has some capacities for self-recovery, however for the assessed scenario these are expected to have a limited effect.

Note: In case of a control function with low capacity for self-recovery is combined with one independent mitigation layer capable of preventing losses regardless of failure cause, the total effectiveness should be considered on a case-by-case basis.

3.3.3 Risk matrix and acceptance criteria

Table 3-5 combines the severity index (Table 3-2) and mitigation effectiveness index (Table 3-4) into what is proposed to be the risk matrix for RBAT, including risk acceptance criteria. As requested by EMSA, it is here recommended that the "as low as is reasonably practicable" (ALARP) principle is applied for risk evaluation[3]:

- High (red region): Risk cannot be justified and must be reduced, irrespectively of costs.
- Medium (yellow ALARP region): Risk is to be reduced to a level as low as is reasonably practicable.
- Low (green region): Risk is negligible, and no risk reduction is required.

The term *reasonable* is interpreted to mean cost-effective. Risk reduction measures should be technically practicable, and the associated costs should not be disproportionate to the benefits gained. How to perform cost-benefit assessments is extensively explained in the FSA guideline and therefore not repeated here.

Table 3-5: Risk matrix based on evaluation of available risk mitigating measures

Effectiveness of risk mitigation layers	Severity					
	No effect	Negligible	Minor	Significant	Severe	Catastrophic
Low	Low	Medium	High	High	High	High
Moderate	Low	Low	Medium	High	High	High
Medium	Low	Low	Medium	Medium	High	High
High	Low	Low	Low	Medium	Medium	High
Very high	Low	Low	Low	Low	Medium	Medium
Extremely high	Low	Low	Low	Low	Low	Medium

An important part of RBAT will be to evaluate whether specific mitigation layers can be considered effective for specific types of failures, in a specific operational context. This *mitigation analysis* is further detailed in section 3.4 below.

When comparing the risk picture associated with a specific function and corresponding risk mitigation measures to relevant acceptance criteria, the following can subsequently be considered:

1. Operational restrictions such as speed limits may be used to reduce the Severity of operational scenarios.
2. It may be possible to follow, e.g., the automotive industry in evaluating exposure rate to the relevant hazard. If it can be argued that the Hazard is relevant less than 10% of the average operational time per year, the required level of mitigations may be reduced by one level. If the hazard is relevant less than 1% of the average operational time per year, the required level of mitigation may be reduced by two levels.
3. If the causal factors behind the initiating event[4] are not related to software, it may be possible to argue for a lower probability than what has been generally anticipated for control functions (ref. rationale made in sub-chapter 3.3.2). In that case fewer independent risk mitigation layers may be required to meet the acceptance criteria. For such events the classical type of risk matrix shown in Table 3-1 can be used as a starting point to determine the initial risk picture before looking at relevant independent mitigation layers.
4. It should be possible to argue that a single mitigation will increase the effectiveness of the mitigation by more than one level. One example may be that if it can be demonstrated than an emergency

[3] MSC-MEPC.2/Circ.12/Rev.2, chapter 4.

[4] Causal factor(s) initiating the event which results in a unsafe condition/ mode

stop function for machinery has a Performance Level (PL) = *d* performance according to the ISO 13849 safety standard for machinery, this would be considered a two-level increase.

5. It should also be possible to demonstrate that safety critical control functions performing more complex functionality than emergency stop has a better performance than what is anticipated in the scheme above. Such claims should be substantiated in an Assurance Case or similar. More advanced forms of risk analysis, carefully selected components and sharper development processes than what traditionally has been applied in the maritime may be required to substantiate such claims.

3.4 Mitigation analysis

Following the re-defined scope and proposed approach for risk evaluation, the methodology *recovery analysis* presented in the RBAT project's second report (DNV, 2021a) has been revised and updated to account for the implications associated with systematic/systemic failures. An outline of the method, now referred to as *mitigation analysis*, is provided below. The step-by-step guidance for the RBAT methodology has also been updated (see chapter 5) to include the proposed changes.

Identify and describe mitigation layers

The first part of the mitigation analysis is to identify and describe the concept's *mitigation layers*. This is initially done as part of writing the ConOps, and a preliminary set of mitigation layers should be developed *prior* to using RBAT. These can be identified by considering what the responses would be in case various loss of control or accident scenarios should occur. EMSAs accident categories can be used as a starting point (DNV, 2021a).

Identified mitigation layers are defined by describing the following:

- ID
- Name
- Short description

Once defined, information necessary to evaluate the mitigation layers risk reducing effects must be gathered. This includes:

- Applicability of the mitigation layer
    - For which unsafe conditions/ modes the mitigation layer is a planned response
    - For which mission phases the mitigation layer is applicable
    - For which mission phases the mitigation layer is NOT applicable
- System and human involvement in the mitigation layer
    - Systems required for executing the mitigation layer
    - How humans are involved in executing the mitigation layer (see sub-chapter 5.3.2.2 for further explanations)
- Limitations to the mitigation layer
    - Environmental and other external limitations in the mitigation layer (e.g., sea state, visibility, day/night, availability of external resources)

  - Resource limitations in the mitigation layer (e.g., time, fuel, energy reserves, manpower, etc.)
  - Limitations in the sequence mitigation layers can be introduced (e.g., a mitigation layer should only be activated after another has been exhausted)
- Transitions between and from mitigation layers (including minimum risk conditions)
  - How to re-enter a normal or as safe-as-possible operational mode (in case the mitigation layer involves entering a minimum risk condition (MRC)
  - What the next mitigation(s) in the sequence is, and how to introduce it ("None" in case the mitigation is a last resort MRC)
  - Emergency response in case there are no other mitigation layers available

Nominate mitigation layers which can prevent losses

For each combination of unsafe condition/ mode and causal factor(s), nominate potential 1st, 2nd, 3rd *independent* mitigation layers:

- Consider if any of the pre-defined mitigations layers are relevant.
- If new mitigation layers are identified, add these to the list, and *then* nominate them in the analysis.

In addition to nominating mitigation layers, the analysis should also consider whether the control function includes any internal capacities for self-recovery from failures (*internal* mitigation layers). These are more difficult to pre-define and are therefore identified and described on a function-by-function basis.

Analyse the mitigation layers to qualify their effectiveness

The mitigation analysis itself is about assessing how effective the mitigation layers are at preventing the unsafe conditions or modes from resulting in losses. Given the technical, environmental, and operational conditions of the scenario, assess whether each of the mitigation layers fulfil the following performance criteria:

- Functional: The mitigation layer's design and intended use makes it effective at preventing the unsafe condition or mode from resulting in (safety) losses.
- Integrity: The mitigation layer is in place, its condition is intact, and it can be relied upon to function under the expected circumstances.
- Robustness: The mitigation layer will remain functional after the unsafe condition or mode has occurred, taking any disturbances and/or accidental loads into account.
- Independence[5] (see sub-chapter 3.4.1.1 for more details):
  - of the causal factors which initiated the unsafe condition/mode
  - of each other (in case a mitigation fails)
- Human involvement: A final assessment is to consider whether the mitigation layer is designed and implemented in such a way that reliable human performance can be expected. This assumes that operator actions are required. See sub-chapter 3.4.1.2 for more details.

[5] Only relevant for independent mitigation layers, and not self-recovery capacities (i.e., internal mitigation layers)

Make note of why nominated mitigations fail to qualify so that actions can be taken to qualify them at a later stage in case the risk turns out as unacceptable.

Rank mitigation effectiveness

After having qualified the nominated mitigation layers, their effectiveness in preventing losses is ranked using the index provided in Table 3-4 above.

3.4.1.1 Independence

When evaluating whether it is possible to take credit for an independent mitigation layer, it is important to evaluate whether the cause of the unsafe condition/mode may also negatively affect the mitigation layer and thereby represent a common cause failure.

Correspondingly when evaluating whether it is possible to take credit for more than a single mitigation layer, it is important to evaluate whether a failure in one mitigation layer also could impact others.

To evaluate to what extent a mitigation layer is sufficiently independent, it is necessary to look at the system/functions involved in the operation from 4 different perspectives. These are:

- Composition,
- Environment,
- Structure, and
- Mechanisms.

The RBAT step-by-step method description has been updated to include specific guidance on how to evaluate independence using the four perspectives (see Table 5-4 in sub-chapter 5.3.2).

3.4.1.2 Human involvement

Similar to how control functions as part of normal operations may require some control actions to be performed by human agents, mitigation layers may also rely on human involvement. Because this involvement is (by definition) characterized by being required *on demand* in case of abnormal situations (e.g., an unsafe condition), they can be said to include the following information processing stages (Parasuraman, Sheridan, & Wickens, 2000; DNV GL, 2020a):

- Information acquisition: Perception of sensory information about the situation
- Information analysis: Making sense of the situation and predicting future events
- Decision-making: Selecting a course of action among several possible alternative options
- Implementation of actions: Executing activities required to achieve desired outcome

For mitigation layers which depend on operator actions, part of the mitigation analysis is to consider whether the systems involved executing the mitigation layer is designed and implemented in such a way that reliable human performance can be expected. This is done through two steps:

- Identify which of the information processing stages are required as part of the mitigation layer. Normally all the stages will be required, but to varying degree depending on how automation is solved.
- Determine whether one or more of the hindrances are present and if their effect(s) on human performance is so negative that the required operator action(s) will fail. If this is the case, the associated mitigation layer fails to qualify as being effective.

The RBAT step-by-step method description has been updated to include specific guidance on how to evaluate presence of potential hinderances against successful human involvement (see Table 5-5 in sub-chapter 5.3.2.2).

4 GAP ANALYSIS AND FURTHER DEVELOPMENT OF RBAT

4.1 Purpose

The main purpose of this activity is to assess if there are any significant gaps between the RBAT framework developed in Part 1 and what is required to address the MASS concepts and sub-functions selected as test cases.

The RBAT framework from Part 1 here refers to:

- The multi-level function map, i.e., the RBAT Mission Model and Function Tree
- Use of automation & remote control
- RBAT accident model
- RBAT risk model, including:
    - Failure analysis (renamed to “Hazard analysis”)
    - Recovery analysis (renamed to “Prevention analysis” and “Mitigation analysis”)
    - Consequence analysis
    - Risk control

4.2 Approach

A standard format was used for the gap analysis:

- Focus areas: Areas for which gaps are to be identified (here: the various parts of the RBAT framework)
- Current state: The state of what is in focus (and should be improved)
- Desired future state: What the state should be after implementing improvement actions
- Actions taken/ planned: Suggestions for how to close the gap between the current and desired state

In addition to identifying gaps by comparing the RBAT framework with the input required to develop the test cases, several gaps have been identified as part of preliminary tests performed by DNV as part of other projects outside of RBAT. This also includes discussions about RBAT with potential user groups external of DNV.

4.3 Results

Results from the gap analysis is documented in Table .

A total of 16 gaps were identified. Gaps for which actions already have been taken to close are marked with the status “Implemented”, while those which will be addressed as part of developing the test cases are labelled as “Planned”.

Overall, most gaps are related to either Use of Automation (4) or Hazard Analysis (5), and fewer to the Mission Model (3), Function Tree (2), and Prevention Analysis (1).

It should be noted that because the gap analysis was requested to be performed and reported prior to developing the test cases, it is expected that additional gaps will be identified at a later stage. DNV will record and implement any gaps as they emerge.

Table 4-1: GAP analysis of RBAT framework

ID	Focus areas	Current state[6]	Desired future state	Actions taken/ planned	Status
G-1	Mission model	Use of Automation & Remote Control is described by breaking down the vessels mission into mission phases, operations, control functions and control actions.	The breakdown of a mission into operational and functional goals should appear logical to the user and guide him/her towards identifying items to be analysed at a level of detail which matches the concepts maturity.	As part of the case studies, evaluate how many layers of the Mission Model and Function Tree need to be represented in the log sheet.	Planned
G-2	Mission model	The Mission Model includes Emergency Response as a separate mission phase, with responses to various abnormal situations (incidents) as operations. Many of the abnormal situations corresponds with key functions in the Function Tree, e.g. “Respond to fire/explosion” (operation) corresponds with “Provide fire protection” (key function). Several of the abnormal situations may however require various minimum risk conditions (MRC), in addition to more dedicated safety functions. MRCs are currently not explicitly covered by the Mission Model. However, the	How to address MRCs should be reflected in the multi-level function map (i.e. Mission Model and Function Tree) and/or RBAT method description.	Compile a list of already known MRCs to be included as part of the case studies. Check to what extent they will overlap with existing contents in the Mission Model and/or Function Tree. Ensure that the final solution does not cause additional confusion.	Planned

[6] ‘Current state’ here refers to how RBAT is described in Report 2/2 from Part 1.

ID	Focus areas	Current state[6]	Desired future state	Actions taken/ planned	Status
		functions required to successfully enter MRCs are to a large extent covered in the Function Tree. Some MRCs are simple and involve few functions, while other are more complex.			
G-3	Mission model	The current Mission Model does not include an operation for transit in enclosed or sheltered waters, such within a bay area or strait.	Mission Model shall include operations for expected kinds of geographical areas.	Include "Navigate through enclosed/sheltered waters" as operations for the "Transit" mission phase.	Implemented
G-4	Function tree	The RBAT log sheet currently includes a column for both Control Functions and Control Actions. The *mission phases* and *operations* are operational goals for which the control functions and actions are required. The log sheet is not set up to include all the function levels between the operations and control actions. It can be argued that it suffices to only include the lowest function level, i.e. control actions and remove the control function column (the link between control actions and operations may still be obvious). This will lower the numbers of columns the user has	The RBAT shall maintain the benefits from a functional approach, and the decomposition of goals. This can be done both through the methodology as well as the tools (log sheet, function maps). This however shall not be on the expense of simplicity and efficiency. All data documented shall serve a specific purpose and benefit for the user. Possible benefits from including control functions are: • Easier to make sense of the control actions	Control functions are defined as: "*Control actions performed by humans or machines for the accomplishment of a functional goal*". I.e., it refers to a common goal for a set of control actions. Control actions are defined as: "*Acquisition or analysis of information, decision-making, or implementation of physical actions performed as part of a control function*." Following the definitions, control functions can therefore be considered "parent" functions to control actions. The definitions and guidance do not however define at what level the	Planned

ID	Focus areas	Current state[6]	Desired future state	Actions taken/ planned	Status
		to populate in RBAT and potentially reduce its complexity.	• May serve the purpose of categorizing/tracing/filtering control actions	control function shall be described. The case studies must test whether there is a need to document control functions in a separate column, or if this can just refer to the collective goal of a set of control actions, which is between the control actions and operations.	
G-5	Function tree	The control function "Charge/receive electrical power (from shore) was found incomplete.	The RBAT function tree shall, to the extent possible, include a complete set of generic functions.	The following sub-functions have been added to the Function tree: • Start charging • Stop charging	Implemented
G-6	Use of automation	The type of supervision and supervising agent may not be constant but can potentially vary across scenarios. For example, an onboard operator may actively supervise an operation, but if a failure happens on one of the systems involved, an alarm is raised in the RCC where the control room operator is responsible for passively supervising the system and act	Principles around Supervision and Mitigations need to be clear, and capture expected technical and operational solutions.	As part of the case studies, check and make note of how different scenarios reveal different aspects of supervision. Compare this with how supervision is defined in RBAT and explained as part of the methodology.	Planned

