



CHALMERS
UNIVERSITY OF TECHNOLOGY



Artificial Intelligence to Improve Situation Awareness for AEW&C Operators

A research through design approach to create guidelines and visual concepts

Master's thesis in Industrial Design Engineering

JULIUS BUNG TIDBLM & ELIAS JANSSON
DEPARTMENT OF INDUSTRIAL AND MATERIALS SCIENCE
CHALMERS UNIVERSITY OF TECHNOLOGY

Gothenburg, Sweden 2021
www.chalmers.se

MASTER'S THESIS 2021

Artificial Intelligence to Improve Situation Awareness for AEW&C Operators

A research through design approach to create guidelines and visual concepts.

JULIUS BUNG TIDBLOM
ELIAS JANSSON

Department of Industrial and Material Science

Division for Design & Human Factors
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2021

Master of Science Thesis (IMSX30)

Artificial Intelligence to Improve Situation Awareness for AEW&C Operators

A research through design approach to create guidelines and visual concepts.

In collaboration with Saab AB

© JULIUS BUNG TIDBLOM & ELIAS JANSSON, 2021.

Department of Industrial and Material Science
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone + 46 (0)31-772 1000

Cover illustration: Operator workstation, by Julius Bung Tidblom

Printed by Chalmers Digitaltryck
Gothenburg, Sweden 2021

Abstract

In the Airborne Early Warning & Control (AEW&C) system, the operator needs to be vigilant, perform multiple tasks, communicate, and simultaneously handle different sources of information during a mission. The operator's main objective is to receive, interpret and distribute information and data provided by the AEW&C system to create a Recognized Air and Sea picture (RASP). The information is then used to, for example, perform fighter jet control and to detect anomalies. Increased capabilities of Command & Control (C2) systems and new threats in the environment make the operators' tasks more extensive, complicated, and complex than ever before, which could potentially impair their Situation Awareness (SA) and in turn their decision making. The domain is associated with high risks and potentially devastating impacts of faulty decisions. Therefore, the importance of finding new technologies to facilitate SA cannot be neglected. The increasing technological sophistication regarding the rapidly growing adoption of Artificial Intelligence (AI) offers new possibilities for creating systems facilitating SA. This motivates conducting studies to explore and analyse the opportunities AI provides to facilitate SA for AEW&C operators, in particular critical roles as the Fighter Controller (FC) and Surveillance Operator (SO). This study aims to answer the questions of *when* and *how* AI could be implemented to facilitate SA.

The questions were answered using a process adapting the *research through design* approach. The *when*-question was answered through system analysis, user studies, concept development and evaluations and the *how*-question were answered by creating guidelines through literature studies and affinity diagrams and testing the guidelines by applying them in the development and evaluation of concepts. The system was divided into four separate subsystems where goals, subjects, tools, and outcomes were specified. Seven categories of SA related challenges were identified for FC and six for the SO respectively. Four concepts were created for each role aiming to improve SA, where concepts of a talk-translation tool, a formation recognition tool, abnormality detection tool, and timeline tool were considered the most promising for further work. Guidelines for three phases in development of AI functions were developed: planning, designing, and evaluating. The guidelines confirmed the utility of *design through research* by being used to explore opportunities to implement AI, ideate and create concepts and evaluate them in an AEW&C context.

Keywords: Artificial Intelligence, Situation Awareness, AEW&C, Research through design, AI guidelines, AI in military operations, Explainable AI, Design, User studies, System analysis

Acknowledgements

This is a master's thesis report by two students at the programme Industrial Design Engineering under the Department of Industrial and Materials Science at Chalmers University of Technology.

Even though there are only two authors of this report, the study it presents would never have existed without the help from others. We always had help and support along the way whenever we asked for it (and sometimes even without asking). Therefore, we would like to express our deepest gratitude to everyone who made this thesis possible.

First of all, we would like to thank our supervisors at Saab, Marielle Preijde and Olivia Bång. You believed in us since day one and always supported us whenever we needed something. Thank you for your guidance and time and thank you for your valuable input and assistance during the operative simulation. It would not have been as successful without you. To all operators and engineers at Saab, a million cinnamon rolls would not be enough to thank you for all your valuable insights and opinions. Thank you for all the precious time you spent helping us understand what it means to be an AEW&C operator. Hopefully, you will find our work as interesting as we found yours. If we can get even close to that, it would be a huge success.

To Lars-Ola Bligård, our examiner and supervisor at Chalmers. You have been a rock, not only during our master's thesis, but throughout all our years at Chalmers. The Department of Industrial and Materials Science can consider themselves lucky to have you as a colleague and a friend. We know we can always count on you, even if it is to read and give feedback on a whole master's thesis report with only a day's notice (sorry about that). Thank you, and good luck in future research!

Finally, we would like to thank the company we did the thesis for, Saab. Thank you for providing us with all possible resources needed and welcoming us with open arms to your office in Kallebäck. We had many memorable coffee breaks on the seventh floor, looking through the huge glass windows at the beautiful city of Gothenburg.

Gothenburg, June 10th, 2021.



Elias Jansson



Julius Bung Tidblom

Terminology

A brief explanation of central concepts of the report.

Airborne Early Warning & Control (AEW&C) – An AEW&C system is an airborne radar picket system designed to provide long-range air-, sea- and land surveillance and support government- and military forces with real time-information.

Command & Control (C2) system – The C2 system is the computer system used by operators to perform mission-specific tasks, for example perform fighter jet control and identification of targets in the environment.

Plot – A plot is target data obtained by the radar visualized as a small dot on the situational display. Plots provide 2-D position information of targets. The plot extractor needs to distinguish actual targets from clutter like terrain, waves, and birds.

Track – A track is a target report estimating what plots are from the same target to create a track with estimated position and speed vector information. Tracks are displayed as icons on the situational display. Provides more information than a plot, for example speed, history (positions of previous plots) and predict vector (based on speed and course).

Fighter Controller – One of the roles in the AEW&C operator team. One of the main tasks of the FC is to guide assets such as fighter jets safely towards certain objectives.

Surveillance Operator – One of the roles in the AEW&C operator team. The SO's tasks involve managing the sensors of the AEW&C and, depending on the mission type, also identify targets.

Cognitive ergonomics – Cognitive ergonomics is a subset of human factors and ergonomics. Focuses on how well the use of a product matches the cognitive capabilities of users. Applies knowledge of human perception, mental processing, and memory.

Mental model – A mental model is how someone understands the world, an internal representation of reality. Humans use mental models to simplify complexity. Mental models play an important role in cognition, reasoning, and decision-making.

Contents

| | |
|--|----|
| 1 Introduction..... | 1 |
| 1.1 Background..... | 2 |
| 1.2 Aim and Objective..... | 2 |
| 1.3 Aim..... | 2 |
| 1.4 Objective..... | 2 |
| 1.5 Research questions..... | 3 |
| 1.6 Deliverables..... | 3 |
| 1.7 Limitations..... | 3 |
| 1.8 Report structure..... | 4 |
| 2 Theory..... | 5 |
| 2.1 AI..... | 6 |
| 2.2 Situation Awareness..... | 7 |
| 2.3 System theory..... | 8 |
| 2.4 Research Through Design..... | 9 |
| 3 Related Work..... | 11 |
| 3.1 AI in military applications..... | 12 |
| 3.2 Explainable AI..... | 12 |
| 3.3 AI in relation to Situation Awareness..... | 12 |
| 4 Process..... | 13 |
| 4.1 Motivation of Process..... | 14 |
| 4.2 The Process..... | 14 |
| 4.3 Pre-Study..... | 15 |
| 4.4 Concept Development..... | 16 |
| 4.5 AI Guidelines phase 1 & 2..... | 17 |
| 5 Application of Methods..... | 18 |
| 5.1 Education in the System..... | 19 |
| 5.2 System image..... | 19 |
| 5.3 Literature studies AI..... | 19 |
| 5.4 Literature studies cognitive ergonomics..... | 19 |
| 5.5 Interviews..... | 19 |
| 5.6 Use Profiles..... | 20 |
| 5.7 Activities in the MTS..... | 20 |
| 5.8 Affinity Diagrams..... | 22 |
| 5.9 Hierarchical Task Analysis..... | 22 |
| 5.10 Generic Task Specification..... | 23 |
| 5.11 Ideation..... | 23 |
| 5.12 Creation of Concepts..... | 24 |
| 5.13 Evaluation of Concepts..... | 24 |
| 5.14 Connections between methods..... | 26 |

| | |
|---|-----|
| 6 System Analysis..... | 27 |
| 6.1 Defining Systems and Environment | 28 |
| 6.2 Environment: The AEW&C environment | 28 |
| 6.3 The whole system: The Joint Military Operation | 29 |
| 6.4 Subsystem 1: The AEW&C System | 30 |
| 6.5 Subsystem 2: The Operator Team..... | 33 |
| 6.6 Subsystem 3: The AEW&C Operator..... | 34 |
| 6.7 Chapter summary: System Analysis | 37 |
| 7 User Studies | 38 |
| 7.1 Findings from Interviews..... | 39 |
| 7.2 Findings from Activities in the MTS | 45 |
| 7.3 Hierarchical Task Analysis | 47 |
| 7.4 Chapter summary: Findings from the MTS activities | 64 |
| 8 Concept Development..... | 65 |
| 8.1 Ideation | 66 |
| 8.2 Application of guidelines in concept development..... | 66 |
| 8.3 Concept Creation | 67 |
| 9 Final Concepts..... | 69 |
| 9.1 Final Concepts for Surveillance Operator..... | 70 |
| 9.2 Final Concepts for Fighter Controller..... | 77 |
| 9.3 Chapter Summary: Final concepts | 82 |
| 10 Evaluation | 83 |
| 10.1 Evaluation of Concepts | 84 |
| 10.2 Summary of Concept Evaluation | 87 |
| 10.3 Evaluation of AI Guidelines | 88 |
| 11 AI Guidelines | 91 |
| 11.1 Planning | 93 |
| 11.2 Designing | 98 |
| 11.3 Evaluating | 104 |
| 11.4 Chapter Summary: AI Guidelines..... | 108 |
| 12 Discussion | 110 |
| 12.1 Recommendations when Developing AI Systems | 111 |
| 12.2 Ethics: AI in Military Operations..... | 111 |
| 12.3 Application of the Results..... | 113 |
| 12.4 Methods | 114 |
| 12.5 Further Work and Improvements in Process..... | 115 |
| 13 Conclusion | 116 |
| 13.1 When to Implement AI | 117 |
| 13.2 How to Implement AI | 117 |
| 14 References..... | 122 |
| 15 Appendices..... | 124 |

1

Introduction

In this chapter the background to this study is presented along with the aim and objective, research questions, deliverables, and limitations.

1.1 Background

In the Airborne Early Warning & Control (AEW&C) system, the operator needs to be vigilant, perform multiple tasks, use communication tools, and simultaneously handle different sources of information during a mission. The main objective for the operators is to receive, interpret and distribute information and data provided by the built-in sensor suite of the AEW&C airplane and the Command & Control (C2) system to create a Recognized Air and Sea picture (RASP). The information is then used in different mission scenarios to execute tasks (sometimes role-specific) for example to perform fighter jet control and to detect, report and act upon detected anomalies. The increasing technological sophistication regarding the rapidly growing adoption of Artificial Intelligence (AI) and Machine Learning offers new possibilities for an AEW&C operator to obtain Situational Awareness (SA). At the same time, the increased capabilities of the C2 system and new threats in the AEW&C environment such as unmanned aerial vehicles, makes their tasks more extensive, complicated, and complex than ever before. Thus, the operators handle large amounts of information and data, which at times, leads to a high mental workload that could affect the human capabilities in terms of SA negatively, which affect the decision-making process. The operators are acting in a domain associated with high risks and potentially devastating impacts of faulty decisions. Therefore, the importance of human factors research is crucial in terms of appropriate visualization and presentation of information and data to make it explainable, understandable, and interpretable by the operators, facilitating SA and decision making.

1.2 Aim and Objective

In this chapter the aim of the study is presented along with the objective, research questions and intended deliverables.

1.3 Aim

The aim of this study is to explore and analyse factors that facilitate Situation Awareness (SA) for AEW&C operators; specifically the Fighter Controller (FC) and Surveillance Operator (SO), find opportunities for *when* AI functions could be beneficial and find *how* AI functionality should be implemented to facilitate AEW&C operators' SA.

Opportunities will be identified for when AI can be used to improve SA and guidelines of how AI should be implemented will be created. In addition to the findings throughout the process, concepts will be designed to exemplify and evaluate the guidelines.

1.4 Objective

The main objective is to create guidelines for when to implement AI and how to design AI in an AEW&C context to help support the operators' decision-making process with improved SA. Due to the complexity of the systems the SA analysis will be limited to a couple of identified tasks for the SO and FC role.

1.5 Research questions

The two research questions research questions of the study are:

- **When** should AI be implemented to improve AEW&C operators' situation awareness?
- **How** should AI be implemented to improve AEW&C operators' situation awareness?

1.6 Deliverables

Related to the research questions, the deliverables of the study are:

- Visual concept(s) suggesting when and how to implement AI to improve the Situation Awareness of AEW&C operators.
- Guidelines of how to implement AI to improve the Situation Awareness of AEW&C operators.

1.7 Limitations

To clarify, the study was limited to the following:

- **Limited accessibility to a diverse set of users:** Due to the specific context, access to a diverse set of users were limited. All users in the study come from Sweden and work within the same organization.
- **No considerations to functional or cognitive variations:** Due to high demands on the users in the specific context of the study, functional or cognitive variations rarely occur. No considerations were taken for such users.
- **No accessibility to the actual context:** Partly due to the limited accessibility to the actual context in general, and partly due to the restrictions due to Covid-19, the actual context was never visited. The study was limited to a simulated context, the Mission Training System (MTS).

1.8 Report structure

The report is structured as follows:

Chapter 1: Introduction – Introduces the background and motivation for the study, as well as aim, objective, research questions, deliverables, and limitations.

Chapter 2: Theory – Describes the theory and concepts used in the study.

Chapter 3: Related Work – Presents interesting findings from work related to the study.

Chapter 4: Process – Describes the process that was used in the study, and why it was designed the way it was.

Chapter 5: Application of Methods – Motivates how the methods were selected and describes how they were applied in the study and related to the process.

Chapter 6: System Analysis – Presents the findings from the System Analysis- an analysis conducted to understand the context in which the AEW&C operators work.

Chapter 7: User Studies – Presents the findings from the user studies, which were divided in two main parts, interviews, and an operative simulation.

Chapter 8: Concept Development – Describes the process of developing concepts and how the concept development related to the evaluations and development of guidelines.

Chapter 9: Final Concepts – Presents the eight final concepts and a justification of their utility.

Chapter 10: Evaluation – Presents the findings from evaluation of concepts and the results from evaluation of the first set of guidelines.

Chapter 11: AI Guidelines – Presents the guidelines developed during the study, which were the main deliverable. The guidelines are divided into three categories: planning, designing, and evaluating.

Chapter 12: Discussion – Discusses the study in terms of: Recommendations when developing AI systems, Ethics: AI in military operations, Application of the results, Methods, and Further work and improvements in process.

Chapter 13: Conclusion – This chapter concludes the findings of the study.

2

Theory

This chapter presents theories relevant to the study and how they were applied.

2.1 AI

Even though Artificial intelligence (AI) is a widely used term, it is challenging to define. AI means different things to different people. Even among AI researchers there seems to be no consensus in the definition of AI. Reasons why the term is hard to define could be because the concept of intelligence itself is vague, and that the concept of AI changes over time. What was considered AI 50 years ago is not considered AI today (The University of Helsinki, 2021). The European Commission (2019) aims to clarify the term AI to avoid misunderstandings and to create a common ground. They propose the following definition:

“Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions. As a scientific discipline, AI includes several approaches and techniques, such as machine learning (of which deep learning and reinforcement learning are specific examples), machine reasoning (which includes planning, scheduling, knowledge representation and reasoning, search, and optimization), and robotics (which includes control, perception, sensors and actuators, as well as the integration of all other techniques into cyber-physical systems).”

The course Elements of AI was created with the aim to teach as many people as possible about AI and over 650 000 students have signed up for the course since the start in 2018 (The University of Helsinki, 2021). They suggest using properties that are characteristic to AI as a definition. The following properties is suggested:

“Autonomy. The ability to perform tasks in complex environments without constant guidance by a user.

Adaptivity. The ability to improve performance by learning from experience.”

To further clarify what AI is, one can look at typical applications of AI and draw conclusions from this. Some commonly accepted AI applications are self-driving cars, content recommendations, image recognition, speech recognition etc.

Based on the definitions and examples of applications of AI it was decided that when AI is mentioned in the study it refers to the study’s interpreted definition of AI:

AI is a human-made computer system with the ability to collect data (input) from the environment, interpret the collected data, act (outcome) to achieve a specific goal, learn from how well the goal was achieved, and adapt accordingly.

In the case of the study, the human-made computer system is the C2 system. The data (input) is collected by the built-in sensor suite of the AEW&C and from human interactions with the system. The interpretation is made by built in AI models in the C2 system. Goal achievement is based on how well the outcome corresponds to the specific goal. The AI can, but does not have to be, a continuous learner when in use. It is still considered an AI system if it has, at some point, been trained on a dataset. Examples of applications can be abnormality detection, situation assessment, predictions, and image recognition.

2.2 Situation Awareness

Situation Awareness (SA) is a commonly used term in operational situations such as in the AEW&C context. The term originates from the military pilot context (Endsley, 2016). The study refers to Endsley's (2016) definition and levels of SA. The formal definition by Endsley is:

“...the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future”
(Endsley, 1988).

What the definition says is basically that SA is being aware of the situation and understanding what the information means in terms of goals now and in the future. Endsley states that SA is a driving component in decision making and performance for operators in complex and dynamic systems. Making systems interfaces that facilitate SA will support the operators' cognitive processes and their ability to make better decisions.

Endsley (2016) categorises SA into three different levels, perception, comprehension, and projection:

- Level 1 SA: Perception of elements in the environment.
- Level 2 SA: Comprehension of the current situation.
- Level 3 SA: Projection of future status.

To achieve level 1 SA, one needs to be able to perceive what is happening in the environment. Perception may come from each sensory modality, or a combination. Errors related to level 1 SA mainly relate to not receiving the information (for example losing radio reception), forgetting the information (for example when switching attention), or not detecting the information (for example not looking at the right display).

To achieve level 2 SA, one needs to understand what the perceived information implies related to goals or objectives. This can be fairly demanding and requires experience and a good mental model. Errors are typically due to not having prior knowledge in new situations, like getting information that an airplane flies at 10000ft, but not understanding whether it is high or low.

To achieve level 3 SA, one needs to be able to predict how the current situation will develop in the near future. Having level 2 SA is a prerequisite for achieving level 3 SA and requires even better understanding of the domain and having a well-developed mental model. It also requires much mental effort. Level 3 SA facilitates proactivity, avoiding undesirable situations and making the operators more prepared when events occur. Errors related to level 3 SA are often due to high workload, insufficient experience, or cognitive biases (like over projecting current trends).

Some aspects that could impair SA according to Endsley (2016) are related to: attention, memory, workload and other stressors, information overload, adding complexity to systems, mental models, and out-of-the-loop syndrome.

The approach to SA in this study is to conduct user studies to identify situations where the AEW&C operators SA is impaired, due to for example high workload. After identifying those situations, concepts using AI will be created that can facilitate the operators SA, for example by detecting abnormalities (Level 1 SA), assess situations (Level 2 SA) and make predictions (Level 3 SA). The hope is that AI can improve SA and assist the operator's decision making and performance.

2.3 System theory

From a general perspective, a system is defined as: A set of things working together as parts of a mechanism or an interconnecting network; a complex whole (Oxford University Press, 2020). The system in which the AEW&C-operator works can be considered as a human-machine system, defined as A combination of one or more human beings and one or more physical components interacting to bring about, from given inputs, some desired output. (Sanders & McCormick, 1993). The definition of a human-machine system implies that the human and machine do not only work together, but they do also so to achieve a desired output, or a system goal. For a complex and technical system, like the one of the study, systems theory can be applied. A key concept of systems theory is that a system is more than the sum of its parts, when working together they could achieve things that would not be possible if each individual part worked alone. Therefore, instead of looking at parts of a system in isolation, one should look at the bigger picture, the system as a whole (Flood & Carson, 1993).

When identifying systems and environments, one must first choose what definition of a system to work according to. The definition used in this study is by Flood & Carson (1993) who claims that a system simply is *an assembly of elements related in an organized whole*. Figure 1 shows an example of a system with relationships between elements, a system boundary and input and outputs. Elements are the parts of the system that contribute to achieving the system goal. Relationships are the flows of materials, information, or energy between the elements. The system boundary distinguishes the system from the environment and other systems. What defines the environment of the system is that, even though it still affects or is affected by the system, there is only a one-way relationship, input, or output (Flood & Carson, 1993).

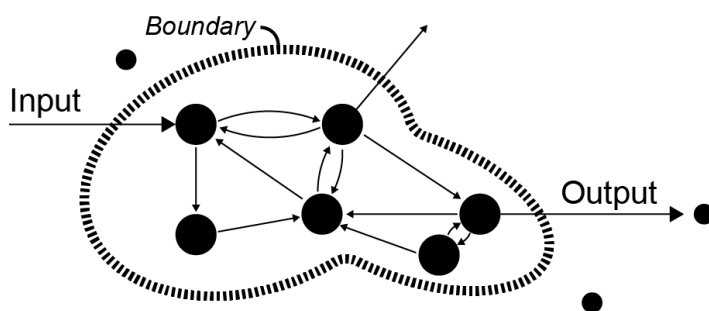


Figure 1. Visualization of a system. Consisting of elements (dots) with relationships (arrows), system boundary, input from the environment and output to the environment.

An important part when identifying systems is to define the appropriate level of resolution for the system-in-focus, where to draw the system boundary. The challenge lies in finding the balance between looking at the system as a whole and understanding the system at a deeper and more detailed level. The system-in-focus is the system the study will put most resources on, in this case the system of the AEW&C operator and their workstation. One approach when drawing the systems boundaries for the system-in-focus is to look at concentration of relationships between elements in relation to the core elements of the system (Flood & Carson, 1993).

In this study the systems will be described drawing inspiration from the First-generation Activity Theory Model (Engeström, 1999) and the system composition proposed by Flood & Carson (1993). The model created to describe systems presented below (See Figure 2).

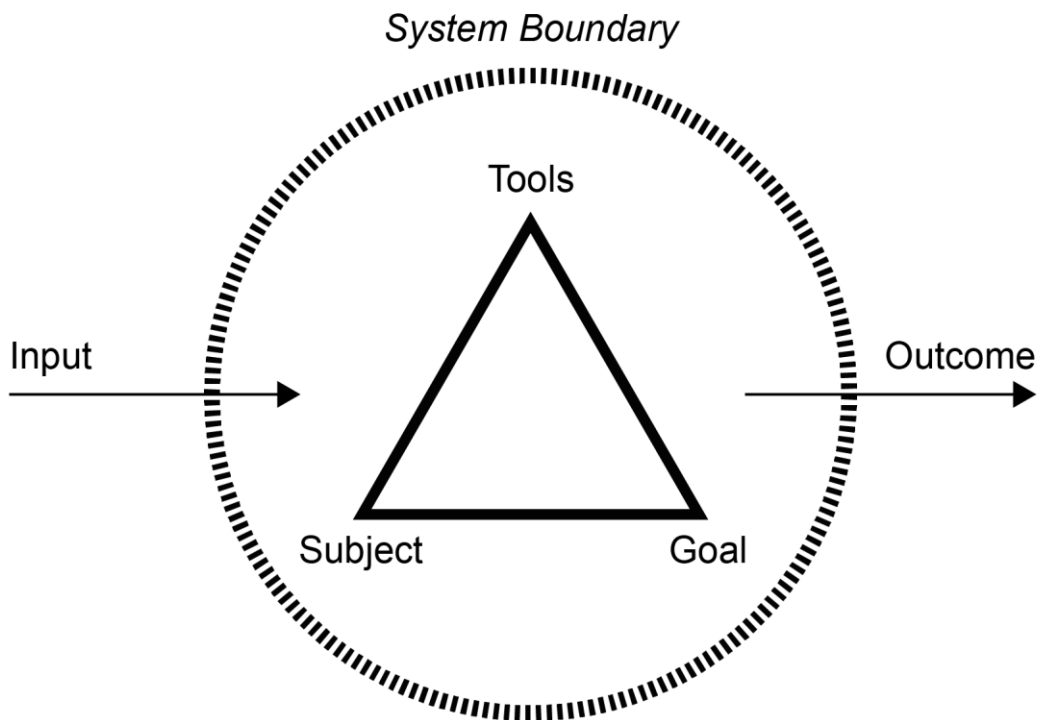


Figure 2. The model used to describe systems in the study.

The subject is considered the human, or thinking element, of the system, which may be an individual user, a team or a subsystem depending on the magnitude and composition of the system. Tools are elements that the subject interacts with to achieve the goal. The goal is the system's goal and the main purpose for the existence of the system. The outcome is the result, which aims to be compliant to the system's goal. The outcome is a direct consequence of how well the subject, tools, and goals match. If there are mismatches, for example if the tools are not suitable to achieve the goal or if the goal is beyond the capabilities of the subject, the outcome will not be in line with the goal. The outcome may affect the environment or another system. The inputs are elements from the environment that affects the system.

2.4 Research Through Design

The objective of the study was to create guidelines for when and how to implement AI to facilitate Situation Awareness (SA) for the AEW&C operators: Fighter Controller (FC) and Surveillance Operator (SO). To create a process that produces these outcomes several unknowns were required to be researched. First, a sufficient understanding of the system needed to be established. Second, the user's needs to achieve a sufficient level of SA in order to reach the system goals needed to be determined. Third, issues within the current system that hampers the operators' ability to achieve SA needed to be discovered. Finally, AI considerations that work within the preconditions determined in the steps taken above need to be researched. This means that the process should facilitate the achievement of AI guidelines that work within the current system and that the guidelines facilitate selection of functionality that aids operators' SA.

A process that achieves the outcomes described above is *research through design* (Zimmerman, Forlizzi & Evenson, 2007), (See Figure 3). An outcome of the study is artifacts, represented in the

study as design concepts and guidelines. The design concepts are created using multiple sources of information from the research community, i.e., engineering literature, anthropology, and behavioural science. In the study of engineering the technological opportunities and challenges of AI are examined. Through anthropology an understanding of how humans perceive, comprehend, and predict is reached. From the field of behavioural sciences theories and models on how to test and create concepts and guidelines are determined. This study takes the perspective of interaction designers using information from the different sources to create guidelines represented through concepts that can be tested and iterated using behavioural science. The concepts and guidelines permeate the research practise barrier to the practice community (Zimmerman, Forlizzi & Evensson, 2007).

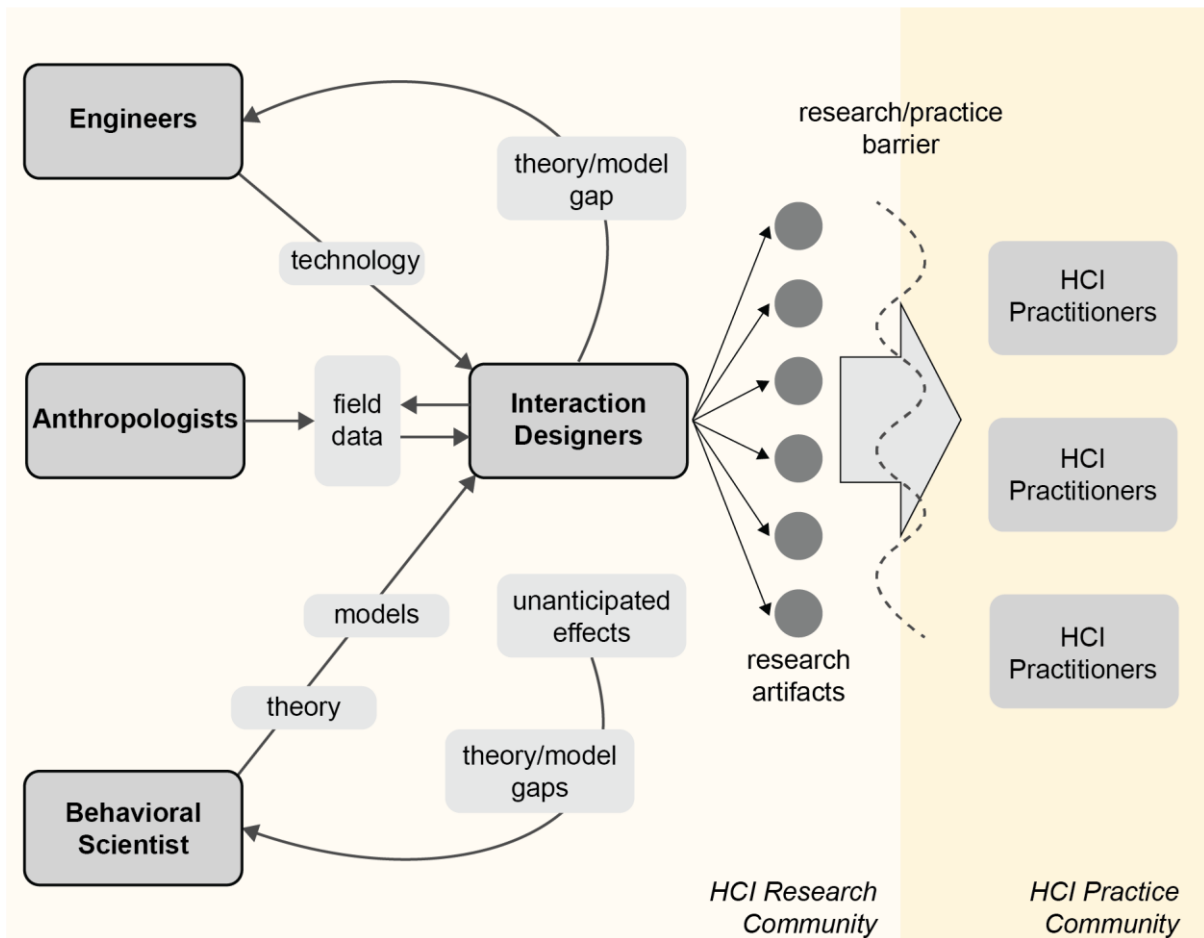


Figure 3. A model of the *research through design* approach by Zimmerman, Forlizzi & Evensson (2007).

3

Related Work

In this chapter previous works are presented in the following order:
AI in military applications, Explainable AI, and AI in relation to situation awareness.

3.1 AI in military applications

In several industries AI has been promising, and the military industry is no exception. AI may for example be helpful in aiding decisions, help perceive threats and learn to act autonomously (Gunning and Aha, 2019). AI based systems can analyse a much larger amount of data than the human team can, along with predicting future events and filter out irrelevant information (van den Bosch & Bronkhorst, 2018). Military operations are often characterised by complex problem-solving using information derived from sensors such as radars and other technologies. A side effect of the growth in sensor technology is the amount of information required to be analysed (van den Bosch & Bronkhorst, 2018). Bosch & Bronkhorst (2018) claims that the increase in information also risks endangering situation awareness in operators because of high workload, thus hampering decision making. In addition, AI algorithms are usually opaque which complicates human operators' abilities to understand and act on AI suggestions effectively (Gunning & Aha, 2019).

3.2 Explainable AI

Gunning & Aha (2019) argue the importance of Explainable AI (XAI) in military applications of AI. The advent of XAI is a consequence of the inherent complexity that AI systems based on Machine Learning (ML) and Deep Learning (DL) have. Traditional AI systems were based on logic, making the AI-inferences easier to understand. ML and DL are powerful models, but paradoxically, the more powerful and performing the AI system is, the harder it becomes to explain its inferences (Gunning & Aha, 2019). Developing AI systems provide designers with two main options, either they select a model that is inherently explainable (logic-based) with a lower performance, or they select a model with a high performance but low explainability and develop an XAI model around it (Gunning & Aha, 2019).

3.3 AI in relation to Situation Awareness

When an AI system is introduced to a system the whole system's performance should be considered. Sanneman & Shah (2020) argue that human-AI team performance should be the main goal of XAI, implying that just focusing on the AI performance in isolation is insufficient. XAI that focuses on facilitating Situation Awareness (SA) through explanation of AI behaviour is an important building block in a human's global SA (Sanneman & Shah, 2020). Sanneman & Shah (2020) propose integrating the three-level framework (Endsley, 2016) for SA when designing XAI systems. Miller (2019) argues that XAI systems too often are developed using the intuitions of the system's developers and calls for a thorough human factors study of what constitutes a good explanation.

4

Process

In this chapter the process used in the study will be presented. First the selection of the process is motivated and then the process is presented by how it relates to the *research through design* framework (Zimmerman, Forlizzi & Evensson, 2007).

4.1 Motivation of Process

The process of the study was designed to answer the two research questions:

When should AI be implemented to improve AEW&C operators' situation awareness?

How should AI be implemented to improve AEW&C operators' situation awareness?

The study aimed to approach the research questions by producing concrete answers in the form of visual concepts mainly answering **when**, and guidelines mainly answering **how**. To answer the questions and produce the deliverables a process was designed inspired by *research through design* (Zimmerman, Forlizzi & Evensson, 2007). What this implied was that a process was designed that adapted multiple research areas to facilitate the creation of guidelines and concepts simultaneously. That means that the guidelines are created to facilitate the creation of concepts. The concepts in their turn evaluate the usefulness of the guidelines they were based on. The guidelines are improved, which in turn facilitate the creation of improved concepts applying the improved guidelines, and so on. This exchange of feedback between guidelines and concepts are the essence of the process designed in the study. The process, and how it relates to the different areas of *research through design* is presented in the following chapter.

4.2 The Process

Figure 4 displays the activities in the study divided into two tracks that were designed to answer the two research questions. The process consisted of four major phases allocated to the two tracks. These phases were: Pre-study, Concept Development and two AI Guidelines phases. The pre-study phase consisted of system analysis and user studies. The concept development phase included ideation, concept creation and evaluation. In the top track, the pre-study and concept development phases mainly concerned the *when*-part of the research questions. The AI guidelines phases consisted of two iterations of AI guidelines which mainly concerned the *how*-part of the research questions. The process was designed to allow some interaction between the two tracks displayed by the arrows in figure 4. The first phase of the AI-guidelines informed how the concepts should be created. By using the guidelines in the concept creation, the utility for design of the guidelines were tested. The two evaluation phases both evaluated the guidelines and the concepts.

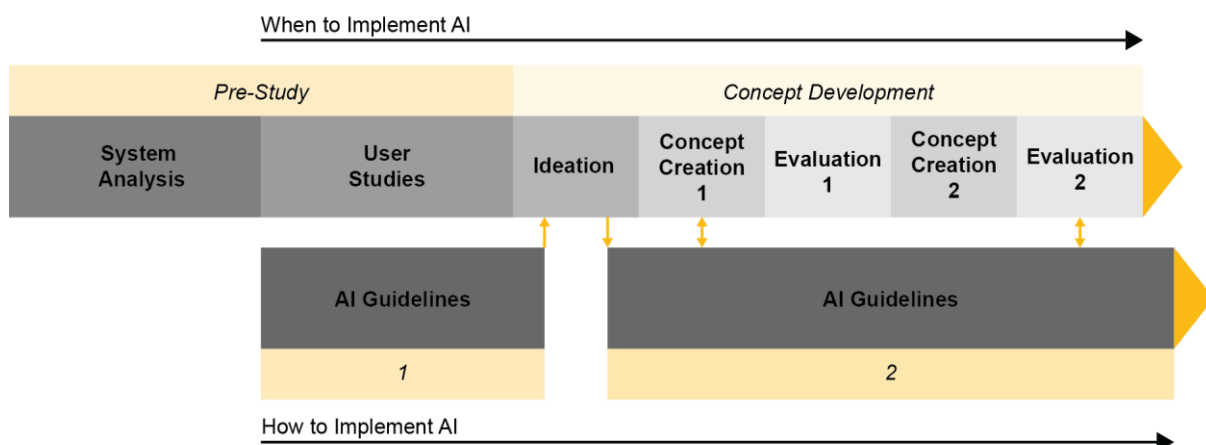


Figure 4. Visualization of the process of the study.

4.3 Pre-Study

Figure 5 shows what aspects the pre-study covered regarding the *research through design* framework. Using methods such as use profiles, NASA-TLX and concepts such as Situation Awareness (SA), the behavioural aspects of the research of the system is covered. To understand what the properties that users act in; a system analysis was performed.

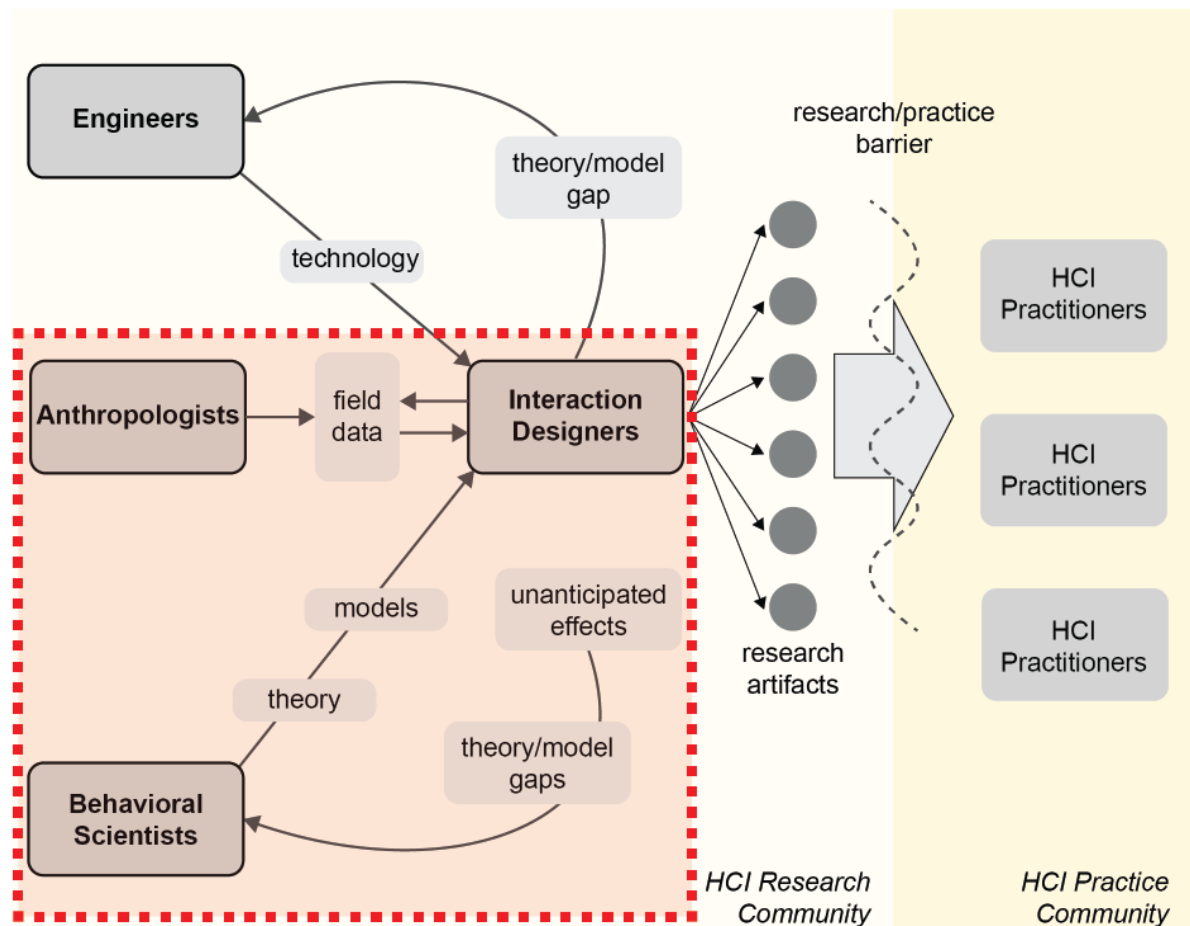


Figure 5. The pre-study phase in relation to the design through research framework.

4.3.1 System analysis

In the system analysis, the system is defined and described in terms of goals, subjects, outcomes, and relationships. Other deliverables from the phase were system images for the system and its three subsystems. The system analysis created a foundation for the *when*-track of the study.

4.3.2 User Studies

The user studies included three major activities that helped to create a basis for the ideation by finding opportunities for when to implement AI. The three activities were interviews, operative simulation in the Mission Training System (MTS, a training system that has similar functionality to the real Airborne Early Warning & Control (AEW&C) system and task analysis using the concept of situation awareness). The interviews provided findings for the use profile, teamwork, errors, and risks, and understanding of work conditions and the users' thoughts on AI.

4.4 Concept Development

Figure 6 shows what aspects the concept development phase covered regarding the research through design framework.

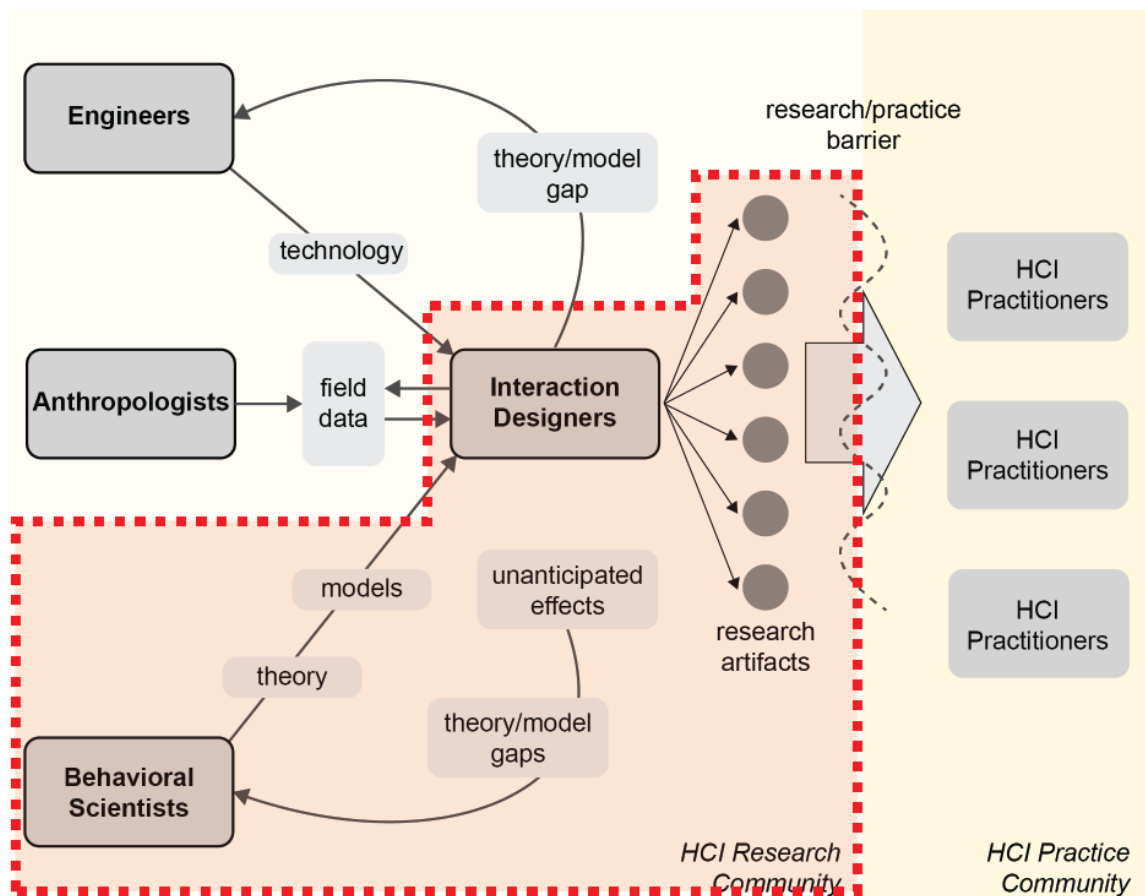


Figure 6. The concept development phase in relation to the design through research framework.

In this phase both the behavioural aspects of the study as well as the creation of artifacts was covered. The ideation and concept creation are connected to the creation of artifacts. The evaluation of the concepts (which were created based on AI guidelines) covered the behavioural side of study once again.

4.4.1 Ideation

In the ideation activity, data and analysis from the previous activities were used to find solutions to issues that the operators have.

4.4.2 Concept creation 1

In this activity the ideas generated in the ideation activities were combined into eight concepts. The concepts were improved using the first iteration of the AI guidelines. By using the guidelines, it could also be seen which guidelines were easy or difficult to use.

4.4.3 Evaluation 1

In the first evaluation, the concepts were evaluated by presenting the concepts to engineers and operators at the company. The guidelines were also evaluated by seeing how well the intended effect that the guidelines should have had in practice.

4.4.4 Concept creation 2

The evaluated concepts in Concept creation 1 were improved using the feedback from the presentation. The concepts were also improved by increasing the fidelity of the concepts. Both by more clearly trying to achieve the guidelines and by making the concepts look more like a real function.

4.4.5 Evaluation 2

The concepts were evaluated by presenting the concepts to six operators. The operators got to say what they liked and disliked for each of the eight concepts.

4.5 AI Guidelines phase 1 & 2

Figure 7 shows what aspects the AI guidelines covered regarding the research through design framework.