ID	Focus areas	Current state[6]	Desired future state	Actions taken/ planned	Status
		upon demand. As such, it may not always be possible to identify who the supervising agent of a control action is before the scenario has been identified.			
G-7	Use of automation	Preliminary testing of RBAT performed as part of the method development has revealed some indications that a separate definition for supervision by a machine agent may be needed.	RBAT shall include a category for machine supervision of other agents (both human and machine).	Update RBAT method description to include a definition of supervision performed by a machine agent.	Implemented
G-8	Use of automation	How to best approach global and continuous functions, i.e. functions which span across several mission phases and/or interact with several other functions, have not been thoroughly explored. Potential challenges may arise related to repeating the same functions across multiple operations, but with diminishing returns from efforts put into performing the analysis. This was also highlighted as part of the recommendations from Part 1 (second report).	Risk contribution from functions such as utility/auxiliary and integrated monitoring and control shall be captured and addressed in a systematic manner.	The Use of Automation part in RBAT has been updated to include a column for capturing other systems/roles involved in performing the control action. This will create an awareness about how failures in such systems (and functions) can act as Causal Factors which triggers the unsafe conditions and modes. Furthermore, they will also create awareness around whether the same systems are required to perform any mitigations, and this will represent potential common	Implemented

ID	Focus areas	Current state[6]	Desired future state	Actions taken/ planned	Status
				cause failures (i.e., lack of independence). Functions associated with integrated monitoring and control are captured either as part of the control actions themselves, or as part of the supervision.	
G-9	Use of automation/ Mitigation analysis	In some cases, there will be opportunities for a performing agent to self-recover. This, however, may be a weak mitigation, due to potential dependencies between the causes of an unsafe condition and the detection or response.	Same as ID G-8 above.	The proposed risk matrix for RBAT allows some risk reduction credit to be taken for self-recoveries (internal mitigation layers).	Planned
G-10	Hazard analysis	RBAT borrows the terminology from Failure Hazard Analysis (FHA) used in the commercial aviation industry. This technique uses the term “Failure conditions/ modes”.	Efforts should be spent on identifying unsafe conditions or operating modes, rather than examining all kinds of failures which can occur but does not necessarily have safety implications. Such a focus should however not be on the expense of failing to scrutinize the potential hazards associated with a system.	Replace the term “Failure conditions/ modes” with “Unsafe conditions/ modes”. This is believed to provide the following benefits: (1) It will trigger and encourage the RBAT user to consider how the control actions potentially can become unsafe, instead of merely describing failure states.	Implemented

ID	Focus areas	Current state[6]	Desired future state	Actions taken/ planned	Status
				(2) It will also promote the need for defining the boundaries and limits of what is considered safe. (3) It will better distinguish RBAT from traditional FMEA/FMECA.	
G-11	Hazard analysis	The RBAT methodology adopts the guidewords for unsafe control actions in STPA and for Failure Conditions/ Modes in Functional Hazard Analysis (FHA) (SAE, 2010). These are however only included in the RBAT method description, and not in the RBAT Excel template.	Descriptions of unsafe conditions/ modes should describe *how* the control action can become unsafe, and not *why*. The reason why (i.e. cause) should be described in the column dedicated for "Causal factors". RBAT users should also be guided to consider all possible conditions and modes, so that the ones that are potentially unsafe can be identified and further analysed.	The guidewords for prompting unsafe conditions/ modes have been made available using a dropdown list in a dedicated column in RBAT.	Implemented
G-12	Hazard analysis	See ID G-11 above. In addition, RBAT considers unsafe conditions/ modes per operation. One of the guidewords is "provided when not required" – when "…not required" is exactly can be a wide variety of situations and will be difficult to pinpoint.	RBAT should incorporate fit-for-purpose guidewords from both STPA and FHA (and other relevant sources), but without unnecessary overlap.	Test guidewords as part of case studies and refine list.	Planned

ID	Focus areas	Current state[6]	Desired future state	Actions taken/ planned	Status
G-13	Hazard analysis	There is no dedicated column in RBAT for describing causes. The argument for why is that RBAT is to be used at an early stage of the concept development and before details about the system architecture has been decided. Instead, the analyst can choose to include description of the causal factors as part of the failure condition/ mode, if this knowledge is available at the time of the analysis.	Knowing what causes the failure conditions/ modes (now referred to as unsafe conditions/ modes) is necessary for identifying and evaluating which mitigations are effective. E.g., unsafe manoeuvring caused by failures in the control system may require a different response than if the cause is a failure in the mechanical propulsion system.	A separate column for describing "Causal factors" have been included in RBAT. This should describe a single cause or a combination of causes which can initiate the unsafe condition/ mode. This was also a recommendation from Part 1 (second report).	Implemented
G-14	Hazard analysis	The RBAT log sheet from Part 1 includes a column for recording Failure Effects and Worst-Case outcomes. When assessing failures on a functional (instead of component level) it can be difficult to distinguish these two outcomes, and they may appear overlapping. Also, by introducing the concept of Causal Factors and Unsafe Conditions, this also captures what would be the Failure Effect.	The accident model in RBAT needs to have a set of definitions which logically can be used to explain an event sequence towards an accident.	The Failure Effect Column has been removed. Now the scenarios are defined as: Causal factor(s) initiates an unsafe condition/mode-> mitigations are successful or fail-> worst case outcome occurs.	Implemented

ID	Focus areas	Current state[6]	Desired future state	Actions taken/ planned	Status
G-15	Prevention analysis	Preventive safeguards are not documented as part of the existing RBAT structure or methodology.	It can be useful to include a column for describing existing *preventive* safeguards. This will provide more insight and understanding of the integrity of the control functions. It will also allow system developers to demonstrate safety measures other than mitigations such as MRCs and other recoveries.	A column with the title “Prevention analysis” has been included to capture performance requirements associated with the control action being analysed. This will help the analyst define the boundaries, limits, and criteria for what is considered safe and unsafe operations (i.e., it will indicate a safe operating envelope). These performance requirements, and the systems performing the associated control functions, will act as preventive safeguards.	Implemented
G-16	Mitigation analysis	The RBAT developed in Part 1 includes a “Recovery analysis” to capture which responses are available in case of unsafe conditions and modes, and to evaluate their success probability.	The index used for determining level of recovery needs to be incorporated with the risk evaluation technique developed for RBAT.	The “Recovery analysis” have been further developed and re-titled to “Mitigation analysis”. See sub-chapters 3.3, 3.4, 5.3, and 5.4.	Implemented

5 STEP-BY-STEP GUIDEANCE TO THE RBAT METHODOLOGY

The current RBAT methodology consists of five main parts:

1. Describe use of automation (and remote control)
2. Perform hazard analysis
3. Perform mitigation analysis
4. Perform risk evaluation
5. Address risk control

The following sub-chapters presents these four main parts as consisting of 15 steps.

5.1 Part 1: Describe use of automation (and remote control)

The purpose of describing the use of automation (UoA) and remote control is to:

- Identify which functions are affected by automation or remote-control
- Understand how these functions are allocated to different *agents* (human or machine)
- Know where the different agents are located (locally on vessel/site or remote)
- Check how the affected functions are supervised, and by which agents

This process should preferably be done as an integrated part of developing and documenting the *Concept of Operations* (ConOps). It is therefore an advantage if the ConOps adopts the terminology and principle of modelling functions using hierarchical goal structures, as explained in Step 1 and 2 below.

The UoA's *context* (e.g., geography, environmental conditions, infrastructure etc.) is expected to be described in the ConOps.

USE OF AUTOMATION/ REMOTE CONTROL					
Control function	**Control action**	**Performing agent**	**Supervision category**	**Supervising agent**	**Other systems and roles involved (onboard, onshore)**
Mission phase: Arrival in port					
Operation: Perform port/harbour manoeuvering					
Perform manoeuvring	Approach dock at low speed	Onboard autonomy system	Active supervision	Onboard safety operator	Thrusters, thruster control system
...	...	...	...	...	...

Figure 2: Use of Automation module in RBAT

5.1.1 Step 1: Describe the vessel's mission (operational goals)

The first step of the process is to describe the vessel or fleet of vessels *mission*. The term mission refers to a set of mission phases, operations and functions the vessels perform to achieve their operational goal(s).

A mission can be described as consisting of three levels:

- The overall mission goal(s), i.e., the commercial, political (e.g., defence) or public intentions which have contributed to and justifies the vessel concept development and operation. An (simplified) example can be "Safe and timely transport of cargo from one Port X to Port Y".

- The mission phases, i.e., subdivisions of the mission typically characterized by a recognizable shift in where the vessel is located in terms of geographical surroundings, or the start and end of one or more operations. An example can be “Arrival in port”.
- The operations, i.e., Activities performed as part of a mission phase in order to achieve the mission goal. An example can be “Perform docking”.

These three levels shall be considered in terms of a *hierarchical goal structure*, e.g.:

Mission: Safe and timely transport of cargo from Port X to Port Y

Mission phase: Arrival in port

Operation: Perform docking

The mission phases and operations are the study nodes under which the functions to be analysed are listed. Together with the details provided in the ConOps, they form the operational context (circumstances) under which the functions are required to perform.

The generic RBAT mission model (Appendix C) can be used as a starting point. Re-phrase and/or add descriptions if needed. Emergency responses should be included as separate Operations.

Figure 2 shows how mission phase (grey row) and operations (golden row) are included as *nodes* in RBAT.

5.1.2 Step 2: Describe the automated and/or remotely controlled functions (functional goals)

The second step of the process is to describe the functions which are subject to or affected by automation and remote control. This includes identifying:

- the *control functions* required to successfully carry out the operations in each mission phase, and
- the *control actions* allocated to various (human or system) agents involved in performing the control function

Control functions and actions make up the functional goals of the hierarchical goal structure:

Mission: Safe and timely transport of cargo from Port X to Port Y

Mission phase: Arrival in port

Operation: Perform docking

Control function: Perform manoeuvring

Control action Y: Adjust speed

Control action Z: Adjust heading

The generic RBAT Function Tree (see Appendix D) can be used as a starting point for this process. For each operation described in Step 1, review and identify which of the (highest level) *key functions*[7] are required to achieve a successful outcome. Then, for each relevant key function, drill down the tree branches to a sub-function level which matches the current maturity of the concept. As a minimum, the functional goals shall be broken down to the level where automation can be made sense of, i.e., it shall be possible distinguish which parts of the function are allocated to different (human or system) agents. The lowest level makes up the control actions, and the parent level makes up the control functions.

[7] In RBAT, key functions are the highest layer of functions in the Function Tree.

The lower-level functions in the RBAT Function Tree should primarily be considered as suggestions. Functions can be re-phrased and/or added on a need-to basis. The list of verbs provided in Appendix E can be useful for this purpose.

When identifying and describing functions it is important to not only include those exerting direct control. Care should be taken to also consider including functions which serve more supportive purposes (often across several other functions), such as auxiliary functions and functions required for system monitoring. If such functions are present across several mission phases and operations, they can be grouped under a separate study node to avoid unnecessary duplication of the assessment.

Functions which involve exchange and interaction with external agents or systems should also be considered for inclusion, such as those provided by surrounding infrastructures, e.g., navigational aids.

Figure 2 shows the columns in RBAT used to describe control functions and actions (i.e., functional goals).

It is helpful if the ConOps includes functional block diagrams (Figure 3) illustrating the relationships and dependencies between the affected control actions (both internal and external).

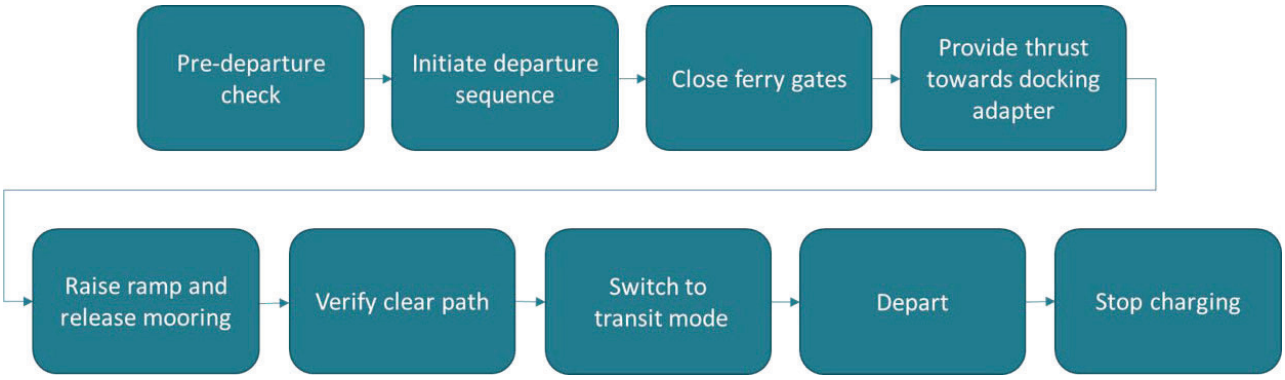


Figure 3: Example of control actions illustrated in a functional block diagram format

5.1.3 Step 3: Describe how control functions are allocated to agents

The third step of the process is to describe how control functions are allocated to different *agents* by indicating who is responsible for performing the various required control actions.

Agents can be a computerized system or a human operator and only one agent can be listed as responsible for performing a control action under normal operations. However, depending on which level of detail control actions are described, cases may come up where more than one agent is involved. In principle, this calls for further decomposing the control action until it can be distinguished which agent is the performing agent. If this appears as being too detailed, the agent making the decision should be nominated. Other agents can then be described in the column titled “Other systems and roles involved”.

The geographical location of the agent shall also be indicated either by using nomenclature pre-fixes, such as “R” for Remote and “O” for onboard (vessel), or by including the actual location of the agent as part of the title. Alternatively, the location can be explained elsewhere, e.g., in the ConOps. This, however, is less preferable as it assumes that the analyst(s) and reviewers are familiar with this content.

Figure 2 shows how the control action “approach dock at low speed” is allocated to the performing agent “Onboard autonomy system”.

5.1.4 Step 4: Assign responsibility for supervision of control actions

The fourth step of the process is to indicate if and how control actions are supervised, and by which agent. *Supervision* is a role with an explicit responsibility to monitor system performance and detect anomalies so

that the desired outcome can be achieved through implementation of corrective responses. Examples of anomalies can be system failures and malfunction, or external conditions which exceed pre-defined criteria for what are considered operational limits (e.g., weather conditions). In case a control system does not have the capacity to self-recover from a failure, the designated supervisory agent is responsible for ensuring that independent mitigation layers are effective, as described in sub-chapter 5.3 (Steps 8-10). *An important principle is that the supervisory agent cannot be the same as the agent performing the control action(s) being supervised.* The supervisor has an overriding authority of the control action performance and is responsible for its outcome.

Supervision can be performed by either a machine or human agent. It is important to consider the strengths and weaknesses of both agents before assigning supervision responsibilities. In cases where humans are the supervising agent of a control action they will often rely on a system for monitoring and detection, while analysis, decision-making and implementation of actions are performed manually. A machine agent will perform all actions. As such, the supervisor is the agent responsible for making decisions about interventions.

Three different categories of supervision are defined in RBAT:

- *Active (human) supervision*: A human agent is responsible for continuously[8] monitoring the performance of a control action with the purpose of being able to successfully intervene at any stage based on judgements about how to best act upon the situation. Because active supervision provides an opportunity for the human agent to continuously create situational awareness, it can be beneficial in cases where there is limited time available to intervene.
- *Passive (human) supervision*: A human agent is responsible for being available[9] to monitor the performance of a control action and successfully intervene upon requests (e.g., an alarm) generated by the system according to pre-defined parameters. Because passive supervision (often) requires the human agent to obtain situational awareness about the events preceding the request, it is best suited for cases where there is sufficient time available to intervene.
- *Machine supervision*: A machine agent is responsible for continuously monitoring the performance of a control action with the purpose of being able to successfully intervene on demand, without involvement of a human agent, for example if pre-defined parameters are exceeded, or there is disagreement in voting between separate functions/components.
- *No supervision*: No agent is responsible for monitoring the performance of a control action.