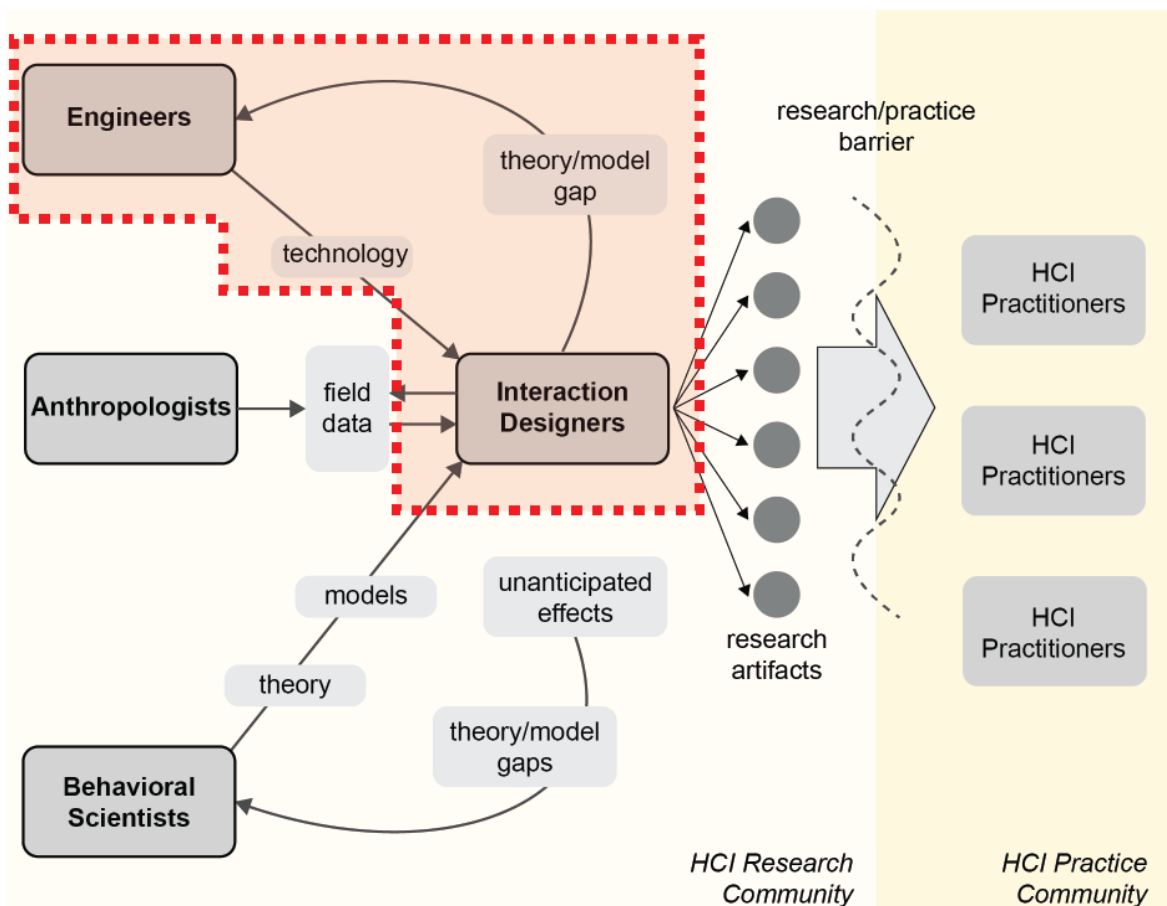


Figure 7. The Ai guidelines phases in relation to the design through research framework.

The creation of AI guidelines was based on engineering research of AI. This research created a basis for finding technological opportunities for how to implement AI, but also partly when to implement AI functionality. The first iteration of the guidelines was created through a literature study. The outcome of the literature studies were multiple quotes that were categorised using affinity diagrams. The categories were written as guidelines and the quotes from the literature motivated the guidelines. The second iteration of the guidelines were informed by the concept creation and the evaluation sessions.

5

Application of Methods

Methods were selected to reach the process goals, i.e., methods that facilitate the creation of the deliverables of the objectives in the study. This meant finding methods from the research *through design* research areas of; anthropology, behavioural sciences and methods that summarizes the necessary data about the technological opportunities of AI from the engineering field.

The motivation for each of the methods selected is presented in the list of methods. They are described based on how they were applied in the study. Finally, the connections between methods in relation to the process is presented.

5.1 Education in the System

To understand the system an education of the system and its parts was requested and attended utilizing operators and personnel knowledgeable in the system. The education involved understanding the system on different levels. The education also contained tactical aspects such as how different types of missions are conducted; how different operator roles and tasks are structured in different missions and what types of functionalities the system has.

5.2 System image

Based on the education of the system, system images were created. System images were created on four different levels, where the lowest level is the focus of the study, *The AEW&C Operator*.

5.3 Literature studies AI

Literature studies searching and filtering articles regarding AI and Situation Awareness (SA) as well as cognitive ergonomics. Literature from the course syllabus in cognitive ergonomics was used (Chalmers Student Portal, 2016). AI literature was found using search terms such as “AI situation awareness”, “AI explainability”, “AI military”, and asking researchers at Chalmers University of Technology.

5.4 Literature studies cognitive ergonomics

For a reasonable collection of methods, the literature from the cognitive ergonomics PM (Chalmers Student Portal, 2016) were consulted. Methods were selected on their merits in similar projects and on the recommendations within the literature explaining the methods.

5.5 Interviews

Two interviews were conducted with Airborne Early Warning & Control (AEW&C) operators with experience in the Swedish military. The interviewees had two different roles, surveillance operator and fighter controller. However, they had knowledge and experience of some of the tasks of the other roles. Both operators had recent operative experience. The interviews were approximately one hour long each. The transcripts of the interviews were sent to the operators after the interview and completed by them. The purpose of the interviews was partly to create a use profile of an AEW&C operator, as well as identifying differences between the roles. Questions that were used to create the use profiles were based on Janhager (2005). In addition to creating use profiles the interview featured questions that helped understand which tasks were relevant for the two roles. The questions were also designed to find critical tasks for the system goals, common mistakes, possible improvements.

Additional interviews were also conducted with engineers and operators to find out more about errors and risks, working conditions and to gain insights about operators' thoughts about AI. The additional interviews were more unstructured than the two operator interviews.

In total, five interviews were conducted before the Mission Training System (MTS) operative simulation (described in *Activities in the MTS*).

5.6 Use Profiles

A use profile is a method where different aspects regarding the user's relation to the product or service is assessed. These aspects are categorised into; use experience, influence and responsibility of use, emotional relationship to the product and degree of interaction with the product. Within each of these categories there are sub-categories such as length of use, frequency of use etc. For each of the sub-categories a value is added pertaining to what degree the category the user adheres to that category from low to high, denoted as "degree of performance" (Janhager, 2005).

Use profiles were created to find differences in needs for the two different roles: fighter controller and surveillance operator. For each of the categories for each of the two operator types an explanation based on the operator interviews were written. The use profiles were subsequently combined into one overarching use profile for both roles where differences between the two roles were explained.

5.7 Activities in the MTS

The Mission Training System (MTS) was utilized to perform operative simulations, observations and interviews concerning Situation Awareness (SA) and NASA-TLX questions.

5.7.1 Operative Simulation

To get a deeper operative knowledge than achieved from previous user research, scenarios were used in an operative simulation and to find how the operators reach SA. A scenario is a fictional, manually created representation of tracks and plots anchored in reality, used when educating new operators. Moreover, the scenario was used to find critical tasks where AI suggestions could be helpful for the two operator types. The scenario was staged in a system called *Mission Training System (MTS)* usually used for training customers' future operators. The system is a representation of the operator room of the AEW&C with a sufficient degree of functionality similar to the one of the actual system. The MTS has five operator workstations with a functionality to control own assets from a separate room. The scenario used was a previously created scenario where fictional tracks had been created to simulate a real critical mission. Two different versions of the scenario were used, one base scenario and the same scenario but with different abnormalities that the operators were not aware of. The abnormalities were based on the guidance of an operator who was not present in the operative simulation. The abnormalities were discussed to represent reasonable events that could happen in operations which are difficult to detect. The first abnormality was rib boats coming from a larger ship standing still outside an island. The rib boats were approaching the island at high speed. The second abnormality was suspicious land activity near a harbour that was supposed to simulate smuggling. The simulation consisted of three consecutive sessions. For each session variations were added. In the first session the base scenario was used, in the second scenario the abnormalities were added and in the third scenario the crew member roles were rotated.

After each session interviews were conducted with the fighter controller and the surveillance operator (described below in: *SA-questions* and *NASA-TLX*).

The sessions were initiated with a mission brief. In the brief the goal of the mission was described including what to look for, what areas to focus on and what type of radar areas to create. The brief was conducted by one of the operators and had a duration of about 30 minutes. Once the brief was finished, operators and the fighter pilot took their seats in the MTS. All the sessions were loosely structured into four parts, which were based on the structure of a real mission. Part 1 was to establish a situation picture which includes starting the sensors, creating areas, and starting links. The

responsibility for these tasks is mainly delegated to the surveillance operator, but since the remaining operators are usually relatively free (during this part of the mission) they assisted in this task. Part one generally took 5-10 minutes. Part 2 included identifying the tracks that appear because of the actions in part 1, which was also performed as a team. Part 2 was intensively focused on for 5-10 minutes but progressed throughout the entirety of the sessions as new tracks and plots appeared. About 30 minutes into the sessions part 3 was initiated when an order was given from the Air Operations Centre (AOC). The order contained information that a specific target needed to be identified by an own fighter jet and where to start the fighter jet from. Part 4 involved the execution of the order in part 3 and progressed for about 30 minutes.

The participants of the scenario were three AEW&C crew members and one fighter pilot. The crew members were trained AEW&C operators. One mission crew commander, one fighter controller and one surveillance operator. The fighter pilot was a SAAB employee with deep knowledge in the technical aspects of the MTS.

All the sessions were audio and video recorded to facilitate following analyses. The audio recordings were taken between the fighter controller and the fighter pilot where most of the communication is present. The video recordings focused on the screens of the fighter controller and the surveillance operator. Two facilitators were responsible for the schedule of the day, preparations and complementing interviews and two additional facilitators were responsible for operation of cameras and aided in taking notes of interesting occurrences.

5.7.2 Observation

Every MTS operative simulation session was observed live and recorded for further analyses. Two observers were present during the sessions, one focused on the tasks of the surveillance operator and the other on the fighter controller. The operators were asked to think out loud when executing tasks as much as possible without jeopardizing the goal of the mission. Time stamps were noted when interesting events occurred which included difficulties in the tasks, errors, missed abnormalities, high workload, stress, and interesting comments from the operators. The notes were also focused on creating a basis for further questions by denoting where the unclear course of events occurred, as well as creating a detailed understanding of the progression of tasks. During the observation questions about these aspects were asked.

5.7.3 Situation Awareness Questions

The questions after each session were based on 8 principles of situation awareness and had the purpose of finding necessary aspects operators need to reach a sufficient level of situation awareness as well as finding deficiencies in the current system (Endsley, 2016). The interviews were conducted on the fighter controller and the surveillance operator separately and lasted for about 10-20 minutes. The interview form was adapted to the current system and can be viewed in Appendix C.

5.7.4 NASA-TLX

Following the situation questions an adapted version of NASA-TLX was conducted separately for each operator.

In NASA-TLX operators' subjective assessment of a task is obtained for six different factors: mental demand, physical demand, temporal demand, performance, effort, and frustration level (Hart & Staveland, 1988). Hart & Staveland (1988) provides a human centred assessment of tasks where workload is viewed as an emergent property of task requirements, the context, operator skills and perceptions. Mental demand involves, among other aspects, the perceptual activity and general

cognitive task demand required. Physical demand involves the physical demand of the task which, in this context, mostly means effects from prolonged static actions of sitting still such as eye strain and back pain. Temporal demand involves time pressure for different tasks. Performance prompts the operators to assess how successful they felt they were for the different tasks and how satisfied they felt. The effort factor combines both physical and mental activity, the operators appreciate the effort required to reach the task goals. Frustration level assesses if the operators feel irritated, insecure etc. about different occurrences during the tasks (Hart & Staveland, 1988).

The six different aspects of workload from NASA-TLX were used as an assessment after each session. NASA-TLX traditionally uses additional ratings for the importance of the different factors which were left out in this study for several reasons. Hart & Staveland (1988) found significant variance in verbal and written operator assessments in their tests and recommended written assessments. Since written assessments were not achievable, due to the participants' work schedule, and to focus on time spent on the sessions to obtain more data from the operative sessions, the ratings were left out. In addition, the discussed variance in the ratings, especially since the operative simulation only featured two operators that were assessed, motivated that the hypothetical result of the ratings would have a low validity in relation to the qualitative data obtained by the discussion obtained through the method.

During the interviews the operators were first introduced to the six different factors that were translated into Swedish. The operators then got to assess which factor that affected their perceived workload during the session the most. Then for each factor the operator got to explain which subtask that affected the factor the most and the least respectively. In total three protocols for each operator were obtained using the method. The questions asked can be viewed in Appendix C.

5.8 Affinity Diagrams

Affinity diagrams were created to sort and find themes in the user studies and for the AI literature review separately. The affinity diagrams for the user studies were created using input from interviews, MTS-activities, function demonstration and education. The diagrams were created using quotes and observations from the different sources concerning issues caused in the system hampering the operators to reach their goals. All the diagrams were then combined into two affinity diagrams for the FC and the SO separately.

The affinity diagram for the AI guidelines were based on quotes from articles concerning AI found in the literature study. The diagram was created using an online based whiteboard software. Similar themes were found, and the quotes were categorised accordingly. In addition, the quotes were categorised based on when in a development process the guideline was relevant, i.e., planning, designing and evaluation.

5.9 Hierarchical Task Analysis

A Hierarchical Task Analysis (HTA) is a structural decomposition of a task; a goal for a task is defined and user actions are specified that are required to reach that goal (Oswalder, Rose & Karlsson, 2015). Hollnagel (2004) argues for the convenience and adaptability of HTA's for task analysis.

HTA's were created for two roles, Fighter Controller (FC) and Surveillance Operator (SO), during simulated missions. Separate HTA's for different roles is also argued by Hollnagel (2004). The purpose of creating the HTA's decompose the structure for the tasks in the two roles and later find where during the progression of a mission AI suggestions may be helpful. The HTA's were created to describe one mission loop that lasts for about one hour and were based on the MTS activities and

previous interviews. The initial HTA's were created with three levels to keep the detail level low. The FC and SO HTA were both expanded for two subtasks, each chosen based on an adapted version of GTS described in the next section. These HTA's employed a higher level of detail to facilitate identification of the tasks crucial to analyse further. This gave the HTA a detail level of up to three levels deeper than the initial HTA's. In total four HTA's were created in addition to the two initial overarching HTA's.

5.10 Generic Task Specification

The *Generic Task Specification* (GTS) is a method to analyse the workload in a system using derivations from fractions of other theories such as Skill-Rule-Knowledge. The model is divided into; task demands, automation levels, mental workload, and physical workload. The aptitude of the method in the study was motivated by the flexibility of the method where aspects may be left out if they are covered using other methods. Or, if the specific aspect is not relevant for the evaluation it can be left out which enables a quick estimation for finding tasks with a high workload, in particular if other completing methods have already been employed (Bligård, 2014). GTS was used to find which subtasks in the HTA's that could be inferred to have a high workload based on the data obtained in previous user studies using the factors in the method. Since the method was not used to quantify workload, but merely find opportunities for AI supporting tools, grading of the aspects was left out.

On the recommendation of Bligård (2014) the most relevant factors for the system were used when assessing the difficulty for each subtask. Relevance of the aspects were assessed by analysing how common the specific issue was in the system using data from previous user studies and how relevant the aspect was in relation to AI. All the different operations/subtasks were evaluated against the following factors: time pressure, information amount, memory demanding, calculation demanding, mission critical information, looking and searching, spatial requirements, high stakes, and risk for errors. The operations that met most of the criteria were two operations for each role.

5.11 Ideation

The ideation of the concepts was performed in multiple iterations. In the first iteration the affinity diagram from the user studies was used to find instances where operators found it difficult to reach their task goals. The issues found were listed and categorised into different themes and rephrased as clearly as issues to solve rather than quotes. All the issues were then assessed regarding what degree the issue could be solved using AI based suggestions. Issues that pertain mainly to usability or, were not relevant to the roles the study focuses on or team dynamics were disregarded. The reduced list of issues was used to generate the first collection of ideas. Ideas of solutions were created using analogue sketches for five minutes for each of the issues in pairs. This resulted in roughly 70 separate ideas with some degree of similarity. All the ideas were discussed and subsequently described in short sentences.

For the second iteration of the idea's similarities between all of them were further specified. This categorisation resulted in 18 different themes containing the 70 ideas. In addition, ideas that could instantly be evaluated as unfeasible or outside the study's scope were disregarded. The result of this stage was a reduced list of ideas categorised based on similarities and the level that the ideas fitted together. This resulted in four groups of ideas per operator role, eight in total, that were precursors to the concepts.

5.12 Creation of Concepts

The creation of concepts was iterated in two steps and were based on the eight ideas described in the previous sections. First after a larger scale presentation and second after an evaluation session with operators.

5.12.1 Use of AI guidelines

Guidelines were used as a foundation in developing the eight ideas into concepts with a higher fidelity. Guidelines that were relevant to the specific idea group were identified and applied to the specific group, meaning that not all guidelines were used for all the concepts but rather the most relevant.

5.12.2 Wireframes

Wireframes on paper templates, depicting a simplified version of a C2 system's interface, were created for each of the eight concepts. The purpose of the wireframes was to create a clear workflow easy to integrate within the current system and easy for operators to follow during presentations of the concepts. Another purpose of the wireframes was to facilitate a discussion about how the concepts worked and detect violations of guidelines and violations against the result from the user research. Approximately five frames per concept were created.

5.12.3 Concept description

The eight concepts were described as a background to what aspects of the operator task goals the concepts help to facilitate as well as the concepts' functionality in chronological order.

5.12.4 Digital concepts

Using the wireframes and concepts descriptions the concepts were visualised in Microsoft PowerPoint (2018). The goal of the visualisations was to create a presentation to evaluate an initial iteration of the concepts to relevant personnel at the company. The Adobe Illustrator (2019) concepts were refined in a higher fidelity showing a more detailed interaction than the previous concepts.

5.13 Evaluation of Concepts

All the concepts were iterated using comments from relevant personnel at the company including AEW&C operators.

5.13.1 Input from concept presentations

Concepts were presented where the functionality and background to the concepts were described. The concepts were presented in two instances. In the first instance the PowerPoint visualisations of the concepts were presented. Comments from the first presentation were used to refine the concepts using illustrator. Especially uncertainties about the functionality, the utility of the concept and the clarity in interactions occur in the concepts were clarified. The concepts were presented exclusively to operators using the refined concepts as visualisations. The concepts were introduced and explained and thereafter operators got to comment about what they liked and disliked about the concepts. Discussions for evaluation of concepts and guidelines were facilitated by asking questions about the guidelines used while creating the concepts.

5.13.2 Evaluation of concepts in relation to guidelines

The evaluation of the concepts was based on the comments from operators. Comments regarding visual aspects were disregarded along with comments not in accordance with the scope of the study. First, all the guidelines that were used when creating each separate concept were listed. Second, the comments operators had for each of the concepts were allocated to which of the guidelines they connected to. Thereby it could be seen how many of the intended guidelines that were achieved or violated, and which guidelines operators did not discern from the visualisations.

5.13.3 Evaluation of guidelines

The guidelines were created and iterated throughout the project. Most of the guidelines were written before the advent of the ideas but were reorganised and rephrased during parallel activities during the study. When creating the concepts, the guidelines were clarified when uncertainties emerged using the guidelines. A subgoal was to minimize the number of guidelines.

In addition to evaluating the concepts the guidelines were evaluated. Guidelines that were mentioned frequently were seen as more important. Guidelines that were not mentioned were either rephrased to increase the clarity of them, combined with other guidelines to facilitate easier selection of guidelines when ideating concepts or removed from the guidelines list.

5.14 Connections between methods

Figure 8 displays how the different methods and processes used in the study relate to each other. In the pre-study phase several methods and processes were employed to understand the users and the system. All the input from the user studies were analysed using affinity diagrams where e.g., main issues in the system were highlighted. NASA-TLX and Situation Awareness questions were used to find factors important from a user's perspective and were subsequently used as factors in the GTS method. The first iteration of HTA's was created using the results from the affinity diagram. These HTA's were expanded using the input from the applied GTS method. Using the updated HTA's and results from the affinity diagram the ideation was initiated. From the ideation concepts were created which were evaluated in two iterations. During the study AI guidelines were derived from AI related literature. The first iteration of the guidelines was created using affinity diagrams of quotes in the literature. These guidelines were evaluated by using them in the creation of concepts and the evaluation of these concepts.

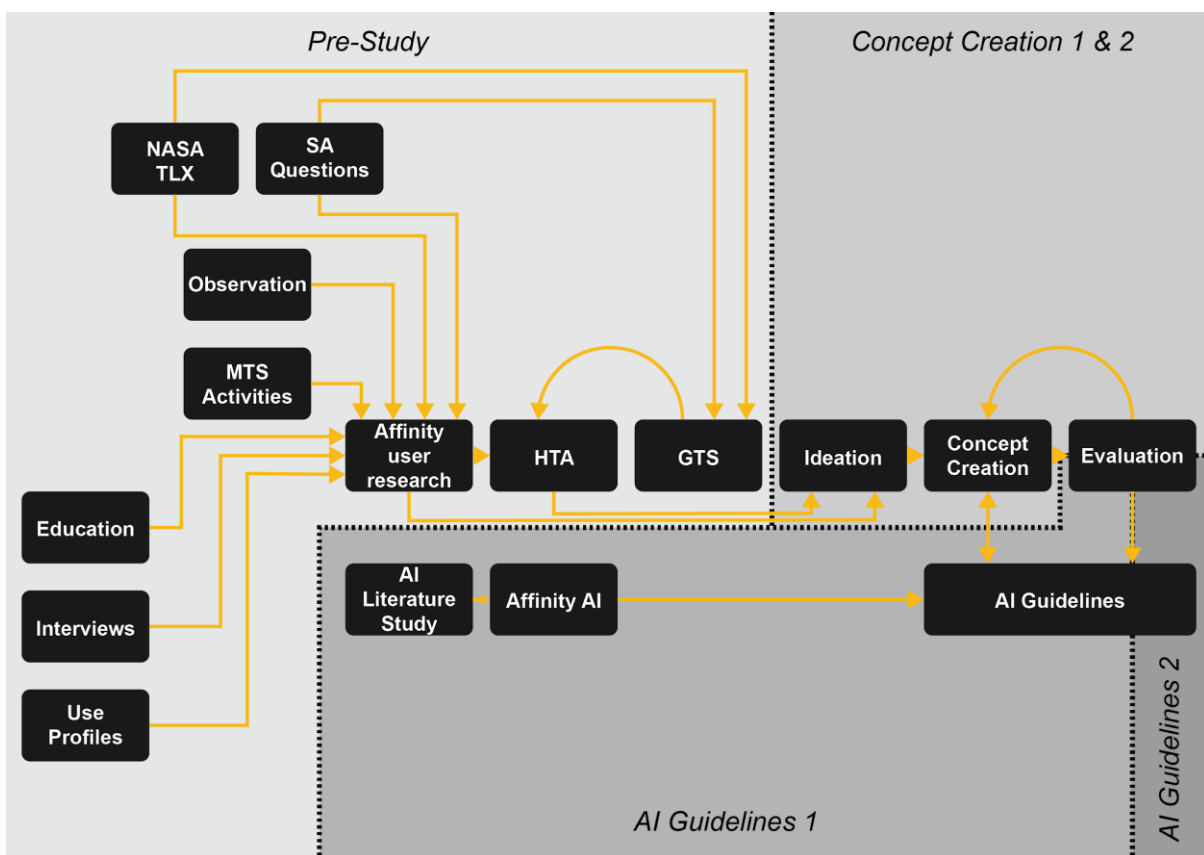


Figure 8: Connection between methods in relation to the process.

6

System Analysis

The focus of the study was to explore and identify when and how AI could be used to improve the Situation Awareness (SA) of operators in the Airborne Early Warning & Control (AEW&C) system. A useful and often necessary starting point when conducting any design project is to establish an understanding of the system. Since the focus of the study was on the operators, the expert users, it was important to understand their role in the human-machine system.

In this chapter, the AEW&C system will be analysed drawing inspiration from systems theory. This implies:

- Identifying systems and environment.
- Describing the systems in terms of goals, subjects, tools, outcomes, and relationships.

6.1 Defining Systems and Environment

The systems analysis started with defining systems. This was done by applying systems theory (see Chapter 2.3 *Systems Theory*). Since the focus of the study was on the AEW&C operators, specifically the Surveillance Operator (SO) and the Fighter Controller (FC), it was natural to start by considering their individual systems, consisting of them and their workstation. The identified systems got the collective name: The AEW&C Operator. The elements with the closest relationships to the SO and FC were the other operators in the operator team and their corresponding workstation. That formed the second identified system: The Operator team. The operator team is in the operator room of an AEW&C. The AEW&C itself was also considered a system, which included the operator team, the pilots, and the airplane with a built-in sensor suite. The AEW&C system was considered a part of an even bigger system, the Joint Military Operation (JMO), which also included other systems like other military assets and the Air Operation Centre (AOC). The JMO system was considered the system of the highest level of the study, the whole system.

When defining the environment, elements were considered that affected or were affected by the AEW&C system. This mainly included target data picked up by the sensor suite.

To conclude, the following environment and systems were identified:

- The AEW&C Environment
- The whole system: The Joint Military Operation
- Subsystem 1: The AEW&C System
- Subsystem 2: The Operator Team
- Subsystem 3: The AEW&C Operator

6.2 Environment: The AEW&C environment

The environment affects the system in the form of inputs. In the Airborne Early Warning & Control (AEW&C) system, the input is the data from targets in the environment. The environment of the AEW&C system can be divided into three areas; air, maritime and land (see Figure 9).

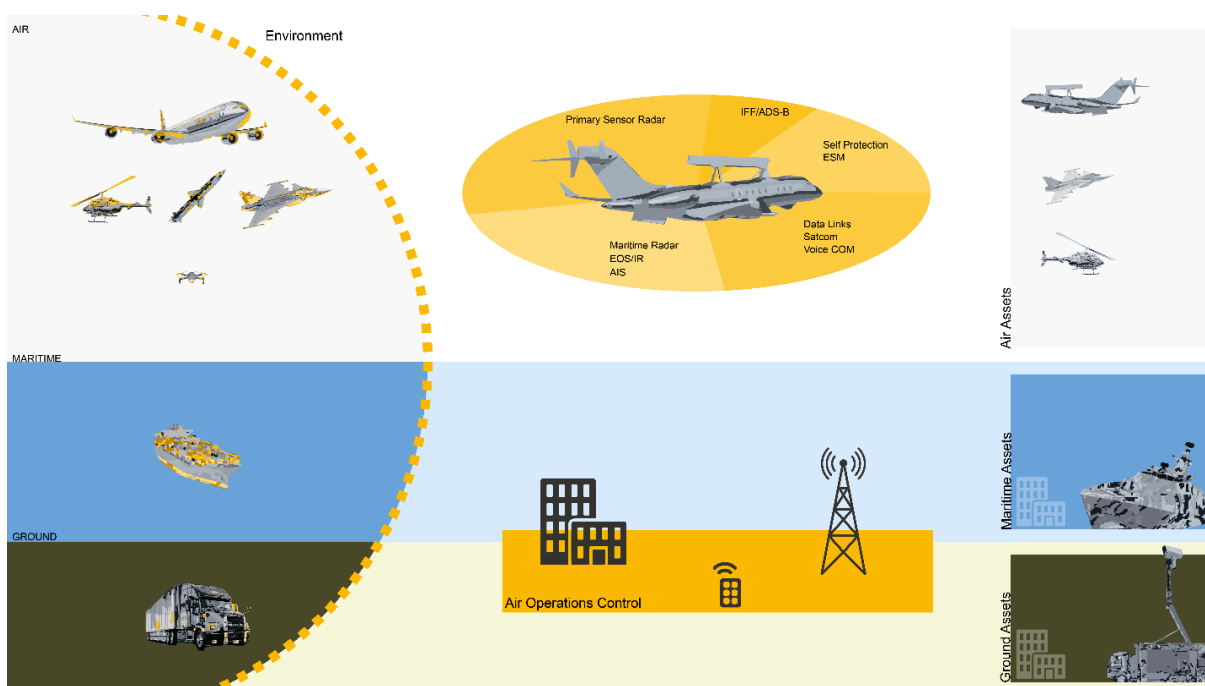


Figure 9: The AEW&C environment displayed to the left.

Each area includes different targets with different characteristics and states. These targets can be vehicles like airplanes, boats, and trucks, but it can also be other moving targets such as drones and missiles. The data from the targets are collected by the sensor suite of the AEW&C and identified as unknown, friend, neutral, suspect, or hostile. The behaviour of the target and the capabilities of the sensor suite are affected by external factors. Examples of these factors are weather conditions, air- and sea routes and national borders. But the capabilities can also be affected by electronic warfare (EW) which for example include hacking and adversarial attacks. The behaviour of the targets in the environment also changes depending on the state of international relations. The states can be divided in three general categories: peace, war, and conflict. The most common state is peace but depending on the location, states of conflict are not unusual.

6.3 The whole system: The Joint Military Operation

The whole system, the Joint Military Operation (JMO) is the largest system the study considers. The system includes all elements that in some way affect the system goal, to *Create a high-resolution image of the air-, land- and sea picture*. The JMO system will be described in terms of the system goal, subject, tools, outcome, and relationships between elements.

6.3.1 System Goal

In the JMO system, the AEW&C system collaborates with other subsystems to achieve the system goal. This large and complex system has the goal to *Create a high-resolution image of the air-, land- and sea picture*. The operators interviewed in the study often excluded land from the term and referred to this as Recognized Air-and Sea Picture (RASP).

6.3.2 Subjects

The subjects in this system consider the crew of the AEW&C and humans in other subsystems that the AEW&C collaborates with. In the JMO, the subjects are divided into two categories:

- Main subject: Crew of the AEW&C.
- Secondary subjects: Humans of the other subsystems within the JMO.

6.3.3 Tools

Tools, as defined as elements the subjects interact with to achieve the system's goal, are considered the assets of the subsystems in the JMO (see Figure 10). This includes fighter jets, AEW&C, land- and sea based radars, Air Operations Centre (AOC), and ground based Remote Control Centre.

6.3.4 Outcome

The outcome of the JMO system depends on each subsystem's individual performance (or outcome) and the collaboration between the subsystems. This requires that each subsystem perform well internally and manage to produce an outcome that contributes to the system's goals and that the communication and collaboration between them is effective and efficient. If the system manages to produce the desired outcome, a high-resolution image of the air- land- and sea picture, the outcome can be used in an even bigger system, whose goal might be to *keep people and society safe*.

6.3.5 Relationships

The system image visualizes relationships between elements of the system (See Figure 10).

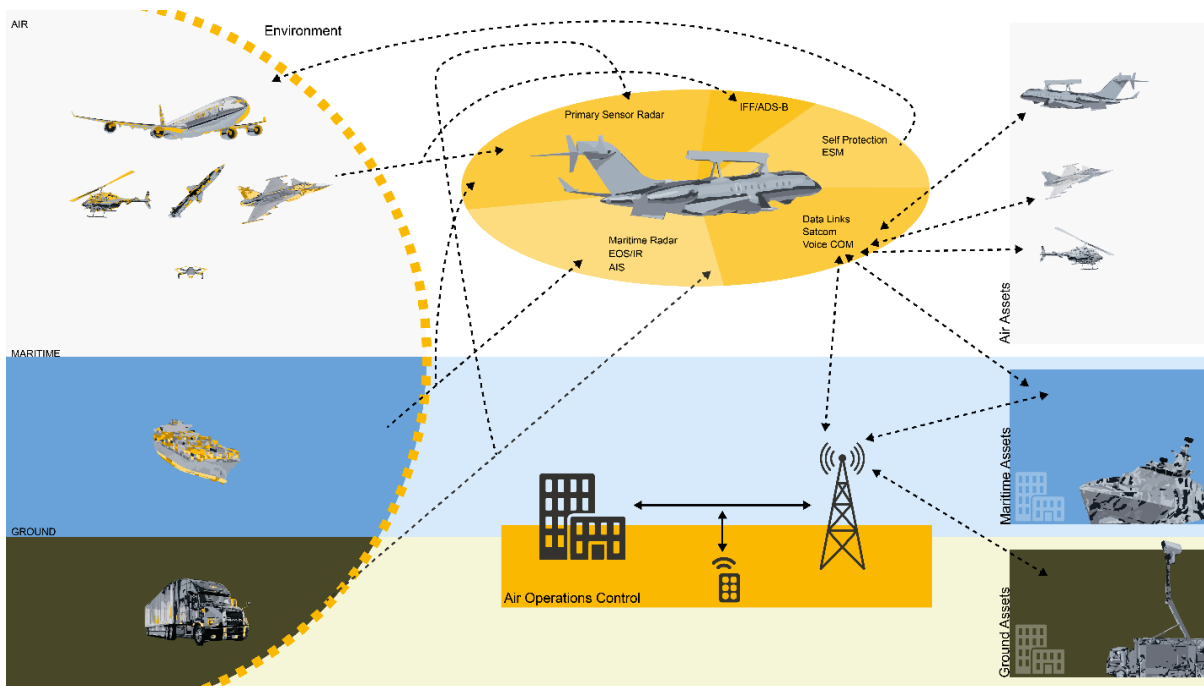


Figure 10. The relationships of elements within the Joint Military Operation system.

The focus of the JMO system is on the AEW&C. The relationships between the subsystems are facilitated by radio links, which in turn are facilitated by radio towers that connect different sensors to the Air Operations Centre (AOC). The AEW&C gathers data from the environment through the sensor suite. The primary sensor radar (PSR) on top of the airplane mainly gathers information about the position, movements, shape, and size of moving targets in the environment. There is also another radar mounted underneath the aircraft, that is used for maritime surveillance. The IFF and ADS-B gather identity information of airborne- and surface targets with transponders, while AIS does roughly the same for maritime targets (and for example rescue helicopters). The EOS is a camera system mainly used to get a picture of maritime and land targets. The data links, SATCOM and Voice COM are tools to distribute information and communicate with other parts of the system. The data picked up by the sensor suite of the AEW&C airplane is sent to the AOC. The AOC analyses the data and identifies the target and distributes the data back to the AEW&C and friendly assets in air, land and maritime. In addition, the AEW&C also identifies data and sends it back to the AOC. The AEW&C can also communicate directly to and control fighter jets that can be used to perform different missions.

6.4 Subsystem 1: The AEW&C System

The system boundary for Subsystem 1: The AEW&C System, is drawn around the AEW&C airplane. This means that everything inside the aircraft, including operators, pilots, observers, and all equipment are included. In addition, all sensors mounted on the airplane also are within the system boundary.

The airplane is divided into four different areas (see Figure 11):

- Cockpit
 - Like an ordinary airplane but also including mission display units (MDU) that display the situation picture.
- Planning & rest area
 - Comfortable seats, tables, MDU.
- Operator room
 - Operator workstations, observer seats.

- Machine room
 - Room for electronics and machinery.
- All rooms include safety tools like fire extinguishers, oxygen masks and life vests.

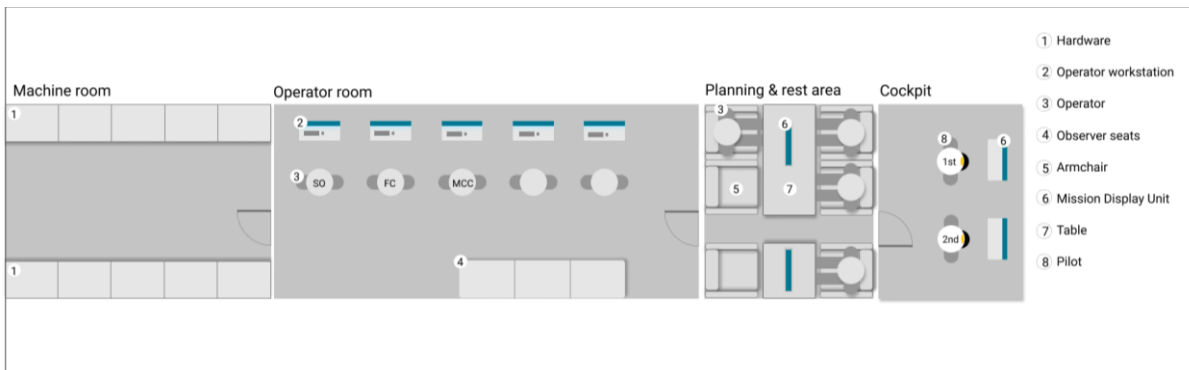


Figure 11. Inside the AEW&C airplane.

6.4.1 System Goal

The system goal of the AEW&C system is to contribute to the larger system’s goal of creating a high-resolution image of the air-, land- and sea picture. The AEW&C contribute to this goal by providing long-range detection and identification of targets in air, at sea and over land. By in principle being an elevated and mobile radar, the AEW&C has a larger radar coverage and more flexibility than ground- or sea-based radars. The goal for the AEW&C system is further elaborated in aspects of effectiveness, efficiency, efficacy, and safety:

Effectiveness

Effectiveness can be assessed based on how well the system is in providing a sufficient image of targets within the environment. Effective use means that resources are used to create a complete and accurate situation picture. This implies that targets are correctly identified and that a high degree of targets are detected and identified.

Efficiency

Efficiency can be assessed on how well all technical elements of the system work. To achieve high efficiency all elements must work flawlessly and jointly. These elements include sensors, i.e., radars and other sensing devices, and communications devices. The efficiency can be measured in metrics like time to detect, understand, communicate, and make decisions on the information provided by the sensors.

Efficacy

The AEW&C system acts within the larger Joint Military Operation (JMO) system with a defined mission. If the AEW&C manages to have a high degree of effectiveness and efficiency, the whole system can work well as well, providing air, naval and army commanders with situation awareness beyond their own capabilities.

Safety

The safety goal concerns keeping all involved personnel away from injuries and all vehicles operative. It also includes detecting potential threats and not making mistakes that can lead to dangerous consequences.

6.4.2 Subjects

The subjects are considered all crewmembers aboard the AEW&C airplane:

- Main subjects: The Operator team.
- Secondary subjects: AEW&C pilots and eventually observers (mainly for educational purposes).

6.4.3 Tools

The tools the subjects interact with to achieve the system's goal depends on if they are an operator or a pilot.

Tools for the operator team:

- The C2 system
- The sensor suite.
 - Primary sensor radar, IFF/ADS-B, Data links, Satcom, Voice communication, EOS/IR sensor, Maritime surveillance radar, AIS, ESM and Self-protection system.

Tools for the AEW&C pilots:

- The cockpit's interface. This involves the equipment for steering, radio, dashboard etc.

6.4.4 Outcome

The desirable outcome of the AEW&C system is to provide long-range detection and identification of targets in the environment. The outcome depends on the performance and relationship of elements within the system. For example, if there is a mismatch in the relationship between the operator (subject) and the C2 system (tool) and the operator uses the wrong settings to achieve the intended goal, the outcome will not be in line with the system's goal.

6.4.5 Relationships

Relationships between elements are defined as flows of materials, information, or energy between the elements (Flood & Carson, 1993). In the AEW&C airplane, information flows were considered most relevant and therefore only those were used in the analysis. The information flow of the AEW&C system is displayed in Figure 12.

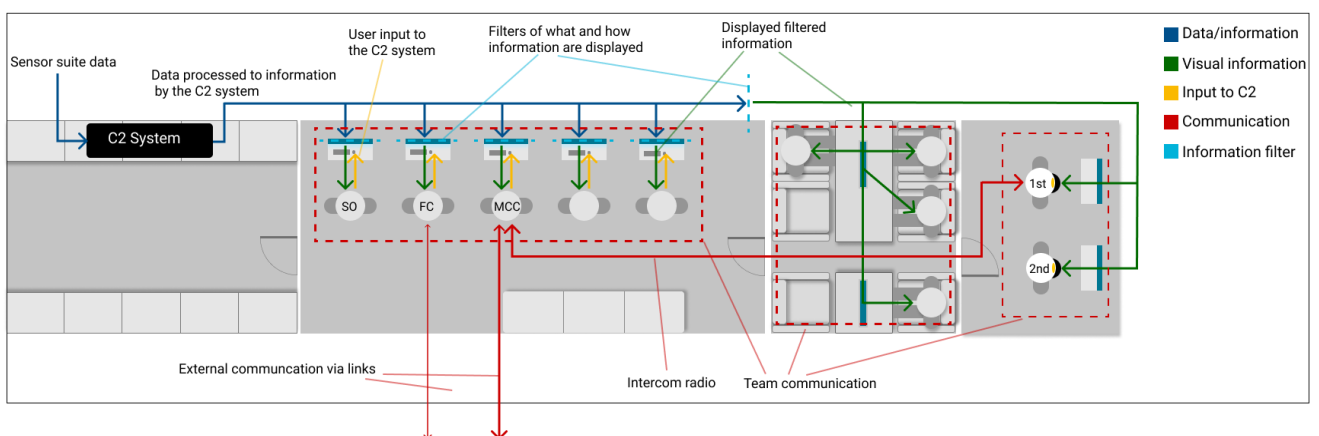


Figure 12. Relationships (information flows) between elements in the AEW&C system.

The data collected by the sensor suite is processed to information by the software system (blue). However, all information available is too much to visually present at the same time due to limited space on the displays and human cognitive limitations. Therefore, the information goes through an

“information filter” (turquoise), which are the settings and modes determining what visual information (green) reaches the human through the C2 system or mission display units (MDU). The operator can interact with the C2 system (yellow) to change the information filter and thereby the visual information displayed to them. Communication relationships are displayed with red arrows. The dotted boxes are team communication boundaries, areas where humans can communicate with each other face-to-face.

6.5 Subsystem 2: The Operator Team

The system boundary for Subsystem 2: The Operator team, is drawn around the operator team in the AEW&C. This includes the operator team of 5 persons with different roles, and the tools they interact with to achieve the system goal. The system is analysed in terms of the system’s goal, subjects, tools, outcome, and relationships.

6.5.1 System Goal

The goal of the operator team is to receive, interpret and distribute information provided by the built-in sensor suite through the Command & Control (C2) system and create the best possible prerequisites to create a high-resolution image of the air, land- and sea picture. The information is then used in different mission scenarios to execute tasks (sometimes role-specific) e.g., perform fighter control and to detect, report, and act upon detected anomalies.

6.5.2 Subjects

The operator team typically consists of five operators with different roles and tasks, there can also be multiple operators with the same role:

- Surveillance Operator (SO): The main tasks of the SO are to manage the radar settings and identify tracks. The SO role is further analysed in Chapter 7 *User studies*
- Fighter Controller (FC): The main tasks of the FC are to communicate with fighter pilots to guide and ensure their safety in different missions. The FC role is further analysed in Chapter 7 *User Studies*.
- Mission Crew Commander (MCC): The main task of the MCC is to communicate with other systems and be responsible for the outcome of the mission.
- Operator 4: Tasks depend on the mission. Can for example be to manage the EOS camera, the radio- and E-links or the self-protection system.
- Operator 5: Tasks depend on the mission. Can for example be to manage the EOS camera, the radio- and E-links or the self-protection system.

6.5.3 Tools

The tools are the elements the subjects interact with to achieve the systems goal:

- The C2 system through the C2 system the operator can interact with:
 - Other subsystems within the Joint Military Operation system (JMO).
 - The sensor suite.

6.5.4 Outcome

The outcome depends on how well each team member performs, but mostly it depends on how well the team can work together to achieve the system goal. For example, the FC is depending on the SO when guiding a fighter. If the FC fails to guide the fighter to identify a suspect track, the situation picture will be incomplete; the outcome will not correspond with the goal.

6.5.5 Relationships

Similarly, to *Subsystem 1: The AEW&C System*, the information flows were considered most relevant and therefore only those were used in the analysis. The information flows of the Operator Team are displayed in Figure 13.

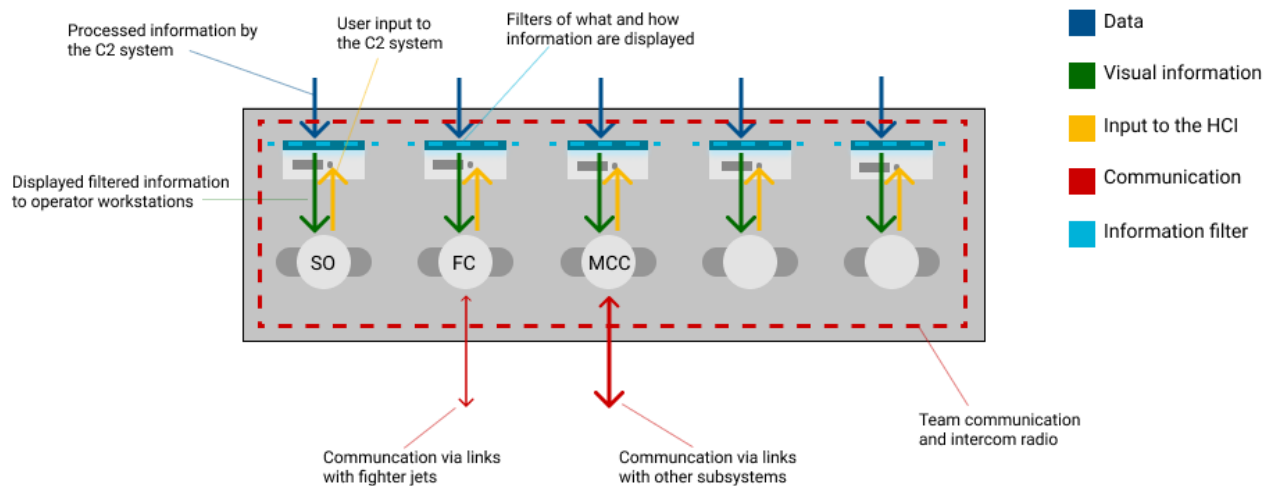


Figure 13. Relationships (information flows) between elements in The Operator Team.

The data collected from the environment by the sensor suite is processed to information (blue) by the Command & Control (C2) system. However, all information available is too much to visually present at the same time due to limited space on the displays and the operator's cognitive limitations. Therefore, the information goes through an “information filter” (turquoise), which are the settings and modes determining what visual information (green) reaches the operator through the C2 system. Depending on the role, the operator interacts with the C2 system (yellow) to change the information filter to display information relevant to their tasks. The Fighter Controller (FC) communicates with the fighter pilot (red) and guides according to the mission. The Mission Crew Commander (MCC) communicates with other subsystems in the JMO system, for example the AOC and 1st pilot of the AEW&C airplane. The dotted box is the team communication boundary where operators can communicate within the team through intercom radio and face-to-face. An example of face-to-face communication is to poke on another operator’s shoulder to get his or her attention and show something of interest on the display.

6.6 Subsystem 3: The AEW&C Operator

At the lowest level, only the Fighter Controller (FC) and Surveillance Operator (SO) was concerned. The subsystem includes the individual operator and their workstation.

The operator workstation, consists of the following components (see Figure 14):

- The C2 system
- Situational display: Displaying the situation picture with tracks and tools like settings, filters etc.
- Input devices.
 - Keyboard.
 - Computer mouse.
- Headset: (To receive information and communicate within the operator team and other subsystems within the JMO).

- Radio: Used for communication.
- Computer chairs.
- Desk.
- Safety equipment
 - Oxygen mask
 - Life vest.



Figure 14. The operator workstation.

6.6.1 System Goal

The general goal for an AEW&C operator is to perform their assigned tasks with effectiveness and efficiency to contribute to the team performance and accomplish the mission.

Surveillance operator: The system goal for the SO is to create the best possible circumstances to create a Recognized Air-and Sea Picture (RASP). In short this implies making sure that the radar collects the right data from the environment according to the mission.

Fighter Controller: The system goal for the fighter controller is to provide the fighter with sufficient information about the situation and ensuring the flight safety.

6.6.2 Subject

At this level only the FC and SO role is concerned.

Surveillance Operator (SO): The main tasks of the SO are to manage the radar settings and identify tracks. The tasks of the SO role are further elaborated in Chapter 5.13 *Hierarchical Task Analysis*.

Fighter Controller (FC): The main tasks of the FC are to communicate with fighter pilots to guide and ensure their safety in different missions. The tasks of the FC role are further elaborated in Chapter 5.13 *Hierarchical Task Analysis*.

6.6.3 Tools

The tools the operator interacts with are:

- The C2 system. Through the C2 system the operator can interact with:
 - Other subsystems within the JMO.
 - The sensor suite.
- Input devices (Keyboard & computer mouse).
 - To interact with the C2 system.
- Communication equipment:
 - Headset.
 - To receive information and communicate within the operator team and other subsystems within the JMO.
 - Radio.
 - Used for communication.
- Role-specific tools: Some tools are only available for certain roles, for example the EOS camera.

6.6.4 Outcome

The outcome depends on the operator's Situation Awareness (SA) and ability to distribute general and role-specific information to other operators and subsystems. The outcome has a direct effect on the operator team's performance, which in turn affects the larger system's goal of creating a high-resolution image of the air-, land- and sea picture.

6.6.5 Relationships

The relationships within the system are exchanges of information. The information flows are visualized in figure 15.

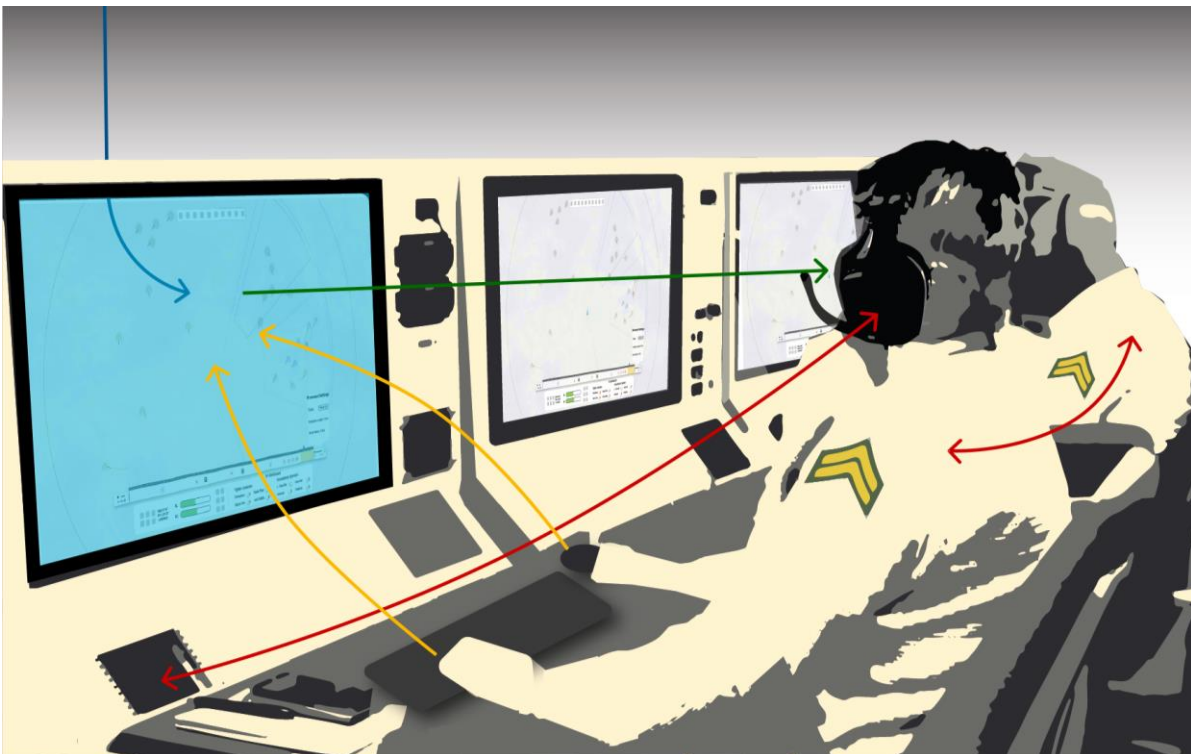


Figure 15: Relationships (information flows) within the AEW&C operator system.

The AEW&C sensor suite collects data (input) from targets in the environment. The data is processed to information by the C2 system (blue). The information goes through an information filter (turquoise) and reaches the operator as visual information through the display of the C2 system (green). By interacting with the C2 system through input devices (yellow) the operator can affect the information filter e.g., what visual information is displayed. By interacting with the C2 system, the operator can also indirectly communicate with other operators and subsystems, for example by identifying targets in the environment. The operator communicates directly to other operators and subsystems through the headset (red). The audial information reaches the other subsystems through the radio. The operator can also communicate face-to-face within the operator team, for example by poking someone's shoulder to show something on their screen.

The main difference of information flow between operators is how they interact with the C2 system to control what is displayed on the display and what subsystem they interact with through the radio. The SO does not communicate as much as the FC, and mainly does so with the Mission Crew Commander (MCC) or the Air Operation Centre (AOC). The FC communicates, on occasions intensely, with both the fighter pilot and the MCC. The FC communicates their own Situation Awareness (SA) to the fighter jet, who does not have the same resources to develop their own SA.

6.7 Chapter summary: System Analysis

Conducting a system analysis is a useful way to start a project related to complex and technical systems, like the one of the study. It is important to establish an understanding of how the systems are composed in terms of environment and subsystems and how they relate to each other. In this study the following systems and environment were identified:

- The AEW&C Environment
- The whole system: The Joint Military Operation
- Subsystem 1: The AEW&C System
- Subsystem 2: The Operator Team
- Subsystem 3: The AEW&C Operator

The system and subsystems were analysed in terms of system goals, subjects, tools, outcomes, and relationships. This facilitated an understanding of each system and clarified the Surveillance Operators (SO) and Fighter Controllers (FC) part in achieving the whole Joint Military Operation's system goal, *to create a high-resolution image of the air-, land- and sea picture*. Most importantly, the system analysis gave the study a foundation to rely on, a common ground, in further studies.

7

User Studies

The main users in the study are the Airborne Early Warning & Control (AEW&C) operators Fighter Controller (FC) and Surveillance Operator (SO). The aim of the study is to look for opportunities to improve their Situation Awareness (SA) with the implementation of Artificial Intelligence (AI) in the Command & Control (C2) System. The study focuses on two major questions: **when** to implement AI, and **how** to implement AI. This chapter, *User studies*, mainly relates to **when** to implement AI.

The study approach to answer this by:

- Conducting interviews which gained findings in:
 - The users' relation to the product.
 - Insights in aspects related to teamwork.
 - SA Errors and risks.
 - How working conditions aboard the AEW&C airplane could affect SA.
 - Users' thoughts about AI.
- Operative simulations in the Mission Training System (MTS) which resulted in:
 - Findings of SA related challenges for the FC and SO role.
 - Understandings of how the operators work at a general level by conducting a general Hierarchical Task Analysis (HTA) for each role.
 - Assessments of workload and SA for steps in the general HTA by using the methods NASA-TLX and GTS.
 - Four detailed Hierarchical Task Analysis (HTA) on tasks with high workload.

The chapter starts by presenting the findings gained from the interviews.

7.1 Findings from Interviews

The main users of the study are the Surveillance Operator (SO) and the Fighter Controller (FC). To establish an understanding of the users and their context, interviews with operators and engineers were conducted. The interviews were transcribed and analysed by making affinity diagrams. This part presents the findings from these interviews. The interviews gained findings in the following categories:

- The Use Profile
- Teamwork
- Errors and risks
- Working conditions aboard the AEW&C Airplane
- Users' thoughts about AI

The following chapters will present the findings of each category.

7.1.1 Use Profile: The AEW&C Operator

In this section, a use profile for the AEW&C Operator is presented (See Table 1). The use profile presented is based on Janhager's (2005) description of use profiles. The aim was to gain insights regarding the users' relation to the product. The roles in focus are the Surveillance Operator (SO) and Fighter Controller (FC). Only one final use profile was created but is considered applicable for both roles. The differences in the two roles use of the product is discussed after the use profile is presented. When the term *product* is used, this implies the AEW&C (described in further detail in Chapter 6.4: *Subsystem 1: The AEW&C system*).

Table 1 below presents the result of the Use Profile for an AEW&C Operator. The degree of performance for each of the categories is marked with bolded text. For some of the categories the degree of performance lies between two values which is marked by two bolded texts.

Table 1: Use Profile for an AEW&C Operator

| Category | Degree of performance | Explanation |
|------------------------------------|---|--|
| Length of use and education | Newcomer Experienced Specialised | Experienced or specialised depending on role. |
| Frequency of use | Rare Occasional Frequent | Daily use and with occasional use during weekends. |
| Influence on the choice of product | No influence Some influence Great influence | Depends on the country. No or some influence in the choice of system. Post development mainly relates to software changes. |
| Influence on the use situation | No influence Some influence Great influence | Depends on the hierarchical structure of the mission. |
| Responsibility in use | No influence Some influence Great influence | Responsibility to use the product according to their individual task during the mission. |

| | | |
|-----------------------------|---|---|
| Ownership | General Rented Owned | Users do not own the system but act as they do to some extent. |
| Social aspects | No influence Some influence Great influence | Proud of being a part of a group selected to do something important. Prestige between roles. |
| Mental influence of product | No influence Some influence Great influence | Affection develops over time. Overall appreciation of the product but can see areas of improvement. |
| Cognitive interaction | No influence Some influence Great influence | High cognitive interaction that varies in intensity over the duration of the mission. |
| Physical interaction | No influence Some influence Great influence | Not as comfortable as working in an office. Long durations and more cumbersome to take breaks. |

Length of use and education: This varies between countries and between the roles, but all operators must have military education to be allowed to use the product. The operators also get additional education of the C2 system used in the product. This education centres around the functionality of the system and in particular the human-computer interface. The tactical education is handled by the customer, i.e., the military organisation within the country that buys the system. In Sweden, additional Fighter Controller (FC) education is mandatory for the role.

Use time and skill between operators varies, but in general the operators have several hundred hours of experience with the product. Some users are substantially more skilful than others. FCs are generally more experienced than Surveillance Operators (SO). The FC also must pass stress tests, spatial tests, and cooperation tests that their peers do not have to take.

Frequency of use: Operators fly mostly every day; the exact time varies between countries. The duration of flights varies but can be up to 11 hours. If an operator has not flown for a couple of months, they must perform a check, and after some years, parts of the education need to be retaken.

The influence on choice of product: This depends on what country the AEW&C is used in, but generally the influence is small or none. In Sweden operators and military procurement officials are in relatively close collaboration compared to other countries. This means that Swedish operators have some influence on the choice of product, in terms of what functionalities it should have. For example, the functionalities within the C2 system have been developed in collaboration between users and engineers. However, once the system is bought, influence is substantially diminished. Future upgrades of the system that the operator may have influence over mainly concern software upgrades.

Influence on the use situation: This depends on the hierarchical structure of the mission. In Sweden there is a flat structure which means that the AEW&C operators have a more evenly distributed influence. In countries where hierarchies are more present some roles, like the Mission Crew Commander (MCC) or the 1st pilot of the AEW&C airplane, have more influence on the use situation than others.

Responsibility in use: The operators are responsible to use the product according to their individual task during the mission. For example, the FC is responsible for ensuring the flight safety of the fighter jet in relation to civil traffic and the SO is responsible for managing the radar settings. The MCC is responsible for the mission overall.

Ownership: Since the operators spend so much time with the product, they treat it as if they own it to some extent, even though they do not. They are proud of the product and feel the need to take care of it. They keep their “office” clean, avoid drinking coffee over the keyboard etc.

Social aspects: Being an AEW&C operator means that you are one of few selected to do something important for society, which adds to the sense of pride. Since the product is considered to be a high value target, this implies that operators are an important asset. There can be some prestige between roles. The FC can be considered a more crucial role than the SO. Generally, operators strive to be FC’s. It involves more responsibility and is often significantly higher paid than the SO role. Another social aspect is cooperation skills. It is important to continually share gained knowledge and communicate. It may also be beneficial for operators to be friends outside the confines of the plane. This since it helps building trust between operators which is important to have during difficult and stressful operative situations.

Mental influence: The operator spends a considerable amount of time with the product and develops an affection to it to some extent. Overall, they appreciate the product and feel like they can trust it, even though they see areas of improvement.

Cognitive interaction: The operator's cognitive interaction with the product is high but the intensity varies over the duration of the mission. For example, the FC’s cognitive interaction with the product is especially intense when leading one or more fighter jets.

Physical interaction: The airborne aspect of the AEW&C operator role results in some additional difficulties. It is more cumbersome to take breaks, it requires more spatial abilities since the plane moves in relation to the targets and even though the environment in the airplane could be considered good it does not beat the comfort of an office.

Differences between roles in the Use Profile:

The use profile is applicable to both the role of FC and SO but have some differences. The most apparent difference between the FC and the SO is within Length of use and education. The SO is considered to have slightly easier tasks which does not require as much education and responsibility. The SO’s tasks are a part of the routine for most missions and are mostly similar independent on mission. The tasks of the FC are more dynamic and unpredictable. But for certain missions, or when links are down, the workload and cognitive demand on the SO increases significantly.

The social aspect may be a considerable factor affecting the team dynamics. Since the FC tasks are more difficult and demands more resources than those of the SO it may lead to undesirable crew resource management related issues. In general, this is not a problem in Swedish crews where the organizational structure is relatively flat. In other countries it can lead to issues where the SO’s opinion may be valued lower than the FC’s.

7.1.2 Teamwork and Communication

For the operator team, teamwork and communication is critical to achieve the system’s goal. The operators are, to a large extent, dependent on each other to be able to perform their role-specific and collective tasks. This chapter presents insights of important aspects and challenges related to teamwork and communication for the operator team.

Findings related to teamwork:

- It is important that the system facilitates the ability to work dynamically and be able to trade tasks and support each other. Since the workload for the roles sometimes varies, there is often someone with more, or less, workload.
- An experienced and cohesive team are better at working dynamically and proactive, they have a better team awareness (awareness of what the other team members do and when they need support).
- Feedback and support should be immediate, and proactive if possible.
- The operators must have trust in the team. Each operator has different responsibilities and must trust that everyone is doing what they are supposed to do to complete the mission.
- It is important that the roles are clear, especially in situations of crisis. Crew resource management is important if disputes occur.

Findings related to communication:

- Communication requires discipline. It should be clear, informative, and concise. Standardized military phrases are used for this purpose.
- Confirmation and repetition are used to ensure understanding from both parts.
- Speak is a quicker means of communication than text, but harder to remember. Pen and paper are used to take notes of things to remember.
- It is a challenge to know when the counterpart is receptive to communication in moments of high workload. It is also hard to assess who has the most important thing to say if several people want to reach the same person at the same time. Emergency words can be used in case of urgency.

7.1.3 SA Errors and Risks

In the interviews the users were asked questions regarding errors and risks related to SA. The purpose of this was to gain knowledge that could be useful when observing and analysing the operative simulation in the Mission Training System (MTS). For example, it could facilitate a better comprehension and prepare what to look for. The SA errors and risks on a general level and for the SO and FC role respectively is presented below.

General SA errors and risks:

- **Stress** can be induced by high workload and time pressure and lead to several risks and mistakes like:
 - Tunnel vision. Losing awareness of the bigger picture and forgetting to close unused windows.
 - Falling behind in the control loop, forcing the operator to work reactive instead of proactively.
 - Impairing reasoning and decision making
 - Making slips, like closing windows by mistake.
- **Tiredness** induced by long working hours and night-time work. May for example lead to impaired vigilance (sustained attention).
- **Risk of forgetting things.** The operators have a lot to keep in mind and remember.
- **Insufficient perception** of tracks and plots: Might miss that two targets are on the same height or not looking at how tracks and plots relate to each other.
- **Filtering:** Risk of filtering out interesting information or not being aware of filters.
- **Window management:** Risk of covering important information on the map with windows.

SA errors and risks for SO:

- **Incorrect identification:** Putting the incorrect identity can have big consequences. It can for example be because of a misjudgement, a rushed decision because of stress or that a target is acting under false identity. When unsure about identity it is better to postpone the identification until more information is available.
- **Losing track of radar status:** The SO might forget to check the status or in the worst case turn off the radar by mistake without noticing.
- **Overloading the radar:** The radar has limited capacity. Putting too much energy in an area affects the performance of the radar in other areas.

SA errors and risks for FC:

- **Not detecting threats:** If the FC does not detect and notify the fighter pilot of threats it can be shot down. It might be that the FC is looking at the wrong area of the map or has a tunnelled vision.
- **Forgetting civil traffic:** If the FC does not detect and notify the fighter pilot of civil traffic it can intrude on their airspace or in worst case collide. This might happen if the FC is too focused on guiding towards a target or when the FC is relaxed when guiding the fighter back after a mission.

7.1.4 How working conditions affect SA

In the interviews the users were asked questions regarding how the working conditions aboard the AEW&C airplane could affect SA. The purpose of this was to get a better understanding of what could possibly be overlooked during the activities in the Mission Training System (MTS). The working conditions in the AEW&C airplane differ from working in an office on the ground, for example in the Air Operations Centre (AOC). Even though the tasks can be similar, for example identifying targets and guiding fighter jets, the working conditions make the tasks more challenging aboard an airplane.

The specific working conditions affecting the users SA are:

- **Turning the airplane:** The airplane must turn occasionally to stay on route. When this happens, the radar loses track of the targets in the environment for a little while. The crew must also adjust their spatial perception of where they are in the air.
- **Long working hours and working night-time:** The missions can last up to 11 hours and are sometimes at night-time. This leads to tiredness, both psychological and physical.
- **Not the same possibilities to take breaks:** Often there are no extra operators aboard the plane that can cover up during breaks.
- **Less comfort:** Turbulence, lack of space and noise leads to less comfort.
- **Flight related stress:** This includes the feeling of being cabined, fear of flight and the risk of being shot down and being considered a high value asset.
- **Less resources:** Less people and not the same capabilities for communication as the AOC.

To deal with the risk caused by the working conditions the crew is adapting Operational Risk Management. Before taking flight, they assess if everyone within the crew is feeling well, have had enough sleep, if someone is having personal issues affecting them etc. If the result of the ORM is not sufficient for one of the operators it may result in the assessment that they are not fit to fly.

7.1.5 Thoughts about AI

In the interviews the users were asked regarding their opinions on AI in general and in the AEW&C system. They were also asked if they had suggestions of how AI could be used in the AEW&C

system to facilitate better SA. The purpose of this was to get insights regarding the user's opinions, knowledge and expectations about AI and be inspired by their ideas.

Trust in the AI's capabilities seemed to differentiate the users in their opinions. One user stated that he would trust the AI to do all the work, provided that the AI had the ability to do so. Another user would not listen to the AI's suggestion in areas where he felt more confident in his own capabilities, and it would make him hesitant if the AI would not agree to a decision in extreme situations. They also discussed the risk of having too much trust in the system, and not being aware of its limitations. Understanding the purpose and functionalities of the AI are considered important.

There were also varied opinions on what type of tasks would be suitable to hand over to an AI system. What seemed to be consensus was not to let the AI identify anything as hostile. In those cases, a human should always have the last opinion. One user acknowledged the need to define a suitable level of automation, so the human could put more focus on more important tasks. One opinion was that the operators should be able to perform all tasks by themselves, the AI should not be allowed to perform tasks beyond the operators' capabilities and comprehension. The risk of knowledge loss due to automation was also mentioned.

Implementation of AI functionalities could also require the operators to adapt and learn new things, making the already complex AEW&C system even more complex. It is important that the operator feels confident in the use of AI systems, that they have control. Maybe the implementation of AI would require a new specific role in the operator team, responsible for AI functionalities in the system.

Other opinions were that AI is mostly developed for show, to sell. AI should not be developed without an underlying need. Considering the complexity and risks associated with the system the AI would require very much training and testing.

Suggestions on what AI could do to improve SA in the AEW&C system:

- Help detecting and notifying FC about civil traffic.
- Filtering to reduce clutter (especially for plots in sea areas).
- Automation of things happening in the periphery.
- Route suggestions for guiding fighter jets, considering multiple targets.
- Threat assessment. Assess the level of danger and notify the user.
- Assessing what plots are actual targets (especially in sea areas).
- Merging unlatched tracks with their corresponding plot.
- Detecting abnormalities.
- Assess situations early, thinking several steps in advance.
- Recognize, and visualize, formations of suspect tracks.

7.1.6 Summary: Findings from the interviews

By conducting interviews and analysing the transcriptions by using affinity diagrams, findings were obtained related to the AEW&C operator. The findings included:

- A Use Profile for the AEW&C Operator. Providing insights regarding the users' relation to the product.
- Insights of important aspects and challenges related to teamwork and communication for the operator team.
- Lists of potential SA errors and risks on a general level and specifically for each role.
- Insights of how the working conditions aboard the AEW&C can affect SA.

- Insights of the potential users' opinions about AI and a list of suggestions on what AI could do to improve SA.

The findings from the interview provided deeper understanding of the users, the operator team, the working conditions, errors and risks and attitude towards AI. This knowledge was useful to get a better comprehension when observing and analysing the operative simulation in the Mission Training System (MTS), which was the next step of the user study. The findings from the activities in the MTS are presented in chapter 7.2: *Findings from activities in the MTS*.

7.2 Findings from Activities in the MTS

The activities in the Mission Training System (MTS) were the most central part of the user studies. The Activities in the MTS is described in detail in Chapter 5 *Application of methods* but are also briefly described below.

The data collection activities were observations of operative simulations in the MTS and interviews regarding SA and workload. The data analysis activities were making affinity diagrams of notes taken, making Hierarchical Task Analysis (HTA) to understand the user's performance of tasks, and assessing workload by making Generic Task Specification (GTS).

In total, three simulated missions were conducted by the operators. The first mission was considered the baseline. This was also an opportunity for the users to warm up. Some operators had not been operating in a while and felt a bit rusty, and the operators needed to get used to working together. However, they were quickly up to speed. During the second mission abnormalities were added. This was to assess whether the operators were able to detect abnormal behaviour of targets, and if they did, assess if and how they detected the abnormality. In the third and final mission the operators swapped roles. This was to get multiple perspectives of the SO and FC role.

During the observation, the operators were asked to think out loud and questions were asked to clarify uncertainties. Notes were taken when interesting events occurred that could be related to SA. For example, difficulties in tasks, errors, workload, stress, and other insightful comments. After each mission, the operators were interviewed and asked questions regarding SA and workload.

The data collected from the observations and interviews were analysed afterwards. This resulted in the following results:

- Identified SA challenges for each role.
- A general HTA for each role.
- An assessment of which two tasks had the highest workload for each role.
- A detailed HTA for the two tasks with the highest workload for each role.

The following chapters will present the findings from the Activities in the MTS.

7.2.1 SA challenges for FC

The affinity diagram of the notes from the observation resulted in seven categories where SA challenges for FC were identified: communication, window management, warnings, working methods & assessments, information filtering, finding & using tools, risk of errors.

Communication: Communication is critical and sometimes intense. The FC needed to communicate the situation both to the fighter pilot and MCC. The communication got increasingly intense when several fighter jets were guided simultaneously. When the fighter pilot has identified a suspect track the communication sequence is the following: Fighter pilot → FC → MCC → External subsystems

(for example AOC). Clarity in communication between FC and fighter pilot is important. The FC makes many handwritten notes to remember crucial information.

Window management: When the FC guides a fighter jet in a mission, information is displayed in a window. When several guiding missions occurred simultaneously the FC could get confused what window belonged to what mission. In moments of stress the FC forgot to close windows that were not used, covering information, and impairing the SA.

Warnings: Pop-up windows with warnings and alerts disrupted the FC without providing information relevant to the mission. Attention was required to dismiss the warnings, potentially impairing the SA.

Working methods & Assessments: There are several ways of working and making assessments as FC. It is challenging to assess the threat levels of different targets and prioritization of routes when there are several moving targets to identify. It is important to keep track of borders. The working methods are proactive, and the system should facilitate proactivity. It is needed to have good awareness of the task at hand and rely on the other operators to take care of the bigger picture.

Information filtering: It is important for the FC to get an as clean picture as possible to be able to put full attention on what is relevant for the task. But at times the picture was messy, with a lot of lines and windows covering the interface. They are aware of the risk of information overload and therefore they filter out some information. The information FC finds interesting is different from the SO, for example they might discard information that is not a part of the mission but increase the information related to the mission. The FC's attention is more focused.

Finding & using tools: There are very many functions included in the complex C2 system. Sometimes the operators had trouble finding the functions they were looking for, or even being aware of the existence of certain functions.

Risk of errors: No serious errors occurred during the simulation. But risks were identified. For example, when the AEW&C airplane turns the FC becomes "blind" for a little while and momentarily loses SA. It is important that the turns are timed so the FC does not become blind in a critical moment. Focused attention to a small area can impair the global SA and spatial ability. Blindness caused by the Doppler effect can come unexpectedly, the FC should be prepared when this happens. When several guiding missions are occurring simultaneously the workload can get to unmanageable levels.

7.2.2 SA challenges for SO

The affinity diagram of the notes from the observation resulted in six categories where SA challenges for SO were identified: communication, window management, radar status, track management, radar areas, identification.

Communication: The SO mainly communicates with MCC directly and the AOC indirectly by identifying tracks. Information about interesting tracks is communicated to the rest of the system. Repetition and confirmation are sometimes required to hear and understand each other.

Window management: It is a common error to forget to close windows after use, which may impair SA. Reasons for the error might be because it feels cumbersome to open the window again when needed, or something else catches the SO's attention and they simply forget the window. However, sometimes the SO wants to keep a window up for a longer duration to be able to make continuous

changes and see additional information. Pop up warning windows may also impair the SA. Even though some information might be useful, it is not possible to minimize the window without dismissing the alerts.

Radar status: The radar status is always visible in the Vital Signs Field (VSF) on the main display. During the simulation the radar was never at risk of overloading. However, it was noticed that there was no clear connection between the radar settings and areas and the status of the radar. Another insight was that there was no indication of critical levels in the bar charts displaying radar status, making it hard to understand how to relate to different levels of load.

Track management: The doppler effect can cause tracks to be unhooked, e.g., losing connection to the corresponding plot. This results in that a new unknown track is created for the plot and the old track continues in the tangent. When this happens, there is a need to manage the unhooked track, something that is not possible if someone else has “locked” the track. There is no, or difficult to access, information on who has locked the track, which may cause confusion. If tracks, for some reason, cannot be identified and are left as unknown it is easy to forget the track.

Radar areas: The SO needs to prioritize the radar’s capacity in different areas, and make sure that they have the optimal settings according to the mission description, but there is no indication of how much radar load an area causes. To be as efficient as possible, different radar areas need to be activated in a certain order. The radar areas can be cumbersome to place due to lacking usability. Another issue related to areas is that for land areas it is hard to differentiate real targets from clutter because the system does not make tracks from plots automatically.

Identification: Identification is hard and time consuming. Often most of the identification is made by the AOC, but for some occasions it is made from the AEW&C. The mission description might include instructions of how to assess the identity. A lot of information is required to identify tracks, sometimes at different places in the system. Experience is crucial to be efficient, and it is important to work at a quick phase in the beginning of the mission to quickly establish a RASP. There is a trade-off between being thorough and quick, often the SO chooses only to look at the most important information (such as transponder, nationality, location etc) when identifying tracks, which could result in incorrect identification. When there are many unknown tracks the SO needs to prioritize what tracks to identify first, mainly based on speed, course, and location. Sea- and land tracks are significantly harder to identify (compared to air tracks) based on the large number of plots, less information, and unpredictable and slow movement patterns. Lastly, some data of tracks are delayed. If the track is left unknown for the time being, it is difficult to remember which it was later.

7.3 Hierarchical Task Analysis

Based on the observations in the MTS, a general Hierarchical Task Analysis (HTA) was made for each role. This was to gain a basic understanding how the operators work and later be able to assess what tasks had the highest workload. In this section a general description of the tasks of the SO and the FC will be described. The most crucial parts of the tasks are highlighted and described in further detail.

7.3.1 General HTA: Fighter Controller

Figure 16 shows what tasks the Fighter Controller (FC) executes in a low resolution.

The structure of a mission for the FC's tasks during a general mission is seen in figure 16. Although many of the tasks are generic for multiple mission types it is important to clarify that there are several use cases not represented in this task breakdown. For example, the FC is able to perform other tasks such as those normally allocated to the SO or MCC depending on the intra-mission workload. A FC does not necessarily have a uniform workload across a flight session. Some sessions can vary between having lengthy breaks and short sprints of high workload. Figure 16 represents one sprint.

The first part of a mission is the preparation phase where target points are placed on the map. At some point after this, depending on what is happening in the outside world, the FC will need to request a fighter jet. The purpose of this can be several different things such as to identify or divert an enemy or suspicious vehicle. This activity starts with someone in the crew detecting that something needs further investigation which also depends on what is in the mission brief. After this the FC requests a fighter jet from the MCC who forwards the request over "radio-link". Thereafter the FC takes over control over the lead of the fighter jet. The lead part that follows depends on the situation, in short, the FC needs to guide the fighter jet to different points or targets in a safe manner. The fighter jet performs their mission, for example to identify a suspicious target, and thereafter the FC reports back to the MCC. After one of these missions has been executed the FC either leads the fighter back to base or to a waiting point called "CAP".

7.3.2 General HTA: Surveillance Operator

In figure 17, an HTA of the general tasks for the Surveillance Operator (SO) is depicted.

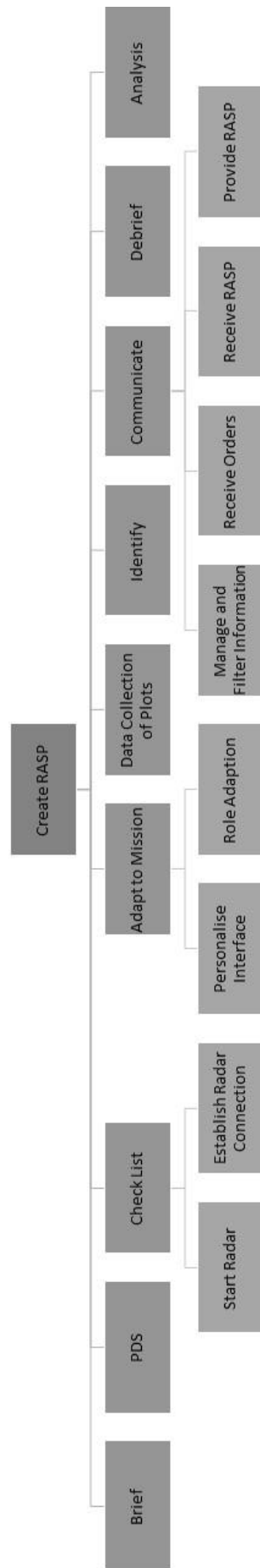


Figure 17. General HTA for the Surveillance Operator.

The main goal of the SO is to provide RASP (Recognized Air- and Sea Picture) for the entire team and interconnected units in the larger Joint Military Operation system. The task of doing this is divided into three major phases: the planning phase, operation, and evaluation. These three phases are both connected to the mission and are mirrored in the Command & Control (C2) interface.

In the planning phase, the mission goals are stated, and this information is used in C2 to prepare as much as possible for the mission. Preparations may include setting up areas and creating points of interest in the interface etc. When inside the AEW&C all the operators have a checklist to follow. For the SO the main task is to start the radar and activate the sensors and areas that are necessary to create a sufficient RASP.

The next task is to adapt the interface to the specific mission. If the mission entails monitoring maritime targets the operator might choose a sea chart map layer and opening windows relevant for those mission goals. In the previous step sensors were activated which enabled the collection of target data which thereby enables identifying tracks that are based on the plots. Usually, the tracks are identified from the AOC and the SO merely enables data collection through the correct sensor settings and prioritizing the correct areas determined by the mission goals.

There are instances where the AEW&C may have identification responsibility, which significantly increases the workload. It could theoretically also happen if there are problems with the link between the AOC and the AEW&C. In either scenario the information gathered by the sensors facilitated by the SO need to be communicated, especially when there are inconsistencies between ground-based operators' and AEW&C operators' assessments of track properties or any uncertainties. These conflicts are generally communicated from the SO through the MCC to the AOC.

The SO also communicates RASP indirectly and directly to the other members of the team depending on the momentary workload of the respective team members. Other communication topics include changes in priorities between areas, need of increased coverage in certain areas and detection of important targets. When a mission is finished events are analysed and findings are used to improve future missions.

7.3.3 Selecting subtasks: Generic Task Specification (GTS)

The approach used in the study was to assess what subtasks of the general HTA had the highest workload and elaborate them into more detailed HTA's. According to Endsley (2016), high workload is a factor that impairs SA. By assessing what tasks have the highest workload, opportunities to improve SA can likely be found in those tasks. Elaboration of the HTA for those tasks will result in a deeper understanding of task performance and the decision-making process. Based on this, the detailed HTA's of the subtasks were then used as a foundation for ideation of AI concepts.

The method used to assess the workload of the subtasks was *Generic Task Specification (GTS)*. The implementation of the implied selecting the factors most contributing to workload and assessing how many factors were relevant for each subtask. In total, ten factors were selected based on findings of the user studies. These were:

1. Time pressure
2. High information load
3. Memory consuming
4. Calculating
5. Critical decisions
6. Searching

7. Spatial ability
8. Attention
9. Performance pressure
10. High risk of error

The results of the GTS for the Fighter Controller are presented in figure 18. The subtasks *Ensure safety* and *Guide fighter to target* were both related to eight factors and selected for further elaboration.

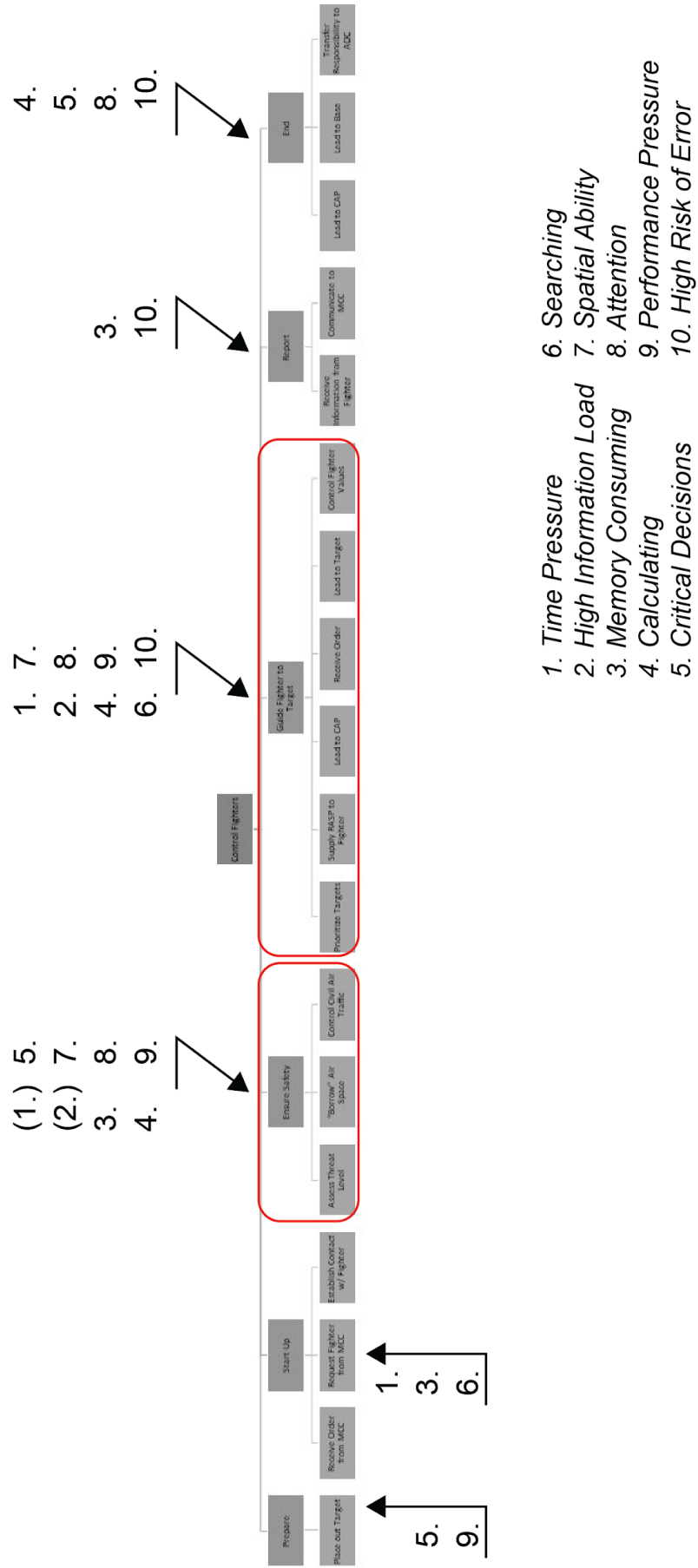


Figure 18: The results of the GTS for the Fighter Controller.

The results of the GTS for the Surveillance Operator are presented in figure 19. The subtask *Identification* was related to eight factors and selected for further elaboration. The subtask *Collect target data* was related to five factors and selected for further elaboration.

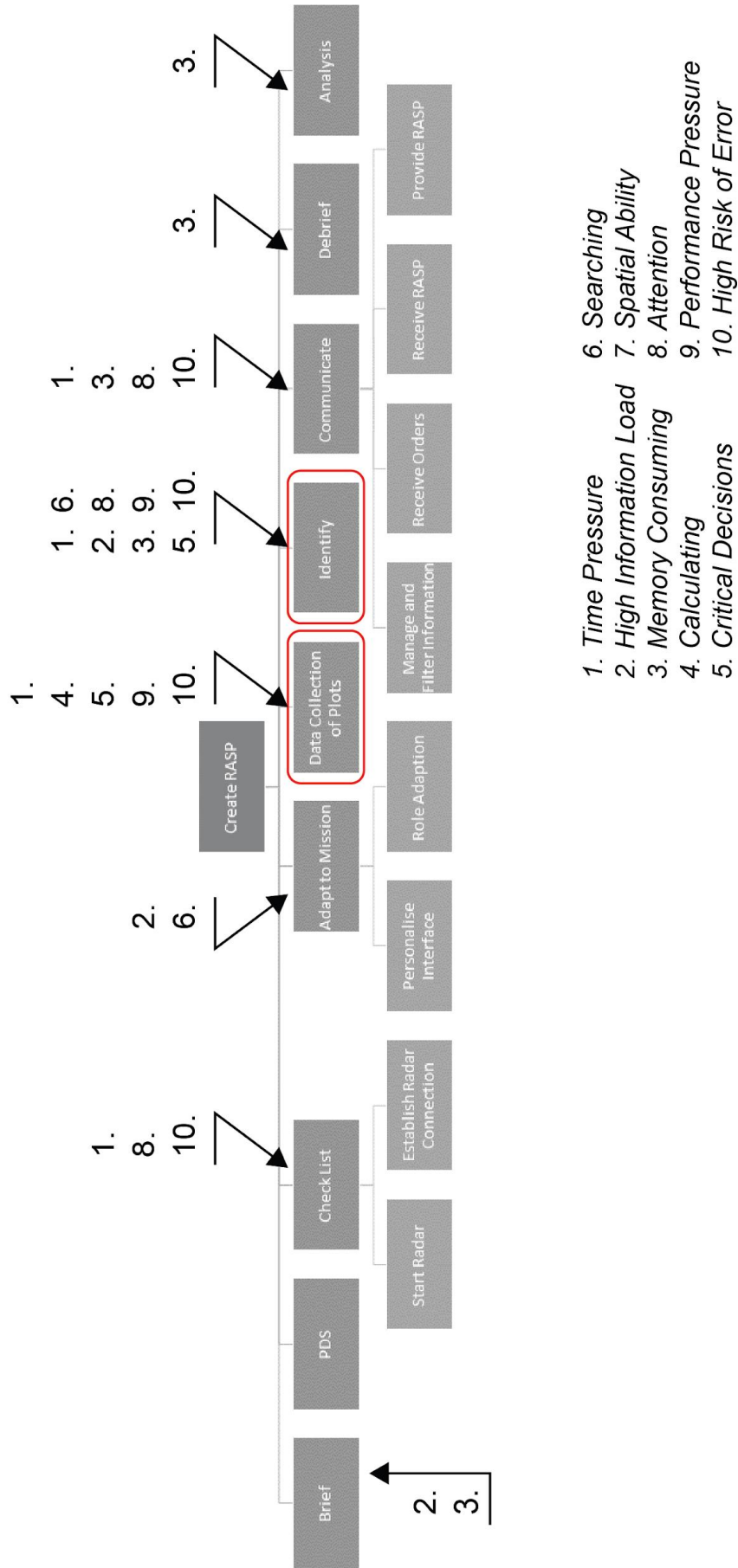


Figure 19: The results of the GTS for the Surveillance Operator.

To conclude, the following four tasks were assessed to having the highest workload, choose to elaborate further:

FC:

- Ensure safety.
- Guide fighter to target.

SO:

- Identification.
- Collect target data.

The chosen subtasks are presented in the following chapters.

7.3.3.1 HTA: Ensure safety

Figure 20 shows “Ensure safety” in a higher resolution.

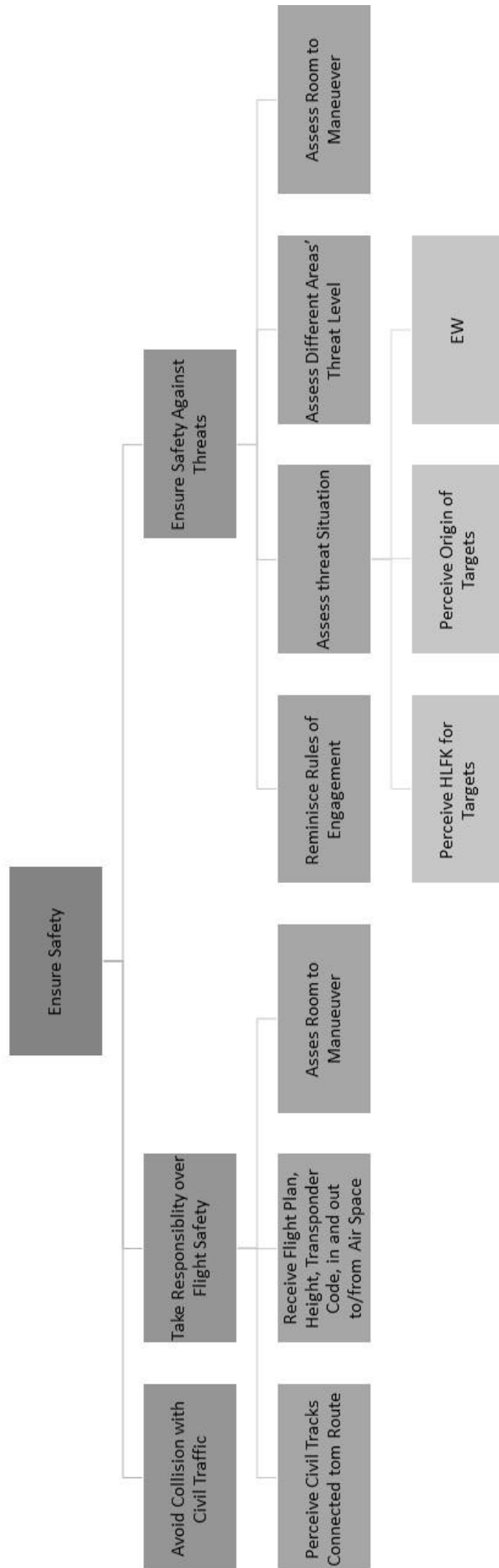


Figure 20: HTA of the task “Ensure safety”.

Ensuring the safety of the fighter jet as well as civil air traffic is the most important task of the FC. The task of ensuring flight safety involves both protecting the assets that the FC controls but can also involve separating own fighters from civil air traffic depending on what country it concerns. The process of ensuring flight safety is continuous over the entire mission and varies in workload over the mission, figure 20 displays this continuous process. But it is important to understand that the process is not linear for all cases.

The left-hand side of figure 20 shows the safety against collision with civil air traffic and the right-hand side shows the safety against threats for the fighter jets. The safety against civil air traffic is more predictable since operators have access to pre-planned flight routes. Based on these flight routes the best path for the fighter jet can be determined. Situations become more intricate when civil air traffic departs from their pre-planned route. Therefore, it becomes a part of the FC's task to compare live routes against flight plans which might change depending on for example weather (which may impact flight routes).

Ensuring safety against enemy threats is more unpredictable than civil air traffic safety. It involves reminiscing the rules of engagement (which are determined by a higher echelon within the military that determines what to do in different scenarios), assessing the current threat situation, determining what areas are dangerous and continuously adapting safe flight routes for the fighters. To be successful in doing this task it requires the operators to take multiple sources of information into account such as height, position, speed, direction, origin, and electronic warfare. Multiple inferences may be conducted based on either one or a combination of these factors and they need to be performed quickly.

7.3.3.2 HTA: Guide fighter to target

Figure 21 shows "Guide fighter to target" in a higher resolution.