It is important to emphasize that the supervision categories represent a specific operational responsibility. This means that if an operator is responsible for actively supervising a control action, this must be reflected in job descriptions, procedures, routines, etc. Selection of supervision categories should therefore be based on the overall philosophy about monitoring and control described in the ConOps, which also include a more detailed description of the supervisory roles. Such descriptions should consider the influence from factors such as fleet size, manning level, competencies, human-machine interfaces (e.g., visual display units) when assigning supervision responsibilities to human agents. A preliminary solution for supervision should therefore be decided upon and described before commencing with the hazard and mitigation analysis (Step 5).

Figure 2 shows how the “Onboard safety operator” is responsible for actively supervising the control action “approach dock at low speed”.

[8] ‘Continuously’ implies that the agent is responsible for, and expected to, direct his/her/its attention to a function for as long as it is being executed.
[9] ‘Available’ implies that the agent is responsible for, and expected to, be in close enough proximity to intervene upon a demand from the system.

5.2 Part 2: Perform hazard analysis

The purpose of the hazard analysis is to:

- Identify unsafe conditions/modes associated with control actions (Step 5)
- Identify causal factors which may initiate the unsafe conditions/modes (Step 6)
- Describe the worst-case outcomes from (unmitigated) unsafe conditions/modes (Step 7)
- Rank the worst-case outcomes severity (Step 8)

HAZARD ANALYSIS					
Unsafe condition/ mode	**Guideword**	**Causal factor(s)**	**Worst-case outcome**	**Event category**	**Severity**
Vessel fails to reduce speed.	Needed but missing/ Not provided when needed	Control system failure	Impact with quay in transit speed	Contact with shore object	Significant - Single serious or multiple injuries
	...	...	...	...	...

Figure 4: Hazard analysis module in RBAT

5.2.1 Step 5: Identify unsafe conditions/ modes associated with control actions

The fifth step of the process is to identify and describe unsafe conditions/ modes associated with the various control actions identified in Step 2. Unsafe conditions/ modes manifest themselves as incidents where a system is operating outside its normal (and safe) operating envelope due degraded performance (e.g., failures) or exceeding its capabilities which, if left unmitigated, has the potential to cause an accident (i.e., losses).

Identification of unsafe conditions/modes can be assisted by using the guidewords provided in Table 5-1 as prompts. The actual descriptions of unsafe conditions/ modes are user-defined and specific for each control action. However, consistent use of terms and expressions will help with the RBAT compilation and review.

All credible and relevant unsafe conditions/modes should be considered. What characterizes a condition or mode as *unsafe* depends on the severity of worst-case outcomes (see Step 8). If it is evident that the worst-case outcome from a potential unsafe condition/mode is *negligible* it does not require mitigation layers to be acceptable (see Step 13). The user can opt to not record such items, as a way of improving the readability and overview of the analysis. On the other hand, recording all items will provide (e.g., reviewers) with transparency and trust that as many as possible unsafe conditions/modes have been addressed. In such a case, the user can save time by not having to include them as part of the mitigation analysis.

Table 5-1: Unsafe condition/mode guidewords

Unsafe conditions/modes	Guidewords
Not providing the control action leads to a hazardous event	Not provided
Providing the control action leads to a hazardous event	Provided when not required
	Incapable/not fit for purpose
Incorrectly provided control actions leads to a hazardous event	Control parameters out of range
	Control parameters are within range but incorrect
	Too early/late or in wrong of order
	Stops too soon
	Applied too long
Control action not being followed leads to a hazardous event	Not followed/Rejected

5.2.2 Step 6: Identify causal factors which can trigger unsafe conditions/ modes

The sixth step of the process is to identify causal factors which can trigger the unsafe condition/ mode. These can be *internal failures* in the vessel's or RCC's systems (e.g., a software issue) or *insufficient capabilities* when it comes to handling external hazards (e.g., unfamiliar objects or strong currents). Hazards external to the vessel, relevant for the operation in question, should therefore always be considered when identifying failures which represents insufficient capabilities.

The following failure categories should be considered when identifying causal factors:

- random (hardware) failures,
- systematic failures,
- systemic failures,
- operator failures,
- failures due to environmental conditions,
- failures due to deliberate actions.

Note that these categories overlap to some extent, yet they are useful as a guide to identify a wide range of failures that may pose risk. See Appendix B for an explanation of these failure categories.

If an unsafe condition/ mode can be initiated by several different causal factors, these should be treated as separate scenarios, as indicated below.

Scenario 1: Causal factors A and B → unsafe condition/ mode Z → outcome X

Scenario 2: Causal factors C and D → unsafe condition/ mode Z → outcome X

The reason is that different causal factors may require different mitigations. For example, an incorrect heading (unsafe mode) caused by a mechanical failure may require a different response than if the cause is software related. Mixing these will therefore make the mitigation analysis difficult.

As a principle, the causal factors descriptions should be made as concise and specific as possible, so that it is possible to make sense of how they can initiate the unsafe condition/mode. This, however, depends on how much information is available about the system, including its planned operations and operating conditions. It is important that the causal factors do not describe the same thing as the unsafe condition/ mode, but with different words. The unsafe condition/mode shall describe *why* the system is unsafe, while the causal factors shall describe *how* the system became unsafe.

5.2.3 Step 7: Describe the worst-case outcomes from (unmitigated) unsafe conditions/ modes

The seventh step of the process is to determine the worst foreseeable outcome of an unsafe condition/mode in case there is no mitigation available (this includes internal mitigation). In RBAT, worst-case outcomes assume the contextual presence of a credible *hazard*. For example, loss of steering (an unsafe condition) close to shore (a hazard) results in a grounding (a worst-case outcome).

The description should include the hazard itself as well as the location, and not just the type of accident. For example, instead of only stating “grounding”, it should also be specified which surface the vessel is grounding onto, such as a reef or sandbank (hazard and location). Or, instead of only stating “fire”, it should be specified what is burning and where, such as diesel fire (hazard) in the machinery (location). This will help deciding (and auditing) which level of severity should be selected (see Step 8).

In case an argument is made that a hazard is not present, e.g., through operational restrictions, this must be clearly stated either as part of the prevention analysis (Step 12) or in the comments for addressing risk control (Step 15).

Finally, an accident category is assigned to each worst-case outcome, using the taxonomy in the list below (Table 5-2).

Table 5-2: Accident categories

General

- No effect on safety
- Injuries/loss of life (general)

Loss of control

- Loss of directional control
- Loss of propulsion power
- Loss of electrical power
- Loss of communication link
- Loss of containment
- Loss of stability
- Loss of control (other)

Collision

- Collision with other ship
- Collision with multiple ships

Contact

- Contact with floating object
- Contact with flying object
- Contact with shore object

Damage to/ loss of ship equipment

Hull failure

Fire/explosion

- Fire
- Explosion

Grounding/stranding

- Grounding
- Stranding

Capsize/listing

- Capsize
- Listing

Flooding/foundering

- Massive flooding
- Progressive flooding
- Foundering

Non-accidental event

- Acts of war
- Criminal acts
- Illegal discharge
- Other

Missing vessel

The accident categories are mutually exclusive and only one shall be assigned to each worst-case outcome. To help with this, the following principles apply:

- *Injuries/loss of life* shall only be used when this happens outside any of the other accident categories. For example, in the case of the crew being exposed to a disease.
- *Loss of control* shall only be used when there are no hazards present which are required to cause an accident.
- *Damage to/ loss of ship equipment* shall only be used when this occurs in absence of the other accident categories.
- *Hull failure* shall only be used in case this occurs without being the direct cause of other accident categories (e.g. capsize or foundering).

5.2.4 Step 8: Rank the worst-case outcome severity

The eight step of the process is to rank the worst-outcome severity. This is done by assigning a degree of severity using the index in Table 5-3.

Table 5-3: Severity index for worst-case outcomes

Severity	Effects on human safety
No effect	No injuries
Negligible	Superficial injury
Minor	Single injury or multiple minor injures
Significant	Single serious or multiple injuries
Severe	Single fatality or multiple serious injuries
Catastrophic	Multiple fatalities (more than one)

5.3 Part 3: Perform mitigation analysis

The purpose of the mitigation analysis is to:

- Identify which mitigations are in place to prevent the unsafe condition or mode from resulting in losses (Step 9).
- Assess and determine whether mitigations can be qualified as effective in achieving their intended purpose (Step 10 and 11).
- Identify measures which are in place to prevent the direct cause of an unsafe condition or mode from occurring (Step 12, optional)

In this context mitigations can involve the following types of responses:

- Withstanding or recovering from a failure before it turns into an unsafe condition/ mode
- Re-entering to a normal (safe) operating envelope by regaining control of an unsafe condition/ mode
- Enter a state of emergency response and abort further operations to prevent escalation

In RBAT mitigations are considered in terms of a system's capacity to self-recover from a failure, here referred so as *internal* mitigation layers, and *independent* mitigation layers implemented to prevent losses in case such self-recoveries should fail. Mitigation layers may involve entering a minimum risk condition (MRC) as a measure to stay as safe as possible while attempting to regain the desired level of control.

The role of mitigation layers is illustrated in the RBAT accident model (Appendix F).

MITIGATION ANALYSIS					
Internal mitigation layer (self-recovery capacities)	1st independent mitigation layer	2nd independent mitigation layer	3rd independent mitigation layer	Mitigation effectiveness	Criticality
Yes	Emergency stop (MRC2)	Drop of emergency anchor (MRC3)	None	High	Medium
...	...	...	...	...	...

Figure 5: Mitigation analysis module in RBAT

5.3.1 Step 9: Nominate mitigation layers which can prevent losses

The ninth step of the process is to identify which internal (self-recovery capacities) and independent mitigation layers are in place to prevent the unsafe condition or mode from resulting in an accident (and losses). This is done by nominating self-recovery capacities[10] and potential 1st, 2nd, and 3rd independent mitigation layer(s) for each combination of unsafe condition/ mode and causal factor(s) (see Figure 5). Preferably, a preliminary set of mitigation layers have already been described *prior* to using RBAT, e.g., as part of drafting the first version of a ConOps. If new mitigation layers are identified as part of the process, these are added to the list of existing ones, and then nominated in the analysis. Self-recovery capacities are expected to be more specific to each function and are more likely to be identified and described as part of the mitigation analysis.

Creating a list of mitigation layers. The mitigation layers can be initially identified by considering what the responses would be in case various failures, loss of control or accident scenarios should occur. The generic

[10] Also referred to as *internal* mitigation layers

accident categories presented in Table 5-2 can be used as a starting point for such considerations. Alternatively, this work is done after performing the hazard analysis as part of RBAT.

Each mitigation layer should be listed with an:

- ID
- Name
- Short description

Furthermore, information necessary to evaluate the mitigation layers risk reducing effectiveness (Step 10) must be gathered. This includes:

- Applicability of the mitigation layer
    - For which incidents the mitigation layer is a planned response
    - For which mission phases the mitigation layer is applicable
    - For which mission phases the mitigation layer is NOT applicable, e.g., due to:
        - Being potentially unsafe
        - Restricting use of other mitigation layers
        - Not being relevant (i.e., effective)
- System and human involvement in the mitigation layer
    - Systems which must function and be available for executing the mitigation layer
    - Human-automation interactions required as part of the mitigation layer (see sub-chapter 5.3.2.2 for further explanations)
- Limitations to the mitigation layer
    - External/ environmental limitations in the mitigation layer (e.g., sea state, visibility, day/night, availability of external resources)
    - Resource limitations in the mitigation layer (e.g., time, fuel, energy reserves, manpower, etc.)
    - Limitations in the sequence mitigation layers can be introduced (e.g., a mitigation layer should only be activated after another has been exhausted)
- Transitions between and from mitigation layers (including minimum risk conditions)
    - How to re-enter a normal or as safe-as-possible operational mode (in case the mitigation layer involves entering a minimum risk condition (MRC)
    - What the next mitigation(s) in the sequence is, and how to introduce it (“None” in case the mitigation is a last resort MRC)
    - Emergency response in case there are no other mitigation layers available

5.3.2 Step 10: Qualify the nominated mitigation layers

The tenth step of the process is to assess and qualify the nominated mitigation layers against a set of performance criteria which characterises them as effective in accident prevention. This includes:

Functionality: The mitigation layer's design and intended use makes it effective at preventing the unsafe condition or mode from resulting in (safety) losses.

Integrity: The mitigation layer is available, its condition is intact, and it can be relied upon to work under the expected circumstances.

Robustness: The mitigation layer will remain functional after the unsafe condition or mode has occurred, taking any disturbances and/or accidental loads into account.

Independence:

- of the causal factors which initiated the unsafe condition/mode
- of each other (in case a mitigation fails)

Note that independence is only relevant for *independent* mitigation layers.

Human involvement: A final criterion is that the mitigation layers are designed and implemented in such a way that it ensures successful human-automation interaction.

Additional guidance for assessing independence and human involvement is provided below.

The qualification itself is qualitative and based on the knowledge available at the time RBAT is used. The conclusions are binary – a mitigation layer is either qualified or disqualified based on the user(s)[11] judgement.

In principle, a mitigation layer can be considered qualified when the user(s) feels confident that all the above-mentioned criteria are fulfilled. If knowledge is available which indicates that one or more of the criteria cannot be met, the mitigation layer is disqualified and shall be removed from the RBAT mitigation analysis (i.e., it shall not be taken credit as part of risk evaluations, Step 11).

It is acknowledged that limited information may be available about the mitigation layers, particularly in the preliminary design stage. In cases where assumptions must be made about the mitigation layers' performance and pre-requisites, these should be noted down so that they can be used to update the concept and included as part of verification and validation (V&V) efforts at a later stage.

In case a mitigation layer disqualifies, a comment should be made about why. If a risk is found unacceptable (see Step 13), disqualified mitigation layers can then be re-visited as the design matures and more knowledge is obtained (Step 15). The approach therefore benefits from being conservative in the early stages, to avoid having to disqualify mitigation layers at a later stage which potentially may result in unacceptable risks.

5.3.2.1 Additional guidance on independence

Additional guidance about how to assess mitigation layer independence is provided in Table 5-4 below.

[11] *Users* here also includes potential reviewers and approvers.

Table 5-4: Perspectives on mitigation layer independence

Perspective	Descriptions	Examples
Composition	This perspective, is used to evaluate whether there are any physical or software components used in the mitigation layer that may be affected by failures in components where an unwanted event has manifested itself.	A mitigation layer that relies on thrusters being reversed will not be independent if the initiating event occurred in the thruster itself or in the thruster control system. Two different types of software applications executed on the same controller will typically be dependent because they will share hardware and software components[12] A system may have several and different types of sensors which can trigger a safety function representing a mitigation layer. However, if the same controller and actuators is used regardless type of initiation, there may only be one fully independent mitigation layer available.
Environment	This perspective evaluates whether there are items outside the system and/or external events that may act upon the system, cause an unsafe condition/ mode and impair the mitigating layer.	• Loss of cooling in control rooms • Fire • Water ingress or flooding • Lightning strike • Radio communication jamming • Electrostatic discharge • Unexpected wind or wave conditions
Structure	This perspective looks at the relationships and bonds among the system constituents and between the system constituents and the environment.	Two systems/functions that are otherwise considered independent may both rely on the Power Management System being operational. An equipment specific protection mechanism may have the authority to reduce capacity to prevent equipment damage in a situation where the mitigation layer requires full capacity from that equipment to be effective. An operator may depend on alarms from the main control system to understand that a failure has occurred, and that activation of a mitigation layer is needed. If an unexpected scenario for which no alarm has been defined should occur (i.e., an un-

[12] Note that there are safety controllers that provide so called logical separation. In such cases the Commercial Of The Shelf (COTS) hardware and software components such as the operating system have been qualified for use in high-integrity systems and designed in such a way that individual software tasks cannot negatively influence each other through timing, memory space or I/O.

Perspective	Descriptions	Examples
		annunciated failure), the mitigation layer may not be activated in time to prevent a mishap.
Mechanisms	This perspective evaluates dependencies that may introduced through systems/functions or components having common requirements, common design, or common implementation*.	The controllers in a redundant control system are typically not independent of each other if a failure has systematic or systemic causes. This is since the two controllers typically will have common requirements, common design, and common implementation. Consequently, they will react in the same way to unexpected input: values, input combinations or input sequences. Two different GPS based positioning references systems may have different design and implementation. However, in case of unexpected input the systems may still fail in the same way as the functional requirements for such systems may be very similar**.

*Avoiding these kinds of dependencies may require some form of diversification, as described below.

** It is common to combine information from positioning references based on different principles to mitigate this kind of common cause through functional diversity as discussed below.

1) Functional diversity involves solving the same problem in different ways.

- This kind of diversity reduces the likelihood, that functional requirements which are inadequate for one or more operational scenarios will lead to dangerous systematic or systemic faults.
- Use of functional diversity may in some cases also lead to use of design diversity as discussed below, but not always.

2) Design diversity involves the use of multiple components, each designed in a different way but implementing the same function. E.g., one may use a CPU in combination with a Field Programmable Gate Array (FPGA).

This kind of diversity may be used to detect, isolate and recover from systematic failures introduced at software design and coding level, as well as in hardware design and manufacturing. It may also be used to detect random hardware faults.

This kind of diversity is not effective against systematic/systemic failures introduced in functional requirements specifications in the same way as functional diversity. It should be noted that software and hardware in controller(s) comparing and/or merging information from diverse functions and/or diverse components may introduce common mode failures.

5.3.2.2 Additional guidance on human involvement

For a mitigation layer to be qualified as effective, it must be designed and implemented in such a way that reliable human-automation interactions can be expected, assuming that operator actions are required.

This is assessed by asking whether it is possible for the operator(s) to:

- Detect and observe (perceive) the situation (information acquisition)?
- Make sense of the situation and predict future outcomes (information analysis)?
- Select a course of action among several alternative options (decision making)?
- Execute activities required to achieve the desired outcome (implementation of actions)?

Answers to these questions are found by determining whether one or more hindrances are present (see Table 5-5) and if their effect(s) on human-automation interaction is so negative that the required operator action(s) can be argued to fail.

During the design process the hindrances will concern technical performance shaping factors (PSFs) such as alarms, control panels and other human-machine interfaces (HMI), communication systems, automation design, equipment performance and tolerances, and more.

Particular attention should be devoted to examining dependencies between the system failures which initiates the unsafe condition/ mode, and the systems operators rely on to perform actions required for mitigation layers to be successful. For example, in case a software-related error causes an un-annunciated failure, the chances for an operator to act diminishes significantly.

Towards and during the operational phase the influence from other non-technical PSFs will emerge, such as procedures, training, and supervision. Although such factors can have a positive effect on human performance, they should not be an excuse to allow sub-optimal solutions at the earlier design stages.

Table 5-5: Hindrances for successful human-automation interaction

Information processing stages	Hindrances
Information acquisition *Perception of sensory information about the situation*	• There is no information available • There is too much information available • Information can easily be missed • Information can easily be misperceived (e.g., misheard, misread) • Information is misleading (e.g., expected but incorrect)
Information analysis *Making sense of the situation and predicting future events*	• Information analysis requires large amounts of information to be interpreted and memorized/recalled • Information analysis requires significant interpretations of uncertainties in parameters (incl. future events) • Information analysis requires understanding complex dependencies between different parameters • Information analysis requires factoring in the impact of unpredictable events (e.g., environment)

Information processing stages	Hindrances
Decision-making *Selecting a course of action among several possible alternative options*	• The decision basis is insufficient and/or unclear • There are too many paths, options, goals and/or they are contradicting, conflicting, or competing • How to prioritize paths, options, goals is unclear • The plan (e.g., a procedure) does not match the situation • Outcomes from decisions are uncertain
Implementation of action(s) *Executing activities required to achieve desired outcome*	• Opportunities for successfully exerting control is limited, e.g., due to being remotely located • There is insufficient time (or other required resources) available to successfully perform the required actions • Expected amount of training and experience is not likely to raise and maintain required skills at an adequate level • There are few or no feasible opportunities to recover and correct an erroneous action.

5.3.3 Step 11: Rank the mitigation layers effectiveness

The eleventh step of the process is to rank how effective the mitigation layer(s) is/are at preventing losses, using the index provided in Table 5-6. For control systems the thinking behind the index is as follows:

For control systems the thinking behind the mitigation scale is as follows:

- For a control function that is not fully redundant, the effectiveness of internal risk mitigation is considered *Low*. There may mitigation measures that can prevent losses from some types of random hardware failures, but the function being analyzed is not fully hardware fault tolerant nor fully tolerant to systematic/systemic faults.
- As discussed in the introduction to this section, a standard critical control system used in maritime is expected to be redundant. This implies that there is least one *internal* mitigation layer that can prevent losses from various types of random hardware failures. There may also be mitigation measures that can prevent losses from some types of systematic faults, but for such systems there will typical be types of systematic/systemic faults that cannot be mitigated without external intervention. Thus, the effectiveness of the internal mitigations in the system should be classified as *Moderate*.
- An *independent* mitigation layer will increase the strength of the mitigating measures by one level. For example, an independent emergency function that can mitigate a control failure in a standard control system will raise the strength from Moderate to Medium. A further strengthening to High will require a second independent mitigation, and so on.

Note: In case of a control function with low capacity for self-recovery is combined with one independent mitigation layer capable of preventing losses regardless of failure cause, the total effectiveness should be considered on a case-by-case basis.

Table 5-6: Effectiveness of Mitigations

Effectiveness	Description
Very high	At least three effective *independent* mitigation layers that for the assessed scenario can prevent losses regardless failure cause.
High	At least two effective *independent* mitigation layers that for the assessed scenario can prevent losses regardless failure cause.
Medium	At least one effective *independent* mitigation layer that for the assessed scenario can prevent losses regardless failure cause.
Moderate	At least one effective *internal* mitigation layer that for the assessed scenario can prevent losses from random *hardware* failures. The control function has additional capacities for self-recovery from other types of failures, however, for the assessed scenario these are not effective regardless failure cause.
Low	The control function has some capacities for self-recovery, however for the assessed scenario these are expected to have a limited effect.

5.3.4 Step 12: Perform prevention analysis (optional)

An (optional) twelfth step of the process is to identify any measures which exist to *prevent* the occurrence of unsafe conditions/ modes. This includes statements about technical and operational performance requirements (limits, boundaries etc.) as well as the activities which provide assurance that the required performance can be expected. This can include maintenance, testing and inspection for technical equipment, or rules about operational restrictions.

As with mitigation layers, only measures which already have been documented prior to the assessment should be included.

5.4 Part 4: Perform risk evaluation

The purpose of performing risk evaluation is to compare the risk level for each assessed scenario against a set of risk acceptance criteria to determine the need for risk control.

5.4.1 Step 13: Determine risk level for each assessed scenario

Step thirteen of RBAT is to determine the risk level for each assessed scenario, i.e., each combination of:

Causal factor-> unsafe condition/ mode-> mitigation layers -> worst-case outcome

As shown in Table 5-7, in RBAT the level of risk is a function of how severe the worst-case outcome of an undesired event is combined with how effective the mitigation layers are at preventing accidental (safety) losses. At this stage in the process, worst-case outcome severity has already been ranked in Step 8 and mitigation layer effectiveness has been ranked in Step 11.

As requested by EMSA, it is here recommended that the "as low as is reasonably practicable" (ALARP) principle is applied for risk evaluation[13]:

- High (red region): Risk cannot be justified and must be reduced, irrespectively of costs.
- Medium (yellow ALARP region): Risk is to be reduced to a level as low as is reasonably practicable.
- Low (green region): Risk is negligible, and no risk reduction is required.

The term *reasonable* is interpreted to mean cost-effective. Risk reduction measures should be technically practicable, and the associated costs should not be disproportionate to the benefits gained. How to perform cost-benefit assessments is extensively explained in the FSA guideline and therefore not repeated here.

Table 5-7: Risk matrix based on evaluation of available risk mitigating measures

Effectiveness of risk mitigation layers	Severity					
	No effect	Negligible	Minor	Significant	Severe	Catastrophic
Low	Low	Medium	High	High	High	High
Moderate	Low	Low	Medium	High	High	High
Medium	Low	Low	Medium	Medium	High	High
High	Low	Low	Low	Medium	Medium	High
Very high	Low	Low	Low	Low	Medium	Medium
Extremely high	Low	Low	Low	Low	Low	Medium

5.4.2 Step 14: Alternative justifications for determining risk levels

The fourteenth step of the process is to explore alternative justifications for determining risk levels. While this is not expected to be a standard part of using RBAT, cases may arise where arguments for lowering the risk level appears to be justifiable.

When comparing the risk picture associated with a specific function and corresponding risk mitigation layers to relevant acceptance criteria, the following alternatives for risk evaluation can be considered:

1. Operational restrictions such as speed limits and weather restrictions may be used to reduce the Severity of operational scenarios. Use of such measures must be clearly stated as an assumption in RBAT and documented in relevant reports (e.g., safety philosophy).
2. It may be possible to follow, e.g., the automotive industry in evaluating exposure rate to the relevant hazard. If it can be argued that the Hazard is relevant less than 10% of the average operational time

[13] MSC-MEPC.2/Circ.12/Rev.2, chapter 4.

per year, the required level of mitigations may be reduced by one level. If the hazard is relevant less than 1% of the average operational time per year, the required level of mitigation may be reduced by two levels.

3. If the initiating event[14] is not related to software, it may be possible to argue for a lower probability than what has been generally anticipated for control functions. In that case fewer independent risk mitigation measures may be required to meet the acceptance criteria. For such events the classical type of risk matrix shown in Table 5-8 can be used as a starting point to determine the initial risk picture before looking at available mitigation layers.

4. It should be possible to argue that a single mitigation will increase the effectiveness of the mitigation by more than one level. One example may be that if it can be demonstrated than an emergency stop function for machinery has a Performance Level (PL) = *d* performance according to the ISO 13849 safety standard for machinery, this would be considered a two-level increase.

5. It should also be possible to demonstrate that safety critical control functions performing more complex functionality than emergency stop has a better performance than what is anticipated in the scheme above. Such claims should be substantiated in an Assurance Case or similar. More advanced forms of risk analysis, carefully selected components and sharper development processes than what traditionally has been applied in the maritime may be required to substantiate such claims.

The pursuit of any such alternative approaches needs to be thoroughly argued for and carefully documented. As it is not within the scope of RBAT to suggest how this is done in practice, each user must determine what is the best possible approach to meet the expectations of approvers and other stakeholders.

Table 5-8: Example of classical risk matrix

			Severity				
Probability of failure per year			Negligible	Minor	Significant	Severe	Catastrophic
Frequent	>=1		Medium	High	High	High	High
Probable	>=1/10	To <1	Low	Medium	High	High	High
Occasional	>=1/100	To <1/10	Low	Medium	Medium	High	High
Remote	>=1/1000	To <1/100	Low	Low	Medium	Medium	High
Very remote	>=1/10000	To <1/1000	Low	Low	Low	Medium	Medium
Improbable		<1/10000	Low	Low	Low	Low	Medium

[14] Causal factor(s) initiating the event which results in an unsafe condition/ mode

5.5 Part 5: Address risk control

The purpose of risk control is to ensure that unacceptable (high) and tolerable (medium) risks are ALARP.

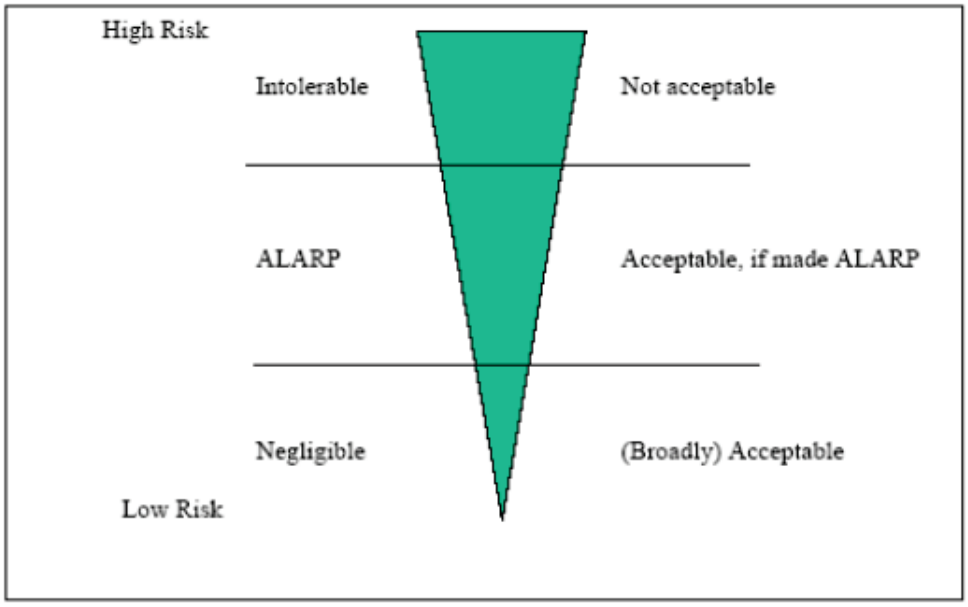


Figure 6: ALARP principle (IMO, 2018)

5.5.1 Step 15: Identify and document risk control measures

The fifteenth and final step of the process is to identify risk and document control measures (RCM). This is done by recording actions and any necessary comments in a column dedicated for this purpose (Figure 7) In the case of RBAT, risk can be reduced by:

- Updating the design so that disqualified internal and/or independent mitigation layers qualifies as effective and can be taken credit as part of the risk evaluation.
- Removing or reducing the hazard associated with the control function, e.g., the fewer or less flammable hazards onboard, the less severe accident outcomes.
- Introduce operational restrictions which reduces the hazards potential impact, e.g., not allowed to sail close to shore in certain weather conditions or in high speed through traffic dense areas.
- Improving the control functions integrity (and thus reducing its failure frequency) through design, component manufacturing and maintenance processes backed up by thorough assurance cases.

An elaborate description of generic RCM attributes (categories) can be found in the FSA guideline's Appendix 6 and is therefore not described in any more detail here.

RISK CONTROL	
Comments (incl. Assumptions)	**Actions**
Dropping the emergency anchor requires manual actions.	Verify that there will be enough time available for the onboard safety operator to drop anchor.
...	...

Figure 7: Comments and actions addressing risk control

REFERENCES

DNV (2021a). Risk-based assessment tool for MASS: Framework for generic risk assessment tool for MASS concepts. Report 2 of 2 (for Part 1 of the RBAT study).

DNV (2021b). Technology Qualification. Recommended Practice: DNV-RP-A203. Edition September 2021.

DNV GL (2018). *Autonomous and Remotely Operated Ships*. Class Guideline: DNVGL-CG-0264. Edition September 2018.

DNV GL (2020a). Framework for generic risk assessment tool for MASS concepts. Report 1 of 2. Report No.: 1, Rev. 0.

DNV GL (2020b). Proposal for A functional study developing a Risk-Based Assessment Tool for MASS (RBAT MASS). DNV GL doc No: 1-1HPDRGR-M-N-ADSS-1.

EMSA (2020). Invitation to tender No. EMSA/OP/10/2020 for the functional study developing a Risk-Based Assessment Tool for MASS (RBAT MASS).