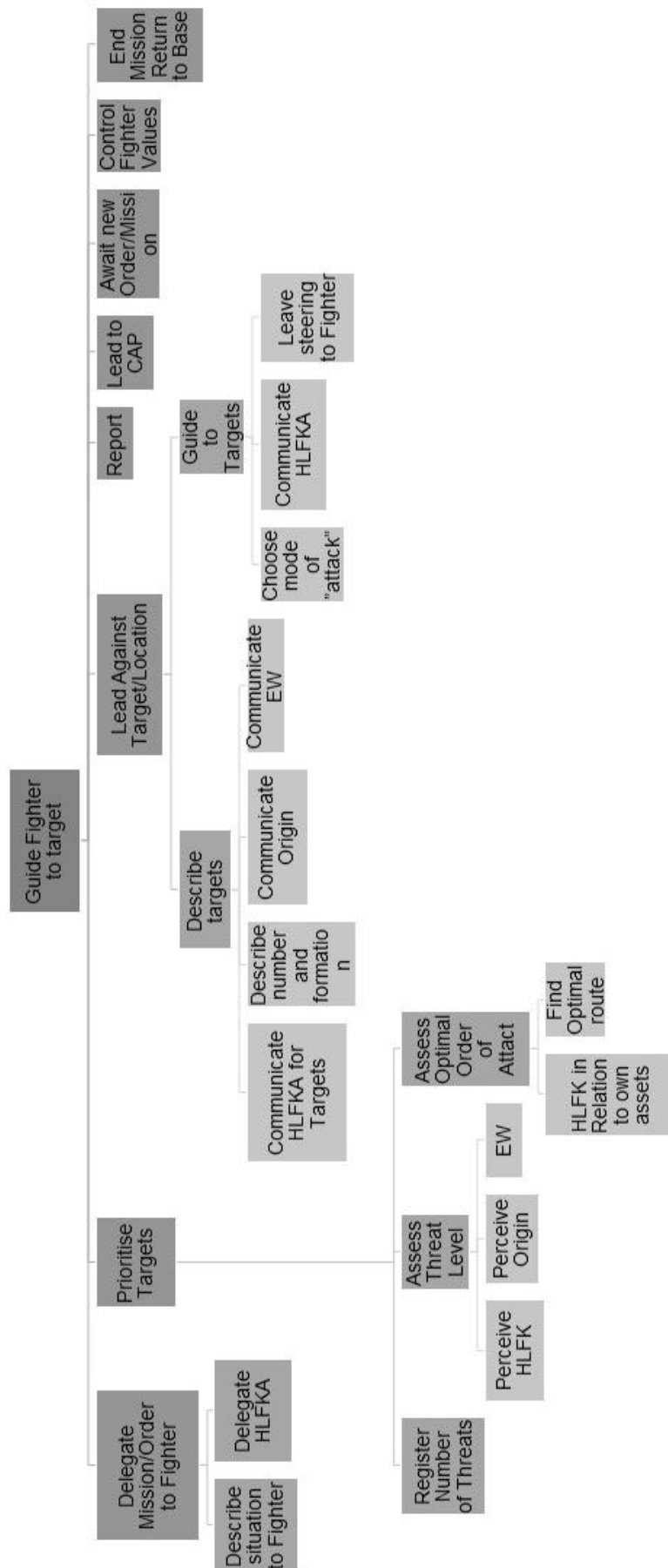


Figure 21: HTA of the task “Guide fighter to target”.

Guide fighter to target in short means to provide fighter jets with the right Situation Awareness (SA) and to lead them from point A to B. In addition to this it involves leading the fighters in a safe manner which is described in the previous section. Providing information that leads to SA involves communicating the SA acquired in the AEW&C operators team in an effective and efficient manner. Figure 21 displays this process which is, as ensuring safety, partly iterative. For example, new information may arrive during a fighter lead mission resulting in that the AEW&C updated their SA and need to communicate this to the fighter pilot.

A fighter lead mission can begin before the AEW&C is airborne, based on the mission brief and the rules of engagement. A mission might be to protect the borders so that enemy aircrafts do not fly over the national airspace, and the rules of engagement may state that if this happens the enemy aircraft needs to be led away. In a mission this may happen or there may be a risk that it may happen soon. Someone in the team detects this and the FC requests a fighter plane to execute the task according to the brief and the rules of engagement. The request is attested by the MCC and the MCC calls in a fighter. The FC starts by describing the general situation to the fighter and then HLFKA, which stands for height, location, speed, course, and intention for the target. The target at this point may be an enemy aircraft or a geographical point where the fighter idles before it gets new orders. The FC prioritizes targets parallel to this if there are multiple tracks. The prioritization relies on multiple factors such as the number of targets, HLFK-info about tracks and EW-information. The prioritization ends with the FC choosing the optimal route to attack multiple targets.

Guiding towards targets starts with describing the targets. Communicating this to the fighter pilot often starts with describing the HLFKA of the target to the fighter pilot. Communication is often auditory in the beginning of a fighter lead mission and transitions into text information when information becomes more specific. If there are multiple targets in a formation the FC needs to find the most efficient way to communicate this which may be to only describe a corner and the formation that the planes are flying in. In combination with this, information about the origin of the targets and EW information (that can indicate the types of planes) is added. After the targets are described, the FC guides the fighter to the targets. The guiding starts by choosing an approach vector towards the target/-s which may be the quickest path or a path that enables effective identification of suspicious tracks. The FC provides the chosen approach vector to the fighter pilot. Once the fighter pilot reaches a point where the target is visible, the FC does not need to provide further steering directions. The fighter pilot gathers information about the suspicious track and reports back to the FC who reports this to the MCC. When the information is gathered, the loop may start again if there are more targets left. The FC also needs to keep track of fighter values such as fuel levels which determines if the fighter needs to return to base or not.

7.3.3.3 HTA: Collect target data

Figure 22 shows “Collect target data” in a higher resolution.

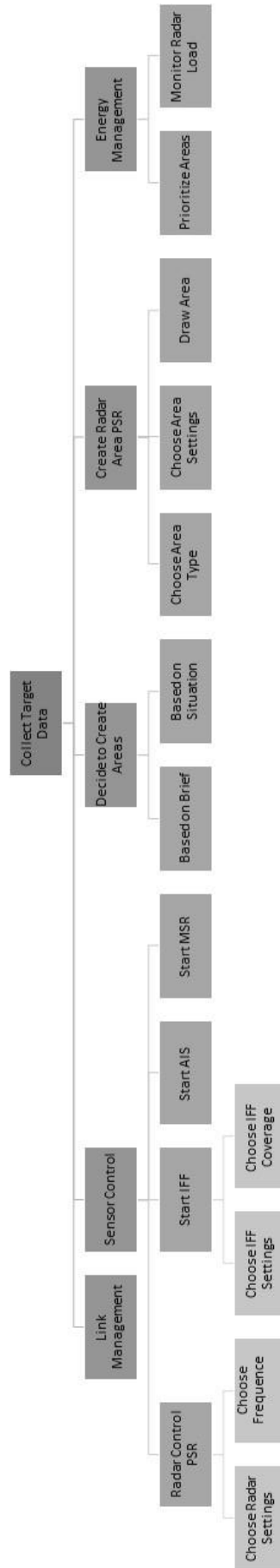


Figure 22: Collect target data.

The goal for this part of the HTA for the SO is to collect information about targets. The AEW&C is equipped with multiple different sensors and is also linked to other radars through the AOC. The different information sources result in plot information that automatically creates tracks. The role of the SO is to manage the different sensors to, on a higher level, create and provide RASP to the team and other units. Interesting targets are based on the brief and multiple other factors. High quality information enables better performance for the entire system.

The first task for the SO is to activate the links that provide the AEW&C operator team with additional radar information from other radar units. It also enables the system to fuse track data for a higher quality picture. Once the links are set up the sensors need to be started. The primary surveillance radar (PSR) is started first followed by IFF, AIS and MSR. The combination of sensors started in this step depends on the prerequisites set in the mission brief. If it for example is important not to be detected by the enemy, the PSR can be left out.

The next step is to decide to create areas which are both created based on the brief but if additional coverage is needed, more areas can be created. This step is also heavily connected to radar energy management. Good radar energy management enables higher quality tracks. The collection of target data is essential for the system goals.

7.3.3.4 HTA: Identification

Figure 23 shows “Identification” in a higher resolution.

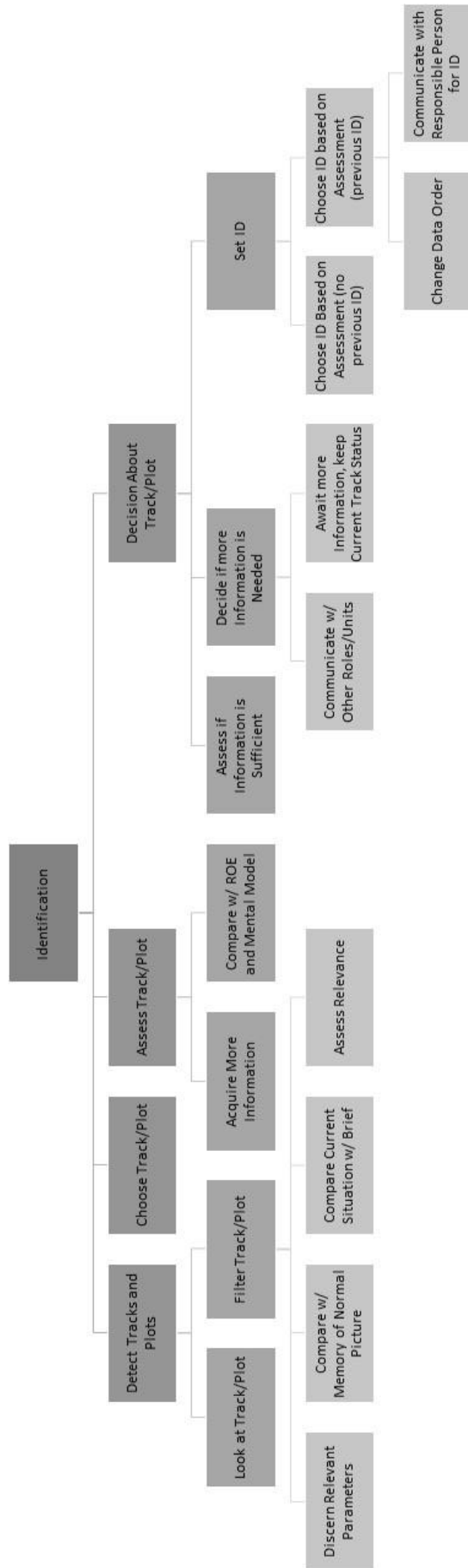


Figure 23: Identification.

As previously determined the identification is mainly done through the AOC. The HTA in Figure 23 represents when AEW&C identify specific tracks or in a situation where all identification is done in the AEW&C. Identification is often considered to be the most difficult thing that the SO does.

In the first step the SO needs to perceive tracks or plots. The tracks are then looked at and relevant information is discerned such as IFF/AIS data, direction, location, plot train, speed, and current ID. The SO compares this information with an internal mental model of the normal situation. The relevance of continuing to assess the track is assessed. Based on the assessment the track is chosen, if it is not relevant the process is repeated.

To further assess a track or plots more information is acquired such as height, EW, origin, and the track route. The new information is integrated and compared with previous information about the track and then compared with ROE and mental models of the situation. The SO determines if the information gathered is sufficient to identify the track or not. If not, the SO awaits, or requests, more information by the AOC. The SO awaits more information and communicates with other units through the MCC to get more information about the track. If the quantity and quality of the information is sufficient the SO sets the adequate identity for the track. If the track already has an identity a CDO needs to be done, i.e., a request to change the identity.

7.4 Chapter summary: Findings from the MTS activities

By conducting activities in the Mission Training System (MTS), observations of operative simulations and interviews regarding Situation Awareness (SA) and workload, data were obtained by using the analysis methods affinity diagram, Hierarchical Task Analysis (HTA) and Generic Task Specification (GTS).

The analysis resulted in the following results:

- Identified SA challenges for each role divided into seven categories for FC and six categories for SO.
- One general HTA for each role.
- An assessment of which two tasks had the highest workload for each role. The assumption was that opportunities to improve SA would be found in these tasks.
- A detailed HTA for the two tasks with the highest workload for each role, providing deeper understanding of task performance and the decision-making process.

The findings from the activities in the MTS provided deeper understanding of how the users work, their challenges, the tasks, and the decision-making processes.

All findings, both from the interviews and MTS activities, laid the foundation for starting to ideate and making concepts. The next chapter will describe the concept development process and the final concepts of AI integrated in the AEW&C system.

8

Concept Development

This chapter presents the results of the concept development phase. The concepts were based on the knowledge gained from the user studies and the first iteration of the AI guidelines (See Appendix B). The concept development had two purposes. First, it was to apply the knowledge gained from the user studies and address the challenges related to Situation Awareness (SA) for the Airborne Early Warning & Control (AEW&C) operator. Second, it was to apply and test the created AI guidelines, both their utility and validity. Thus, the concepts created in the concept development phase addressed both research questions, **when** and **how** to use AI to improve the Situation Awareness of the AEW&C operators.

The process and steps in the Concept Development are presented in chapter 4.4 *Concept development*, and the results that lead up to the final AI concepts described briefly in this chapter.

8.1 Ideation

The first step of the ideation phase was to ideate on functions using AI that could improve the operators Situation Awareness (SA) in the tasks identified as having the highest workload. The selected tasks Hierarchical Task Analysis' (HTA) were used as mediating tools and the findings from the interviews and activities in the Mission Training System (MTS) as a basis and motivation for the ideas. In total, the first step of the ideation phase resulted in roughly 70 separate ideas with some degree of similarities (See Appendix D).

The second step of the ideation phase was to identify similar themes of the ideas. The ideas were discussed regarding similarities and in total 18 different themes were identified.

The third and last step of the ideation phase was to evaluate the ideas and combine them into fewer groups of ideas, laying the foundation for concept creation. The criteria for evaluation were the estimated potential of improving SA and whether they were in the scope of the study, e.g., using AI technology. In total, this resulted in eight concept ideas, four ideas for each role. These were:

For Surveillance Operator (SO):

- Abnormality detection
- Area suggestion
- Timeline
- Pattern areas

For Fighter Controller (FC):

- Collision warning
- Communication (talk-to-text)
- Formation recognition
- Route planner

These concept ideas, together with the first iteration of AI guidelines, laid the foundation for creating the first iteration of concepts. The process of the ideation phase is described further in Chapter 5.11: *Ideation*.

8.2 Application of guidelines in concept development

In parallel with the user studies and ideation phase, the first iteration of AI guidelines was created. The process of making the AI guidelines was to first conduct a literature study. The literature included both general studies of AI, for example definitions of AI, ethical considerations of AI, and capabilities and limitations of AI. The literature study also included literature more applied to the AEW&C context, for example AI in military operations, explainable AI, and Situation Awareness related to AI. The main sources used are listed in Appendix A.

Quotes were extracted from the literature and analysed by the method affinity diagram. The affinity diagrams resulted in three larger categories of AI guidelines related to when they were most applicable in the development process. The larger categories were also divided into subcategories. The first draft of AI guidelines is available in Appendix B.

The first iteration of the AI guidelines was used as a foundation in developing the eight concept ideas into concepts with a higher fidelity. Guidelines that were relevant to the specific idea group were identified and applied to the specific group, meaning that not all guidelines were used for all the concepts but rather the most relevant.

8.3 Concept Creation

In this phase the ideas generated in the ideation activities were combined into eight concepts. The concepts were improved using the first iteration of the AI guidelines. By using the guidelines, it could also be seen which guidelines were easy or difficult to use.

In this phase the concept ideas gained from the ideation phase were further elaborated and visualized. The findings from the user studies and the first iteration of the AI guidelines laid the foundation of the concept creation. By applying the AI guidelines to the concept creation, it was possible to assess whether the guidelines were useful in the design process of AI systems.

The first step of the concept creation phase was to make sketches of wireframes for each concept. This resulted in roughly 5 wireframes per concept with low fidelity (See Figure 24).



Figure 24: Examples of wireframes sketches of concepts.

The second step was to make digital visualizations of the concepts, slightly more detailed but still at a low fidelity level. The tool mainly used for this phase was Microsoft PowerPoint (2018). An example of a concept made in this step is presented in figure 25.

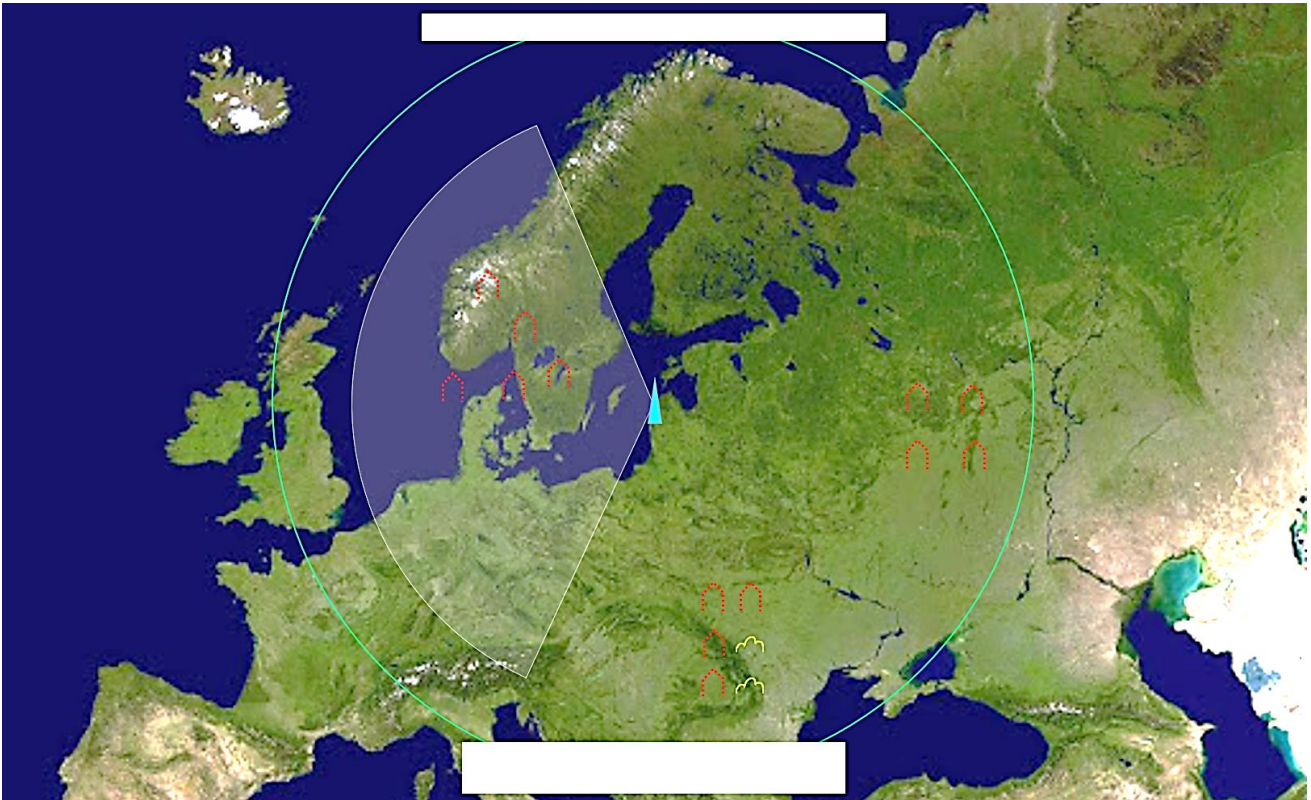


Figure 25: An example of the low fidelity concepts made in Microsoft PowerPoint.

The third step was to evaluate the concepts. By directly evaluating the concepts, the AI guidelines were also indirectly evaluated by seeing how well the intended effect of the guidelines had in practice. The evaluation was in the form of a digital presentation, mainly to engineers and project managers within the organization the study was conducted at.

The fourth and final step of the concept development phase was to improve the concepts based on the feedback from the first evaluation. The concepts were further elaborated in image editor tools like Adobe Illustrator (2019) by increasing the fidelity of the concepts, both by more clearly trying to achieve the guidelines and fulfilling the needs of the users and by making the concepts look more authentic.

The final concepts of how and when AI can be used to improve the SA is presented in the chapter 9: *Final concepts*.

9

Final Concepts

In this section the eight concepts produced in the study are presented. First the background that motivates the existence of the concepts is described and then the most important functionality is described.

9.1 Final Concepts for Surveillance Operator

In this section the four final concepts created for the Surveillance Operator (SO) are presented.

9.1.1 Abnormality Detection

The following sections present the background that motivated the concept and the description of the concept.

9.1.1.1 Motivation Background

An important aspect of the user studies was to examine if the operators were able to find abnormalities. In scenario 2, two abnormalities were added. The result was that one abnormality was partly found and the other one was found after reminding the operators to search in a specific area. A conclusion of the results is that it is very difficult to detect abnormalities if you do not know where to look or what you are looking for. It is like looking for a needle in a haystack, only you do not know where the haystack is and that you are looking for a needle.

It is especially challenging to detect abnormalities for sea- and land tracks. This is e.g., due to:

- Large amount of information, there are much more plots and more clutter.
- Less restrictions and more unpredictable behaviour.
- A wider array of types of vehicles including size, owners etc.
- Less information about the tracks, like transponders and flight routes.

Abnormalities of air tracks can also be hard to detect. The operator often looks for the most apparent information like height, location, speed, course and if they have a transponder. Then they compare the information to their view of the normal situation picture and the rules of engagement. In most cases, that is fine. However, it is still hard to detect air tracks that are using false transponders or starting to deviate from their route, which in rare cases may happen with air tracks. Another challenge is that operators pay less attention on already identified tracks. A track can behave normal, be identified as neutral, and then start to behave abnormally unnoticed.

9.1.1.2 Description of Concept

Figure 26 below shows the final visualization of the concept *Abnormality Detection*. All visualized frames of the concept are available in Appendix E.

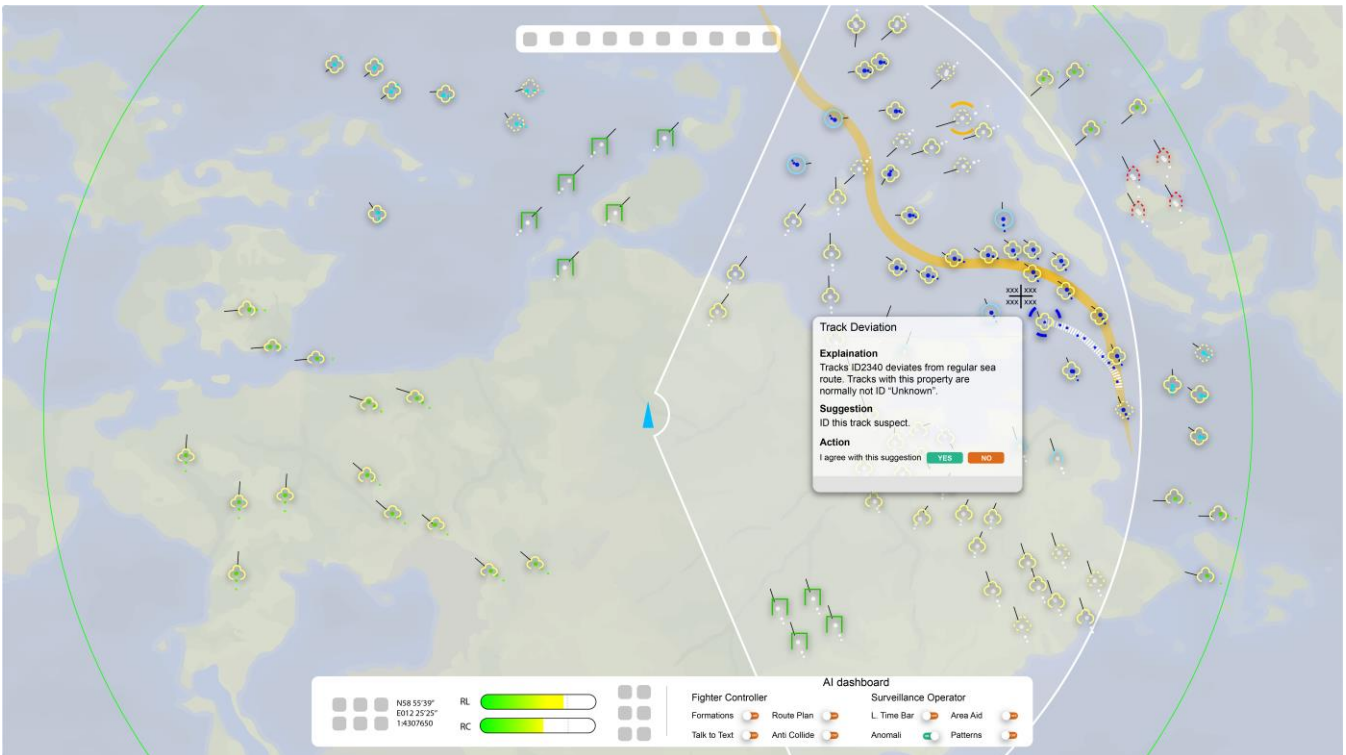


Figure 26: The final visualization of the concept Abnormality Detection.

When the operator activates the abnormality detection function the AI starts searching for abnormalities. What abnormality is based on could for example be:

- **The mission description (brief):** The AI can read the brief and make conclusions based on what is written. For example, it might be stated that boats in a specific area from a specific country should be considered suspect, or to be extra attentive to flying targets with a specific origin. The AI will then consider this when highlighting abnormalities.
- **Normal situation picture (Height, location, speed, and course):** Based on previous data the AI has learned what the normal situation picture looks like and can detect deviations from this. It might be that a track moves abnormally fast, at a low height, in an unusual or restricted area or deviates from the assumed route. However, for the AI to be useful it would be necessary for it to also take into consideration events that would cause many tracks to behave in an abnormal way, for example storms, blocked canals, or closed roads etc.
- **Missing or suspicious transponder information:** All planes and most larger boats have transponders that send information. This information could for example include what type of vehicle it is, how large the vehicle is, the owner, the origin etc. The AI can detect if planes or larger boats do not have a transponder or if the transponder information corresponds with the actual data collected by the radar.

When the AI detects an abnormality, the track or plot gets highlighted with two moving orange indications. The visualization of the highlights is based on the need for it to stand out in the existing interface. Orange is not used in the existing interface, and in addition, it lies between yellow (the colour of unknown tracks) and red (the colour for suspect and hostile tracks) which suggest it could be in accordance with the operators' mental model of how an abnormal, and potentially suspect or hostile, track would look like.

When the operator selects an abnormal track an information window appears. The information window is structured the same for all abnormalities. It consists of four parts:

- **Status:** Explains what type of abnormality it is. For example, "Track deviates from route".

- **Explanation:** The AI explains how it made its conclusion about the status. For example: “Flight XXX deviated from the predicted/reported flight route from X to Y at XX:XX.”. The explanation is also visualized on the map by displaying the predicted flight route and the route the track took.
- **Suggestion:** The AI suggests an identity for the track based on the type of abnormality and how tracks with similar behaviour have been identified historically.
- **Action:** Here the operator can take an action. He/she can choose to agree or disagree with the suggested identity by the AI. The AI can use this feedback to learn and improve further suggestions. If the operator is unsure, he/she can choose to snooze, to postpone the action until later. Then the abnormality indication disappears for a couple of minutes and reappears again if the track still behaves abnormal. If the operator just chooses to close the window, the indications will remain.

9.1.2 Area Suggestion

The following sections present the background that motivated the concept and the description of the concept.

9.1.2.1 Motivation Background

The area suggestion tool was based on the challenges related to radar management, especially radar areas and radar status. The time pressure to quickly establish a RASP in the beginning of the mission in combination with the fact that radar areas were mostly placed based on the mission brief implied that it could be beneficial for the SO’s SA to suggest areas. The concept also considers the challenges related to managing radar capacity in an optimal way. This is achieved both by visualizing the connection between radar areas and radar load and suggesting changes of existing areas based on the current situation picture.

9.1.2.2 Description of Concept

Figure 27 below shows the final visualization of the concept *Area Suggestion*. All visualized frames of the concept are available in Appendix F.

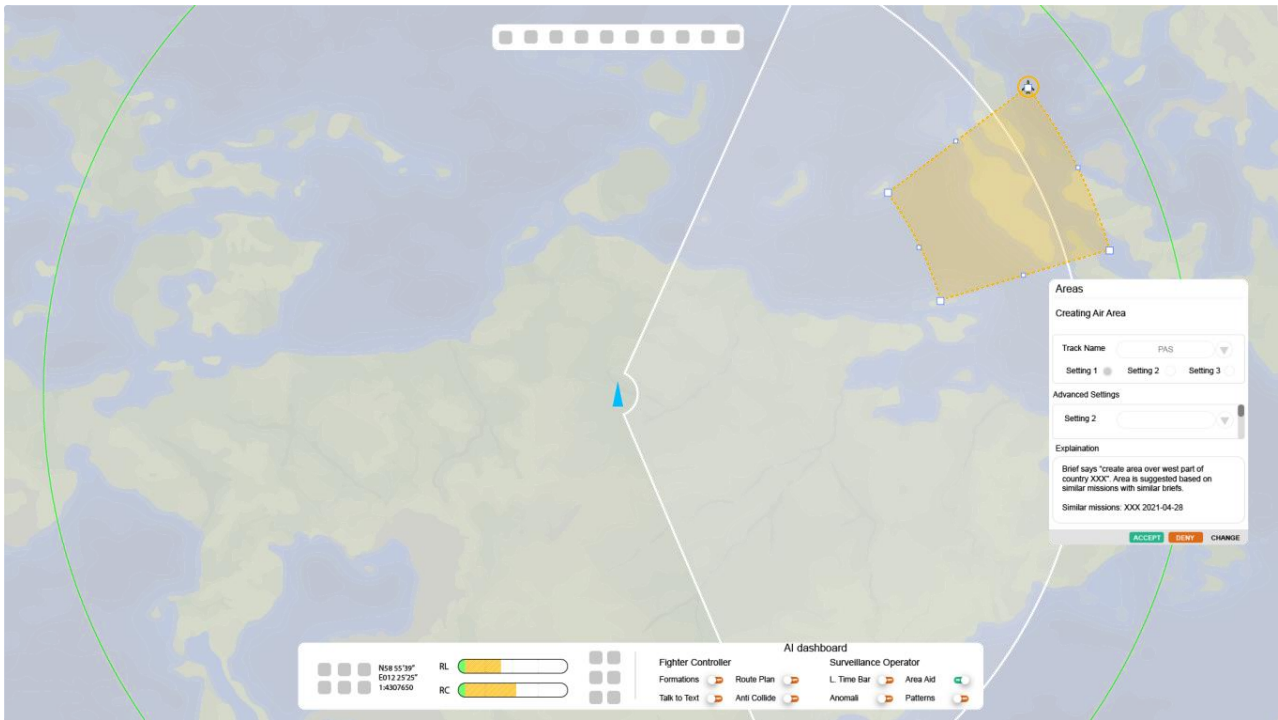


Figure 27: The final visualization of the concept *Area Suggestion*.

When the area aid tool is activated, the AI suggests radar areas to place on the map. Initially, the suggestions are based on what is written in the brief (mission data). The AI reads the brief and suggests areas. For example, if the brief says: “look for planes from country X in area X” the AI suggests placing an air area in the area mentioned in the brief. The AI also takes into consideration the current brief in comparison with previous briefs of similar missions and suggests areas similar to the ones used before.

When an area is suggested it is displayed with an orange dotted shape on the map. An icon indicates the type of area. In the vital signs field, there is a preview of how much the suggested areas would affect the radar load and radar cycle time.

When the operator selects the area, it is highlighted, and an information window appears. The selected area’s impact on radar load and radar cycle time is also highlighted. In the information window the area’s suggested settings are displayed as well as an explanation to why the area was suggested. The operator gets the choice to accept, deny or change the suggested area. The AI will learn based on the feedback to give better suggestions in the future.

The AI is not only limited to suggesting new areas, but it could also suggest changes or to remove already existing areas. This is based on the current situation picture. For example, the AI might suggest increasing the priority of an area with high activity and suspects, expand areas to get better coverage or to remove areas that have not been looked at for a while.

9.1.3 Timeline

The following sections present the background that motivated the concept and the description of the concept.

9.1.3.1 Motivation Background

A timeline functionality was requested by the operators who wanted to be able to look at “replays” during the mission to be able to, for example, see where tracks originated from and get a better understanding of movement patterns. The need was also confirmed by the results of the user studies. For identification and detecting abnormalities it was significantly more challenging with sea-and land tracks, partly due to their slow movements. The ability to pan back and forth in a timeline facilitates seeing movement patterns that happen over longer periods of time and being able to detect tracks or plots with abnormal behaviour.

The concept also deals with the risk of forgetting past events and makes it possible to look at past events that were not paid attention to at the time being. This can be useful in communication, to be able to go back and look at events to get a better understanding of what the counterpart describes. Proactive work was considered important to stay in the loop and having a sufficient level of situation awareness. AI can learn from historic data to predict likely future events. Therefore, it was considered interesting to suggest the ability to extend the time bar into the future, to be able to see what the most likely situation picture would look like in advance.

9.1.3.2 Description of Concept

Figure 28 below shows the final visualization of the concept *Timeline*. All visualized frames of the concept are available in Appendix G.

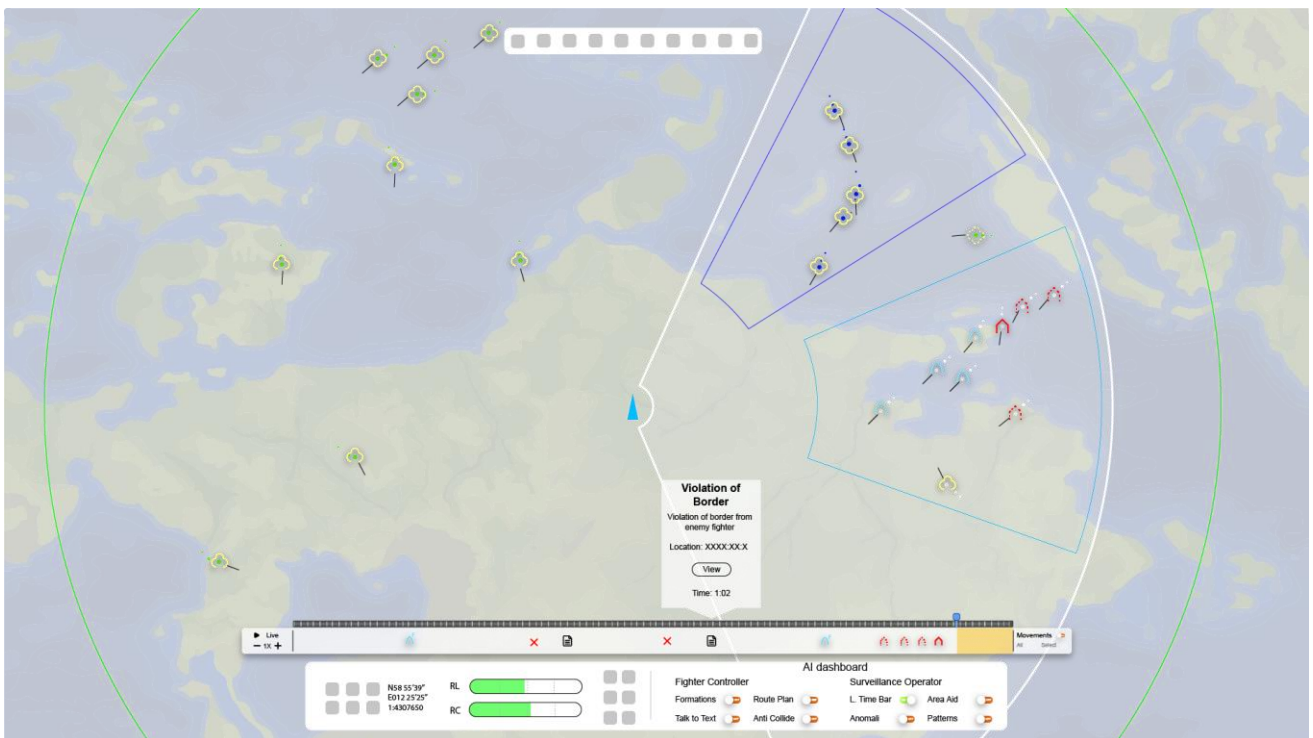


Figure 28: The final visualization of the concept *Timeline*.

When the operator selects the timeline function, he/she enters a “timeline mode”. The system has recorded the situation picture from the beginning of the mission, which allows the operator to look at replays of different sequences of the mission. The operator navigates the recording using a slider. The operator can play and pause the recording and change the play speed. The movements of plots and tracks are often quite slow, and it can be useful to be able to look at events at a higher speed. It is important to keep in mind that if the operator spends too much time looking at previous events, he might lose track of what is happening in real time. This might harm his situation awareness and force him to try to catch up and work reactively instead of proactively.

Along the timeline, small icons are displayed to indicate events. This makes it easier and faster to find events of interest. The events might be, for example, when a track is identified as suspect, when you start a mission with a friendly fighter, violations of borders and reports. When the operator hovers with the cursor over an event symbol he gets some brief information about the event, like the name of the friendly fighter, where it started from and the mission-and real time. The event is also shown on the map by displaying the track at the location the event occurred.

Not only can the operator look at previous events, but the timeline function can also be used to look at the predicted situation picture in the near future. The AI makes an estimation of how the plots and tracks will move based on their previous movement pattern, speed, and course. By doing this the operator can predict potential events like violations of borders and pay extra attention to the tracks related to such potential events. The operator can choose to display the movement of tracks, where historic movement is displayed as a white line and the predicted movement as an orange dotted line. This can be done for a selection of tracks or for all tracks.

9.1.4 Pattern Areas

The following sections present the background that motivated the concept and the description of the concept.

9.1.4.1 *Motivation Background*

This concept mainly addressed the need to have a well-developed understanding of the normal picture. The operators compare the current situation picture to their mental model of the normal picture when detecting abnormalities and identifying tracks. Visualizations of patterns derived by AI was an idea to facilitate better understanding of the normal situation picture. It also draws inspiration from old methods used to make pictures of movement patterns by taking long exposure photos on the radar display.

9.1.4.2 *Description of Concept*

Figure 29 below shows the final visualization of the concept *Pattern areas*. All visualized frames of the concept are available in Appendix H.

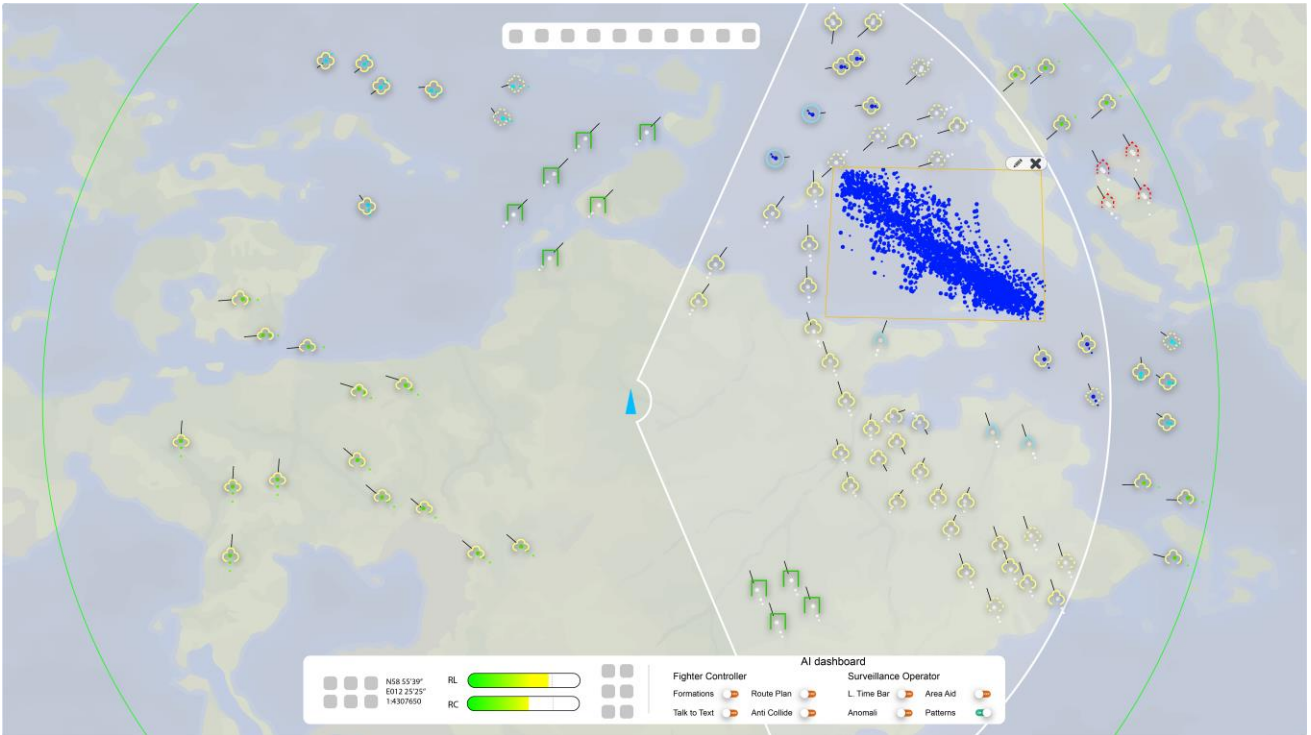


Figure 29: The final visualization of the concept *Pattern areas*.

When the operator activates the pattern function, the pattern window appears. Preferably, the operator starts with selecting in what area they want to apply the pattern. This is preferred because the area adapts to the type and settings the operator applies to the area, giving them direct feedback on their choices. Placing a pattern area works like placing a radar area, the operator selects the corners of the area and finishes the area by pressing “enter”.

Then the operator gets to choose what type of pattern they want to use. They can choose between:

- **Heat map:** Shows activity of areas by using colours, ranging from blue (low activity) to red (high activity).
- **Routes:** Shows what routes the plots typically take, can be flight routes, sea-lanes, and roads.
- **Long exposure camera:** Shows an image of the situation picture over time, like a camera with long exposure time.
- **Connecting plots:** The AI estimates what plots belong to the same target and connect the plots with lines.

The operator can choose between making the area **map fixed** (the area will be fixed to an area on the map, independent of the window) or **window fixed** (the area will be applied on the window as a lens and adapt when the operator pans and zooms on the map). When the type of pattern is selected, the operator gets to choose settings for the pattern. They choose the timeframe of the data the pattern will be based on, ranging from live to minutes or all the way to the beginning of the data collection. The operator can choose what quick command they want to use to show or hide the pattern area, for example by clicking “ctrl” + “x”. This makes it easier for the operator to toggle between displaying the area and hiding it to get a cleaner image. The operator chooses what type of data (air, sea, and land) they want to use to create a pattern. For example, they might only be interested in displaying land movement. Then they can choose to only select the “land”-box, and no air-or sea data will contribute to creating the pattern. In the pattern window, the areas are organized in tabs (like web browsers). To create an additional pattern, click on the “+” button next to the tab. If an area is hidden on the map, and the operator has forgotten the quick command to show it, they can find the area by selecting the pattern function and go to the tab of the pattern area.

9.2 Final Concepts for Fighter Controller

In this section the four final concepts created for the Fighter Controller (FC) are presented.

9.2.1 Collision Warning

The following sections present the background that motivated the concept and the description of the concept.

9.2.1.1 Motivation Background

The idea behind this concept originated from the importance of ensuring safety towards civil traffic. Findings from the user studies suggested that there were risks related to this task. The risks mainly concerned having too much focused attention on the object to guide towards, impairing the overall situation awareness. Another risk was being too relaxed and not paying enough attention when guiding the fighter jet back from a mission. AI is suitable to handle these kinds of data intensive tasks and the model used for this kind of functionality could be quite simple.

9.2.1.2 Description of Concept

Figure 30 below shows the final visualization of the concept *Collision warning*. All visualized frames of the concept are available in Appendix I.

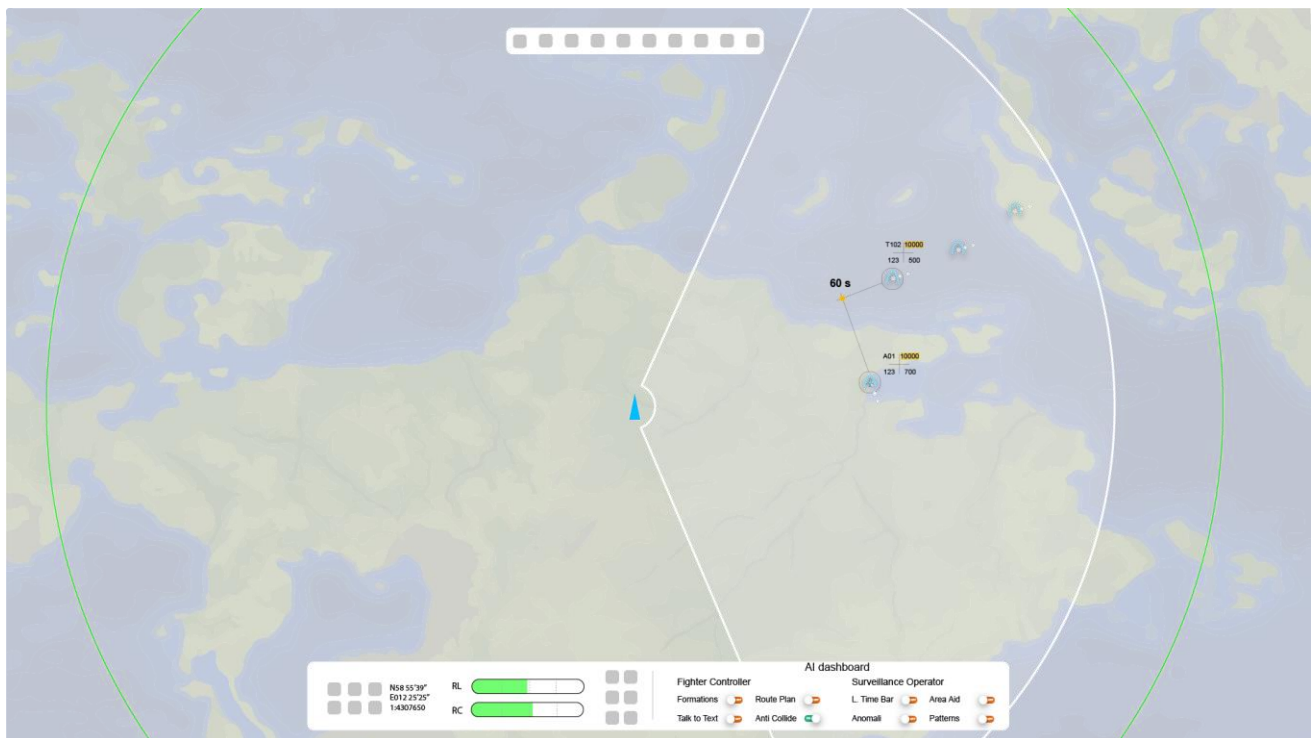


Figure 30: The final visualization of the concept *Collision warning*.