Endsley, M.R. (1995). "Toward a theory of situation awareness in dynamic systems". Human Factors. 37 (1)

International Electrotechnical Commission, IEC (2000). IEC 61839 Nuclear power plants – Design of control rooms – Functional analysis and assignment. First edition.

International Electrotechnical Commission, IEC (2009). IEC 60964 Nuclear power plants – Control rooms – Designs. Edition 2.0.

International Electrotechnical Commission, IEC (2010). IEC 61508 Functional safety of electrical/ electronic/ programmable electronic safety-relate systems.

International Electrotechnical Commission, IEC (2013). IEC 60050-351 International Electrotechnical Vocabulary (IEV) - Part 351: Control technology.

International Electrotechnical Commission, IEC (2016). IEC 61511 Functional safety – Safety instrumented systems for the process industry. Part 1: Framework, definitions, system, hardware and application programming requirements.

International Electrotechnical Commission, IEC (2018). IEC 60812 Failure modes and effects analysis (FMEA and FMECA).

International Electrotechnical Commission, IEC (2020). IEC 61226 Nuclear power plants – Instrumentation, control and electrical power systems important to safety – Categorization of functions and classification of systems. Edition 4.0.

International Maritime Organization, IMO (2018). MSC-MEPC.2/Circ.12/Rev.2 – Revised guidelines for formal safety assessment (FSA) for use in the IMO rule-making process.

International Standard Organisation, ISO (2000). ISO 11064 Ergonomic design of control centres – Part 1: principles for the design of control centres. First edition.

International Standard Organisation, ISO (2009). ISO 31000:2009(E) Risk management – Principles and guidelines. First edition.

International Standard Organisation, ISO (2011). Road vehicles – Functional safety, 2011.

International Standard Organisation, ISO (2015). Safety of machinery - Safety-related parts of control systems - Part 1: General principles for design.

ISO/IEC/IEEE 15288:2015 Systems and software engineering -- System life cycle processes.

Kritzinger, D. (2016). Aircraft System Safety: Assessments for Initial Airworthiness Certification. Woodhead Publishing.

Leveson, N.G. & Thomas, J.P. (2018). STPA Handbook. Downloaded from: https://psas.scripts.mit.edu/home/get_file.php?name=STPA_handbook.pdf

MUNIN (2015). Research in maritime autonomous systems. Project Results and technology potential.

Parasuraman, R., Riley, V. (1997) Humans and automation: use, misuse, disuse, abuse. Human Factors: The Journal of the Human Factors and Ergonomics Society 39, 230–253. https://doi.org/10.1518/001872097778543886 .

Parasuraman, R., Sheridan, T.B., Wickens, C.D. (2000). A model for types and levels of human interaction with automation. IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans 30, 286–297. https://doi.org/10.1109/3468.844354.

SAE Aerospace (1996). Guidelines and methods for conduction the safety assessment process on civil airborne systems and equipment. Aerospace Recommended Practice ARP4761. First edition.

Sheridan T. B., Parasuraman R. (2006). Human-automation interaction. In Nickerson R. S. (Ed.), Reviews of human factors and ergonomics (Vol. 1, pp. 89–129). Santa Monica, CA: Human Factors and Ergonomics Society.

SAE Aerospace (2010). (R) Guidelines for Development of Civil Aircraft and Systems. Aerospace Recommended Practice ARP4754. Rev. A.

APPENDIX A
MASS concepts

Project	Project partners	Description	Type of project	Status
Small passenger ferry				
Zeabus	- Zeabus - NTNU - Trondheim havn - Trondheim kommune - AtB	Zeabuz is a highly ambitious spin-off from the progressive research center for Autonomous Marine Operations and Systems, at the Norwegian University of Science and Technology (NTNU).	Commercial	Construction
Callboats	- Callboats	Small on-demand autonomous unmanned passenger ferry ordered by mobile app.	Commercial	Commissioning/Testing
Hydrolift Smart City Ferries	- Hydrolift Smart City Ferries AS - Hydrolift AS - Eker/Eker design - SAMS Norway	Aims to provide a smart and sustainable unmanned MASS transit solution to alleviate congestion through effective use of underutilized waterways.	Commercial	Concept/Design
Fjordbus	- Hauschildt Marine A/S	Hauschildt Marine A/S has developed the new innovative EL-ferry tender design and specification with focus on no-emission, low energy consumption and future autonomous operation.	Commercial	Concept/Design
Milli-Ampere	- NTNU	A small and unmanned autonomous passenger ferry which is intended for sue in the port area of Trondheim. The ferry will be on-demand.	Research	Commissioning/Testing

Project	Project partners	Description	Type of project	Status
Roboat	- Amsterdam Institute for Advanced Metropolitan Solutions - Massachussets Institute of Technology	In developing the world's first fleet of autonomous floating vessels for the city of Amsterdam, it investigates the potential of self-driving technology to change our cities and their waterways. Roboat is a new kind of on-demand infrastructure: autonomous platforms will combine to form floating bridges and stages, collect waste, deliver goods, and transport people, all while collecting data about the city.	Research	Commissioning/Testing
Ballstad small passenger ferry		The purpose of the project is to clarify the possibility of a ferry (floatman) at Ballstad, based on future technology around autonomous (driverless) vessels.	Commercial	Unknown
Short sea cargo				
Yara Brikeland	- Yara - Konsberg Maritime - Massterly	The world’s first fully electric and autonomous container ship, with zero emissions. The vessel will carry containers for Yara between Larvik, Brevik and Herøya.	Commercial	Commissioning/Testing
NYK	- NYK	Navigation using sherpa system for real ship (SSH). Obtained approval in principle for an autonomous ship framework.	Commercial	Commissioning/Testing
ASKO	- ASKO - Kongsberg Maritime - Massterly	Norwegian grocery distributor ASKO has agreed a deal with Kongsberg Maritime and Massterly, to equip two new vessels with autonomous technology and manage their operations at sea.	Commercial	Commissioning/Testing
Seashuttle	- Samskip - Flowchange - Kongsberg Maritime - Hyon - Massterly	The Seashuttle project aims to develop two smaller autonomous short sea container ships powered by hydrogen fuel cells.	Commercial	Concept/Design

Project	Project partners	Description	Type of project	Status
Zhi Fei / Zhi Teng Chinese box ship	- Navigation Brilliance (Qingdao) Technology - Dalian Maritime University - China Waterborne Transport Research Institute	300 teu cargo vessel is set to enter service next month on a short-sea route between Dongjiakou and Qingdao. The 5,000 dwt Zhi Fei, 117 m in length with a maximum speed of 12 knots.	Commercial	Commissioning/Testing
ASTAT - Autonomous Ship Transport at Trondheimsfjorden	- Kongsberg Maritime - Allskog - SINTEF - et al.	[...] the ASTAT project will investigate one particular new type of unmanned ships: A small and fully electric shuttle type vessel. The ASTAT project had its kick-off at May 16th 2017.	Research	Unknown
USVs and remotely controlled				
Swarm	- Spatial Integrated Systems, Inc. - US Navy	Five smaller USVs demonstrated. SAS is the foundational software behind all of SIS autonomous vehicles. Based on the NASA-JPL CARACaS (Control Architecture for Robotic Agent Command and Sensing) software designed for the Mars Rover program, SAS provides an intelligent, goal-oriented vehicle control system that can turn any vehicle into a smart robot.	Other	Concept/Design
Unmanned Surface Vehicles (USV)	- Kongsberg Maritime - Norsafe AS (Hull)	Small USV for hydroacoustic applications. The Sounder USV System is designed to perform for hydroacoustic applications, from fish finding and fishery research to seabed mapping. It is capable of direct remote control, supervised or autonomous operation.	Commercial	In operation

Project	Project partners	Description	Type of project	Status
Fugro Blue Shadow/Essence and SEAKIT	- Fugro - L3HARRIS	Fugro has deployed the world's most ambitious uncrewed surface vessel fleet to support subsea inspection, construction support, and hydrographic and geophysical surveys. We're proud to be at the forefront of the maritime industry's mission for safer, faster and more sustainable operations.	Commercial	In operation
Reach Remote	- Reach subsea - Kongsberg Maritime - Massterly	Innovative concepts for remote and autonomous operations of subsea vessels.	Commercial	In operation
Maritime Robotics	- Maritime Robotics	Maritime Robotics is a leading provider of innovative Unmanned Vehicle Systems (UVS) for maritime operations in harsh environments.	Commercial	In operation
Armada	- Ocean Infinity - Grovfjord Mek	21 and 36m multipurpose work/supply boats. Grovfjord Mek. Verksted will build an Armada of large, cutting-edge robot ships for Ocean Infinity, the world leading marine robotics company. The fleet of ultra-low emission marine robots will be deployed from shorelines across the globe.	Commercial	Construction
ROSS	- SeaOwl - Marilink - ADEME - Bureauv Veritas	Remotely Operated Service at Sea (ROSS). SeaOwl intends to order a fleet of unmanned vessels for underwater inspection of offshore oil, gas and renewables infrastructure between 2023 and 2028 following completion of the Remotely Operated Service at Sea (ROSS) development project.	Commercial	Concept/Design
Saildrone Surveyor	- Saildrone	USV powered by wind and solar technology. Carries sonar equipment capable of seafloor mapping.	Commercial	Commissioning/Testing
Tugboats				

Project	Project partners	Description	Type of project	Status
IntelliTug	- Wartsilla - PSA Marine - MPA Singapore - Lloyd's register - TCOMS	The PSA Polaris is a 27-metre harbour tug with dual azimuth thruster controls. It has been fitted with a sensor suite, including Wärtsilä's RS24 near-field high resolution radar and Wärtsilä's Dynamic Positioning (DP) system, to enable autonomous capabilities.	Research	Finished
RECOTUG	- Svitzer - Kongsberg Maritime	Aims to provide a commercially viable tugboat that can be used in operations worldwide.	Commercial	Concept/Design
Car ferry				
Bastø Fosen	- Bastø-Fosen - Kongsberg Maritime	Kongsberg Maritime will develop a system for automating the crossing between Horten and Moss in Norway, and for moving ferries to and from the quay.	Commercial	In operation
Autonomous functions system suppliers				
Buffalo Automation		Buffalo Automation is an artificial intelligence start-up that makes products for commercial vessels, recreational boats, ports and water taxis with the goal to improve maritime safety and automation.	Commercial	Concept/Design
uSEA	- Blue Logic - NTNU - GCE Ocean Technology	uSEA works to step-change autonomy for marine and underwater robotics to offer seabed surveys and mapping, pipeline and subsea cable inspection as well as environmental monitoring. [...] an unmanned, all-weather subsea survey system.	Commercial	Unknown

Project	Project partners	Description	Type of project	Status
Sea Machine Robotics	- Sea Machine - Damen Shipyards	Provides various services for ship autonomy, such as remote-control equipment and situational awareness systems. Products/solutions: Situational Awareness/Decision support System: - SM400 - SM300	Commercial	On the market
Robosys Automation		We use cutting edge Artificial Intelligence to provide scalable levels of autonomy to existing and new-build vessels. We enjoy excellent links with leading equipment manufacturers, offering sophisticated control solutions for Smart bridge systems, machinery and sensors. Our Robosys Voyager software can transform an existing vessel into a fully autonomous USV, capable of independent navigation and collision avoidance.	Commercial	On the market
Wartsilla		Offer a range of software/hardware products that provide hazard detection and situational awareness as well as sensors that can enable autodocking and auto-manoeuvres. Products/Solutions: Product vendor: Situational Awareness/Decision support System and autodocking etc. - SmartQuay - SceneScan - Smart family: SmartDrive, SmartPredict, SmartDock, SmartTransit, SmartCommand -SmartMove suite - Etc.	Commercial	On the market

Project	Project partners	Description	Type of project	Status
Kongsberg Maritime		Offer a range of software/hardware products that provide hazard detection and situational awareness as well as sensors that can enable autodocking and auto-manoeuvres. Products/Solutions: Control systems/Situational Awareness/Decision support System/Performance monitoring and autodocking etc. - SEAAWARE - SEAEYE - PROXIMITYVIEW - Autocrossing/Autodocking - Etc.	Commercial	On the market
ABB		Provides various services advocating increased autonomy through their ABB Ability product portfolio. One of their visions is the bridge zero (B0) concept. Products/solutions: Situational Awareness/Decision support Systems/Performance monitoring: - ABB Ability™ Marine Pilot	Commercial	On the market
Shone		Provides situational awareness/decision support solutions.	Commercial	Concept/Design
ORCA AI		Provides situational awareness/decision support solutions for individual ships and (navigation) performance monitoring for the fleet. Products/Solutions: Situational awareness/Decision support System/Performance monitoring - ORCA for the ship - ORCA for the fleet	Commercial	On the market

Project	Project partners	Description	Type of project	Status
L3Harris		Provides misc. Systems for autonomous navigation. Products/Solutions: Misc. autonomous surface vehicles, hardware and software solutions: - 360-DEGREE SYSTEMS, NAVAL SURFACE IMAGING - ASVIEW™ CONTROL SYSTEM	Commercial	In operation
Leidos		Provides misc. Systems for autonomous navigation	Other	Concept/Design
Military				
Madfox	- L3Harris - Royal navy - NavyX	With Madfox now directly in the hands of NavyX, the team will be able to explore a multitude of issues such as safety, regulatory compliance, new missions, new payloads and the role that a USV can play in complex operations and within the future fleet.	Other	Commissioning/Testing
Sea Hunter	- Leidos - U.S. Department of Defence (DoD) - Navy - Intelligence Community	Medium displacement Unmanned Surface Vessel. The Sea Hunter program is leading the world in unmanned, fully autonomous naval ship design and production.	Other	In operation
HEBE	- UK Royal Navy - Atlas Elektronik UK	Autonomous vehicle for mine countermeasures.	Other	In operation
The Watcher	- Metalcraft Marine - US Coast Guard	A high-speed autonomous surface vehicle designed for long-endurance patrol, surveillance and interception missions. The craft can operate under full autonomous mode for up to 30 days.	Other	In operation
Line shipping				

Project	Project partners	Description	Type of project	Status
None, rather utilize automated functions, support systems, remote options from the development in other segments				
Other vessels				
RALamander 2000	- Robert Allan Ltd. - Kongsberg Maritime	Collaboration on the development of a radically new remotely operated fireboat that will allow first responders to attack dangerous port fires more aggressively and safer than ever before. with conventional firefighting assets or be deployed on its own.	Commercial	Concept/Design
Remote pilotage	- Wartsila	Remote pilotage service	Commercial	Unknown
Hronn	- Kongsberg Maritime - Automated Ships - BORBON	Hrönn is a light-duty, offshore autonomous utility ship servicing the offshore energy, scientific/hydrographic and offshore fish-farming industries. Scheduled to enter service in 2018.	Commercial	Unknown
Research, consortium, framework development				
ROMAS	- DNV - Fjord1 - NMA - Høglund	Remote operations of machinery and autonomous systems. The ROMAS project aims to establish a framework of regulations, rules and verification methods for remote (shore-based) operations of ship machinery and automation systems, enabling improved operations and cost-efficiency without compromising safety of ship operations.	Research	Finished
Aegis	- SINTEF - DFDS - NCL - et al.	Three logistics concepts. The AEGIS consortium will design Europe's next generation sustainable and highly competitive waterborne logistics system comprising more autonomous ships and automated cargo handling. Standardized cargo units and digital connectivity are key elements in the AEGIS system.	Research	Ongoing

Project	Project partners	Description	Type of project	Status
AUTOSHIP	- Ciaotech - Kongsberg - SINTEF - Blue Line Logistics - et al.	AUTOSHIP – Autonomous Shipping Initiative for European Waters – aims at speeding-up the transition towards a next generation of autonomous ships in EU. The project will build and operate 2 different autonomous vessels, demonstrating their operative capabilities in Short Sea Shipping and Inland Water Ways scenarios, with a focus on goods mobility.	Research	Ongoing
SFI AUTOSHIP	- NTNU - SINTEF - IFE - UiO - et al.	SFI Autoship is a 8-years research-based innovation center that will contribute to Norwegian players taking a leading role in the development of autonomous ships for safe and sustainable operations.	Research	Ongoing
MAGPIE	- Wartsilla - Port of Rotterdam Authority	Development of a grand plan that sets out how transport within, to and from ports can be made carbon-free by 2050. Wartsila is to demonstrate a commercially viable autonomous intra-port inter-terminal container shuttle to address an emerging capacity bottleneck for internal container transportation.	Research	Ongoing

APPENDIX B
Failure categories

A modern system may be subject to many different types of failures. Failures can be classified as:

- Random (hardware) failures,
- Systematic failures,
- Systemic failures,
- Operator failures,
- Failures due to environmental causes
- Failures due to deliberate actions.