The collision warning function is a tool that supports the FC to ensure flight safety of friendly fighters in relation to civil aircraft. When the function is activated it starts to scan the environment for any potential collisions between friendly fighters and civil aircraft. Based on their movement pattern and current height, location, speed, and course the AI makes a prediction of where their paths might intersect, and collisions might occur.

When there is a risk for a potential collision it is highlighted on the map by displaying the predicted movement of the related tracks as lines and the intersection of the lines as an orange flashing dot.

When the operator clicks on one of the lines or the dot, information about the tracks is displayed as well as the time remaining for intersection. The operator can use this information to guide the friendly fighter to a safer flight route, for example by telling the fighter to decline or rise.

9.2.2 Communication (talk-to-text)

The following sections present the background that motivated the concept and the description of the concept.

9.2.2.1 Motivation background

This concept originated from the challenges related to communication. Especially the communication between FC and fighter jet needs to be clear, concise, and up to date. The FC took handwritten notes to cope with the challenges of remembering critical information and that repetition was required to confirm. The importance of teamwork was a recurring theme during the user studies. Since the MCC communicated with most persons and was responsible for the success of the mission and safety aboard it made sense to make a concept that supported that role as well. High workload, stress and tiredness were factors affecting the operators during a mission, impairing their situation awareness and performance overall. Talk translation using AI is well tested and should not be too challenging to implement. Eye tracking is an established method to measure workload, stress, and tiredness.

9.2.2.2 Description of concept

Figure 31 below shows the final visualization of the concept *Communication (talk-to-text)*. All visualized frames of the concept are available in Appendix J.

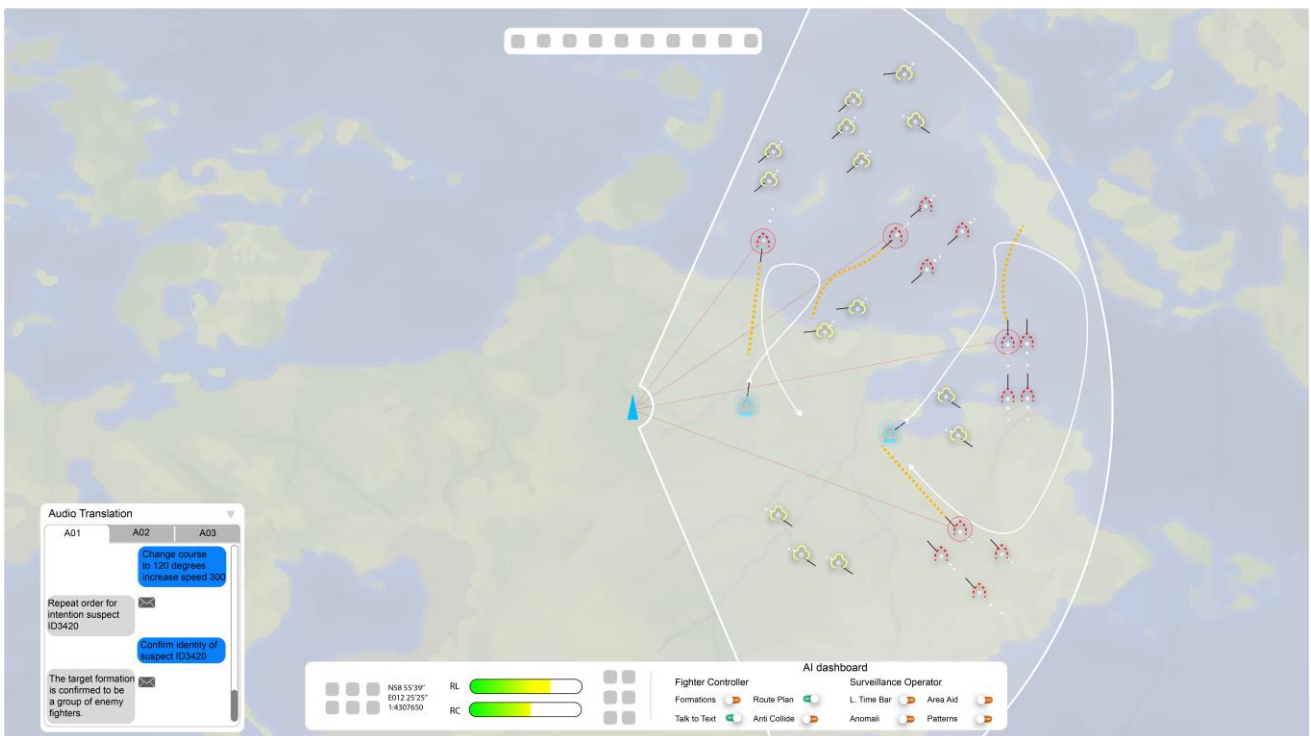


Figure 31: The final visualization of the concept *Communication (talk-to-text)*.

During a mission, when the FC and a fighter pilot is communicating, it is important that they can hear and understand each other. There are also moments when it is necessary for the FC to remember what the fighter pilot says, for example when reporting an identity of a suspect to the mission crew commander.

When on a mission, the FC can activate the talk-to-text function. When this function is active, the AI translates what the fighter pilot says to text. The translated text is displayed in real-time in a bar, in this example to the bottom left of the FC's screen. This implies that the communication from the fighter pilot reaches the FC in dual modalities, audio, and text.

The text bar can be expanded to display the whole conversation between FC and fighter pilot. This might be useful when the FC needs to remember something in the conversation, for example when reporting something to the Mission Crew Commander (MCC). It is also possible to click on a sentence and send it forward directly to the MCC. In case of guiding several fighter pilots at the same time, each conversation is in its own tab that the FC can toggle between.

Implication for the Mission crew commander:

Among all crew members, the MCC communicates the most. The MCC communicates with team members, the air operation centre, the AEW&C pilots etc. He is also responsible for the success of the mission and the safety on board. Therefore, it would make sense if he could make use of a talk-to-text function as well.

The talk-to-text function for the MCC is like the one for the FC. The difference is that he has one text bar for each operator, and he can see the current workload of the operators. Seeing the operators' workload could make it possible to help other crew members proactively and having workload in consideration when delegating tasks. Eye tracking could also be used to measure tiredness. If an operator gets tired an icon could indicate that and the mission crew commander could suggest the operator to take a break or a coffee.

9.2.3 Formation Recognition

The following sections present the background that motivated the concept and the description of the concept.

9.2.3.1 Motivation background

Formation recognition of suspect tracks was a requested AI functionality. User studies suggested that communicating the situation picture to the fighter pilot is an important and challenging task. AI could help assess situations early and explain it to the FC who could communicate it to the fighter pilot. A formation recognition tool could also be useful in the identification of suspect tracks by assigning an unknown track to a formation of suspects, which would imply the unknown track is also suspect. AI is recognized for being well adapted to recognizing images and patterns, which could promote the users trust in the AI's capabilities.

9.2.3.2 Description of concept

Figure 32 below shows the final visualization of the concept *Formation recognition*. All visualized frames of the concept are available in Appendix K.



Figure 32: The final visualization of the concept *Formation recognition*.

The formation recognition function mainly has two purposes:

1. To recognize which tracks that belong to a formation.
2. To recognize the shape of a formation that suspect tracks flies in to support the FC in describing the situation picture to the fighter pilot.

When the formation recognition function is activated, it starts scanning the situation picture for possible formations. When a track is identified as a suspect the system automatically puts more radar resources around the track. This is to improve the systems capabilities to find tracks close to the suspect. For example, a track is identified as suspect, and a moment later five other unknown tracks emerge close to the suspect track.

Based on the behaviour of the tracks close to the suspect track, the AI can conclude what tracks are a part of the formation. The AI highlights the formation by outlining the shape of the formation with an orange line as well as filling the shape with a transparent orange colour. When the operator clicks on the formation, a window appears with information about the type of shape, the number of tracks in the formation and the height, location, speed, and course of a point in one of the corners of the formation (indicated with an orange dot).

In the information window the operator can click “more information” to expand the window and get additional information. Here the name and identity of each track is displayed, possibility to move the data point and to give feedback to the AI whether the formation was helpful and being able to hide the formation.

9.2.4 Route Planner

The following sections present the background that motivated the concept and the description of the concept.

9.2.4.1 Motivation background

A route planner for FC was one of the requested AI functionalities by the users. During the user studies the FC had to prioritize objects to identify and routes to take. Since the situation was changing the FC had to be quick in assessing the situation, prioritize and choose a route. This is cognitively demanding, especially when there are several objects to consider. The FC had to work dynamically due to the movements of the objects. AI has the possibility to evaluate a tremendous number of alternatives and suggest the routes based on the goals of the mission. It could also consider threats and guide accordingly.

9.2.4.2 Description of concept

Figure 33 below shows the final visualization of the concept *Route Planner*. All visualized frames of the concept are available in Appendix L.

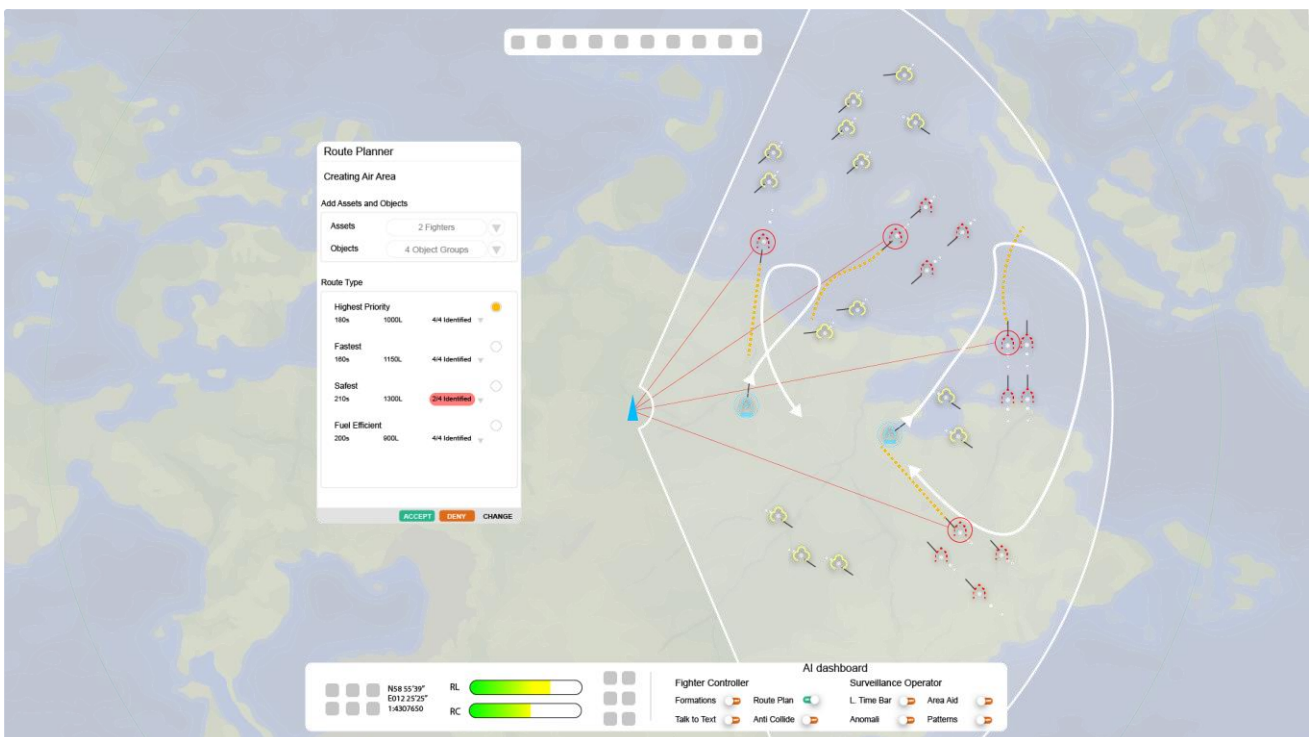


Figure 33: The final visualization of the concept *Route Planner*.

When the operator activates the route planner function the route planner window appears on the screen. First, the operators select the assets (friendly fighters) he wants to guide, and the objects (suspect or hostile tracks) he wants to guide the assets to. In this case, there are two assets and four objects.

Then the operator chooses what type of guiding he wants to do. The type of guiding depends on the mission. If the mission is to identify a suspect flying target, a stern might be suitable (to approach the object in a curve from behind and fly beside the object until identification is possible). For another mission a cut-off might be more suitable, where you fly in a straight line to where the assets and object will intersect. In advanced settings, the operator can choose what type of guiding for each object. In this case the operator chooses sterns for all objects.

When assets, objects and guiding types are selected, the AI suggests possible routes. The suggestions are based on height, location, speed and course of assets and objects but also considers prioritisation of objects, mission time, safety & threats, and fuel consumption. The route suggestions are listed in the route planner window. For each suggestion the total mission time, fuel consumption and goal

achievement (how many objects possible to guide to) are presented. Each suggestion in the list can be expanded to display more information where the operator can see mission data for each asset. The four route suggestions are:

- **Highest priority:** In this suggestion, the AI assesses the prioritisation of the objects based on their threat level. The threat level is mainly based on the location, speed, course, and possible EW-data of the object. When expanding this suggestion in the list, the operator can change the prioritisation of objects if he disagrees with the AI. The AI can use this feedback to learn and improve further suggestions.
- **Fastest:** In this suggestion, the AI calculates and suggests the route with the lowest possible mission time.
- **Cautious:** In this suggestion, the AI considers the threat of objects and areas and guides assets to avoid any possibly dangerous situations. Danger zones are displayed on the map. If the operator disagrees about the danger zones, he can disregard them. The AI can use this feedback to learn and improve further suggestions.
- **Fuel-efficient:** In this suggestion, the AI calculates and suggests the route with the lowest possible fuel consumption.

When selected, the route suggestions are displayed on the map, showing what assets are assigned to what objects and what route they should take, the predicted movements of the objects, the mission time and fuel consumption.

9.3 Chapter Summary: Final concepts

In this chapter, eight different concepts were described. Four concepts were aiming to improve SA for the Fighter Controller and four concepts were aiming to improve SA for the Surveillance Operators. The concepts were the result of the second iteration of concepts in the concept iteration phase. These concepts were also based on the first iteration of AI guidelines created in the first AI guidelines phase.

In the next chapter, *Evaluation*, both the concepts and the guidelines are evaluated.

10

Evaluation

In this chapter the result from the evaluation of the eight concepts and the AI guidelines will be presented. First, the evaluation of the eight concepts is presented. Second, the evaluation of the guidelines is presented.

10.1 Evaluation of Concepts

In this chapter the results from the evaluation of the eight concepts are presented. First, a short recollection of the basic functionality is presented. Second, the comments from the operators are summarised. Finally, the main issues with the concepts are summarised as further works.

10.1.1 Collision Warning

Collision warning is a concept that helps fighter controllers to see if their assets run a risk of colliding with civilian planes through data-analysis of flight patterns.

10.1.1.1 Comments from Operators

The collision warning system is a functionality that is requested by the operators. However, the operators required more information pertaining to what factors the suggestions for collisions were based on. The operators' expectations of the concept were that it did not have AI functionality meaning that it was merely based on passively aggregated information. Therefore, the concept's need to adhere to the two requirements "manage user expectations" and "make capabilities clear" more distinctly.

10.1.1.2 Further Works

Since there are already similar systems in use for other industries such as in air traffic control more benchmarking is required.

10.1.2 Communication (talk-to-text)

This concept helps the fighter controller to keep track of what has been said by the fighter jet pilot.

10.1.2.1 Comments from Operators

Providing that the tool works well the operators found it to complement human skills well since remembering earlier conversations is difficult and takes up attention. It was clear to the operators what the functionality of the tool was, meaning that the guideline "make capabilities clear" was achieved.

10.1.2.2 Further Works

The visual representation of the tool needs to be determined more clearly.

10.1.3 Formation Recognition

This concept helps operators distinguish formations that potential enemy threats fly in. This facilitates the communication with their own assets and may help identifying unknown targets using guilt by association.

10.1.3.1 Comments from Operators

The formation recognition concept was well received by the operators. One of the reasons mentioned was that the task of recognizing formations is a data intensive task because cautious filtration of irrelevant tracks is needed. There are often several tracks in the vicinity of formations making it cumbersome to distinguish formations among the clutter. This also means that the AI aids humans in decisions and complements where humans may be inferior to machines. While formations may be difficult to find, operators still get a chance to confirm or deny AI-suggested formations, thereby keeping humans in the loop. In addition, the concept is directly influenced by user suggestions

derived from user studies making the concept clearly connected to user norms. This also enables the function to communicate information on a high level that the users find familiar. The functionality where the AI suggests that unknown tracks to be a part of a formation is closely related to operators' current mode of operation where tracks adjacent to a formation could be assigned as threats, which is called "guilt by association".

10.1.3.2 Further Works

Operators stated that the concept could be using additional information such as radial speed when assessing if tracks are a part of a formation. Further work with this concept involves determining which additional information that is required. In addition, as with all concepts, the system should be tested to ensure that humans and AI are collaborating effectively.

10.1.4 Route planner

The route planner is helping fighter controllers to plan the optimal route for identifying potential enemy threats.

10.1.4.1 Comments from Operators

This concept bears some resemblance to what operators have seen before. However, the functionality was not clearly displayed during the demonstration and thereby all the guidelines could not be spotted by the operators. Operators liked the functionality where fuel levels of their assets are displayed.

10.1.4.2 Further Works

The concept needs to be more clearly visualised.

10.1.5 Area Suggestion

The area suggestion concept helps the surveillance operator to create and manage effective and efficient radar areas.

10.1.5.1 Comments from Operators

Operators found this function to not efficiently complement human skills since the aid received from the AI functionality did not help with a difficult task. Using a better interpretation of expert knowledge may have facilitated a functionality more relevant for an increased system performance, i.e., finding what is actually difficult in this step of the process. A better use of systems thinking is necessary since operators found that the concept did not help them with something that is difficult in the current system. In addition, it may be cumbersome to use the brief as a foundation for the area suggestion since there is not always a brief. Other parts of area management may utilize AI functionality to improve system performance. There are also other factors apart from the brief, if there is a brief, that are relevant as input to the AI. Since the context is dynamic the AI needs to take dynamic data into account as well. From the discussion about this concept, operators found it unclear what the information was based on which suggests that the guideline about traceable decision processes were not reached. This is also connected to the reliability and trust of the AI because the information that the area suggestion is based on is influenced by the quality of the brief. Utilizing users' conceptual framework in explanations was also a guideline that was partly violated since operators did not understand the term "brief" in the same way as the concept communicated. Operators had questions on how the AI will learn in addition to the factors presented in the visualization.

10.1.5.2 Further Works

Based on the feedback the AI functionality requested was more related to area management during operation rather than area setup.

10.1.6 Abnormality Detection

Abnormality detection helps operators detect if tracks behave in a way that makes them a candidate for being a threat.

10.1.6.1 Comments from Operators

In this concept the operators found it difficult to assess exactly what the anomaly detection would include in the visualisation used. Although, if the function were used effectively, anomaly detection was viewed as a useful tool. The fidelity of the visualisation presented was lower than others which may have affected the validity of the critique received. The concept was presented as a combination of the previously rejected idea of Auto ID and the first iteration of the anomaly detection concept. By using low-fi sketches and explanations of the concept, a discussion of the concept's utility was facilitated.

Operators alluded to that the functionality of the AI tool was unclear which suggests that one of the guidelines used were not properly achieved. Also, the guideline of making the capabilities clear was partially violated. Some of the functions in this tool have already been seen by the operators though, which made the discussion easier.

10.1.6.2 Further Works

The functionality of the tool needs to be visualised more clearly. And the concept needs to show what it can do and what information is used in the suggestions for anomalies.

10.1.7 Timeline

This concept helps operators go back and assess previous movements and events about tracks and predict future states.

10.1.7.1 Comments from Operators

Operators were critical of the utility of the predictions in the timeline concept; predictions may give a false sense of security depending on how AI warnings are calibrated. For example, if something historically occurred with a probability of 1% the AI may inadvertently contribute to a false sense of security by implying that the event is very unlikely to occur.

In addition to calibrating warnings correctly, operators worried about the underlying data used in AI predictions. For example, if an AI were to predict future behaviours of planes, an overwhelming majority will provide a prediction of a trajectory similar or the same as the current one. The AI must therefore be efficient in how to predict changes that are unexpected rather than those that most certainly will happen, the AI will be more useful when it provides information about possibly detrimental scenarios. In addition, the prediction should be based on adequate data sets that enable such predictions. This means that the AI needs to base its predictions on data beyond the most obvious such as speed, course, height, and previous movement patterns. Additional data for prediction may for example be radial speed, historical behaviour around certain areas, weather data (that may affect flight routes) and EW-data. There are certainly more factors that may be relevant for the predictions of future locations of tracks which need to be evaluated individually and in interaction between each other.

They were also critical of the specific track type selected for the visualisation during the evaluation. The functionality visualised was not the most data intensive task concerning track and plot prediction. During the idea generation, plot predictions for sea tracks were one of the main use cases. This since the sea plots move slowly and are substantially cluttering the display. Since this functionality was not properly evaluated, it needs to be further researched to assess the utility of the time bar, especially for predictions.

The main issues that operators reacted to in the visualisation are connected either directly or indirectly to the developing trust to the AI. Facilitating a trust for the AI include parameters discussed above such as clarifying what data that were used in predictions. Obvious connections between data and predictions will be easier to communicate in a human centred manner, where the reasoning for a certain prediction is opaquer the requirement for presenting the rationale becomes more challenging. It needs to be clear how the AI came to its conclusion for the trust to be properly developed.

In conclusion, the operators unanimously agreed that the timeline function of past events was good to have. The predictions however need to be clearer and risks of a too high level of trust need to be further evaluated.

10.1.7.2 Further Works

Further works include adding more parameters for the in-data used in the AI predictions. The “thought”-process of the AI for the AI predictions need to be visualised more clearly. The visualisations of the concept need to include other tracks and plots such as sea plots.

10.1.8 Pattern Areas

This concept helps the operators using filters over the current situation picture in four different ways using for example heat maps or connecting track plots.

10.1.8.1 Comments from Operators

Since the pattern areas concept was one of the concepts visualised only in sketch form it did not receive much criticism. Operators found it difficult to distinguish the functionality of the concept in relation to the others which suggest that operators did not understand the concept well enough to comment on it. However, it was unclear if the filters were based on AI generated information or not. Subsequently they commented on how useful the tool would be if it were not.

10.1.8.2 Further Works

This concept needs to be visualised better to reach more guidelines when evaluating it with operators.

10.2 Summary of Concept Evaluation

In this section the comments from the eight different concepts were summarised and a suggestion for what further work needs to be done on the concepts. One finding that was apparent for several concepts, was that concepts need to be clearly visualised. Operators were most positive regarding the Formation recognition and the Communication (talk-to-text) concepts. Collision warning, Route Planner, Abnormality detection, Timeline and Area Patterns got mixed feedback. The Area suggestion concept was viewed to have a low utility as it was currently specified and was the concept that operators liked the least.

10.3 Evaluation of AI Guidelines

In this section the guidelines that were used when designing the concept and the guidelines discussed during the concept evaluation will be presented. While ideating the concepts it became apparent that the guidelines in the planning phase were easier to use. This since the concepts are not developed with a level of fidelity necessary for the utility of the guidelines from the two subsequent phases. The result from the evaluation suggests what guidelines to change, remove or combine with other guidelines. First, the total number of instances that each of the guidelines were used is presented. Second, the number of times each of the guidelines was discussed during the evaluation of the concepts is presented.

10.3.1 Guidelines: Planning

Figure 34 presents the result from the evaluation of the guidelines for planning.

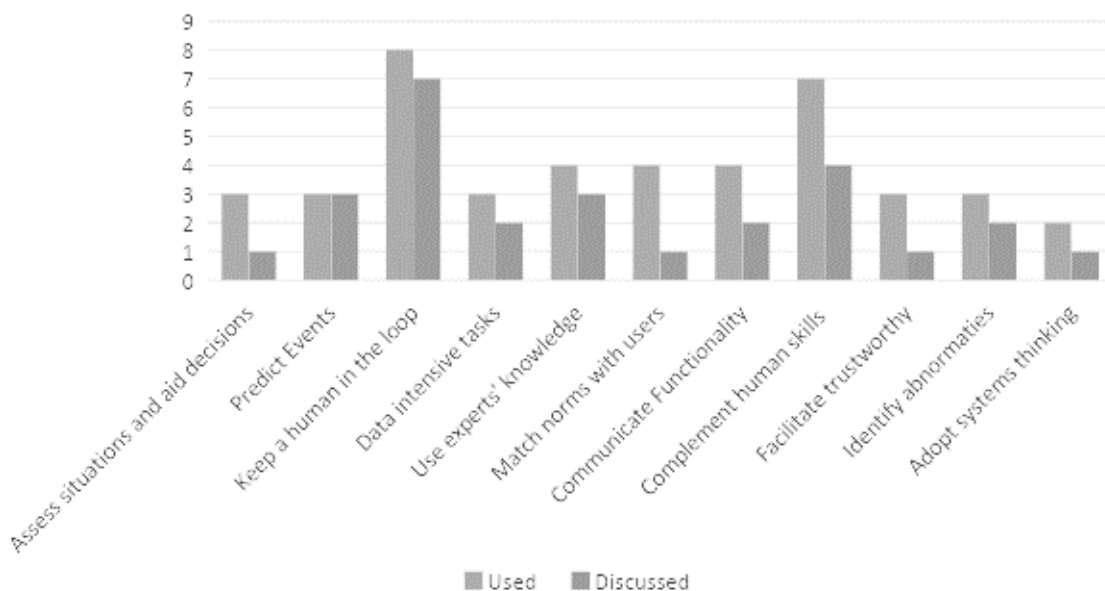


Figure 34: The result from the evaluation of the guidelines for planning.

The guidelines not used when developing the concepts were the following:

- Choose the simplest model.
- Choose explainable models.
- Train on appropriate data sets
- Identify ethical trade-offs.

Out of the four guidelines not used three of them were deemed more useful for developers rather than designers. This was choosing the simplest model, choosing explainable models, and training on appropriate data sets. This does not mean that the guidelines are not necessary but rather that they are not as useful for designers. Identifying ethical trade-offs was seen as important but not useful when creating visual concepts.

Concerning the remaining guidelines, used in the concepts during ideating, it was seen that the guidelines used also were discussed by operators which validates the guidelines to an extent. Keeping the human in the loop was both easy to design for and was seen as important for operators. The same

can be said for several of the guidelines, especially; complement human skills, use experts' knowledge, and predict events.

10.3.2 Guidelines: Designing

Figure 35 presents the result from the evaluation of the guidelines for designing.

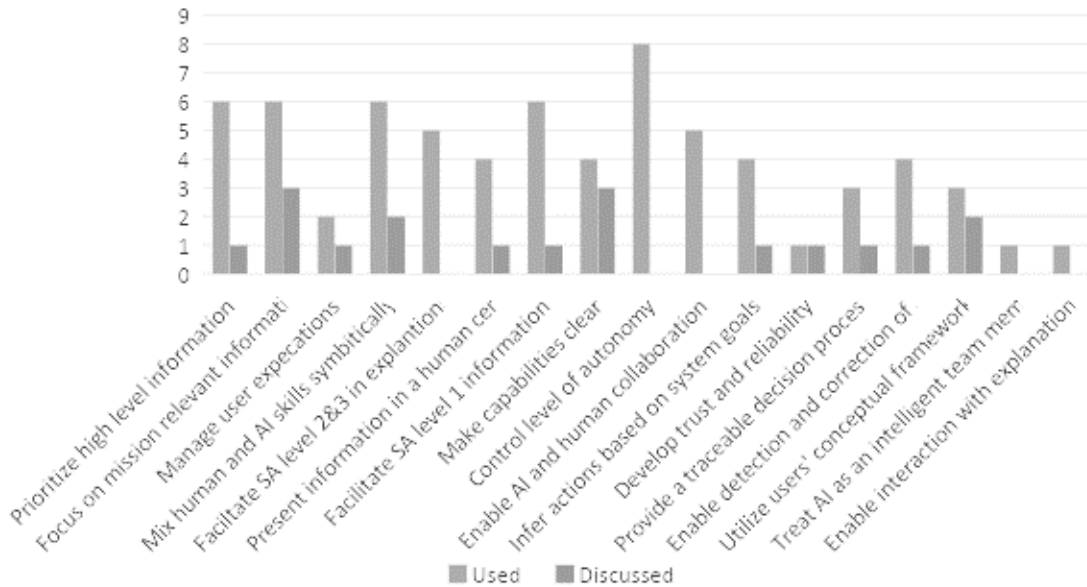


Figure 35: The result from the evaluation of the guidelines for designing.

The guidelines not used when developing the concepts were the following:

- Educate about system.
- Ensure awareness of team status.
- Safeguard against hacking
- Analyse potential risks of the AI.
- Detect user biases.
- Use contrastive explanations instead complete explanation.
- Combine statistical explanations with causal explanations.
- Enable interaction with explanations.

In this phase a higher percentage of the total amount of the guidelines were not used. About a third of the guidelines were not used at all which can be seen in the list above. The connection between the intended guidelines in the concepts and the visualizations were not as clear for the operators. For example, controlling the level of autonomy was used in all concepts but not discussed for any. Figure 35 shows four other guidelines that were used but not discussed. The two guidelines discussed the most were “focus on mission relevant information” and “make capabilities clear”. Safeguarding against hacking and analysing potential risks are important guidelines but difficult to visualize in pictures of the concepts. The same goes for some of the other guidelines. Three of the guidelines not used were connected to explanations of the AI suggestions. For example, enabling interaction with explanations which needs to be tested using interactive prototypes rather than showing frames of the concept to operators. Detecting user biases and educating about the system were not mentioned since

it is also difficult to visualize in a concept. There may be several reasons why guidelines in the use phase were not used and discussed to the same extent as in the prepare phase. One of the most obvious is because the concepts were not very developed concerning depth, functionality, and ability to interact with the concepts.

10.3.3 Guidelines: Evaluating

Figure 36 presents the result from the evaluation of the guidelines for evaluating.

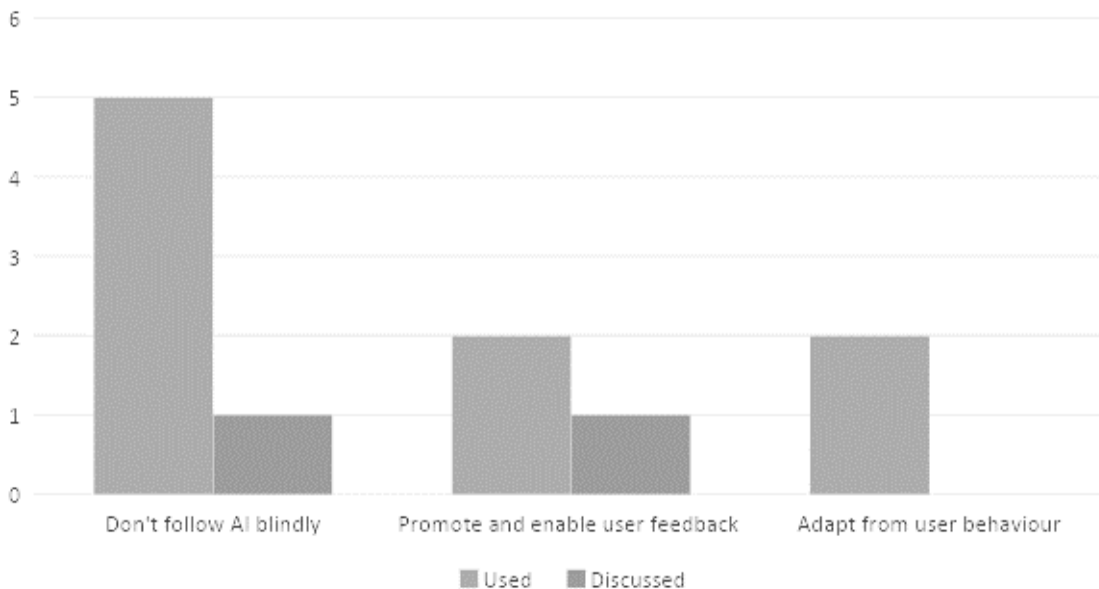


Figure 36: The result from the evaluation of the guidelines for evaluating.

The following guidelines were not used.

- Notify users about user generated adaptations.
- Be conservative in adaptations.
- Assess limitations of the AI.
- Understand that attacks on opaque AI is hard to detect.
- Use multiple evaluation measures.
- Evaluate explanations effectiveness.
- Use agents to agents to simulate human operators.

In the evaluation phase more than half of the guidelines were not used in the creation of the concepts. The three guidelines that were used did not get much attention during the evaluation. The reason for why they were not used is mostly because the guidelines are for evaluation of developed functions.

10.3.4 Summary of Guideline Evaluation

The evaluation of the guidelines shows that the guidelines in the planning phase were both easier to apply in the study and more often discussed by operators. A higher percentage of the planning phase guidelines were used than both the design and evaluation phase guidelines. In addition, a higher percentage of the guidelines used were also discussed in the planning phase compared to the other two phases.

11

AI Guidelines

This chapter presents the AI guidelines. The final AI guidelines are the result of the two AI guidelines phases in the study.

The first AI guidelines phase resulted in the first iteration of AI guidelines. These were made based on the findings derived from an AI literature study. The literature included both general studies of AI, for example definitions of AI, ethical considerations of AI, and capabilities and limitations of AI. The literature study also included literature more applied to the AEW&C context. For example, AI in military operations, explainable AI, Situation Awareness of AI. See Appendix A for the list of sources.

Quotes were extracted from the literature and analysed by the method affinity diagram. The affinity diagrams resulted in three larger categories of AI guidelines related to when they were most applicable in the development process (See Figure 37). The larger categories were further divided into subcategories, which each included a set of guidelines. The first iteration of AI guidelines is available in Appendix B.

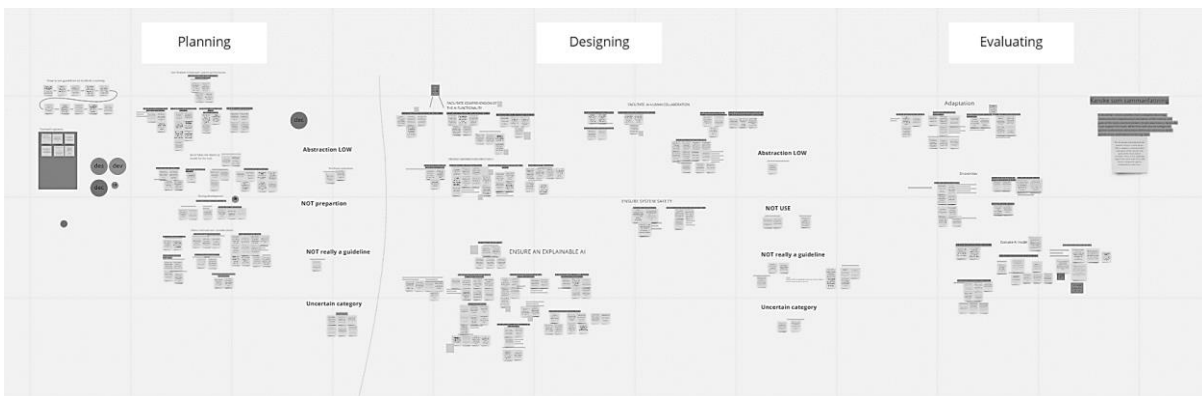


Figure 37: Low resolution picture of the affinity diagram used to make the first iteration of AI guidelines.

The first AI guidelines phase was in parallel with user studies and concept development. The result from the first phase, the first iteration of AI guidelines, was used as input to the first concept creation phase, which in turn gave input to the second AI guidelines phase by evaluating their utility and validity in the concept creation phase (See figure 38).

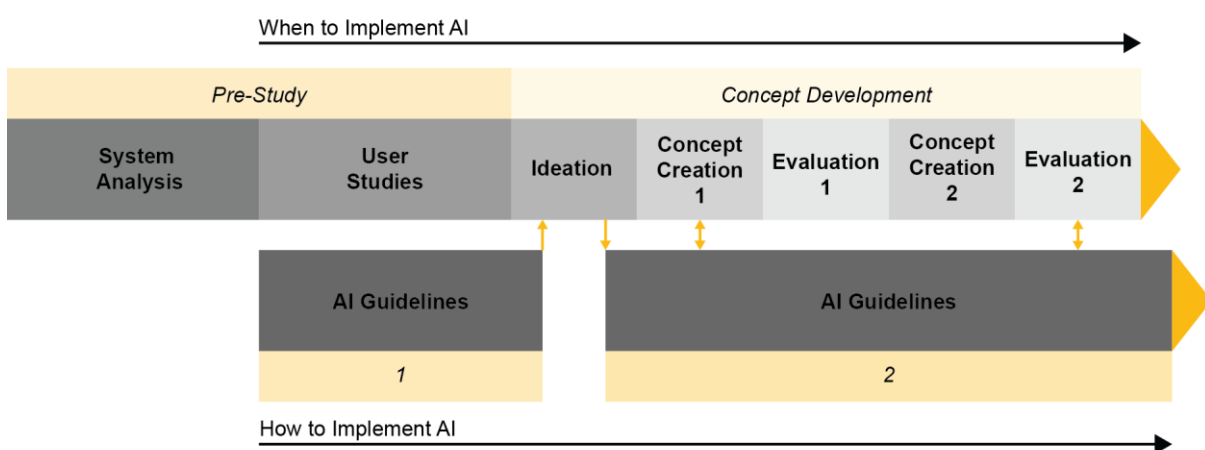


Figure 38: The process model of the study.

In the second AI guidelines phase was in parallel the evaluations and the second concept creation phase. Based on the evaluation of the utility and validity in the first concept creation phase the AI guidelines were further elaborated. During the second phase the AI guidelines both were evaluated

based on their utility and validity for concept creation while at the same time used as a tool to evaluate to create concepts against. This way of working was inspired by *research through design* (Zimmerman, Forlizzi & Evenson, 2007).

The AI guidelines presented in this chapter are the final results of the study. The AI guidelines mainly aims to answer the research question of **how** to implement AI to improve Situation Awareness for AEW&C operators, but also touches upon the question **when**.

The AI guidelines are presented in three categories based on when the guidelines are most relevant in the development process. The three categories are further divided into subcategories, which each included a set of guidelines. The categories are the following:

- **Planning:** Mainly applicable when planning to implement AI systems. For example, when exploring opportunities of AI for a specific context.
- **Designing:** Mainly applicable when designing AI systems. For example, by being used as a basis for ideation and creation of user-centred, ethical, and safe concepts.
- **Evaluating:** Mainly applicable when considering how to evaluate AI systems, both before implementation and over time.

However, it is important to point out that all guidelines should be considered during the whole development process, even though they target specific phases. For example, ethics are something that always should be considered but is a part of the Planning category because it is crucial to consider ethics as soon as possible in every project.

In the following chapters the guidelines for each category, *Planning*, *designing*, and *Evaluating* will be presented.

11.1 Planning

Implementation of AI systems is a large commitment for a company, depending on the purpose of the system it can involve large costs in both time and investments. Not only can it be costly, but there is also an obvious risk that the project fails to achieve the intended result. Planning is important for all kinds of design projects, but due to the potentially high cost and risk of designing AI systems it could be considered extra important in this domain.

The guidelines in Planning are suitable to use when exploring and identifying opportunities for AI implementation and when evaluating already existing AI systems on a higher abstraction level. The guidelines derived for pre-studies relate to; how AI can complement human skills, how to select the appropriate AI model, how to adapt systems thinking and ethical considerations. Figure 39 presents the guidelines developed for *Planning*.

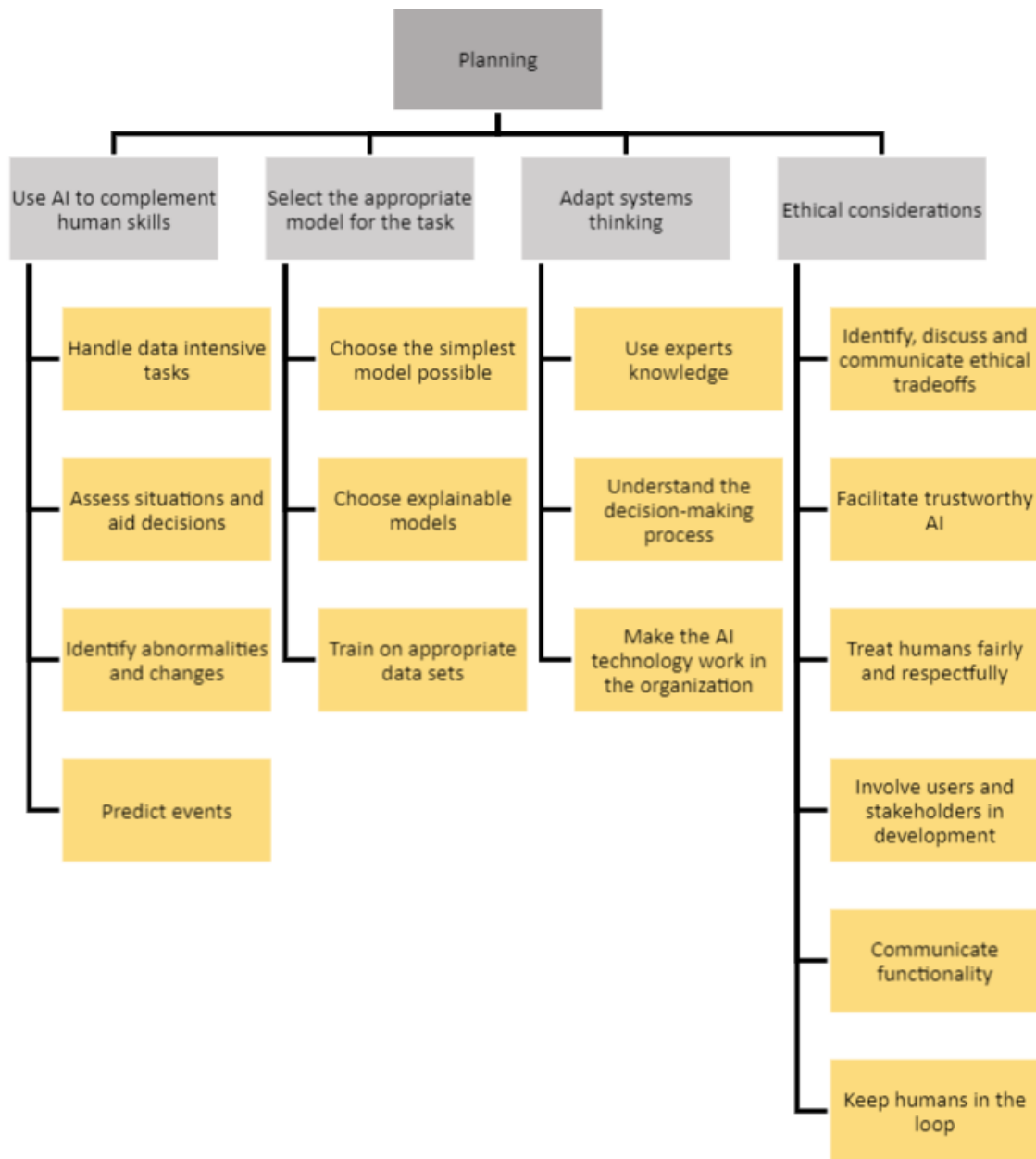


Figure 39: The guidelines developed for Planning.

11.1.1 Use AI to Complement Human skills

In human-AI systems the goal should be to augment, complement and empower human skills to support them in making better choices and, to consider human limitations and design AI systems that complement humans' perceptual and cognitive abilities in tasks they find difficult.

Generally, humans have challenges in handling large amounts of data, keeping track of uncertain information, making predictions through mental simulations, and staying unbiased in high-value situations. Due to this, the following guidelines of which tasks the AI systems can complement the human in are suggested.

11.1.1.1 Handle Data Intensive Tasks

Information overload is an apparent problem for the operators. Due to human cognitive limitations, it is impossible for the operator to handle all information at the same time. Computers and AI however can handle tremendous amounts of data at a rapid speed.

11.1.1.2 Assess Situations and Aid Decisions

Even though operators may try to act rational they are limited to their cognitive processing abilities when they are dealing with complex problems. This is referred to as bounded rationality. Humans are also affected by different biases, for example confirmation bias. AI can assist the operator by assessing complex situations and aid unbiased decisions.

11.1.1.3 Identify abnormalities and Changes

Experienced operators have a well-developed mental model of the normal image and are trained to spot abnormalities and changes. However, due to cognitive and perceptive limitations they can have a hard time detecting less obvious abnormalities and slower changes. AI can be trained to discover familiar behaviour patterns and alert the operator when it finds abnormalities and changes that deviate from those patterns.

11.1.1.4 Predict Events

Similar to the challenges for the operators to assess situations and aid decisions the operators are limited to cognitive limitations and affected by bias when predicting future scenarios. The amounts of scenarios humans can project are limited and affected by bias. AI can consider a tremendous number of scenarios and assess the likeliness and impact of those in an unbiased manner. However, it is important to consider that AI might have trouble in predicting scenarios that have not yet happened. Just like humans are limited to their cognitive abilities, AI systems are limited to the data they are trained on. Additionally, predictions should be presented with caution. The expert evaluation implied that they could give a false sense of safety and strengthen confirmation bias.

11.1.2 Select the Appropriate Model for the Task

From the outside, AI can seem like a magic formula that can solve whatever problem it faces. While AI might be able to perform better than humans for some instances (see guideline *11.1.1 Use AI to complement human skills*), the AI system will never be better than the models it uses and the data it is trained on. While humans are limited to their cognitive abilities, AI systems are limited to the models used and data available. Therefore, the selection of the right AI model is critical. There is no best model for all cases, it depends on the problem and often a mix of models are required. Below are some general guidelines to use when selecting an AI model.

11.1.2.1 Choose the Simplest Model Possible

According to Occam's razor principle: "entities should not be multiplied without necessity". To choose the simplest theory that explains the data also applies when choosing an AI model. A learner should be designed such that it prefers simpler theories to complex theories, except in cases where complex theories are proven better. One of the reasons why simpler theories are preferable is because

they tend to be more testable. Related to choosing the simplest model is to identify and try to avoid considering a broad range of possibilities unlikely to be beneficial.

11.1.2.2 Choose Explainable Models

Similar to choosing simple models, one should choose the most interpretable models whose performance is easier to explain to people. However, it is important to be aware of the trade-off between performance and explainability. Often the highest-performing methods are the least explainable, and the most explainable are the least accurate. There are strategies that aim to tackle this dilemma, for example deep explanation, interpretable models, and model induction (Gunning, D., & Aha, D., 2019).

11.1.2.3 Train on Appropriate Data Sets

Training on appropriate data sets implies “avoiding learning the wrong lesson”. For example, an image classifier trained only on pictures of black horses and white cats might conclude that all black patches are likely to be horses. Other cases might be that the data set is outdated or not capturing the complete and real-life environment. This can be avoided by choosing up to date and more diverse data sets to train on, better capturing the environment in which the AI will act.

11.1.3 Adapt Systems Thinking

When designing new products, it is important to consider the bigger picture, often the product is a part of a bigger system. A system is more than the sum of its parts, changing one part may affect other parts or the whole system. Therefore, it is important to understand the context in which the product acts and the needs and goals of the users of the product. The following guidelines address this issue.

11.1.3.1 Use Experts' Knowledge

Adapting a user-centred design process is recommended when developing systems, especially complex systems with expert users. The expert users can provide valuable insights that would be difficult to derive without them. Examples of expert user insights when designing AI systems could be to assess the required level of accuracy for different use cases, the values of strengths of arguments when weighing options and the cost and risk of mistakes when tuning model parameters. An open dialogue with the expert users across the whole design process facilitates a better understanding of how they work and their needs and requirements.

11.1.3.2 Understand the decision-making process

For AI systems designed to support users in making decisions it is important to establish an understanding of the decision-making process. This can be done by conducting different task analysis methods, like Hierarchical Task Analysis (HTA) or Goal-Directed Task Analysis (GDTA). These methods can help to map out what steps the operators take when making decisions and what they base their decisions on. It can also be used to find opportunities for improvements and in what steps the operator is in most need of support. AI systems should be made in parallel with development of decision-making processes to make sure that the result matches the actual decision-making process.

11.1.3.3 Make the AI technology Work in the Organization

The AI technology will not be useful unless it works in the organization it is in. Do not only focus on the technology itself, consider what prerequisites, possibilities, and limitations there are in the organization and how the AI technology can enhance the outcomes on an organizational level.

11.1.4 Ethical Considerations

Before implementing new AI systems, ethical considerations should be considered. AI raises a lot of ethical questions, especially when used in high-risk environments directly affecting people's lives. There is a lot to be said about ethical considerations, the guidelines below give some general advice to consider when designing AI systems. However, the guidelines here are not comprehensive for all cases and designers should actively search for further considerations when designing new AI systems.

11.1.4.1 Identify, Discuss, and Communicate Ethical Trade-offs

For most projects, ethical dilemmas and trade-offs will occur where there is no single right or wrong decision. The designer should try to identify these cases early and approach them via reasoned, evidence-based reflection. The reasoning and decisions taken in ethical trade-offs should be properly documented. On a higher level, the company can consider forums to discuss AI ethics related to the company, like establishing an “ethical AI review board”.

11.1.4.2 Facilitate Trustworthy AI

Not only does trust improve effectiveness and use of AI systems, but it is also a prerequisite for people and societies to develop, deploy and use AI systems. According to the European Commission (2019) trustworthy AI has three components: it should be lawful, ethical, and robust. This implies that it should comply with laws and regulations, ensure adherence to ethical principles and values and be robust both from a technical and social perspective.

11.1.4.3 Treat Humans Fairly and Respectfully

All humans should be treated fairly and with respect. Even with good intentions, AI systems can cause unintentional harm. When designing AI systems, it is important to consider how humans are treated and having respect for their integrity. Ensure that data collected about the users are not used unlawfully or that it unfairly discriminates against them. Make sure that the model has a comparable performance for different demographic groups. Match the mental models of the users given their social and cultural context while at the same time making sure not to reinforce unfair stereotypes and biases.

11.1.4.4 Involve Users and Stakeholders in Development

User involvement can help to ensure that the diversity of the users is considered. The teams involved in making the systems should reflect the diversity of the end users. Not only the users should be considered, but also other stakeholders who may directly or indirectly be affected by the system throughout its life cycle.

11.1.4.5 Communicate Functionality

Make sure that the users are aware of the potential impact of the AI systems and what it can and cannot do. Here, communication, education, and training play an important role. There should be no unclarity for the users that they are interacting with an AI system, a system that represents itself as a human could cause costly misunderstandings. It is also important to consider that most people do not read documentation, you cannot rely solely on manuals. The AI system should by itself be able to communicate its functionality.

11.1.4.6 Keep Humans in the Loop

Maintain the respect for human autonomy. Allow human oversight over the AI system and give them the ability to contest and make autonomous decisions. Assess the appropriate level of human control for the AI system and use case. Take steps to counteract skill loss and job loss. Remember that the less human oversight over an AI system, the more extensive testing and stricter governance is required.

11.2 Designing

If the guidelines for the previous category, *Planning*, were more related to “what” on a high abstraction level, this category, *Designing*, is more related to “how” on a lower abstraction level. The guidelines for Designing are suitable to use for ideation and creating concepts, trying to concretise the identified opportunities for AI implementation. The guidelines can also be used to evaluate and improve already existing AI systems.

The guidelines derived for the usage of AI systems relate to; how to present information, how to facilitate comprehension and collaboration, how to ensure safety in the use phase and how to ensure and ensure explainability. Figure 41 presents the guidelines developed for *Designing*.

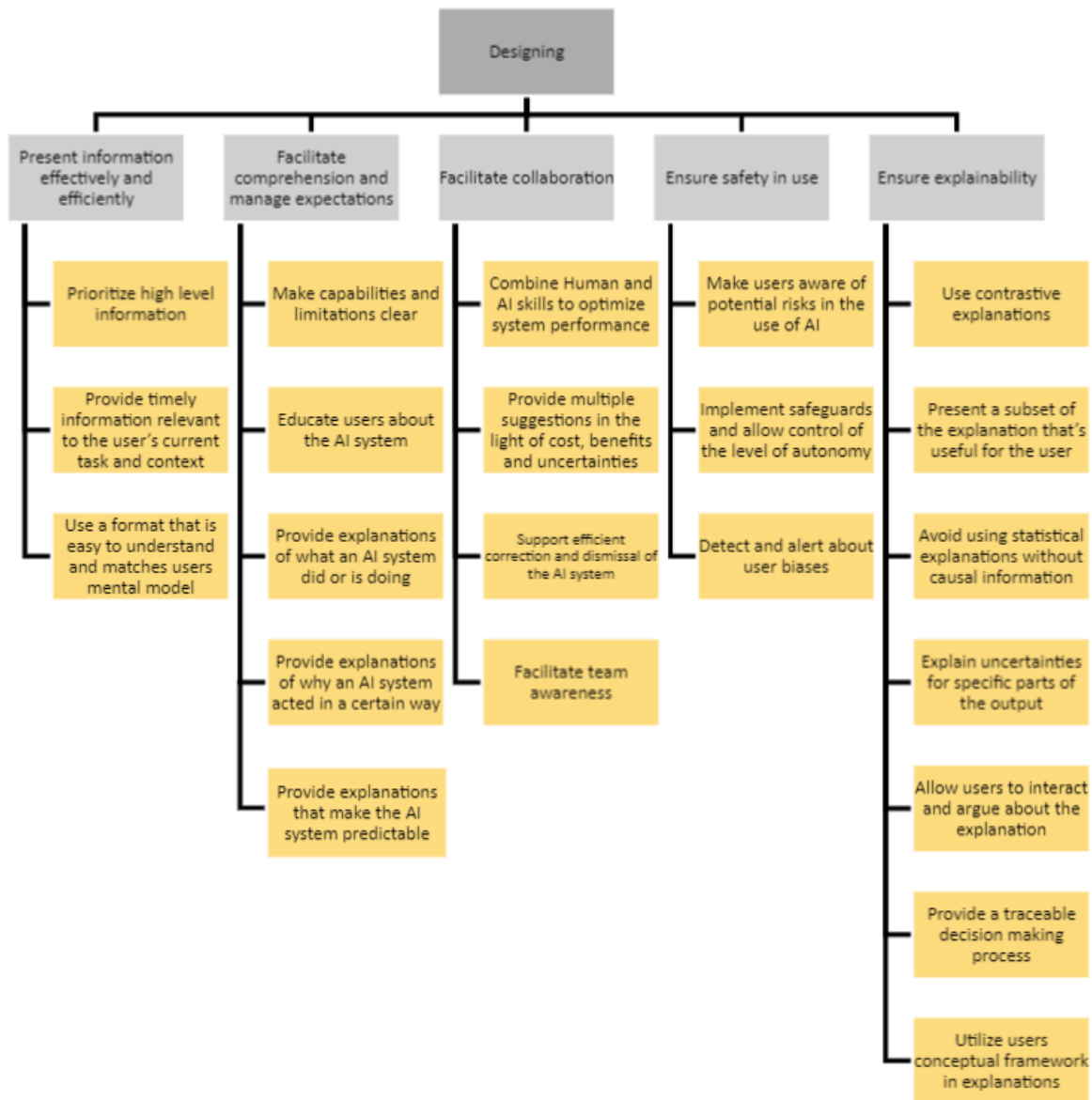


Figure 40: The guidelines developed for Designing.

11.2.1 Present information effectively and efficiently

The information derived by the AI systems needs to be presented in an effective and efficient way. This implies that the operator gets the most relevant information at the right time according to their current task and context. The information should be easy to understand and match the users' mental model. The users should also be provided information of what the AI system did or is doing.

11.2.1.1 Prioritize high level information

In an AEW&C system, the operator often works under time pressure and at times they are overloaded with information. Therefore, it is important not to expose them to more information than necessary. Keep the user informed only on the high-level aspects that are of immediate interest, but also give them the opportunity to find more detailed information if they wish.

11.2.1.2 Provide timely information relevant to the task and context

Both the users' tasks and context may vary during a mission. It can go from calmly monitoring the situation picture to an intense guiding mission of a fighter. The information should be presented in accordance with the users' current task and context. Display mission-relevant information in correlation with the points reached in the decision-making process. Support the user to focus attention on key attributes and relations. Ensure that the information is timely and up to date and make it easy for the user to invoke the AI tools when needed.

11.2.1.3 Use a format that is easy to understand and matches mental models

Present information in a format that is easy to understand, matches the user's mental model and requires minimal amounts of translation and cognitive processing. For example, today's AEW&C system uses colour coding to indicate the identity of tracks. The colour matches the users' mental models and makes the information easier to understand and require less cognitive processing. This should be taken into consideration when developing new AI systems, that the new systems should match already existing mental models.

11.2.2 Facilitate Comprehension and Manage Expectations

Facilitating comprehension and managing expectations of AI systems may be particularly important because people often have unrealistic expectations about the capabilities. This, combined with a lack of comprehension of how the AI system functions, can lead to dissatisfaction, distrust, product abandonment, or even injury and harm. The following guidelines address this issue and suggest ways to facilitate comprehension and manage expectations of the AI system.

11.2.2.1 Make Capabilities and Limitations Clear

Unrealistic user expectations are common for AI systems. Inflated expectations can be the source for misunderstandings with serious consequences. Therefore, it is important that the capabilities and purpose of the AI system is clearly communicated to the users. Make it clear what the system can do but also make it clear how well the system performs. Explain uncertainty and help users understand the error boundary, how often the AI system may make mistakes and if there are situations where it performs better or worse.

11.2.2.2 Educate Users About the AI System

A way to facilitate comprehension and managing expectations is to educate users about AI in general and, more specifically, the AI system in question. Instruct and train the users to make them receptive to understanding an AI system. Let users experiment with AI systems to explore capabilities and limitations. Specify use cases and communicate these in an understandable and appropriate way. Consider how to best inform users who might know little about AI.

11.2.2.3 Provide Explanations of what the AI System did or is doing

Facilitate perception of AI by providing the users with information about what the AI system did or is doing and the decisions it made. This includes information about the system's inputs and outputs and decisions and actions taken by the system. An example might be to display what track (input) the AI has assigned what identity (output). The challenge is to determine what information is relevant for

users, complex systems may take decisions over many input factors and produce many outputs, where far from all are relevant to the user. For example, in a collision warning system, the user might not be interested in knowing all input the AI systems use, only the input which might result in a collision (output) is relevant.

11.2.2.4 Provide Explanations of why the AI System Acted in a Certain way

Facilitate comprehension of AI by providing users with explanations about why an AI system did what it did and what this means in terms of the system's goals. Explain causality. For example, why did the AI system indicate that a track behaved abnormally? How does a suggested route relate to mission time and fuel consumption? The challenge is to determine what causal information is most relevant for the user to understand the AI's behaviour. For example, the AI might alert a track deviating from a route, but why? What was the predicted route and by how much did the track have to deviate from it to trigger the alert?

11.2.2.5 Provide Explanations that make the AI System Predictable

Facilitate predictability of AI by providing explanations about what the AI system will do next, or what it would do in a different scenario. Make the user understand what would happen to the outcome if inputs, model parameters or constraints were changed (forward simulation). But also make them understand what would be necessary to achieve a desired outcome (backwards reasoning). Explanations like this could help the user predict the future behaviour of the AI system.

11.2.3 Facilitate collaboration

Rather than working as separate entities, the human and AI should collaborate to achieve the system's goals. The goal is to make human-AI teams perform better together as a unity than either alone. Let the human and AI collaborate in a complementary manner, provide users with choices rather than one optimal solution, support efficient correction and dismissal of the AI system, and facilitate team awareness.

11.2.3.1 Combine Human and AI Skills to Optimize System Performance

As stated in *1.1 Use AI to complement human skills*, AI should be used to augment, complement, and empower human skills to support them making better choices. This however goes both ways. The human should also help the AI overcome limitations, for example ability to handle new situations and rare events and vulnerability to hacking (see *3.2.1 Assess limitations of the AI system*). The challenge is to make the combination of skills balanced and design a human-AI system that is seamless in the collaboration. This can be achieved by establishing an understanding about the system's goal and the capabilities and limitations of both humans and AI.

11.2.3.2 Provide Multiple Suggestions in the Light of Costs, Benefits and Uncertainties

When operators are facing situations where they must make decisions, there is seldom just one optimal choice. The goals may vary for different missions and contexts. For example, when guiding fighters, the route depends on the type of mission and the priorities of the tracks etc. Humans expect contrastive explanations (see *2.5.1 Use contrastive explanations*) and want to compare and weigh options. When AI systems make suggestions, instead of just suggesting one optimal solution, the AI

system should help the user explore the space of possibilities. Present the possibilities in the light of cost, benefits, and uncertainties so the operator can make well-informed and rational choices.

11.2.3.3 Support Efficient Correction and Dismissal of the AI System

No matter how well the AI system is designed and developed there is always a risk that situations will occur where the AI system is wrong or does not function as intended. Provided that the AI system is transparent and comprehensive enough for the operator to detect the errors, the AI systems should be designed to support efficient correction and dismissal. For example, if the AI system suggests a formation of tracks the operator should be able to correct the suggestion if it is wrong. And if the context changes to the extent that it exceeds the AI systems capabilities, the operator should be able to temporarily dismiss or ignore the AI systems services.

11.2.3.4 Facilitate Team Awareness

In the AEW&C system the operators work as a team to achieve the system goals. When incorporating an AI system in the same environment it should be considered, at least to some extent, a part of that team. When working together, team awareness is crucial for the success of the mission. Humans should be aware of the AI system and its capabilities and limitations, but this also goes the other way. The AI system should be aware of what the users are doing, their current workload, preferences, proneness to bias etc. All members of the human-AI team should have knowledge of their own role and the other members' roles, and what that implies to the decision-making process. The AI system should be designed to facilitate observability, predictability and directability. This implies facilitating the ability to observe the status of the team, the task, and the environment. making the actions of the team members predictable to some extent and ability to take over and delegate tasks, both reactively and proactively.

11.2.4 Ensure safety in use

Implementing new tools for users in a high-risk environment such as the AEW&C system always involves some type of risk. It is important to consider those risks in the use phase and take measures to reduce them and ensure safety. The guidelines presented in this category are associated with ensuring safety in the use of AI systems.

11.2.4.1 Make users Aware of Potential Risks in the use of AI

Not only should the users be aware of their own limitations, but they should also be aware of the limitations of AI and the potential risks it imposes. This is closely related to *11.2.2 Facilitate comprehension and manage expectations*. For example, the users should be aware that AI can make errors in incomprehensible ways, that there is a risk of blindly following the AI system and that it is vulnerable to adversarial attacks, especially in the defence industry. Making the users aware of potential risks in the use of AI systems should be considered an important part of the operators training.

11.2.4.2 Implement Safeguards and Allow Controlled Level of Autonomy

The AI system should have a fallback plan in case of problems. This can be achieved by implementing safeguards like stop buttons and verification steps and the ability to control the level of

autonomy. An aspect that is important to consider for these kinds of safeguards is the delegation of control between humans and AI. Consider whether the task the AI system was responsible for should be aborted or partly or completely delegated to the human.

11.2.4.3 Detect and Alert About Biases

Humans mostly make decisions unconsciously and based on intuition. While that can be effective and efficient, it can also lead to biased judgement and suboptimal decisions (Kahneman, 2011). The AI system should check for biased behaviour and alert the users about these in a timely and proportionally manner.

11.2.5 Ensure Explainability

AI systems, especially those using machine learning techniques like deep learning, offer almost no clues how they work or how the results are obtained. The lack of transparency hampers the users trust in the AI system and makes it harder for them to detect and act on errors. Trust and comprehension are also important for collaboration and acceptance of the AI system. The problem with the lack of transparency grows as today's AI systems get increasingly more complex. A way to deal with the problem is to build more explainable AI systems which will facilitate comprehension and trust. The following guidelines relate to how explanations by the AI system should be structured and presented.

11.2.5.1 Use Contrastive Explanations

People request contrastive explanations. When they hear or give an explanation they do so with a foil, a counterfactual case, in mind. People have a hard time understanding the meaning of an explanation if they do not have anything to relate to, therefore they evaluate the explanation based on the contrast between fact and foil. What foil people use may vary and depend on for example previous experiences. For example, a person without knowledge in the AEW&C area might think 10,000ft is a high altitude for an airplane, when in fact it is quite low. The user's assessment of the explanation depends on what foil the explainee compares the explanation with. The challenge with providing contrastive explanations is to determine the most likely foil the users have in mind, this could be something to consider in *1.3.1 Use experts' knowledge*.

11.2.5.2 Present a Subset of the Explanation that is Useful for the User

An event may have close to infinity causes. Since the cognitive burden of complete explanations is too great, the user chooses a small subset of causes that they find most useful and relevant. However, this is done in a biased manner and it is not always the most likely explanation the user finds most useful. Try to identify which casual information is most useful and relevant for the user to understand the behaviour of the AI system or an event.

11.2.5.3 Avoid Statistical Explanations Without Causal Information

Statistical explanations might be the first that comes to mind when considering how to explain AI system performance. For example, "the AI is 90% accurate in recognizing formations". However, these explanations alone are not very effective. People do not perceive statistics as computers, their confidence does not change linearly. While they might not make a difference between 75% and 76%, their decisions might change drastically between a confidence of 75% and 90%. Instead of expressing

probability with values, people use contrastive terms like "high", "possible" etc. For the example mentioned earlier, instead of just providing a statistical explanation, it could be accompanied with a term and the shape of the formation outlined on the map.

11.2.5.4 Explain Uncertainties for Specific Parts of the Output

When AI systems explain uncertainties or accuracy for the whole output, questions arise. For example, if the AI system says it is 80% confident that it sees 5 tracks in a box formation, how should the user interpret the confidence? How confident is the AI system for the different parts of the explanation? How did it assess the confidence? To deal with this dilemma, the AI system could explain or highlight uncertainties for specific parts of the output.

11.2.5.5 Allow Users to Interact and Argue about the Explanation

When humans hear and give explanations to each other, they expect to be able to interact and argue about the explanation. They might want to ask questions about the explanation like “why” and “what if” and ask for more information or try to change the explainer's beliefs. The interaction creates a shared meaning of the explanation, engenders trust, and enables teamwork. Therefore, if possible, the users should be allowed to interact and argue about the explanation provided by the AI system.

11.2.5.6 Provide a Traceable Decision-Making Process

During expert evaluation of the AI concepts, the expert often asked for the source of the output or explanation. For example, they wanted to know what type of data the AI system used for movement prediction and what the model considered. They felt unsure if the prediction was just based on a continual movement of the current path, or if it considered other variables as well. This suggests there is a need to provide a traceable decision-making process to facilitate trust and understanding. The working of the system needs to be understood and traceable. Explain the system output by following the steps of arguments and make the decision-making process explicit.

11.2.5.7 Utilize Users' Conceptual Framework in Explanations

Humans often assign human-like traits to AI. To design collaborative human-AI systems the explanations by AI should be using the same conceptual framework as when humans explain to each other. To make this possible it is necessary to identify what is required to give an explanation according to the users' conceptual framework. Conduct user studies to understand how users define, generate, select, evaluate, and present explanations.

11.3 Evaluating

When the guidelines from the categories *Planning* and *Design* have been applied, hopefully a concept of an AI system or AI functionality have been developed. The guidelines in this category, *Evaluating*, is suitable to use when it is time to consider how the AI system should be evaluated, both before and after implementation and over time.

The guidelines derived for the evaluation of AI systems relate to; how the AI system should adapt from user feedback and behaviour, how to evaluate potential errors and risks, and how to evaluate the AI model. Figure 41 presents the guidelines developed for *Evaluating*.

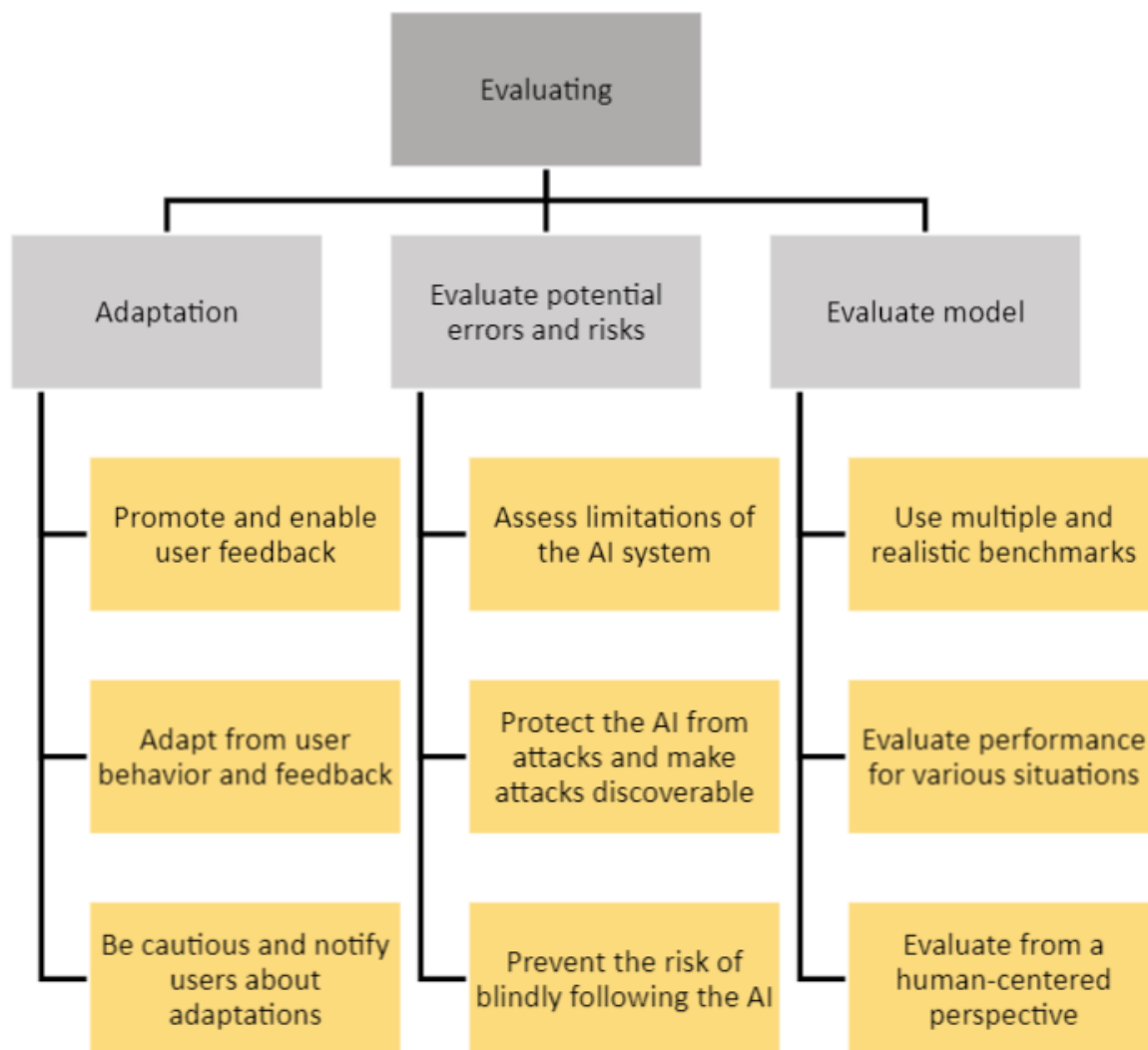


Figure 41: The guidelines developed for Evaluating.

11.3.1 Adaptation

A learning AI system should be able to adapt and adjust over time. The input to the adaptation can be both from the environment and the users. The AI system should continuously learn from the data gathered from the environment. An AI system without the ability to adapt to the environment could become irrelevant if the environment changes, for example if flight routes change or new roads get built. The AI should also enable and adapt from user feedback and behaviour to make sure the outputs by the AI system are useful to the users. However, adaptation does not come without risks. Therefore, it is important to notify users about changes and be cautious with adaptations.

11.3.1.1 Promote and enable user feedback

The feedback from users is valuable and should both be enabled and promoted by the system. Establish processes that consider user feedback and enable users to give feedback indicating their preferences during interaction with the AI system. Ensure that the users can report on negative actions and decisions made by the AI system, and allow users to respond to the errors to minimize negative impact.

11.3.1.2 Adapt from User Behaviour and Feedback

The AI system should learn from the user and adapt in an intelligent manner, e.g., being aware of the users' actions and intentions. This does not only imply adapting to active feedback by the user, but also their behaviour when using the system. The AI system should adapt based on users' objectives, recent interactions, preferences, and history of making biased decisions. For some cases the AI system should be able to adapt to users playing different roles in the team, providing each with the specific information relevant to his or her own task and potentially at different levels of abstraction.

11.3.1.3 Be Cautious and Notify Users about Adaptations

Adaptation should be done with caution. When the AI system adapts, it is important that they do so in a way that benefits the users. There is a trade-off between adaptation and consistency. The user's awareness of capabilities becomes increasingly more challenged with adoption. Changes due to adoption can make the user perceive the AI system as unpredictable and inconsistent, causing confusion and irritation. It is important that the user trust and understand how the system adapts. The AI system should notify users about changes and when the AI system adds or updates capabilities and convey how actions will impact future behaviours of the AI system.

11.3.2 Evaluate potential errors and risks

Implementing AI systems in high-risk environments such as the AEW&C environment always involves some type of risk. It is important to consider those risks and take measures to reduce them and ensure safety. The higher the risks and capabilities are, the higher the safety measures should be. The following guidelines provide some guidance on what to do to evaluate potential errors and risks associated with AI systems.

11.3.2.1 Assess Limitations of the AI System

A crucial part of implementing an AI system is to assess the limitations and potential errors and risks of the AI system in question. Get insights from stakeholders, future users, and developers to make sure that the assessment is thorough. The assessment should be unique for each case, but some general limitations to consider about AI is:

- AI often makes different mistakes than humans make, sometimes in an incomprehensible way.
- Just as humans only achieve bounded rationality due to cognitive limitations, AI has bounded rationality due to other limitations. Uncomplete training data, flaws in the model and limited resources like time and computational power might result in suboptimal actions and suggestions.
- Rare events or new contexts can lead to worse performance or errors due to lack of relatable data.
- The AI system may filter out important information.
- Risk that the AI and human operate separately, and AI becomes no more than an add-on to the system.
- Risk of overconfidence in the AI system, making users blindly follow suggestions.
- Vulnerability to adversarial attacks.

The AI system should be built with knowledge of its limitations and actively check for situations for which it may not be as fully capable.

11.3.2.2 Protect the AI System from Attacks and Make Attacks Discoverable

AI can be vulnerable to attacks. This is especially important to consider in the defence industry where hostile actors may try to sabotage the AI systems. Hacking, or the use of adversarial examples, mischievous data, can cause the AI system to make errors. Measures should be taken to protect the AI system from attacks. Another aspect to consider regarding attacks is that the opaquer and more incomprehensible the AI system is, the harder the attacks are to discover. Facilitate comprehension and increase the transparency of the AI system to make attacks easier to discover.

11.3.2.3 Prevent the risk of Blindly Following the AI System

In evaluation of the AI concepts, the operators mentioned the risk of blindly following the AI. For example, if predictions are based on the most common scenario, it will likely also be the safest. This can cause a false sense of safety and strengthen confirmation bias. The risk increases as algorithms become better, more independent, and opaquer, in combination with skill loss of operators. Ways to deal with the risk of blindly following is to keep the human in the loop and increase the transparency of the AI system.

11.3.3 Evaluate model

The models used in the AI systems should be evaluated to ensure that it has a high and reliable performance. It should be evaluated against multiple and realistic benchmarks and performance differences should be measured for various situations. Also, human-centred metrics should be evaluated such as the model's explanation effectiveness.

11.3.3.1 Use Multiple and Realistic Benchmarks

Similar to training on appropriate data sets, the AI should also be evaluated using appropriate benchmarks. Use multiple benchmarks to reduce risks of overfitting the model to one specific benchmark. Include some real-world data from the environment the AI system will act in to reduce the risk of missing important conditions for the specific environment.

11.3.3.2 Evaluate Performance for Various Situations

When the context of the AI system changes there is a risk of performance drop. Unpredictable changes of the context are to be expected in the defence domain. When situations where performance drop may occur cannot be avoided it is important to evaluate how likely these drops are. It is crucial to maintain as high accuracy as possible in situations where AI systems directly affect human lives. Therefore, it is important to test whether the AI system behaves differently in specific contexts or situations. Identify these situations and take measures to reduce potential negative impacts.

11.3.3.3 Evaluate from a Human-Centred Perspective

The behaviour and performance of the AI must also be evaluated from a human-centred perspective. Evaluation metrics to consider could be user's satisfaction, task performance, accordance to mental

model, level of trust and explanation effectiveness. The models explanation effectiveness can be measured in terms situation awareness-based levels of explainable AI which implies:

1. Measure perception of AI, e.g., how well users understand inputs and outputs in the contexts of their intended goals.
2. Measure comprehension of AI, e.g., how well users understand the meaning of the AI system's behaviour and what it implies in terms of progress towards the goals.
3. Measure predictability of AI, e.g., how well users can predict what the AI system will do next or in an alternate scenario.

Since explanations should utilize the user's conceptual frameworks, experts of AI systems and decision-making models are not suitable to evaluate explanation effectiveness. Make sure that explanation effectiveness is measured with potential end-users.

11.4 Chapter Summary: AI Guidelines

The AI guidelines were created through literature studies where findings were sorted through an affinity diagram. The guidelines were then used in the first iteration of the concept development phase and then evaluated. The guidelines in this chapter were the result of this process which resulted in guidelines for how to implement AI in three phases: planning, designing, and evaluating.

Guidelines for planning include four categories of guidelines; Use AI to complement human skills, Select the appropriate model for the task, Adapt systems thinking and Ethical considerations. These guidelines are suitable when exploring new opportunities for AI. These guidelines could for example help designers and developers to:

- Create functions where AI will complement human skills by for example aiding in data intensive tasks, identifying abnormalities and changes, and predicting future events.
- Consider the user's perspective and use their expert knowledge before developing new AI-functionalities.
- Encourage designers to consider ethical issues such as keeping humans in the loop and facilitating trust.

Guidelines for designing include five categories; Present information effectively and efficiently, Facilitate comprehension and manage expectations, Facilitate collaboration, Ensure safety in use, and Ensure explainability. The guidelines could for example help designers and developers to create functions that:

- Provides the information that operators need for their tasks.
- Makes it clear to the users what the AI can and cannot do.
- Strives for an optimal division of labour between humans and AI.
- Ensures that users are aware of potential risks using the AI system.
- Facilitate explainability of the AI systems decisions and suggestions,

Guidelines for evaluation include three categories; Adaption, Evaluate potential errors and risk, and Evaluate model. The guidelines could for example help designers and developers to ensure that:

- The AI system utilizes feedback from users effectively over time.
- The AI system is continuously evaluated concerning outside threats as well as other limitations of the functions.
- The model that the AI relies on is continuously evaluated concerning performance in different situations and that the functions meet criteria from a human centred perspective.

Further recommendations on how to use the guidelines could be used are discussed in *Chapter 12: Discussion*.

12

Discussion

The following chapter will discuss: Recommendations when developing AI systems, Ethics: AI in military operations, Application of the results, Methods, and Further work and improvements in process.

12.1 Recommendations when Developing AI Systems

Find an underlying need. This is general advice for development of new tools in the C2 system. The system is already complex, adding additional and unnecessary tools can potentially harm the user's situational awareness. However, it is even more apparent for tools related to AI due to buzz and potentially inflated expectations about AI capabilities. Companies might feel an urge to implement AI for the purpose of marketing and keeping up with competition. For the Timeline concept, the prediction functionality was not appreciated during the evaluation. Without the prediction functionality, the concept would barely use AI technology. This suggests that AI is not always equal to usefulness and should not be implemented by force. When conducting user studies to find opportunities to make improvements one should not limit themselves to only AI-based solutions, even if that was the preliminary purpose. AI should be considered as a technology providing additional possibilities and broadening the solution space.

Use cross-disciplinary teams when developing AI. Knowledge of AI, understanding the users and the context and knowing the capabilities and limitations within the organization are all important considerations when developing tools related to AI. Human factors experts, developers, users, and other relevant stakeholders can all contribute with knowledge and insights that can help the development of useful and sustainable products. During the expert evaluations the users brought up important questions that had been partly overlooked in the study, for example what data sources should be used for the AI concepts, if they should be bought from an external organization or collected by the organization itself.

Using low fidelity concepts and visualizations as mediating tools for discussion proved to be a very useful method to extract user insights, opinions, and ideas for AI. It creates a common ground to base the discussion on and triggers ideas and opinions that would be hard to derive using words alone. Based on the success of the method it could be assumed that the process could benefit even more from it if it were used to a larger extent. Co-creation sessions could potentially be a successful method. However, it requires careful consideration of how to execute such a session to exploit the full potential. Using low fidelity concepts, or guidelines, as a stimulus for such sessions might be necessary to guide attention and facilitate creativity.

Establish a common ground when talking about AI. Knowledge about AI varies among people in general, the AEW&C context is no exception. Even among people with knowledge about AI there seems to be no consensus regarding the definition of the concept. During the study the users were not told what definition of AI was used. How this may have affected the study is unclear, but one cannot overlook that it may have affected the users answers on AI-related questions. Therefore, we suggest that when conducting studies related to AI, one should choose, or make, a definition of AI for the study and inform users, and other stakeholders, about this. For example, in this context it was not necessary that AI was continuously learning during use. This is related to the high risk associated with the context. AI with continuous learning could, eventually, lead to unexpected behaviour that can be difficult to predict. But considering the possibilities with continuous learning, it would be disadvantageous to discard it completely. When developing AI systems, one should assess the risk and rewards for each specific case and assess whether it is worth it or not to let the AI learn continuously.

12.2 Ethics: AI in Military Operations

AI raises a lot of ethical questions, especially when used in high-risk environments directly affecting people's lives, like military operations. The AI guidelines from the study provide some general guidance but every case needs to be treated uniquely. The AEW&C itself does not have any weapon

equipped, but since it is used to guide fighter jets it does indirectly. The decisions the operators take can have an impact on a whole nation's security and be the difference between life and death. This means developing new tools in the military context leads to great responsibility. A confusing solution could lead to faulty decisions and devastating consequences. This makes it important that the study is honest and transparent in the limitations of the findings. We must be clear about the validity of our findings and not exaggerate the possible implications of the results.

What should the AI be allowed to do in a military context? Our recommendation is to be cautious in letting the AI make its own decisions, especially if the decisions are critical. The safest recommendation would be to not let the AI make decisions on their own without human oversight and acceptance. The military context is simply associated with too much risk. One should also be cautious with nudging the users. Letting the AI make suggestions could potentially nudge users to make certain critical decisions. Inflated expectations about AI capabilities may lead the user to follow the suggestions blindly.

Knowledge about concepts of AI in general, and the AI implemented in the C2 system is important. AI does not have one single established definition and understanding of what it is may vary between people. If AI tools were to become implemented in the C2 system, education about AI and the tools should be a part of the operators training. The users should have a clear picture of capabilities and limitations of the tools they use. Transparency and explainability is important to facilitate comprehension and trust.

Regarding automation. Some automation is already included in the C2 system with proven success. This includes tasks like making tracks of plots. How does one assess an acceptable level of automation? And what tasks would be suitable to automate? Routine tasks that are cumbersome and time consuming to perform and do not imply great risks if not working could be suitable to automate. For example, it could be tasks that merely enhance humans Situation Awareness, like detecting and highlighting abnormalities or formations. When automating tasks, one should ask the question: What could be the consequences if the automation stopped working? This triggers questions like: Would it be associated with risks? Would the operator be able to take over? Could there be risks associated with skill loss? Even though the AI is properly trained and well tested there are always risks with AI making errors or dropping in performance, for example due to new unknown situations and adversarial attacks.

But what about tasks associated with risks where AI could potentially perform better than humans? Self-driving cars are an example of where AI is used to automate tasks typically performed by humans. Driving is associated with risks and is affecting human lives, but self-driving cars and other vehicles is still considered worthwhile. However, tremendous amounts of resources are spent on AI technology for self-driving cars. As suggested in the guidelines proposed in the study, higher risks require higher safety measures (*11.3.2. Evaluate potential errors and risks*).

To summarize the ethical discussion about AI in military operations, a conclusion is that it is important to make the distinction about AI in military operations on a general level, and AI in military operations on an applied level, like the case of the study. Military is a large and varied domain. This makes it hard to discuss on a general level, there are an overwhelming number of aspects to take into consideration. Therefore, we recommend discussing ethics on an applied level

instead. Looking back at the aim of the study, the aim was to explore when and how AI could be implied in the C2 system to improve operators' situation awareness. SA is a key driver for decision making. As previously discussed, the operators make decisions of high importance and potentially large impact. Wouldn't it then almost be considered unethical **not** to explore how new technologies, such as AI, could be used to improve their SA and assist them in their decision making?

12.3 Application of the Results

The guidelines were derived from articles mainly concerning AI from a user-centred perspective, for example, human-AI interaction and collaboration, explainable AI, trustworthy AI etc. The guidelines were also interpreted and written by engineers with human factors expertise. Therefore, the guidelines are mainly covering areas related to the users, the human aspect, of AI implementation. The guidelines are very brief regarding technological aspects of AI, providing general guidelines on how to select appropriate models and how to evaluate them. They also provide guidelines on what to consider regarding ethics. However, the ethical guidelines are not comprehensive for all cases, one should actively search for further considerations when designing new AI systems.

Even though the guidelines mainly cover user-related areas, they are not only meant for designers to follow. They provide guidance on how to consider the user in a human-AI system, what AI models to prioritize and ethical considerations, which are useful for all stakeholders taking part in development of AI tools.

Another relevant question to ask is when they can be used, and how. Our recommendation is to implement the AI guidelines as early as possible in the development process. In early stages of the development process the ability to make changes is higher and the cost of making changes lower. We believe the AI guidelines can be used to:

- *Explore AI opportunities:* This was mainly how the guidelines were used in this study. User studies were conducted to identify AI opportunities, mainly by adapting systems thinking and finding areas where AI could complement human skills.
- *Ideate and create concepts:* If opportunities to implement AI have been identified the guidelines can be used to ideate and create concepts that are user-centred, ethical, and safe.
- *Evaluate AI concepts:* If AI concepts have been created, they can be evaluated against the guidelines, to assess utility and if changes should be made.

However, the AI guidelines in their current state should be used with caution. They have not been evaluated to the extent that they can be considered having a high validity. Further research could include more thorough evaluations and iterations of the guidelines to increase their validity. The guidelines do also vary in both importance and abstraction. Some guidelines are indeed a must for all kinds of projects, while others are more situational. The more detailed the guidelines are, the more situational they tend to be. It is most likely not possible to follow all guidelines for a single project, all projects must be treated differently, and one must assess what guidelines are most relevant for each unique project.

12.4 Methods

Methods in the study were predominately based on situation awareness and workload to discern important user considerations for the guidelines and the concepts. The user studies methods were mainly used to find requirements for when to implement AI functionality and the how-parts were covered through literature studies and evaluations. In this section methods will be discussed concerning how they impacted the result of the study. In addition, adaptations in how methods were used in the study will be discussed.

12.4.1 NASA TLX and Situation Awareness Questions

NASA TLX was used to qualitatively measure workload for the operators. This enabled an objective way to find instances where workload was high. Since high workload impacts Situation Awareness, it was inferred that these tasks were candidates for AI-functionality that alleviate workload. The rankings in the methods were disregarded and operators were instead asked which sub-tasks impacted their workload the most for each of the six factors. This enabled a quick way to find difficult tasks since the study's main tasks were not selected prior to using the method. The system has vast functionality for the mission type simulated during user studies, that has only been partly treated in this report, and even more beyond the mission type. If the NASA-TLX would have been used the entire mission it would have been much more challenging to find subtasks that contribute to workload issues in the system. One of the preconditions for disregarding metrics in this method without affecting the utility of the results may have been the user studies prior to using the method and the combination with SA questions. Since some knowledge of difficult tasks were already obtained through interviews with operators, the study was safeguarded against missing important tasks, activities, and actions.

By using questions concerning SA provided by Endsley (2016) after each session in the MTS, information required for the tasks for the operators to improve their SA were found. The combination of finding the most critical tasks in terms of workload using NASA TLX and then finding out through interviews what information needed for perceiving, comprehending, and predicting regarding those tasks proved to be a useful strategy. In conclusion, the two methods worked well together in selecting tasks that could be facilitated through AI-functionality.

12.4.2 GTS and HTA

The selection of HTA's to represent task decompositions was motivated by the convenience and adaptability of the method. There were also additional positive effects that were found during various interactions with the company's employees. First, it was a way to verify how well the system was understood with supervisors and operators. Second, it proved to be useful in combination with other methods such as GTS, ideation, and concept creation. In ideation, especially in a large complex system, the HTA's that were expanded using the GTS, were used as a template to develop the ideas.

12.4.3 MTS Activities

The MTS activities were the most crucial activities in the project where much of the user research data comes from.

12.4.4 Evaluation guidelines and concepts together

The concepts and the guidelines were evaluated together. This strategy was deemed feasible since the concepts and the guidelines were created in parallel throughout the project. The strategy of doing this was based on the concept of *research through design* (Zimmerman, Forlizzi & Evensson, 2007).

12.4.5 Affinity diagrams

Affinity diagrams were a very useful strategy to categorize and sort data gathered from user studies as well as creating guidelines from the literature study.

12.5 Further Work and Improvements in Process

There are several methods that could have provided data that motivate certain guidelines or help designing new concepts. Various types of biometrics such as pupil dilation, blood pressure, sweat, heart rate and eye tracking. Eye tracking seems to be especially promising for both evaluating new concepts and could also be part of the concepts.

It became clear in the evaluation of the guidelines created in the study that guidelines for the planning phase were easier to use. The most obvious reason for this is that the study created concepts from scratch and that it may be easier to apply the guidelines from the design phase and evaluation phase once the concepts are further developed. However, it could also potentially mean that guidelines in the design and evaluation phase need to be further revised. Once AI-guidelines are further developed it will be important to apply and test the guidelines from the aforementioned phases.

The study mainly used qualitative measurement in evaluations and user studies. There are several interesting and important quantitative metrics that could be part of further research. For example, measuring workload and situation awareness. Workload and SA could thereafter be used to verify if concepts do improve operators SA. Gunning & Aha (2019) suggest testing AI functions by first creating a baseline where no AI-functionality is present in the system, second with AI functionality but without explainability, third with a low degree explainability and finally with a high degree of explainability. The same structure for evaluating guidelines and concepts in a continuation of this study could be useful to obtain a verification of the utility of concepts. In these verifications it would be beneficial to add interactability to the concepts to let operators test the concepts. To save time, the most liked concepts could be further developed and tested.

13

Conclusion

This chapter presents what was concluded in the study.

13.1 When to Implement AI

Below the conclusions related to the first research question: *When should AI be implemented to improve AEW&C operators' situation awareness?* are presented.

Conducting a system analysis was a useful way to start a project related to complex and technical systems like the one of the study. The Surveillance Operator (SO) and Fighter Controller (FC) are a part of a large complex system and to really understand their tasks and challenges, one must consider the whole system of which they are a part of. This includes both understanding how they are affected by the system, and how the system affects them.

Situation Awareness (SA) challenges were identified for each role, seven related to the FC role and six related to the SO role. These were:

- For FC: communication, window management, warnings, working methods & assessments, information filtering, finding & using tools, risk of errors.
- For SO: communication, window management, radar status, track management, radar areas, identification.

The two tasks with the highest workload for each role were identified as:

- For FC: Ensure safety and Guide to target.
- For SO: Identification and Collect target data.

Eight visual concepts suggesting when and how to implement AI to improve the SA for AEW&C operators were created, four for each role. These were:

- For FC:
 - A collision warning system to detect if assets run a risk of colliding through data-analysis of flight patterns.
 - A communication tool translating talk to text.
 - Ability to recognize formations of groups of suspect targets.
 - An adaptive route planner that suggests optimal routes based on intentions.
- For SO:
 - Abnormality detection.
 - Area suggestion to create and manage effective and efficient radar areas.
 - A timeline to display past and predicted events.
 - A tool to display patterns of behaviour in selected areas.

The four concepts considered having the most potential for further work were:

- For FC: Formation Recognition and Communication (talk-to-text)
- For SO: Abnormality Detection and Timeline (provided the prediction functionality is elaborated).

13.2 How to Implement AI

Below are the conclusions related to the second research question: *How should AI be implemented to improve AEW&C operators' situation awareness?* Or more specifically, conclusions regarding the created AI guidelines.