Note that these categories overlap to some extent, yet they are useful as a guide to identify a wide range of failures that may pose risk.

Random hardware failures are linked to the physical properties of components. The term random is used because the exact moment a specific component will fail is unknown and does not imply that the failure happens arbitrarily. Typical failure rates for a large group of the same component can be predicted through analysis of statistics from field experience, and this makes it possible to perform Quantitative Risk Analysis (QRA) that takes into account the probability of failure for the different components in a system.

The degradation mechanisms that lead to random failures can to some extent be controlled by adjusting how components are designed produced, transported, installed, operated, and maintained. Thus, the failure rates for specific components will partly depend on the quality, operational and maintenance regimes applied. In this regard, it is important to be aware that generic failure rates for specific type of components consider all employed quality regimes equal, which is a simplification that represents an uncertainty in the calculations. Furthermore, it should be noted that the failure rates used in QRA typically excludes the run-in and wear-out periods, and therefore failures experienced in usage inside of these periods may be considered systematic failure events rather than random.

Systematic failure events are the consequence of inadequate work processes and may be introduced at all stages in the system lifecycle. Some examples are incomplete risk analysis, inadequate development of barrier strategies, incomplete requirement specifications, weaknesses in software design, programming errors, quality problems in hardware production, and inadequate planning of maintenance. It is difficult to quantify the probability of systematic failure events as they typically will be present in a system from day one, or introduced through modification, but be hidden until specific circumstances occur. This makes it difficult to compare the risks associated with different systems quantitatively, and necessitates broader risk descriptions if a comparison is to be made

A **systemic failure** is an event which occurs even if no individual component in the system has failed. This may be caused e.g., by overlooked dependencies among the technical, operational, human, and organisational elements of systems, specifications that are based on inadequate understanding of physical processes, or unexpected inputs for which no specific response has been specified. Increasing system complexity may increase the risk of systemic failures, and this is particularly relevant for systems containing software functions. It can be related to intricate dependencies and feed-back mechanisms among system components leading to nonlinear and unpredictable system behaviour. Lack of knowledge and understanding of interactions in a system increase the risk of systemic failures as it makes it difficult to implement robust barrier strategies to prevent them. Choice of simple solutions with few interacting or interdependent elements may reduce the risk of systemic failures and make systems more robust.

Operator failures occur when an operator fails to perform appropriate actions or performs an inappropriate action. The ability of an operator to perform appropriate actions and avoid inappropriate actions depends on the availability and quality of information to act on, the availability of sufficient time to act, and possession of knowledge of how to act. Therefore, the underlying causes of an operator failure may be systematic or systemic failures that involve technical, operational and organisational elements. In particular, operator failures may be dependent on system designs, operational procedures, training of the operator, and assumptions made in the risk treatment strategy. The latter includes availability of measures that realistically can be used to mitigate the risk under relevant operational conditions.

Failures due to environmental causes are caused by physical processes having negative influence on the control system. Some examples are: Lightning strike, water ingress, fire, electrostatic discharge from personnel, sensors covered by salt, and electromagnetic interference affecting communications. What is considered the environment depends on the boundaries of the system being analysed. E.g., loss of cooling in a control room may in some risk analyses be seen as an environmental cause, but not if the cooling system is a part of the system being analysed.

Failures due to deliberate actions may be caused for example by hacking, data viruses, physical sabotage, deliberate jamming of radio signals, GPS spoofing (false signals).

Regarding evaluation of possible mitigations, it should be considered that a systemic failure reflects inadequate identification of relevant requirements. Thus, systemic failure may be seen as a form of systematic failure introduced in the requirement specification phase. Mitigation of a failure scenario caused by inadequate requirements typically requires some level of functional diversity between the control functions affected by the failure and the mitigating measure.

In general, all software failures are systematic or systemic in nature, although the occurrence of the input conditions revealing the weakness in the software may in some cases may be perceived as being random-like in nature. Local detection mechanisms, e.g., range checking and plausibility checks may be used to detect some of these. Other failures can only be detected at higher levels in the system that have a broader overview of the system state and the current operational mode, e.g., by comparing output from different controllers in functionally diverse subsystems, or through operator observation of system behaviour.

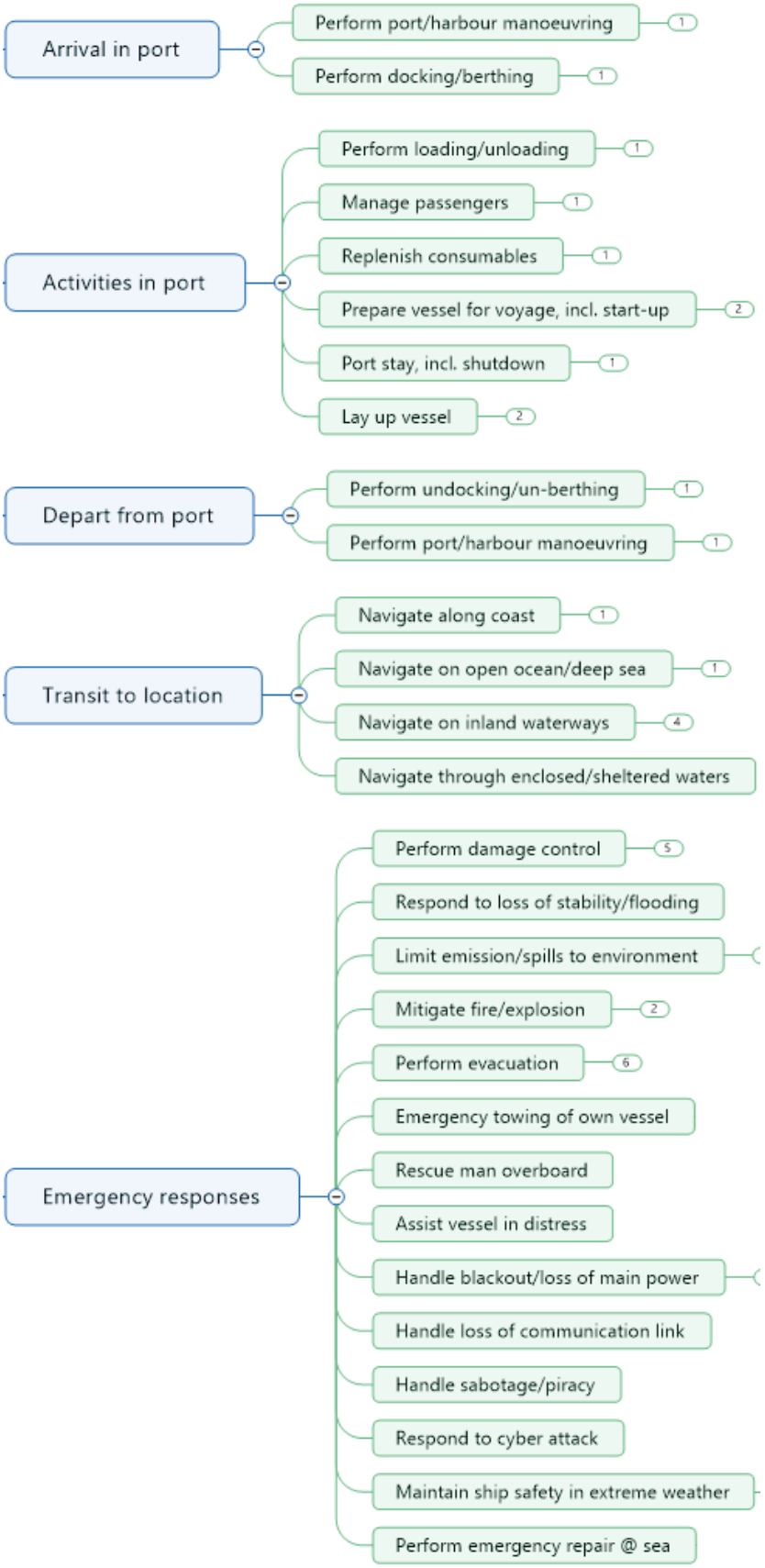
It will not always be possible to test a system under all relevant use scenarios, and it may even be that the test scenarios that are feasible to check are not realistic. In addition, for software functions within a system, the number of possible input combinations and possible execution paths typically prevents exhaustive testing even when using a simulated environment. This means that testing typically can only demonstrate the presence of conditions that can lead to failures and not their absence. A cautionary approach is therefore warranted to make systems robust to unforeseen conditions that it may experience. This may include fall-back solutions and use of safety margins considering worst-case scenarios.

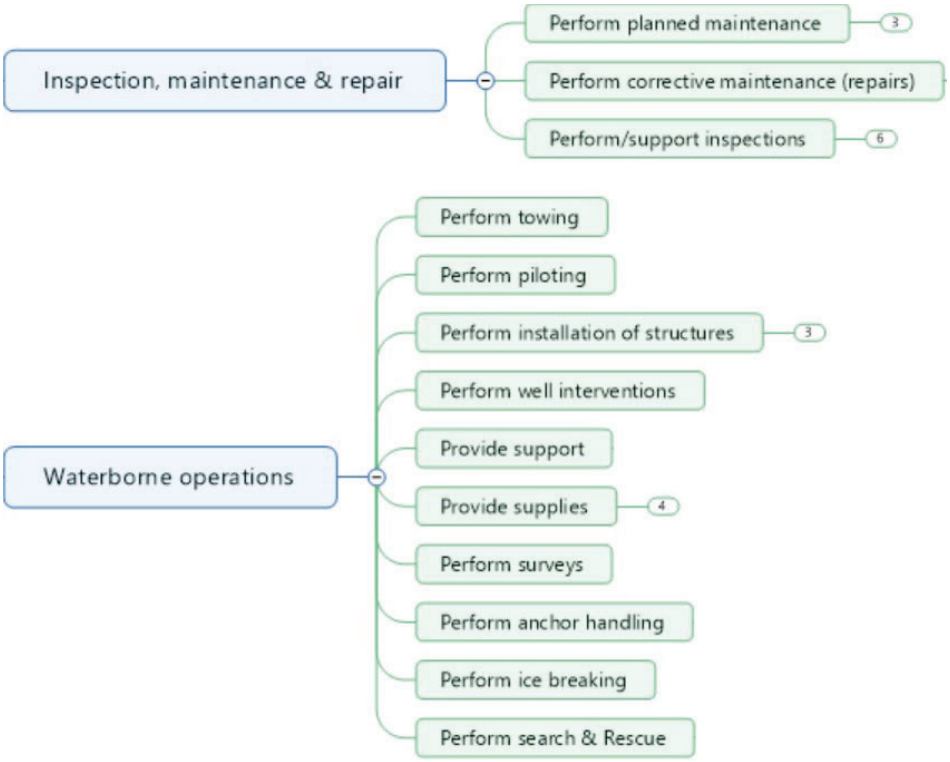
It will in many cases not be possible to implement detection for all types of systematic/systemic failures. E.g., incomplete analysis of systems, operations, interfaces, and risks may lead to omissions in specifications evading all detection mechanisms. For safety-critical systems, there must either be an efficient fallback chain, or it must be possible to argue that activities associated with analyses, development, verification, and validation have reduced the likelihood of systematic and systemic failures to a tolerable level.

The latter approach may be challenging, e.g., the number of possible combinations of inputs to the system, and the number of possible sequences of input combinations can make it difficult to know whether specifications are complete. Thus, in practice, one often uses a combination where both a fallback chain and a rigorous development process are used to reduce residual risk to a tolerable level.

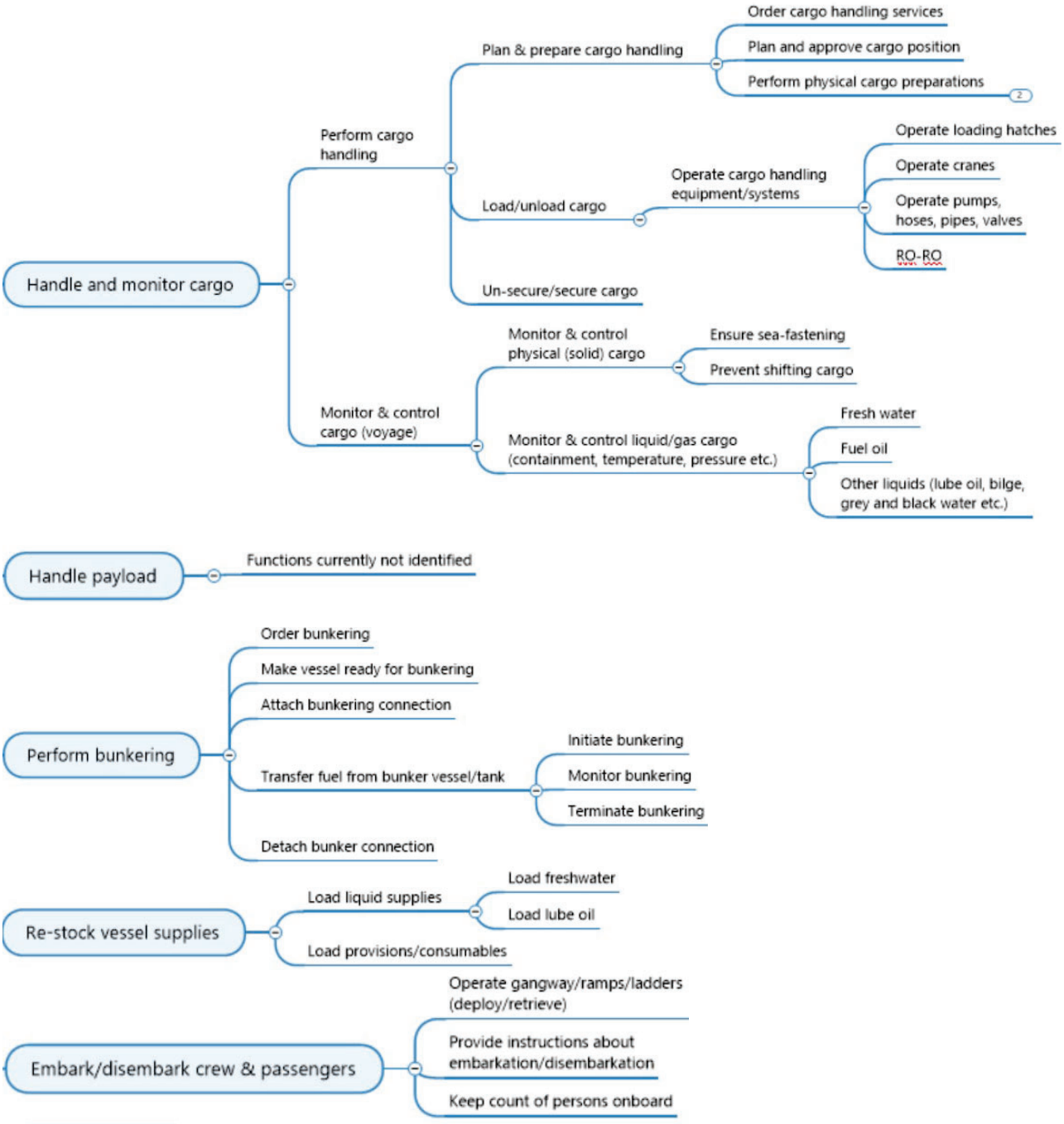
Since the effectiveness of mitigation measures varies with the type of cause, it is important to consider all failure categories mentioned at the start of this section when performing risk analysis and developing risk treatment strategies. For example, hardware redundancy in combination with voting may be an efficient mitigation against random hardware failures, but it will not be efficient if the cause is systematic or systemic. Furthermore, the use of functional diverse supporting functions may reduce risks related to systematic failures in those functions, but it may not be efficient against systematic failures in the top-level function. Operator intervention through independent means may be efficient against systematic failures in the top-level function, but additional measures may be necessary if the cause is an operator failure, fire and flooding, or deliberate actions like hacking or sabotage.