The guidelines created in the study were mainly applicable in three phases: *Planning*, *Designing* and *Evaluating*. Based on how the guidelines were applied in the study it was concluded that they can be used to:

- Explore opportunities to implement AI in an AEW&C context.

- Ideate and create AI concepts in an AEW&C context.
- Evaluate AI concepts in an AEW&C context

Even though the guidelines were mainly applicable in different phases, it was concluded that they should be considered during the whole development process. An example which strengthens the conclusion is the guidelines related to ethics. Even though they are a part of the *Planning* category it is crucial to consider ethics throughout the entirety of the project.

The study confirmed the utility and validity of several guidelines in the first category: *Planning*. However, further works are required to confirm the utility and validity of the guidelines in the subsequent categories *Designing* and *Evaluating*.

Lastly, the process designed in the study, inspired by *research through design*, was useful for both creating guidelines and concepts simultaneously, and thereby to answer the two research questions of **how** and **when** to implement AI to improve the Situation Awareness of the AEW&C Operator.

14 References

- Adobe Inc. (2019). Adobe Illustrator. Retrieved from <https://adobe.com/products/illustrator>
- Ainsworth, L. K. (2004). Task analysis. *Human Factors for Engineers*, 81.
- Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., ... & Horvitz, E. (2019, May). Guidelines for human-AI interaction. In *Proceedings of the 2019 chi conference on human factors in computing systems* (pp. 1-13).
- Bligård, L. O., & Osvalder, A. L. (2014). CCPE: methodology for a combined evaluation of cognitive and physical ergonomics in the interaction between human and machine. *Human Factors and Ergonomics in Manufacturing & Service Industries*, 24(6), 685-711.
- Brynielsson, J. (2007). Using AI and games for decision support in command and control. *Decision Support Systems*, 43(4), 1454-1463.
- Chalmers Student Portal. (2016). Retrieved from https://student.portal.chalmers.se/en/chalmersstudies/programme-information/Pages/SearchProgram.aspx?course_id=32318&parsergrp=2
- Engeström, Y. (1999). Activity theory and individual and social transformation. *Perspectives on activity theory*, 19(38), 19-30.
- Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. In *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 97-101). Santa Monica, CA: Human Factors Society.
- Endsley, M. R. (2016). *Designing for situation awareness: An approach to user-centered design*. CRC press.
- European Commission. (2019, 8 April). A DEFINITION OF AI: MAIN CAPABILITIES AND DISCIPLINES. Brussels
- Flood, R. L., & Carson, E. (1993). *Dealing with Complexity: An Introduction to the Theory and Application of Systems Science*. Springer Science & Business Media.
- Gaillard, A. W. K. (1993). Comparing the concepts of mental load and stress. *Ergonomics*, 36(9), 991-1005.
- Gerrig, R. J., Zimbardo, P. G., Campbell, A. J., Cumming, S. R., & Wilkes, F. J. (2015). *Psychology and life*. Pearson Higher Education AU.
- Gunning, D., & Aha, D. (2019). DARPA's Explainable Artificial Intelligence (XAI) Program. *AI Magazine*, 40(2), 44-58. <https://doi.org/10.1609/aimag.v40i2.2850>
- Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology* (Vol. 52, pp. 139-183). North-Holland.
- Hollan, J., Hutchins, E., & Kirsh, D. (2000). Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 7(2), 174-196.
- Hollnagel, E. (2001). Time and control in joint human-machine systems.
- Hollnagel, E. (2004). Automation and human work. *Human factors for engineers. The Institution of Engineering and Technology, Herts*, 113-136.
- Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.
- van den Bosch, K., & Bronkhorst, A. (2018). Human-AI cooperation to benefit military decision making. NATO.
- Microsoft Corporation. (2018). Microsoft PowerPoint. Retrieved from <https://office.microsoft.com/PowerPoint>
- Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial intelligence*, 267, 1-38.
- Osvalder, A L., Rose, L., Karlsson, S. (2015) Arbete och teknik på människans villkor. Chapter 9, Metoder. (3:1) Oxford University Press. (2020). *Definition of system*. Lexico.com. <https://www.lexico.com/definition/system>

Sanders, M. S., & McCormick, E. J. (1993). *Human factors in engineering and design* (7th ed.). McGraw-Hill Book Company.

Sanneman, L., & Shah, J. A. (2020, May). A Situation Awareness-Based Framework for Design and Evaluation of Explainable AI. In *International Workshop on Explainable, Transparent Autonomous Agents and Multi-Agent Systems* (pp. 94-110). Springer, Cham.

Das, S., & Grecu, D. (2000, June). COGENT: Cognitive agent to amplify human perception and cognition. In *Proceedings of the fourth international conference on Autonomous agents* (pp. 443-450).

Gunning, D., & Aha, D. (2019). DARPA's explainable artificial intelligence (XAI) program. *AI Magazine*, 40(2), 44-58.

Smuha, N. A. (2019). The EU approach to ethics guidelines for trustworthy artificial intelligence. *Computer Law Review International*, 20(4), 97-106.

Tsang, P. S., & Vidulich, M. A. (2006). Mental workload and situation awareness.

Janhager, J. (2005). *User consideration in early stages of Product Development: Theories and methods* (Doctoral dissertation, KTH).

The University of Helsinki. (2021). *How should we define AI?*. <https://course.elementsofai.com/1/1>

Young, M. S., & Stanton, N. A. (2002). Attention and automation: new perspectives on mental underload and performance. *Theoretical issues in ergonomics science*, 3(2), 178-194.

Zimmerman, J., Forlizzi, J., & Evenson, S. (2007, April). Research through design as a method for interaction design research in HCI. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 493-502).

Warm, J. S., Parasuraman, R., & Matthews, G. (2008). Vigilance requires hard mental work and is stressful. *Human factors*, 50(3), 433-441.

Wickens, C. D., Hollands, J. G., Banbury, S., & Parasuraman, R. (2015). Mental workload, stress, and individual differences: cognitive and neuroergonomic perspectives. *Engineering Psychology and Human Performance (International Edition)*, 346-376.

15 Appendices

In this section the appendices in the study are presented. Appendix A lists the sources of articles used to create the guidelines. Appendix B shows the first iteration of the guidelines. Appendix shows the NASA-TLX and Situation awareness questions used in the MTS activities. Appendix D shows the first iteration of ideas generated in the project. Appendix E-L shows the eight different final concepts including the frames not displayed in the chapter *Final Concepts*.

A. Appendix - Sources of AI guidelines

This appendix includes the main sources used for the first iteration of AI guidelines.

Amershi, S., Weld, D., Vorvoreanu, M., Fournery, A., Nushi, B., Collisson, P., ... & Horvitz, E. (2019, May). Guidelines for human-AI interaction. In *Proceedings of the 2019 chi conference on human factors in computing systems* (pp. 1-13).

Artificial intelligence. (2021, 3 April). In *Wikipedia*. https://en.wikipedia.org/wiki/Artificial_intelligence

Brynielsson, J. (2007). Using AI and games for decision support in command and control. *Decision Support Systems*, 43(4), 1454-1463.

Collisson, P.(2019, 7 march). *AI Guidelines in the Creative Process*. Medium. <https://medium.com/microsoft-design/ai-guidelines-in-the-creative-process-807b6d31cda2>

Das, S., & Grecu, D. (2000, June). COGENT: Cognitive agent to amplify human perception and cognition. In *Proceedings of the fourth international conference on Autonomous agents* (pp. 443-450).

European Commission. (2019, 8 April). *A DEFINITION OF AI: MAIN CAPABILITIES AND DISCIPLINES*. Brussels

European Commission. (2019, 8 April). *ETHICS GUIDELINES FOR TRUSTWORTHY AI*. Brussels

Gunning, D., & Aha, D. (2019). DARPA's explainable artificial intelligence (XAI) program. *AI Magazine*, 40(2), 44-58.

Microsoft. (2020, 7 February). *How to build effective human-AI interaction: Considerations for machine learning and software engineering*. <https://www.microsoft.com/en-us/research/project/guidelines-for-human-ai-interaction/articles/how-to-build-effective-human-ai-interaction-considerations-for-machine-learning-and-software-engineering/>

Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial intelligence*, 267, 1-38.

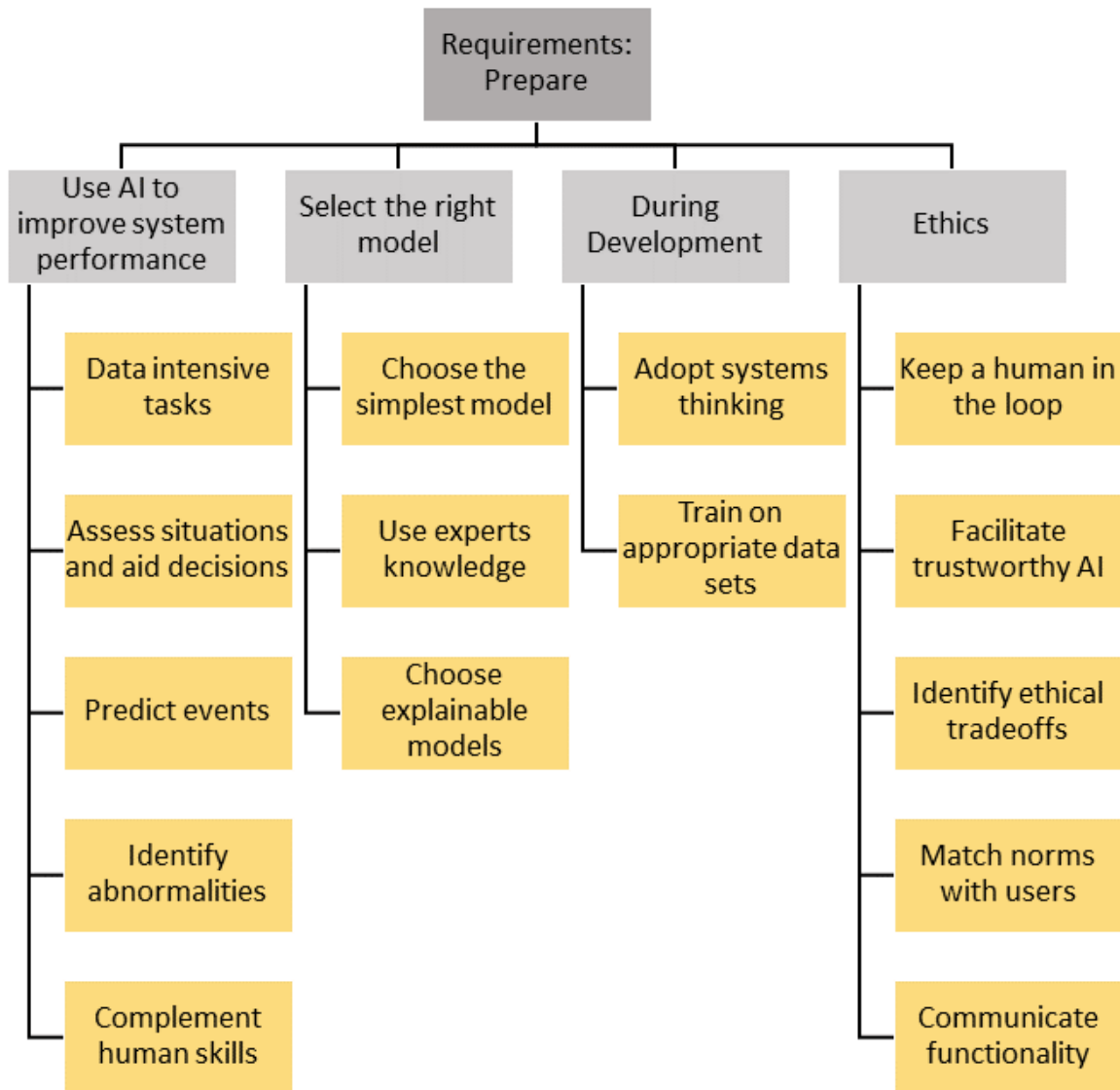
Sanneman, L., & Shah, J. A. (2020, May). A Situation Awareness-Based Framework for Design and Evaluation of Explainable AI. In *International Workshop on Explainable, Transparent Autonomous Agents and Multi-Agent Systems* (pp. 94-110). Springer, Cham.

Valdez, A. C., & Ziefle, M. (2018, July). Human factors in the age of algorithms. understanding the human-in-the-loop using agent-based modeling. In *International Conference on Social Computing and Social Media* (pp. 357-371). Springer, Cham.

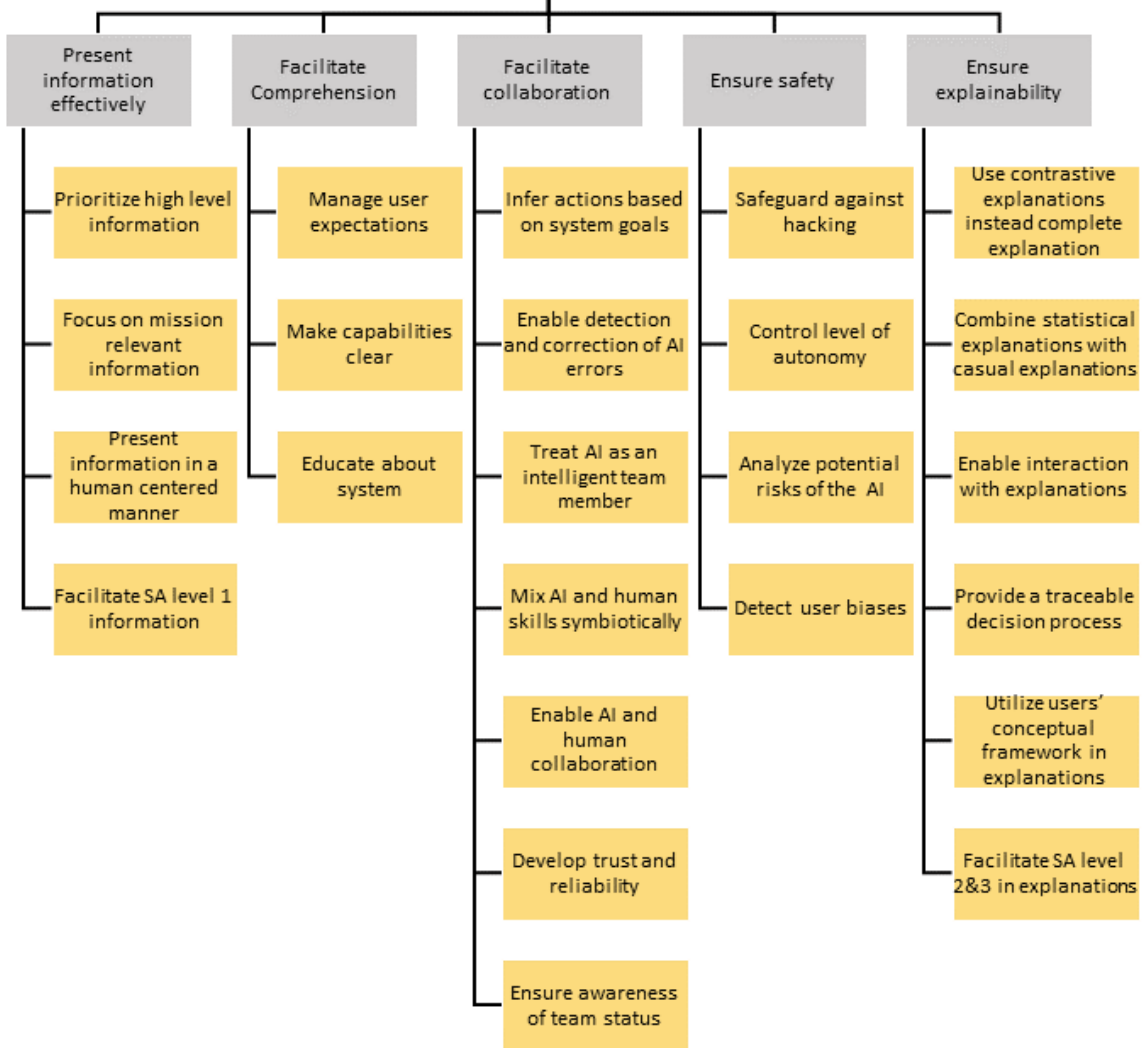
van den Bosch, K., & Bronkhorst, A. (2018). Human-AI cooperation to benefit military decision making. NATO.

B. Appendix - First version of AI guidelines

This appendix includes the first version of AI guidelines.



Requirements: Use



Requirements: Evaluation

Adaptation

Promote and enable user feedback

Adapt from user behavior

Notify users about user generated adaptations

Be conservative in adaptations

Errors/Risks

Assess limitations of the AI

Understand that attacks on opaque AI is hard to detect

Don't follow AI blindly

Evaluate model

Use multiple evaluation measures

Evaluate explanations effectiveness

Use agents to simulate human operators

C. Appendix - NASA-TLX and SA questionnaire

This appendix includes the questionnaire used in between the operative simulations.

NASA-TLX frågor:

Följande frågor är baserade på en väldokumenterad metod för att utvärdera arbetsbelastning, metoden heter NASA-TLX.

I metoden finns det 6 olika faktorer som anses bidra till arbetsbelastning: hur mentalt krävande uppgiften var, hur fysiskt krävande den var, hur hög tidspressen det var, din egen prestation, hur mycket du ansträngning som krävdes och hur frustrationsnivån var.

I metoden ska användaren egentligen betygsätta varje sån här faktor på skalor från lågt till högt, för varje enskild uppgift. Men för att spara tid och för att vi inte är helt säkra på vilka uppgifter är mest intressanta att titta på har vi en mer generell approach till metoden.

- Till att börja med, för hela scenariot du precis gick igenom, vilken av dessa sex faktorer anser du bidrog mest respektive minst till den totala arbetsbelastningen? (mentalt krävande, fysiskt krävande, tidspress, din prestation, ansträngning, frustration).

Svar

- Vilka moment/uppgifter anser du vara minst respektive mest mentalt krävande? Varför?
- Hur påverkade den mentala belastningen i dessa uppgifter din situation awareness (lägesbild)? Varför?

Beskrivning mentalt krävande: Hur mycket mental och perceptuell aktivitet krävs (t.ex. tänkande, beslutande, beräkning, minne, tittande, sökande, osv.)? Var uppgiften enkel eller krävande, simpel eller komplex, exakt eller förlåtande?

Svar

- Vilka moment/uppgifter anser du vara minst respektive mest fysiskt krävande? Varför?
- Inget.
- Hur påverkade den fysiska belastningen i dessa uppgifter din situation awareness (lägesbild)? Varför?

Beskrivning fysiskt krävande: Hur mycket fysisk aktivitet krävdes (t.ex. trycka, dra, vrida, kontrollera, aktivera. etc.)? Var uppgiften lätt eller krävande. långsam eller snabb, slapp eller slitsam, vilsam eller mödosam?

Svar

-
- Vilka moment/uppgifter anser du ha lägst respektive högst tidspress? Varför?
 - Hur påverkade tidspressen i dessa uppgifter din situation awareness (lägesbild)? Varför?

Beskrivning tidspress: Hur mycket tidspress kände du på grund av den hastighet eller takt som uppgiften hade eller element av uppgiften inträffade? Var tempot långsamt och lugnt eller snabbt och häftigt?

Svar

- I vilka moment/uppgifter anser du att din prestation var sämst respektive bäst? Varför?
- Hur påverkar din prestation i dessa uppgifter din situation awareness (lägesbild)? Varför?

Beskrivning prestation: Hur framgångsrik tycker du att du var i att uppnå målen för uppgiften? Hur nöjd var du med din prestation för att uppnå dessa mål?

Svar

- Vilka moment/uppgifter anser du vara minst respektive mest ansträngande? Varför?
- Hur påverkade ansträngningen i dessa uppgifter din situation awareness (lägesbild)? Varför?

Beskrivning ansträngning: Hur mycket behöver du anstränga dig (mentalt och fysiskt) för att uppnå din nivå av prestanda?

Svar

- Vilka moment/uppgifter anser du vara minst respektive mest frustrerande? Varför?
- Hur påverkade frustrationen i dessa moment/uppgifter din situation awareness (lägesbild)?
Varför?

Beskrivning frustration: Hur osäker, avskräckt, uppretad, stressad, och irriterad kontra säker, tillfredsställt, belåten, avslappnad och självbelåten kände du dig under uppgiften?

Svar

D. Appendix - First list of ideas

This appendix includes the ideas derived in the first brainstorming session.

AI reports automatically when tracks cross borders etc.

AI monitors the team members and assess their workload. The status is communicated to the rest of the team (or only MCC).

Recognize if someone is getting tired and needs to take a break.

AI tracks what everyone is looking at (through following screen or eye tracking) and recognize if some areas have not been covered in a while.

AI recognize who is speaking to who and when people is available to speak to

Collision warning system. Show areas where collisions may happen based on predictions of flight paths. Indicate what track it relates to. Display when the collision may happen (ex 30s) at what height (+100 feet).

Suggest safe flight routes for fighter, based on civil flight plans (or predictions of paths). Show the safe route with different colour, green, yellow, red for example. Maybe consider height and suggest going down or up (maybe some kind of height indicator, like 3D or sommarjobb 2019).

Assess threat level of tracks. Indicate with a line with multi-coloured indications what the safety levels are. Highlight when moving towards dangerous distance.

Display threat levels of different areas, areas around tracks and locations (K for example).

AI can make use of “unused” radar capacity. –When RL and RC is not high, radar can look at other areas (passive scanning). What other areas to look at can be determined by for example: heatmaps where operators look (search in unlooked areas), previous interesting areas, tracks that are in suspicious areas. Indicate when something is abnormal, even when it is outside mission area.

Display RL and RC in a way that facilitates SA. Like graphs with indications of levels.

Translation of talking to text. Like an in-game chat, fades out after a while and have time stamps. Or movie subs, might require some storage of the conversation (or not?). The fighter could have a standalone display with the chat, or maybe it is not necessary that they have the same support. (Keep in mind that the information decisions are made based on must be up to date, every second counts. Is not suitable to record voice messages in most cases.)

Pattern recognition. (Like champagne glass, box etc). Could be used to create groups and suggest what tracks are a part of that group (even if they have different ID).

Send HLFKA information directly to fighter (by marking it as an object for example).

Target prio assessment. Based on for example EW, speed, path (towards border or high prio asset), position, etc.

Provide global SA of suspects. Provide easy navigation to suspects. Could be a minimap, displayed in VSF, indications on the edges of the screen, a list of how many of each ID (Can be used as filter and linking & brushing).

Adaptive/dynamic route planner for fighter. Considers moving targets and prio. Like an advanced stern support tool. (How should it behave if it cannot ID everyone? Ask for help?)

Route prediction of a track, suggest the most probable route and other possible routes.

Help to provide global SA in intensive FC missions. Close down windows that have not been used in a while. Let windows fade away slowly.

Auto pilot leading. AI takes care of the communication and guides the fighter. FC have oversight, chooses what targets to ID and can interrupt if AI performs bad.

Push to record and send recording + text. When the responder to the message is available it plays in audio and text.

Push to translate voice to text (when having a lot to remember).

Link mission windows to corresponding fighter (for example colour coding or position)

AI estimate fight value (stridsvärde) of fighter. Based on time in air, speed etc.

Mirror fighter radar display on FC screen (or extra screen). Better understanding of fighters SA.

Dynamic indication on how much the radar settings and size of area contributes to RL and RC.

AI makes EEW (high priority area) nearby of high priority suspects, to make it easier to find other targets in the group.

Voice control. Voice activated areas. “Activate area XXX”, “Deactivate area XXX” “Make high priority area over blabla”. (Voice control might not be suitable in C2 environment when communication with other operators is a crucial part for mission success).

Store use profiles/mission profiles. Based on use profiles (or mission profiles) suggest areas or settings. Remind if something has been forgotten. Have your own profile so it is easier to take over a role. Suggest most used tools for that role. Or suggest favourite tools etc. Plot train settings.

AI reads the brief (like resumes are read and autofills forms) and suggest areas and settings. AI learns when suggestions are accepted or denied. (Explain why areas were suggested, refer to parts of the brief etc). Have information about the areas. The user can finetune the suggestions, the AI learns by the changes the user makes.

AI highlight abnormal tracks. Users set intervals to indicate what is abnormal or normal. AI could learn from previous actions of users when they have set suspect on a track.

Filter tracks. Set intervals to filter out interesting tracks. Like prisjakt. Could also work as dynamic feedback, give feedback on the tracks when you adjust parameters.

Linking and brushing. When hovering filter or characteristics or similar (for example nations, or fishing boats etc.), highlight related tracks. When hovering a track, highlight tracks with similar characteristics. (Look at the IDE project)

AI derived flight patterns (as map layer, or quick command).

Time bar. To help get an understanding of the movements of plots and tracks over time. The get a normal picture (normalbild). Could extend in future to show predicted behaviour.

High resolution time camera (maybe not the right word for it but inspired by old cameras that took long to take the picture). Get a “smushed” picture of the situation. Like map layer pattern.

AI read the brief and suggest ID based on that. Accept/change/ignore to teach AI. Explainable AI can be important.

Suggest ID based on previous identification of tracks. AI learns over time.

Automate ID, or multiple ID based on characteristics.

Watch the AI work with ID in real-time, like an autopilot. The AI works like an operator would do. Indicate where the AI is working with a moving eye, show what parameters the AI looks at. The operator can interrupt when necessary. The operator can accept/change/deny the AI's decisions. The operator could get the choice to decide how fast the AI should identify tracks, instant or X/minute. The operator can identify or work with other stuff in parallel.

AI can recognize when a plot and track gets separated. A track without a plot should be highlighted and the AI can suggest fusing the track with another plot/track or deleting it.

Indicate who is responsible for a tracks ID (or who has marked it as an object, high priority etc.)

Help the users remember tracks (okänd tills vidare). If an unknown track is unknown for too long, indicate that. If a track has been selected and left without ID, remind of that track after a while. Maybe have some indication if a track has been “opened” (like with emails).

Show where people have been working/looked with heatmap. (Could be quick command, minimap e.g.)

AI that compares reported data from transponder (IFF or AIS) to actual data. Check if the data fits to the radar estimation (with size, nationality etc.)

Smushed plot trains. To easier see the movement of sea- and land plots.

Setting to only display reappearing plots, not displaying plots only appearing one time (might be a bad idea).

AI suggest type of vehicle (type of car, boat, plane etc) based on movements, location, size etc.

Heatmaps to indicate activity of plots/tracks.

Set the AI to look for certain behaviour in a certain area.

Follow/highlight tracks from suspicions areas. Select locations of high priority and follow plots that appears from that place/enter that place. Notify when tracks stop/disappear or starts/appear from there.

Summary of activity in an area. Summary report. Like Spotify review of the year, youtube rewind, or fast forward of a growing flower. Summarize track behaviour, heatmaps, patters, number of suspects etc.

Suggest in what order to identify tracks, based on prio.

Ideas:

FC ideas

1. Status window for all relevant information about the mission. For example, fighter fuel levels, active fighters, targets, estimated time to take care of all targets. The status is the aggregated into a number that tells the mission health, in some cases actions can be suggested such as: rtb fighter, add new fighter. If no actions are suggested the mission health can be assumed sufficient to reach the goals by a certain margin.
2. Place out cap-point. The AI could suggest where to place out a cap based on how strategic the point is for the mission.
3. Describing situation to fighter by giving suggestions for shapes derived from AI pattern analysis. Need to be trained for common patterns that planes fly in. The tool is enabling quicker descriptions of the situation for the fighter. The tool also helps with giving a reference point to describe HLFK.
 - a. Fill formations, if there are a bunch of planes flying in a formation and one or more are set to a certain identity the AI will suggest identities for the others based of their participation in the formation.
 - b. This could probably be done for ships too.
4. Reports of the situation will be speech-text analysed and prepared into a document.
5. Collision warning system with civil air traffic. AI will also suggest the best route to avoid collisions.
6. AI will show the best route to approach several hostile or threats.
7. Flight safety against enemy planes could be done through displaying the area around a hostile plane where enemy robots are able to hit from. If these areas are clicked additional information about the EW/HLFK and origin is showed with a probability that the plane is using a particular weapon.

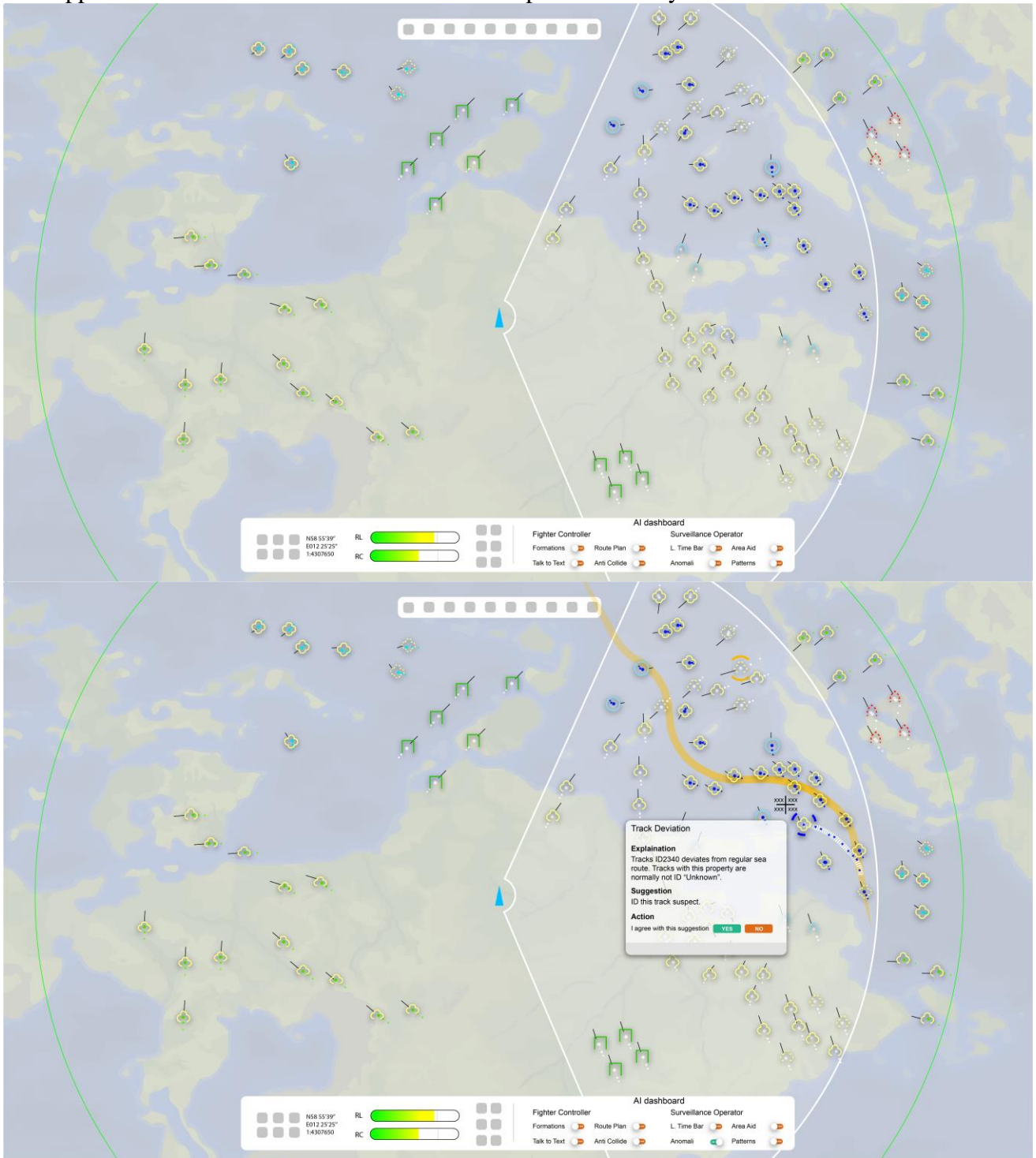
SO

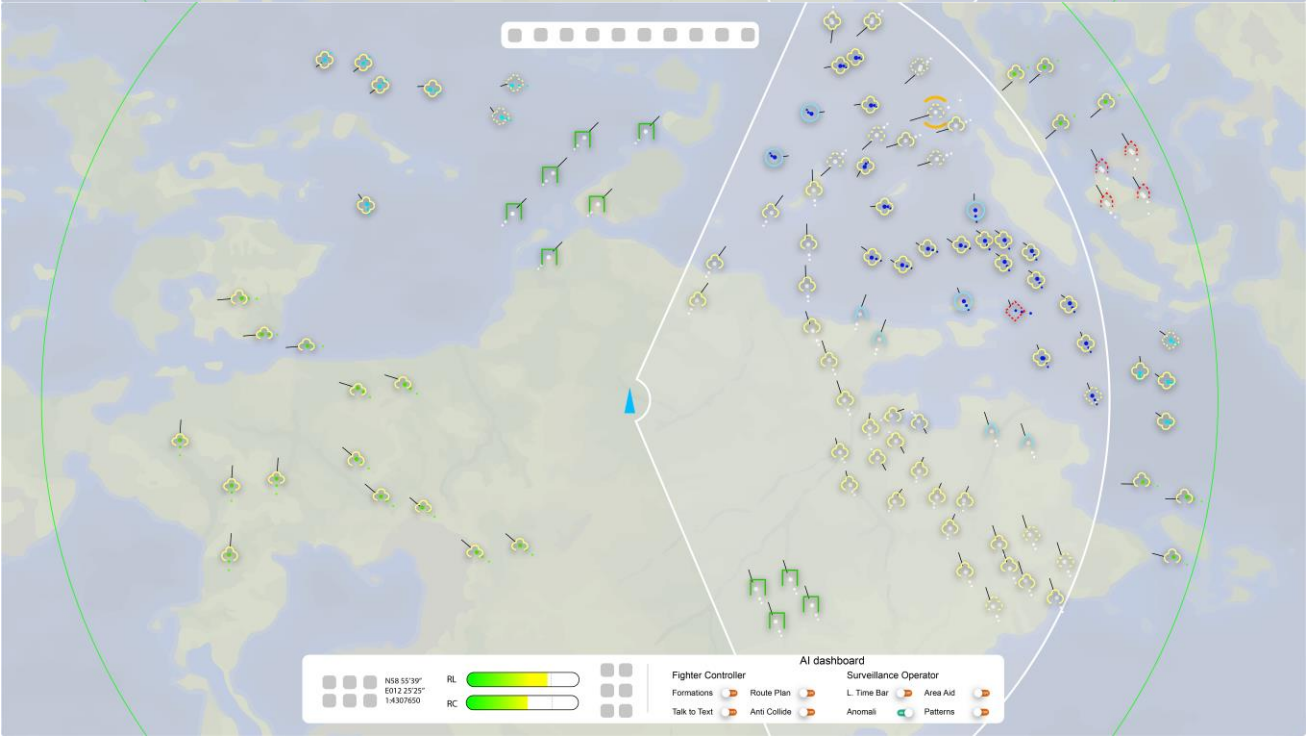
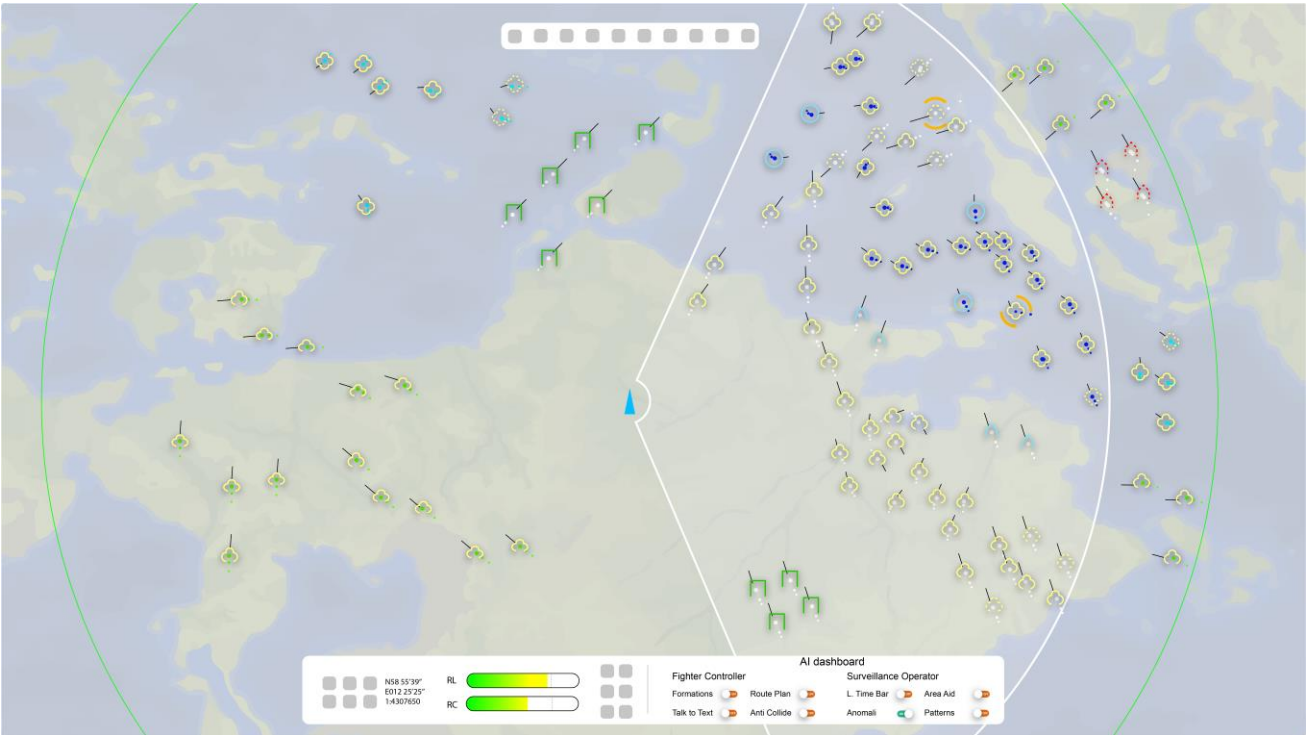
1. Helping operators to detect important tracks. In this idea the tracks get prioritised by their HLFK, origin and EW to show the operator which tracks to focus on.
 - a. By clicking on the track, the AI thought process will be displayed in a decision tree or similar.
2. Auto ID using the same priorities as above will be done. Auto-identified tracks will have a different symbol over them that indicates that they are automatically identified. If the operators click on the track the previously discussed decision tree will be showed.

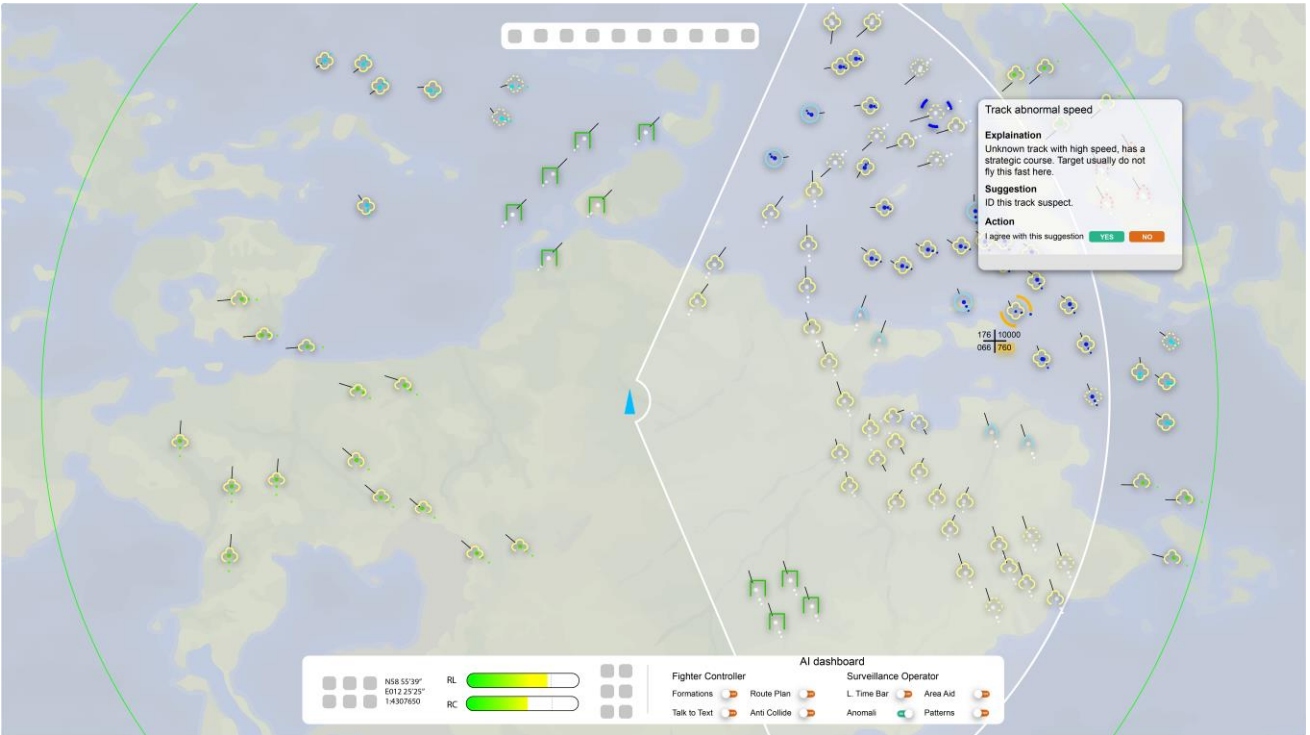
3. Tracks that are of high interest outside the current zoom level will be highlighted in the edge of the display.
4. Showing the behaviour of tracks as a line after them, if the line looks like a threat-line the usual threat behaviour will be showed.
5. Anomalies will be showed by displaying the flight routes, if a track succeeds from the usual flight path that day the track is highlighted.
6. Heat maps of the current situation where most tracks move can be showed. This could be divided into different categories where sea, land and air could be distinguished. This will also be an overlay that will be accessed through a quick command.
7. Heat maps could also be used to showed where deviations from a normal picture occurs.
8. Time bar deluxe where you not only could view past pictures of the situation but also predictions of where tracks could be in the future based on their HLFK, origin and EW. The user will be able to slide the bar of time using a quick command.
 - a. Tracks that will probably pass a certain border or other areas of interest or get close to own assets will be highlighted.
9. Ground plots will be connected will lines where the most likely connection is showed. This is an overlay function.
10. For ground tracks the environment can be considered in predictions for future positions, for example where there is a road or if there are obstacles that are difficult or impossible to pass with a ground-based vehicle.
11. Tracks on the sea can be connected to common sea lanes. Tracks that deviate can be highlighted.
12. Predictions of radar areas' load on the radar capacity can be showed explicitly.
13. Current track information will be showed in the same place such as the number of tracks, how many tracks that are identified, AI-agreement with track ID, which types of tracks (air, land, or sea).
14. See all areas in a list, AI suggest optimal priorities for the different areas. And how much load there are for each area.
15. Suggest where and how and area should look like.
16. Predictions on how an area will look like based on operator clicks, lower level that the idea above. Saves time.
17. AI suggests changes in priorities for different created areas to optimise radar efficiency and effectiveness.
18. AI reads the brief and setups everything, operator will only approve or disapprove.

E. Appendix - Final concept: Abnormality detection

This appendix includes all frames of the final concept Abnormality detection.







F. Appendix - Final concept: Area suggestions

This appendix includes all frames of the final concept Area suggestions.



Creating Air Area

Track Name: PAS

Setting 1 Setting 2 Setting 3

Advanced Settings: Setting 2

Explanation

Brief says: "create area over west part of country XXX". Area is suggested based on similar missions with similar briefs.