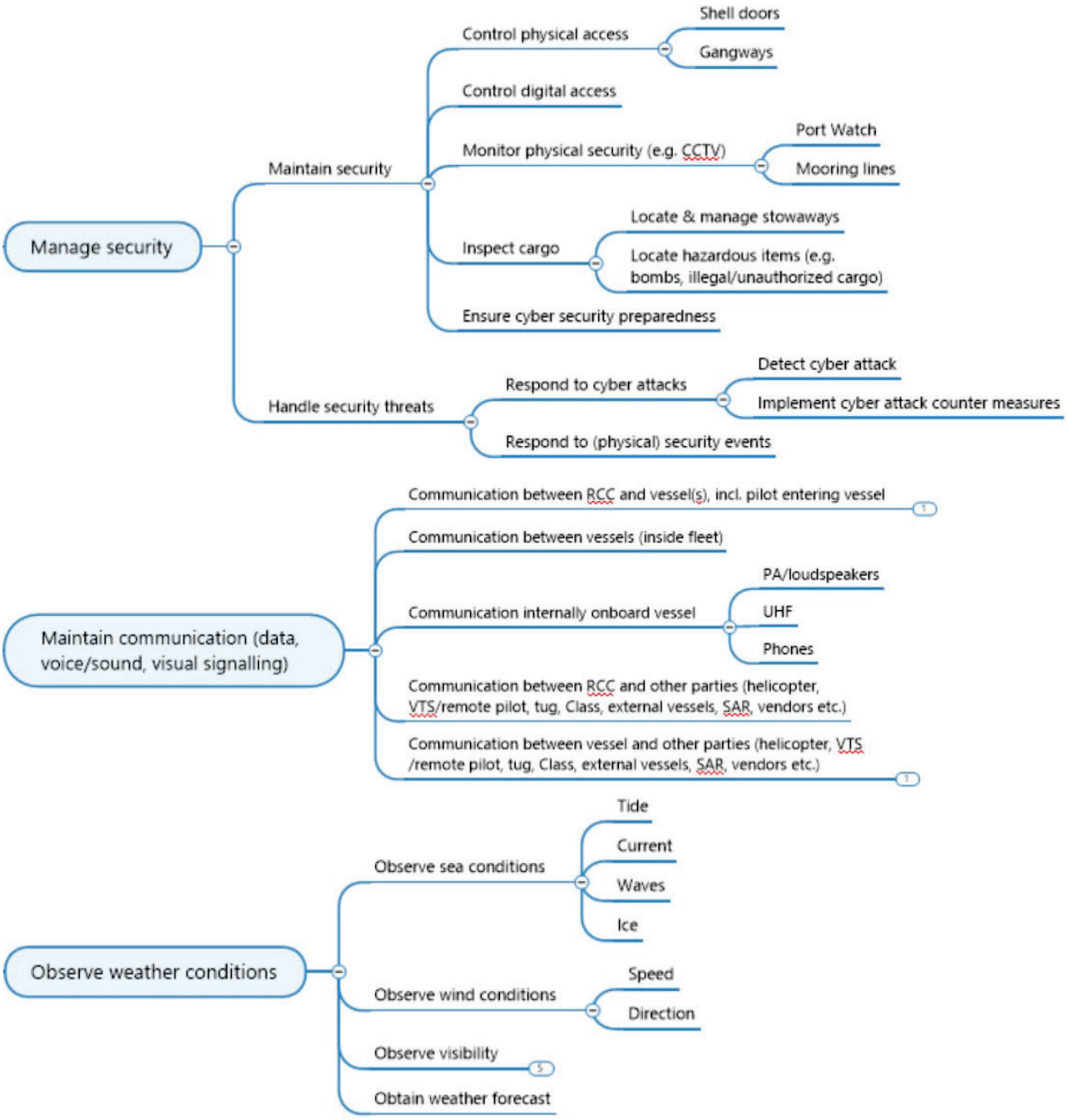
APPENDIX C
Mission Model

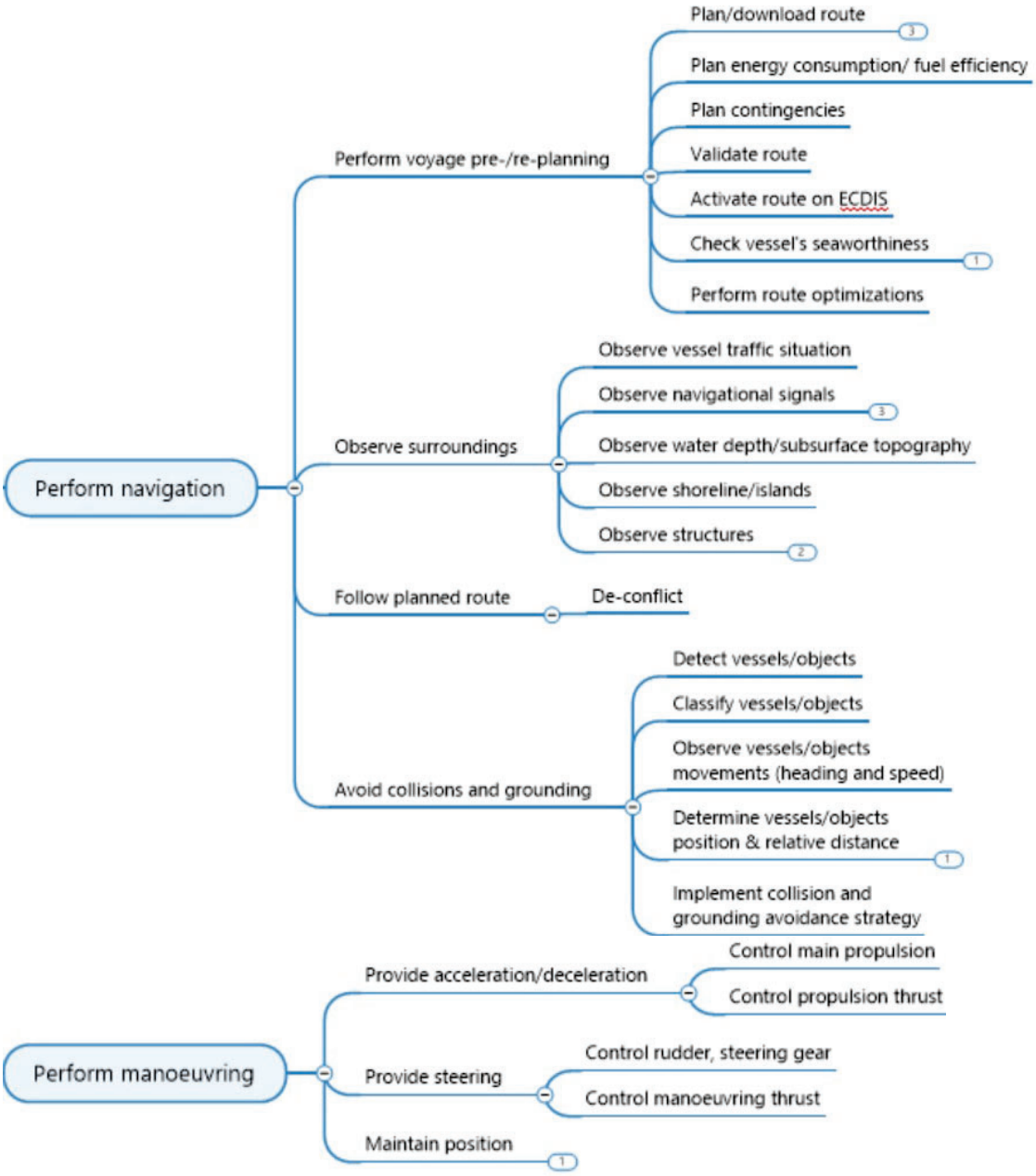


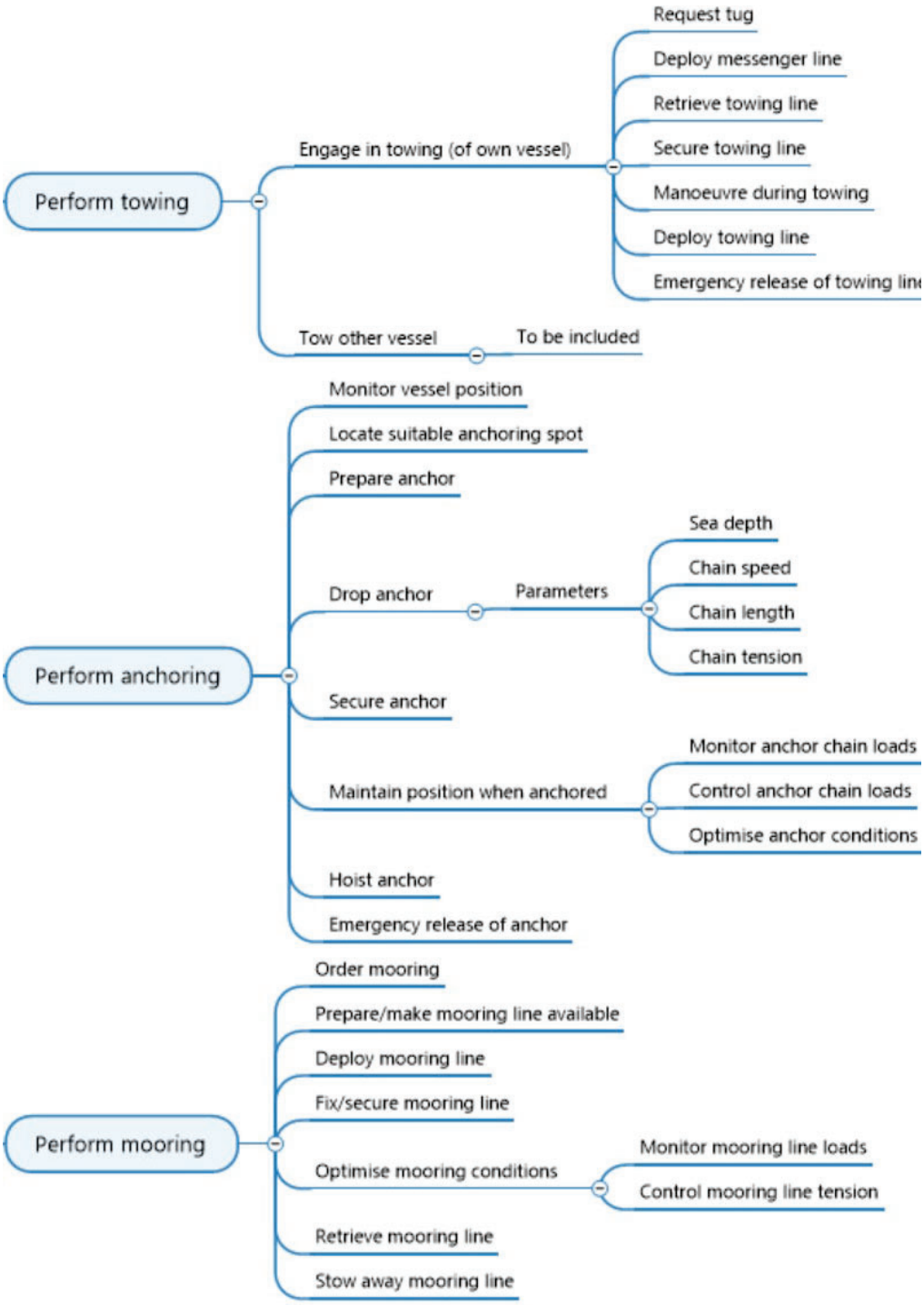

Inspection, maintenance & repair
Perform planned maintenance
3
Perform corrective maintenance (repairs)
Perform/support inspections
6
Waterborne operations
Perform towing
Perform piloting
Perform installation of structures
3
Perform well interventions
Provide support
Provide supplies
4
Perform surveys
Perform anchor handling
Perform ice breaking
Perform search & Rescue

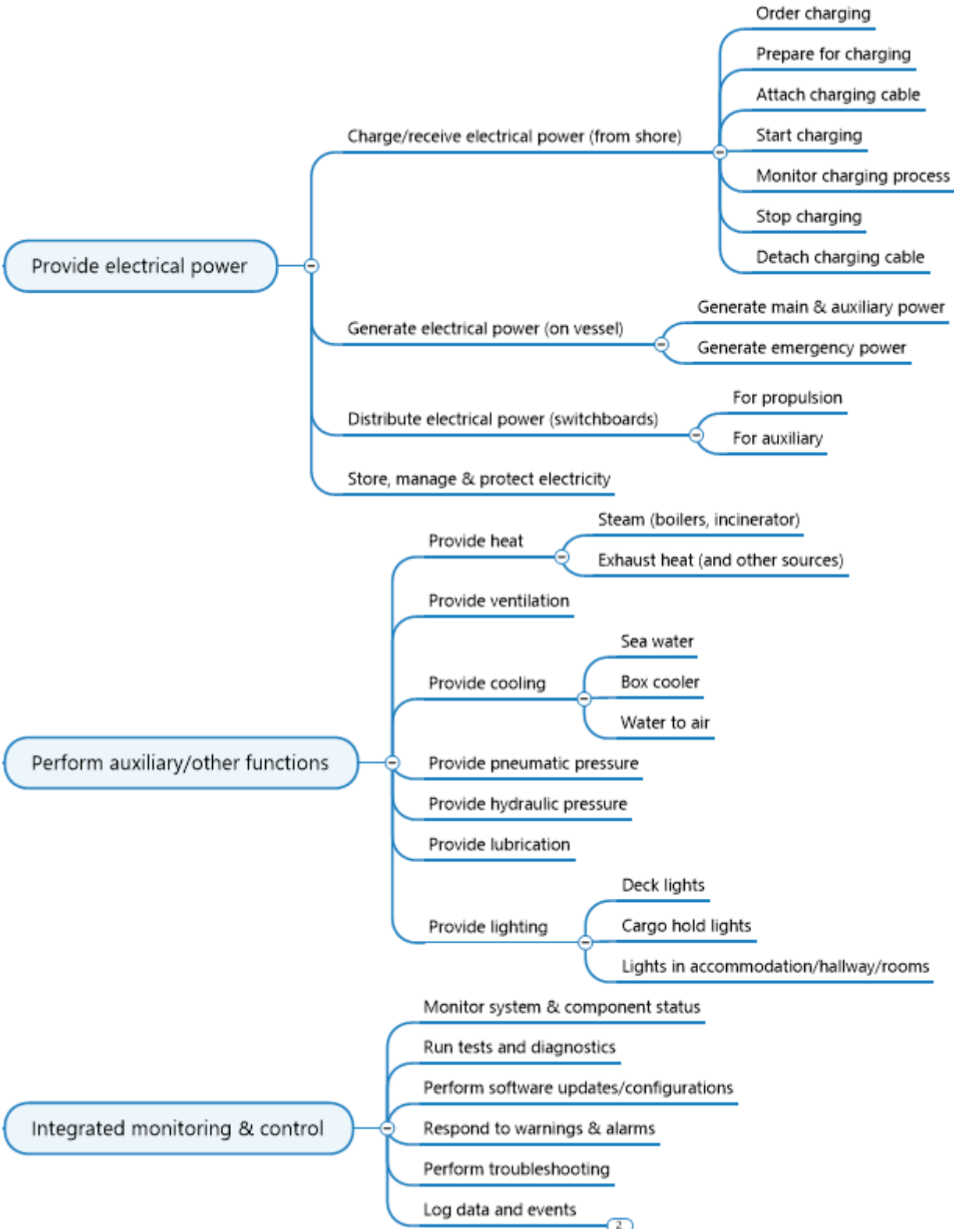
APPENDIX D
Function Tree

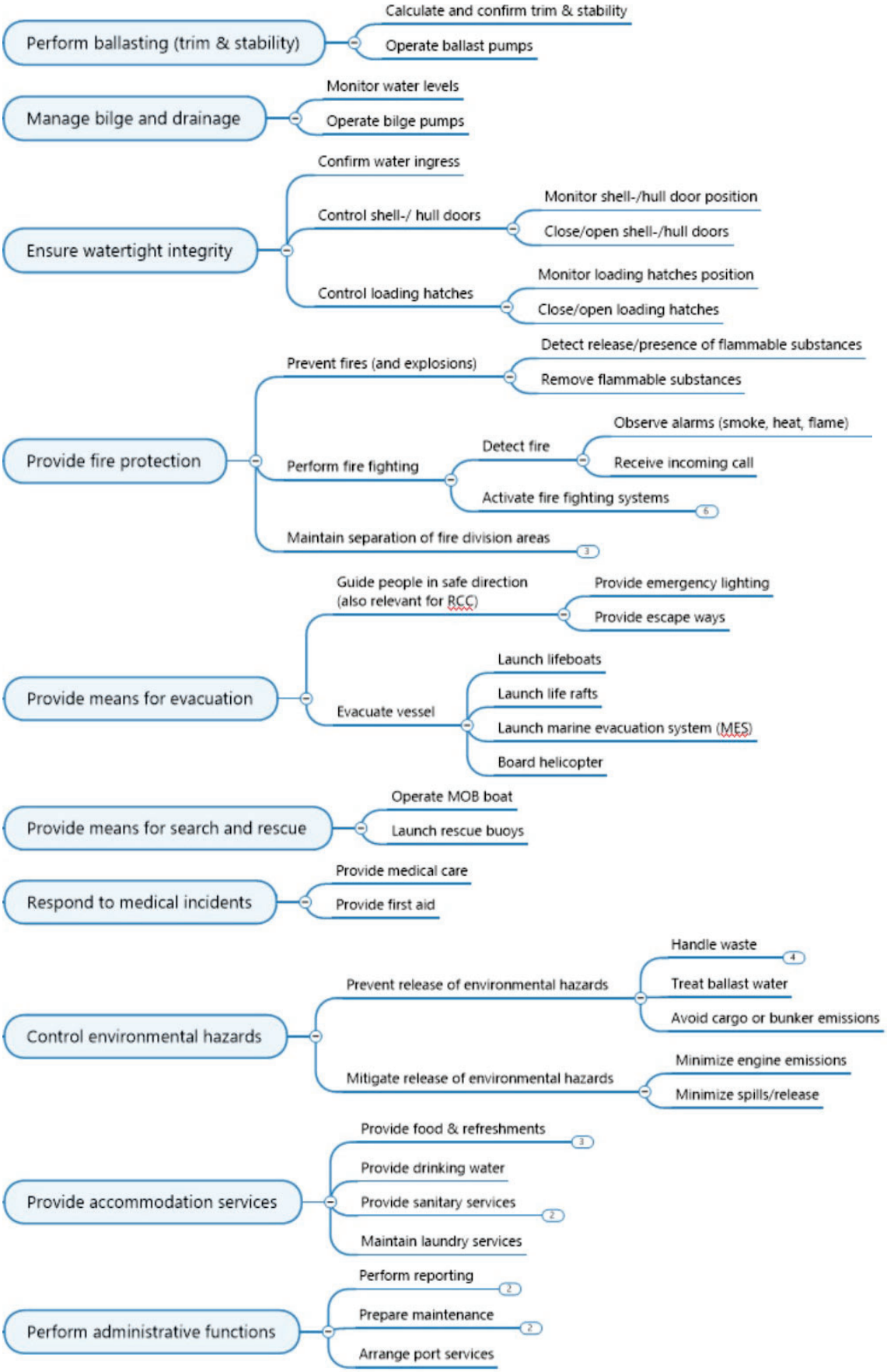



Manage security
Maintain security
Control physical access
Shell doors
Gangways
Control digital access
Monitor physical security (e.g. CCTV)
Port Watch
Mooring lines
Inspect cargo
Locate & manage stowaways
Locate hazardous items (e.g. bombs, illegal/unauthorized cargo)
Ensure cyber security preparedness
Handle security threats
Respond to cyber attacks
Detect cyber attack
Implement cyber attack counter measures
Respond to (physical) security events
Maintain communication (data, voice/sound, visual signalling)
Communication between RCC and vessel(s), incl. pilot entering vessel
Communication between vessels (inside fleet)
Communication internally onboard vessel
PA/loudspeakers
UHF
Phones
Communication between RCC and other parties (helicopter, VTS/remote pilot, tug, Class, external vessels, SAR, vendors etc.)
Communication between vessel and other parties (helicopter, VTS /remote pilot, tug, Class, external vessels, SAR, vendors etc.)
Observe weather conditions
Observe sea conditions
Tide
Current
Waves
Ice
Observe wind conditions
Speed
Direction
Observe visibility
Obtain weather forecast


Perform navigation
Perform voyage pre-/re-planning
Plan/download route
3
Plan energy consumption/ fuel efficiency
Plan contingencies
Validate route
Activate route on ECDIS
Check vessel's seaworthiness
1
Perform route optimizations
Observe surroundings
Observe vessel traffic situation
Observe navigational signals
3
Observe water depth/subsurface topography
Observe shoreline/islands
Observe structures
2
Follow planned route
De-conflict
Avoid collisions and grounding
Detect vessels/objects
Classify vessels/objects
Observe vessels/objects movements (heading and speed)
Determine vessels/objects position & relative distance
1
Implement collision and grounding avoidance strategy
Perform manoeuvring
Provide acceleration/deceleration
Control main propulsion
Control propulsion thrust
Provide steering
Control rudder, steering gear
Control manoeuvring thrust
Maintain position
1


Perform towing
Engage in towing (of own vessel)
Request tug
Deploy messenger line
Retrieve towing line
Secure towing line
Manoeuvre during towing
Deploy towing line
Emergency release of towing line
Tow other vessel
To be included
Perform anchoring
Monitor vessel position
Locate suitable anchoring spot
Prepare anchor
Drop anchor
Parameters
Sea depth
Chain speed
Chain length
Chain tension
Secure anchor
Maintain position when anchored
Monitor anchor chain loads
Control anchor chain loads
Optimise anchor conditions
Hoist anchor
Emergency release of anchor
Perform mooring
Order mooring
Prepare/make mooring line available
Deploy mooring line
Fix/secure mooring line
Optimise mooring conditions
Monitor mooring line loads
Control mooring line tension
Retrieve mooring line
Stow away mooring line

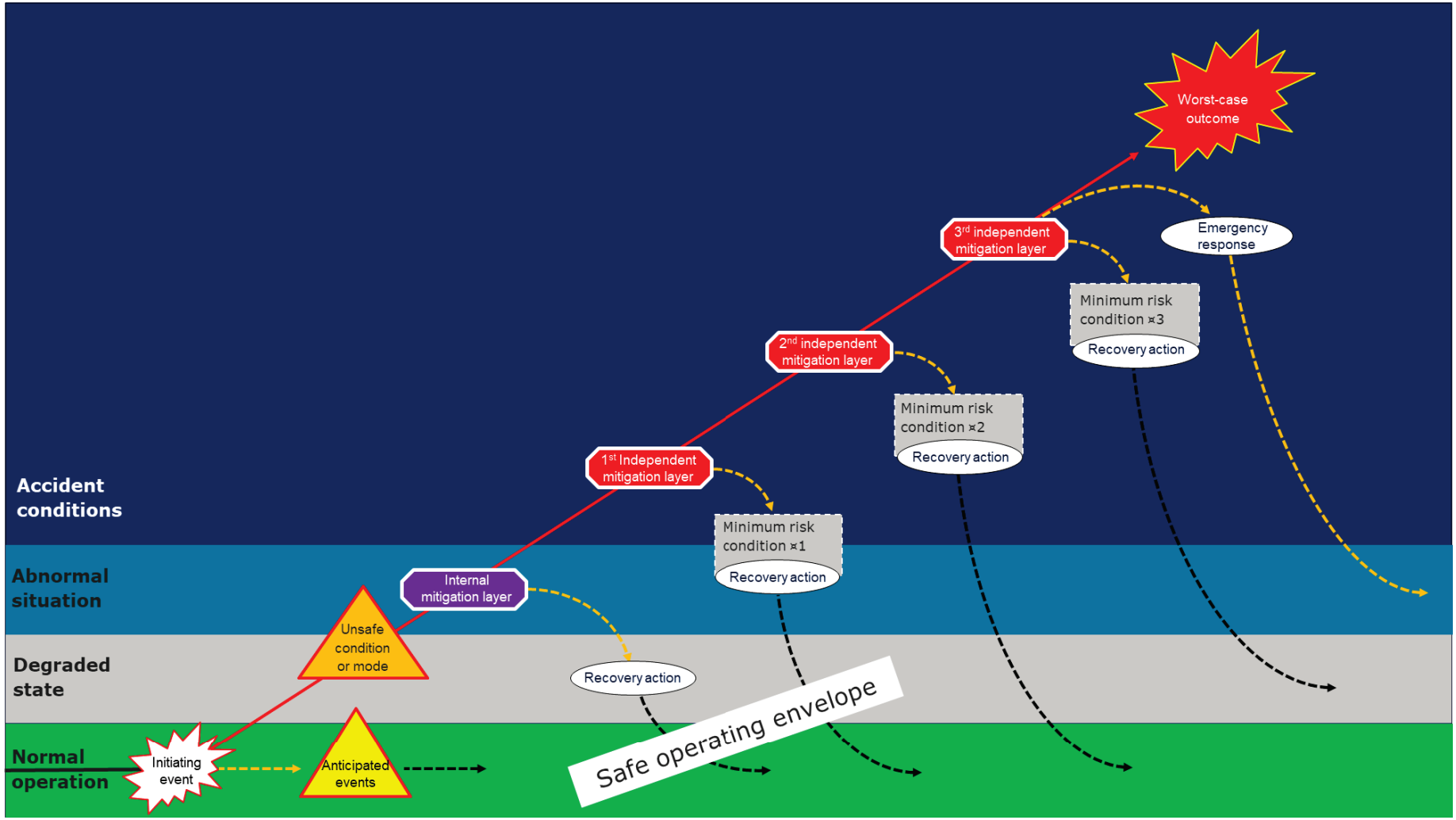

Provide electrical power
Charge/receive electrical power (from shore)
Order charging
Prepare for charging
Attach charging cable
Start charging
Monitor charging process
Stop charging
Detach charging cable
Generate electrical power (on vessel)
Generate main & auxiliary power
Generate emergency power
Distribute electrical power (switchboards)
For propulsion
For auxiliary
Store, manage & protect electricity
Perform auxiliary/other functions
Provide heat
Steam (boilers, incinerator)
Exhaust heat (and other sources)
Provide ventilation
Provide cooling
Sea water
Box cooler
Water to air
Provide pneumatic pressure
Provide hydraulic pressure
Provide lubrication
Provide lighting
Deck lights
Cargo hold lights
Lights in accommodation/hallway/rooms
Integrated monitoring & control
Monitor system & component status
Run tests and diagnostics
Perform software updates/configurations
Respond to warnings & alarms
Perform troubleshooting
Log data and events


Perform ballasting (trim & stability)
Calculate and confirm trim & stability
Operate ballast pumps
Manage bilge and drainage
Monitor water levels
Operate bilge pumps
Ensure watertight integrity
Confirm water ingress
Control shell-/ hull doors
Monitor shell-/hull door position
Close/open shell-/hull doors
Control loading hatches
Monitor loading hatches position
Close/open loading hatches
Provide fire protection
Prevent fires (and explosions)
Detect release/presence of flammable substances
Remove flammable substances
Perform fire fighting
Detect fire
Observe alarms (smoke, heat, flame)
Receive incoming call
Activate fire fighting systems
6
Maintain separation of fire division areas
3
Provide means for evacuation
Guide people in safe direction (also relevant for RCC)
Provide emergency lighting
Provide escape ways
Evacuate vessel
Launch lifeboats
Launch life rafts
Launch marine evacuation system (MES)
Board helicopter
Provide means for search and rescue
Operate MOB boat
Launch rescue buoys
Respond to medical incidents
Provide medical care
Provide first aid
Control environmental hazards
Prevent release of environmental hazards
Handle waste
4
Treat ballast water
Avoid cargo or bunker emissions
Mitigate release of environmental hazards
Minimize engine emissions
Minimize spills/release
Provide accommodation services
Provide food & refreshments
3
Provide drinking water
Provide sanitary services
2
Maintain laundry services
Perform administrative functions
Perform reporting
2
Prepare maintenance
2
Arrange port services

APPENDIX E
List of verbs

Information acquisition	Information analysis	Decision making	Action implementation
Access	Calculate	Command	Acknowledge
Detect	Classify	Conclude	Activate
Hear	Compare	Determine	Alert
Observe	Consider	Generate	Align
Read	Define	Plan	Announce
Receive	Identify	Select	Approve
Record	Integrate		Attach
Registrate	Interpret		Attain
Review	Organize		Brief
Scan	Predict		Close
Sense	Prioritize		Communicate
	Trend		Compute
	Verify		Configure
Action implementation cont.			
Continue	Extinguish	Monitor	Reset
Control	Fasten	Open	Respond
Coordinate	Fill	Operate	Secure
Cycle	Follow	Order	Stabilize
Deactivate	Guard	Perform	Start
Debrief	Illuminate	Position	Steer
Decelerate	Increase	Prepare	Stop
Decrease	Initialize	Pressurize	Stow
Depressurize	Initiate	Prevent	Test
Detach	Inspect	Proceed	Transmit
Deviate	Intercept	Program	Trim
Discharge	Interrogation	Provide	Tune
Eliminate	Isolate	Recover	Turn
Enter	Load	Remove	Unfasten
Evacuate	Maintain	Repeat	Unload
Exit	Manoeuvre	Report	Unsecure
Extend	Modify	Request	Update

APPENDIX F
Accident model



About DNV

DNV is the independent expert in risk management and assurance, operating in more than 100 countries. Through its broad experience and deep expertise DNV advances safety and sustainable performance, sets industry benchmarks, and inspires and invents solutions.

Whether assessing a new ship design, optimizing the performance of a wind farm, analyzing sensor data from a gas pipeline or certifying a food company's supply chain, DNV enables its customers and their stakeholders to make critical decisions with confidence.

Driven by its purpose, to safeguard life, property, and the environment, DNV helps tackle the challenges and global transformations facing its customers and the world today and is a trusted voice for many of the world's most successful and forward-thinking companies.