Similar missions: XXX 2021-04-28

ACCEPT DENY CHANGE

AI dashboard

N58 55'39" E012 25'25" 14307650
 RL
 RC

Fighter Controller
 Formations Talk to Text

Surveillance Operator
 Route Plan L. Time Bar Area Aid
 Anti Collide Anomaly Patterns

AI dashboard

N58 55'39" E012 25'25" 14307650
 RL
 RC

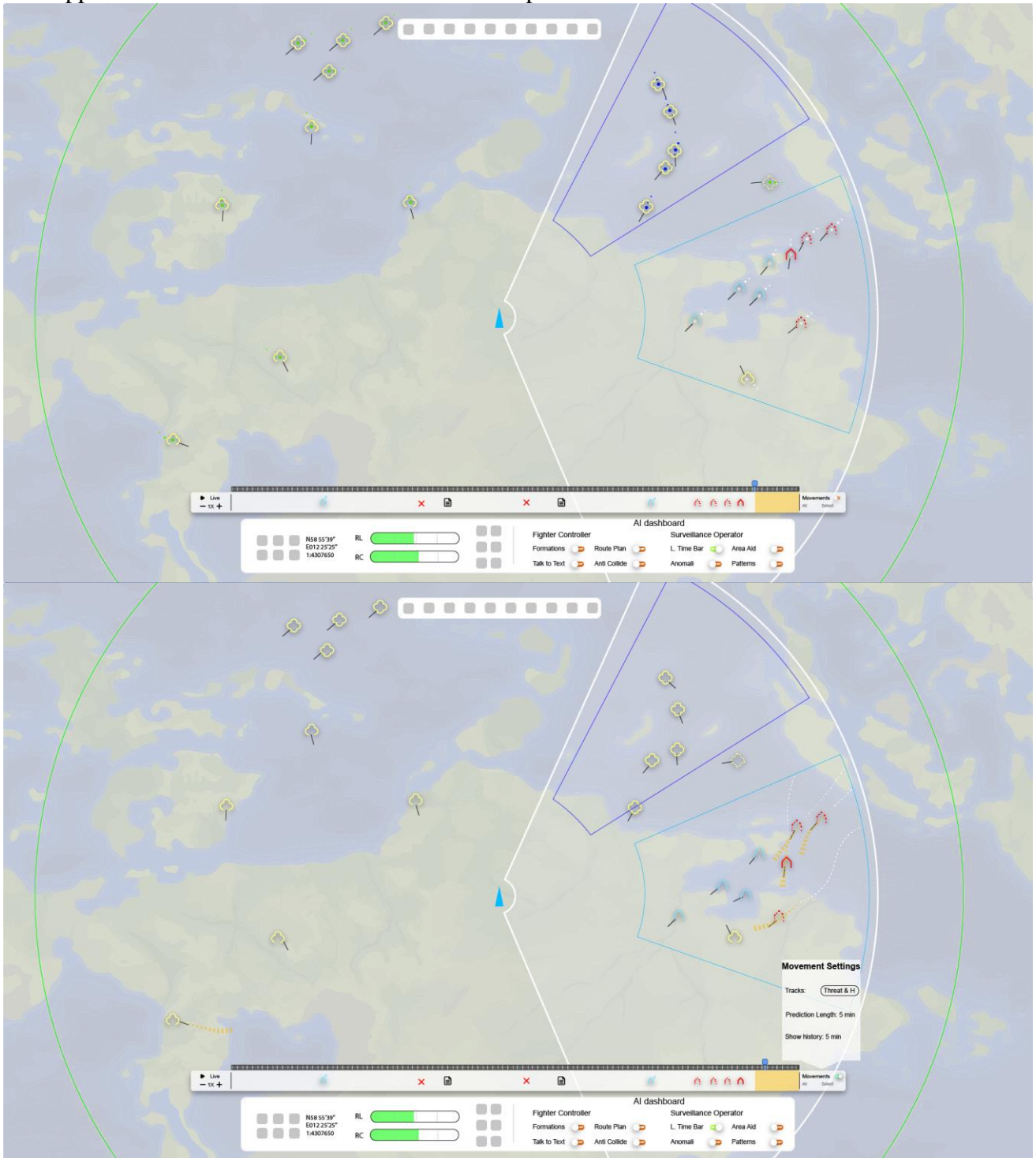
Fighter Controller
 Formations Talk to Text

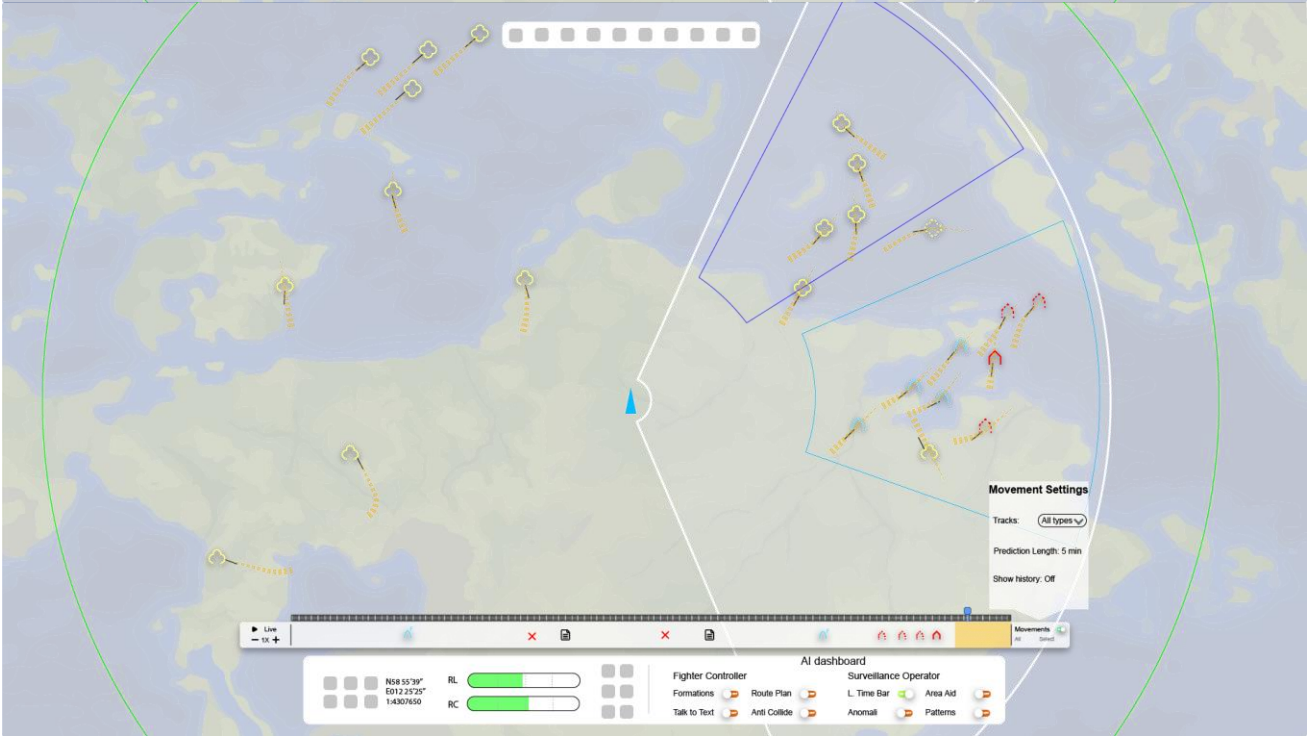
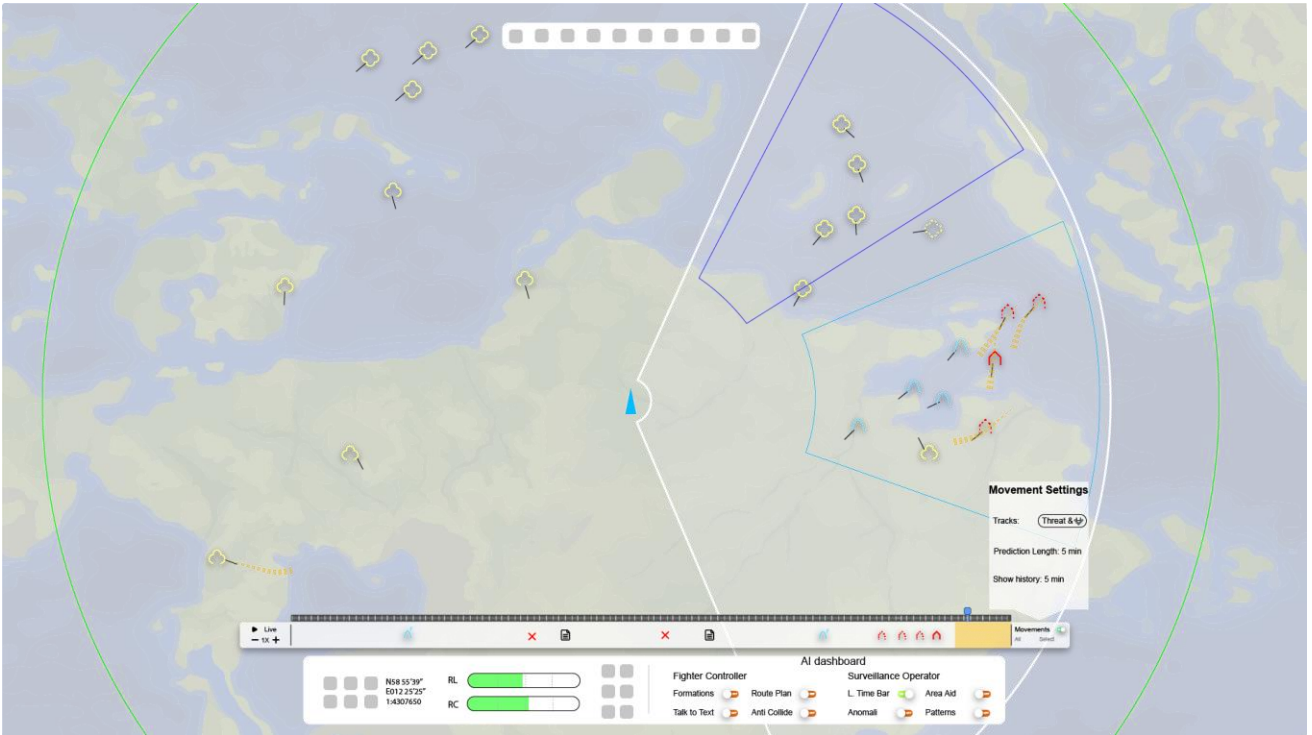
Surveillance Operator
 Route Plan L. Time Bar Area Aid
 Anti Collide Anomaly Patterns

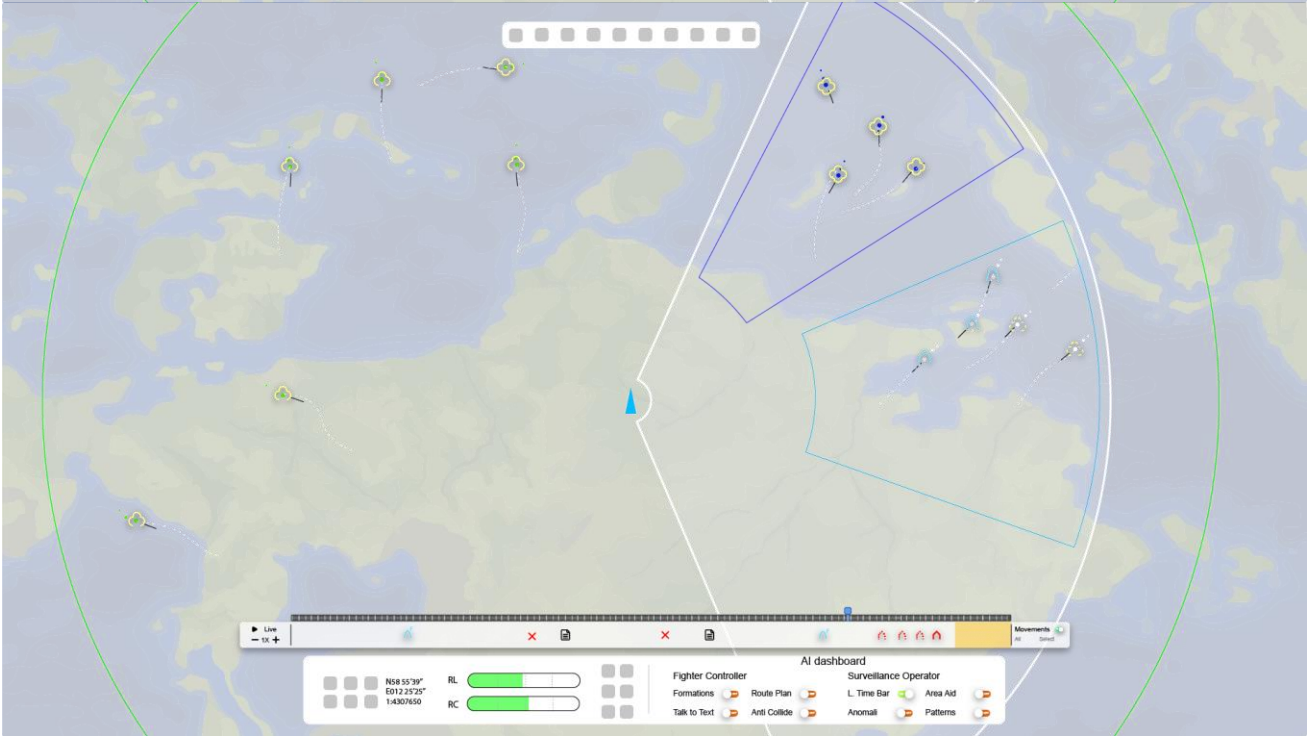
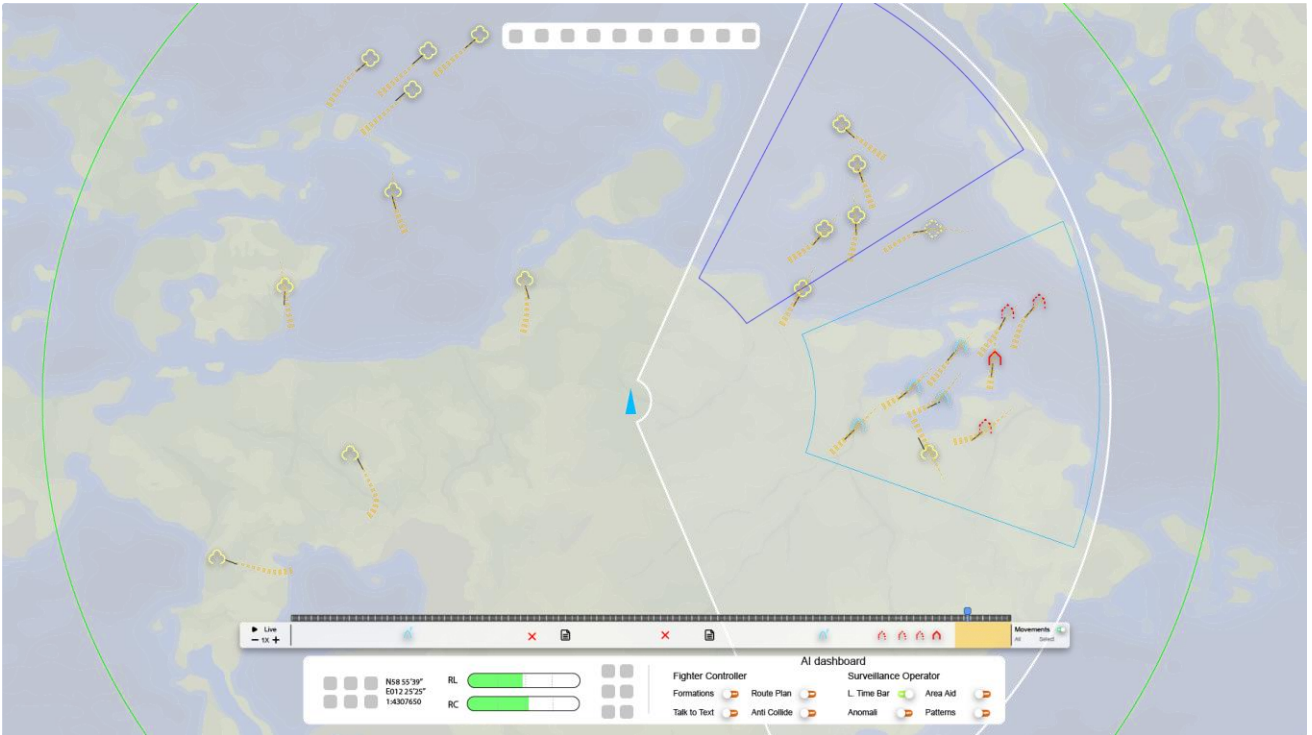


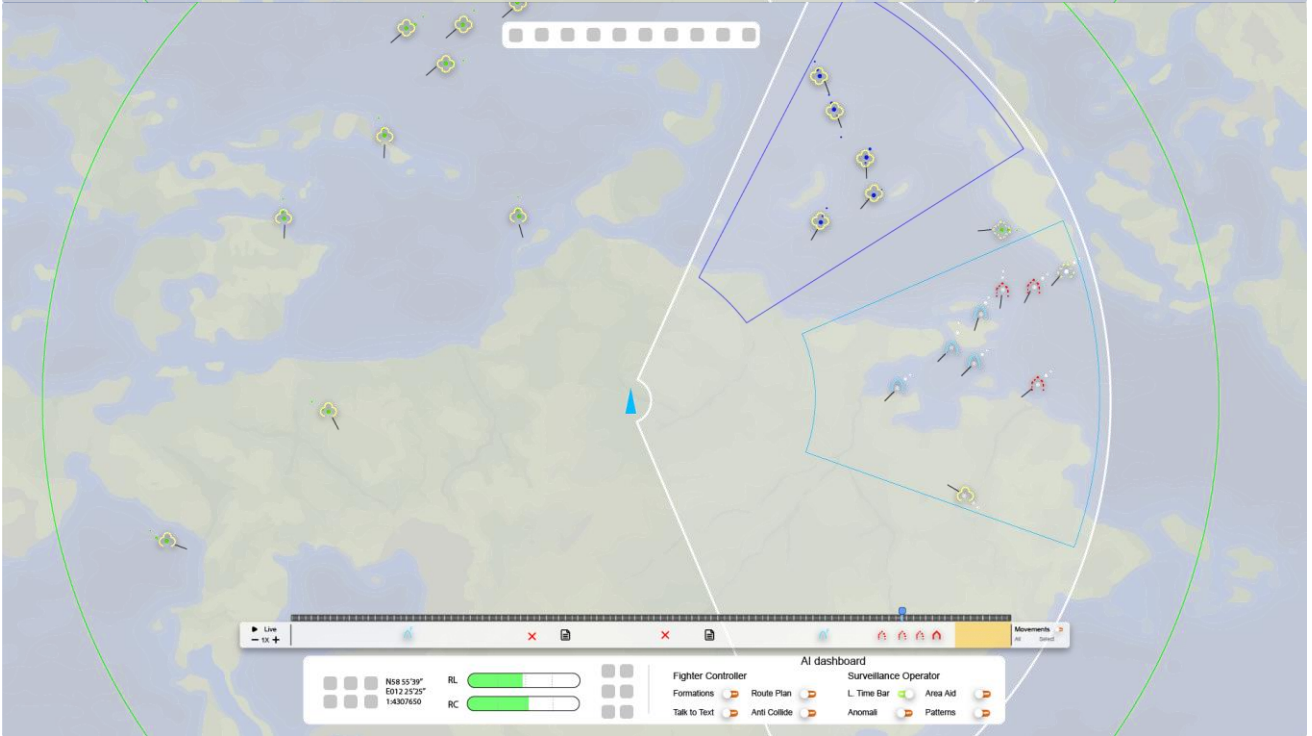
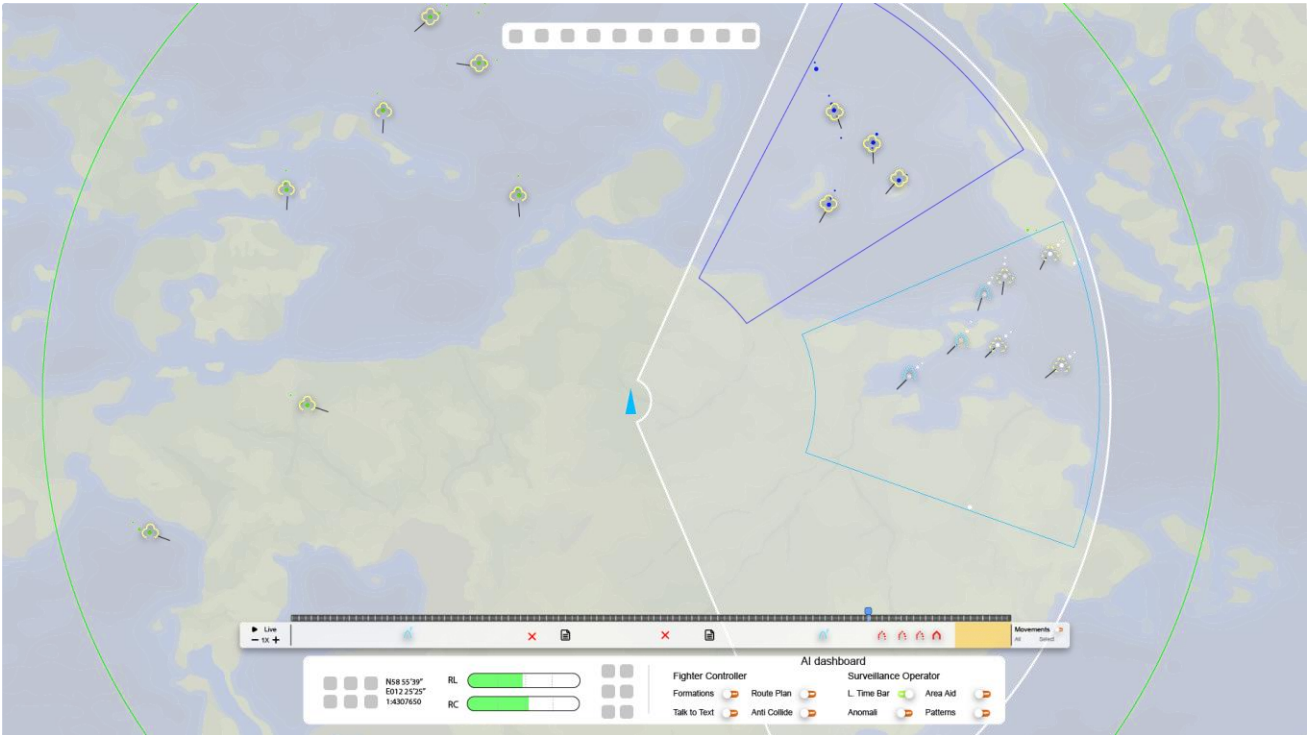
G. Appendix - Final concept: Timeline

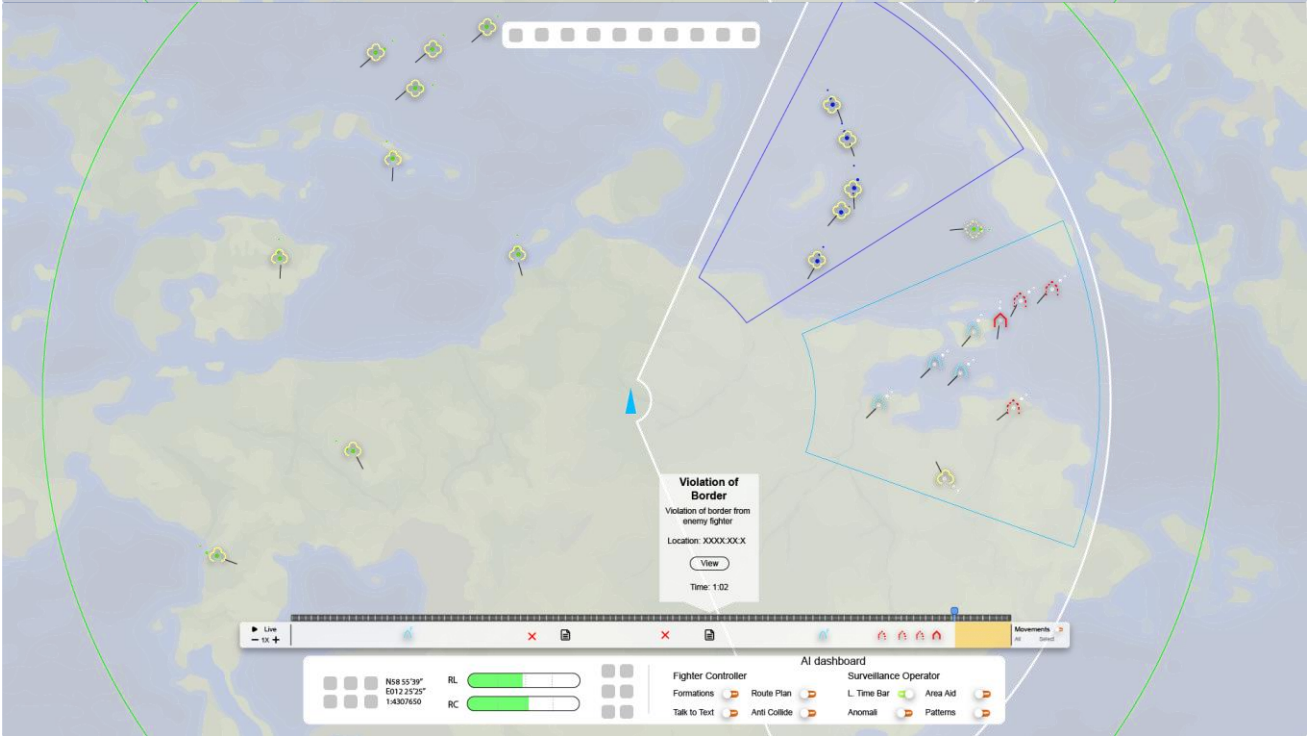
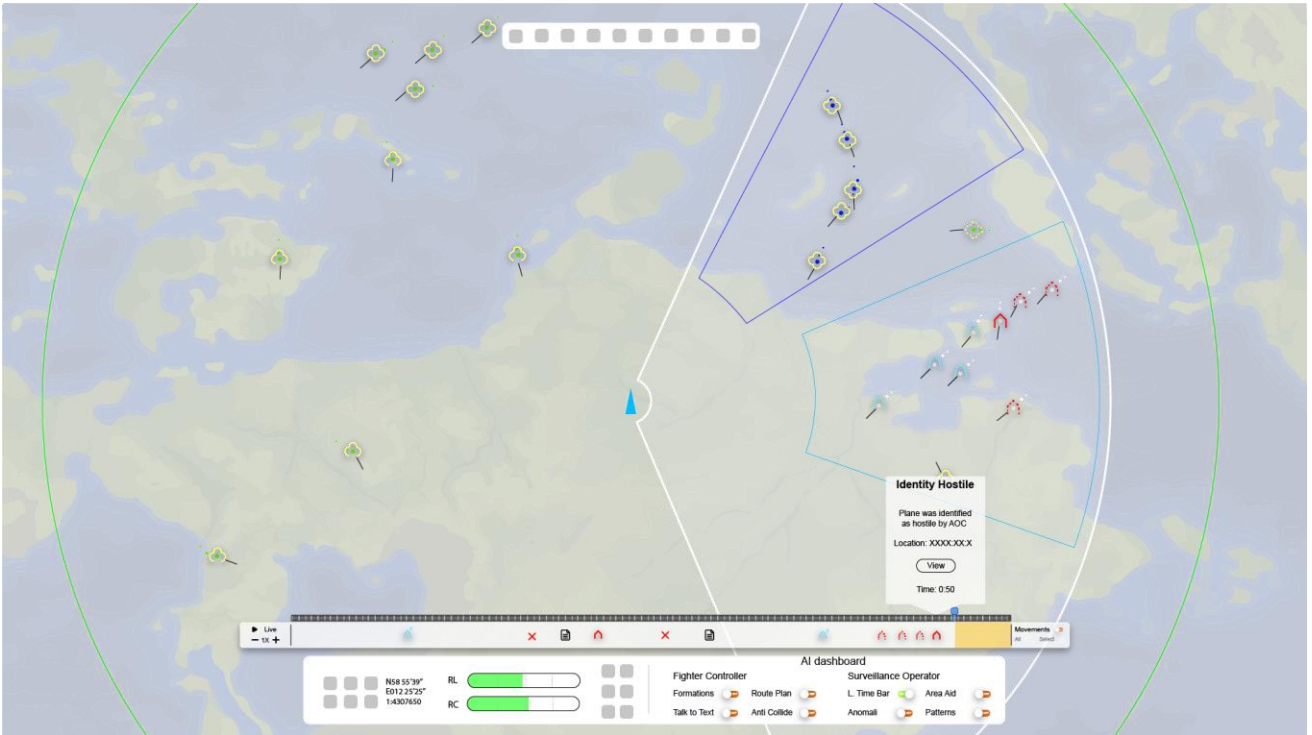
This appendix includes all frames of the final concept Timeline.

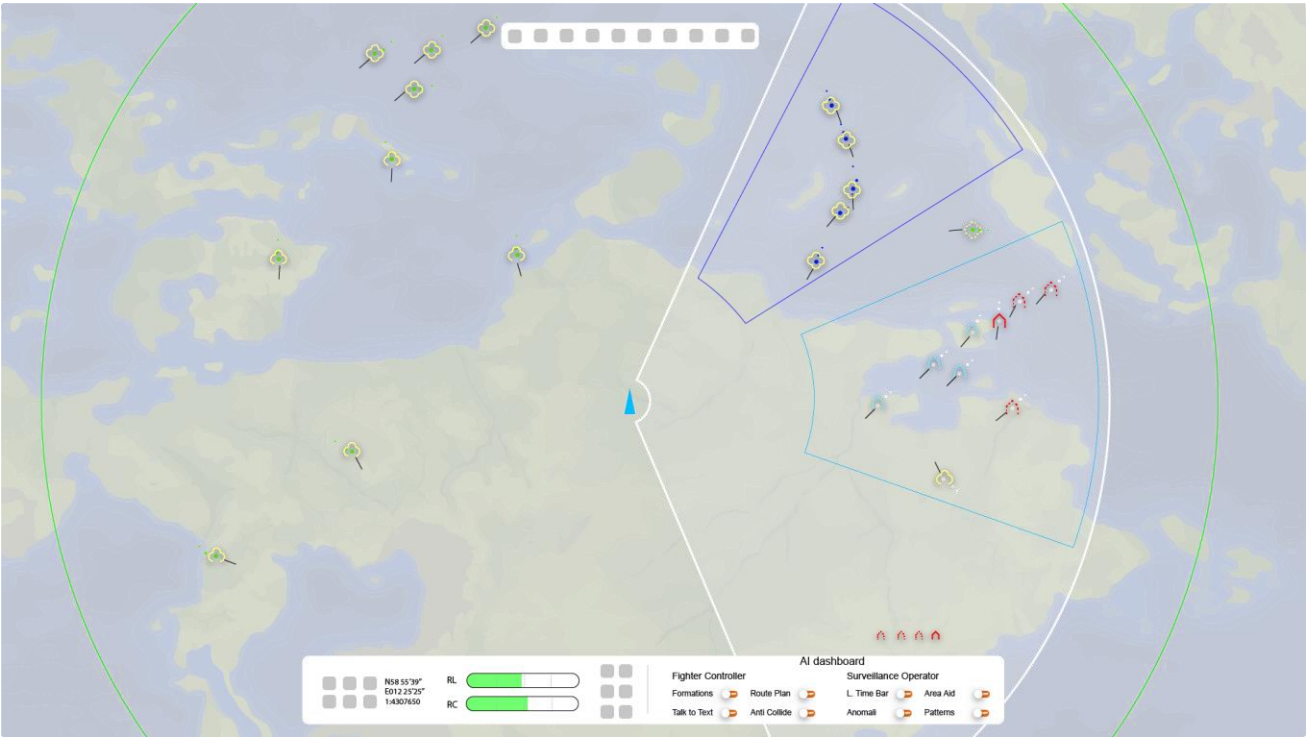






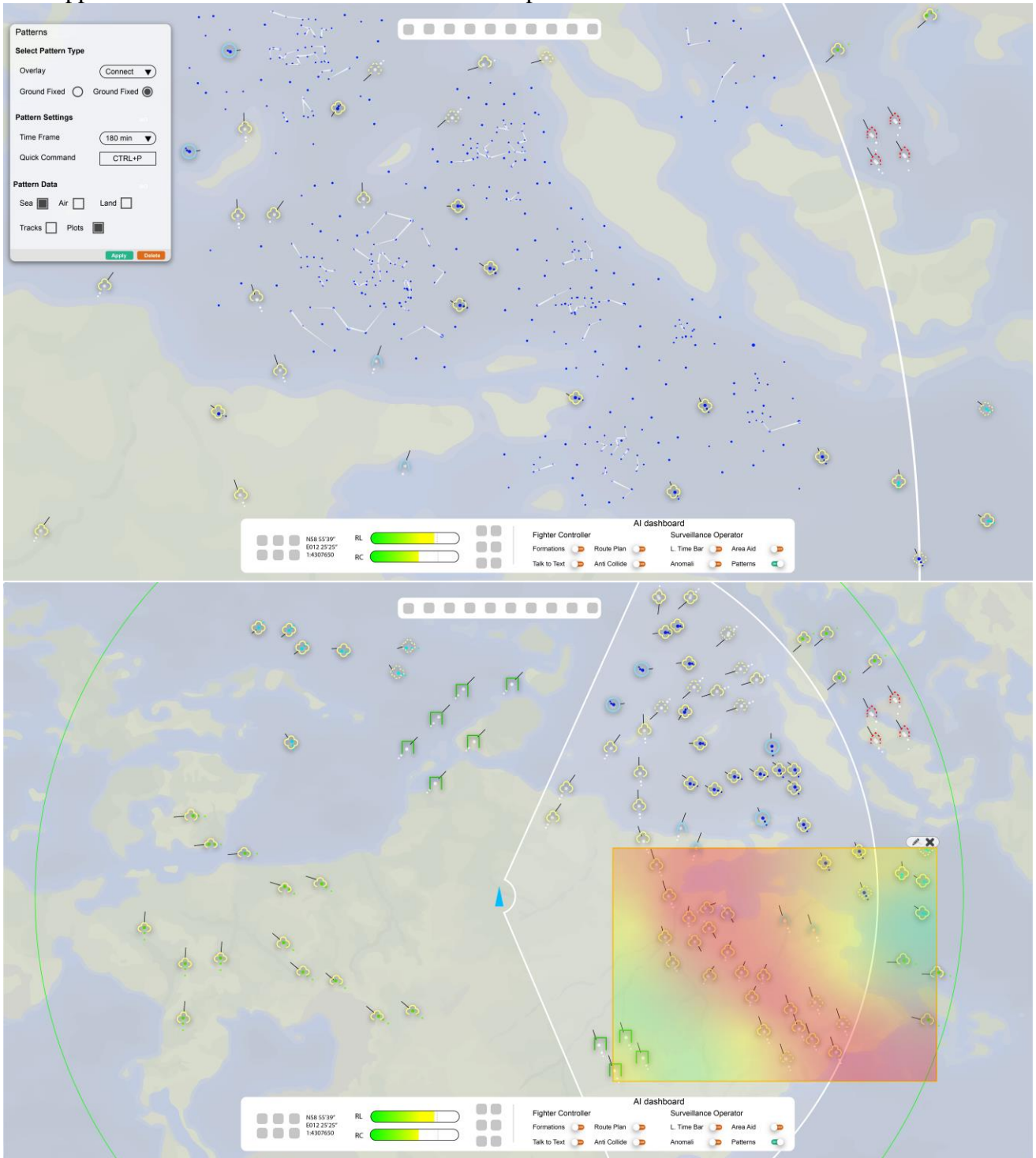


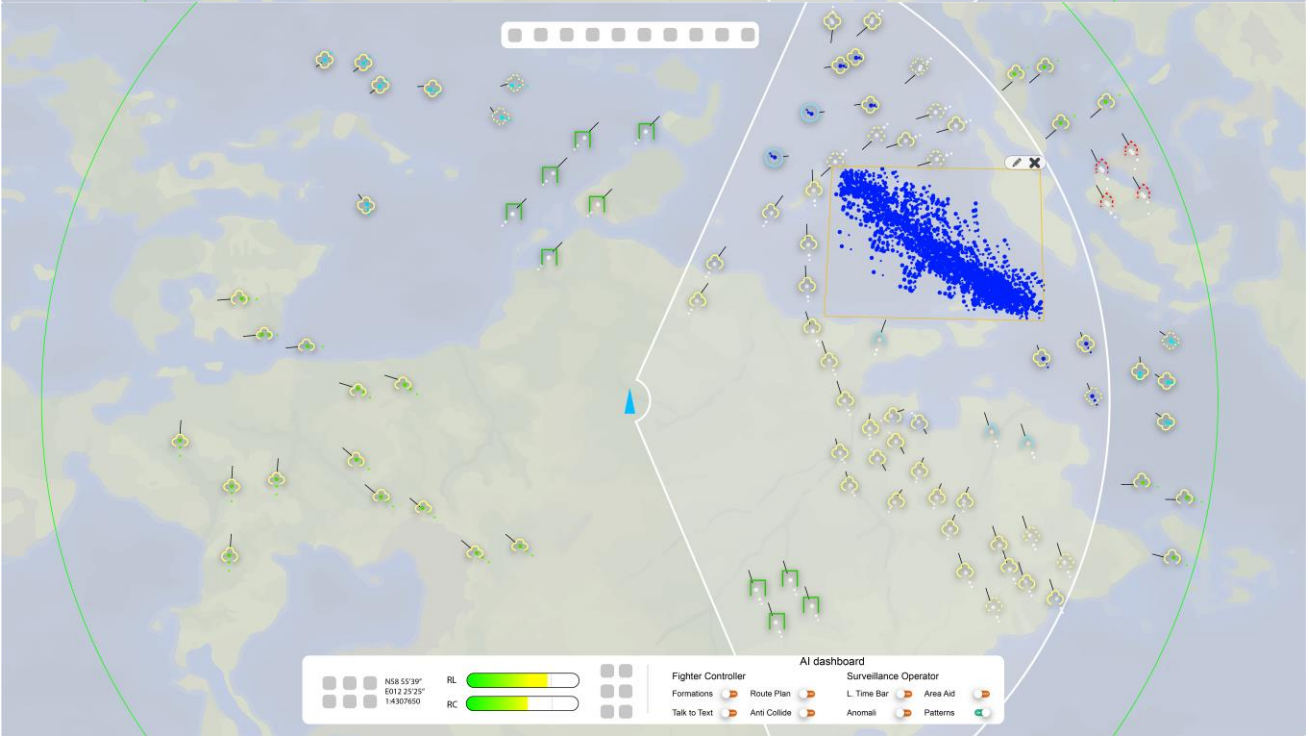
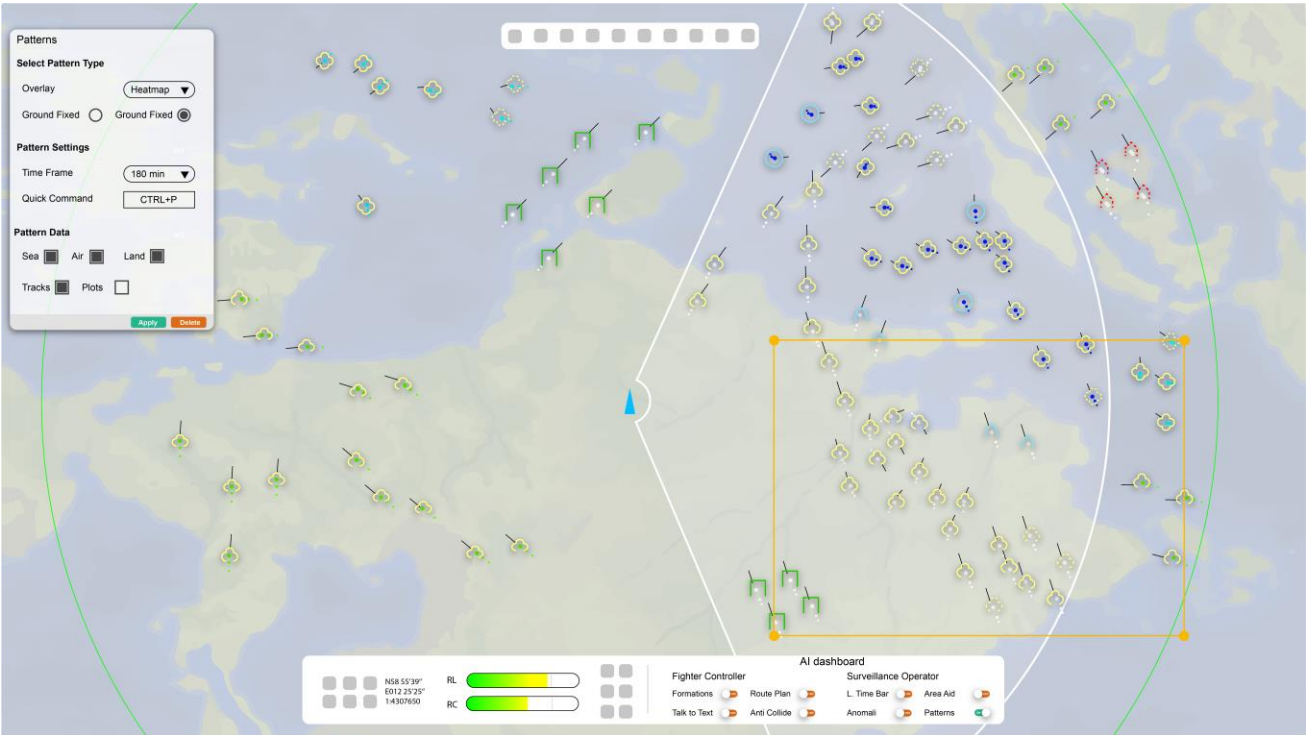




H. Appendix - Final concept: Pattern areas

This appendix includes all frames of the final concept Pattern areas.





Patterns

Select Pattern Type

Overlay: Exposure

Ground Fixed: Ground Fixed

Pattern Settings

Time Frame: 180 min

Quick Command: CTRL+P

Pattern Data

Sea: Air: Land:

Tracks: Plots:

Apply Delete

AI dashboard

| | | | | |
|--|---|--|--|--|
| <p>NS8 55°39' E012 25'25" 14307650</p> <p>RL <input type="range"/></p> <p>RC <input type="range"/></p> | <p>Fighter Controller</p> <p>Formations <input type="checkbox"/></p> <p>Talk to Text <input type="checkbox"/></p> | <p>Surveillance Operator</p> <p>Route Plan <input type="checkbox"/></p> <p>Anti Collide <input type="checkbox"/></p> | <p>L Time Bar <input type="checkbox"/></p> <p>Anomali <input type="checkbox"/></p> | <p>Area Aid <input type="checkbox"/></p> <p>Patterns <input checked="" type="checkbox"/></p> |
|--|---|--|--|--|

Patterns

Select Pattern Type

Overlay: Exposure

Ground Fixed: Ground Fixed

Pattern Settings

Time Frame: 180 min

Quick Command: CTRL+P

Pattern Data

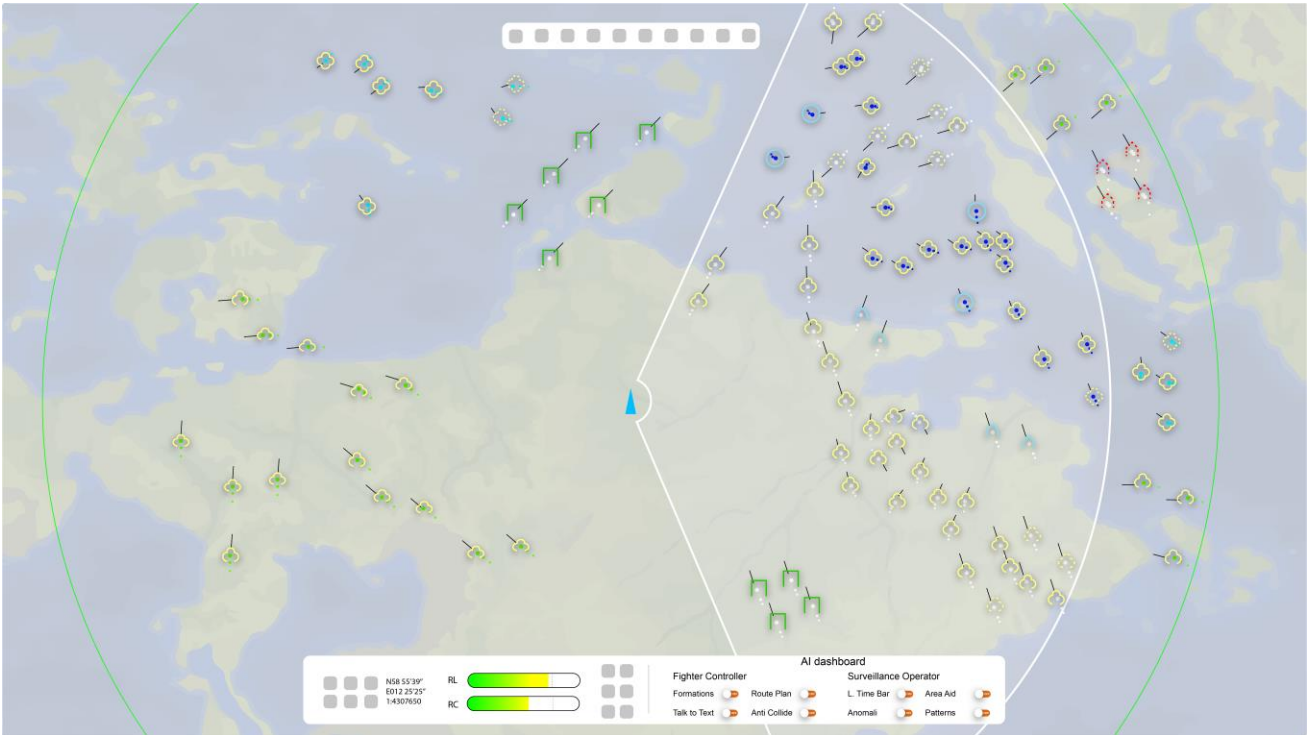
Sea: Air: Land:

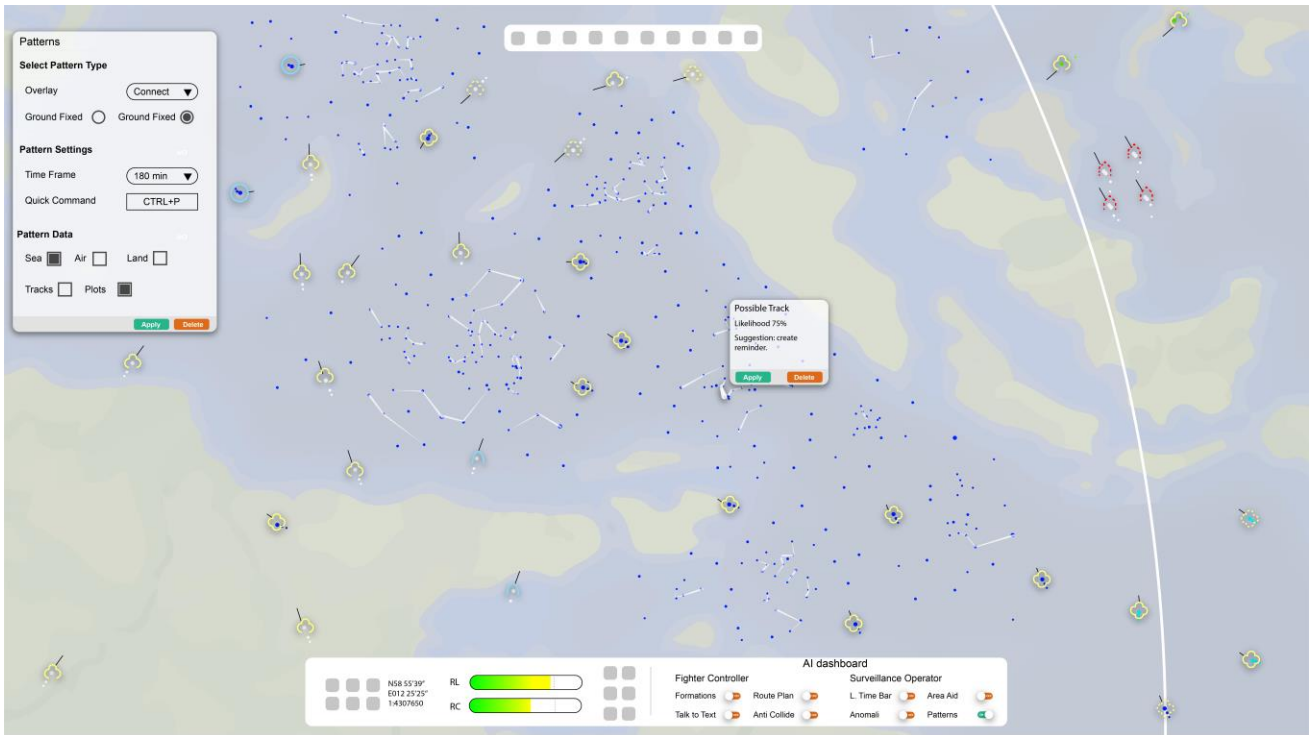
Tracks: Plots:

Apply Delete

AI dashboard

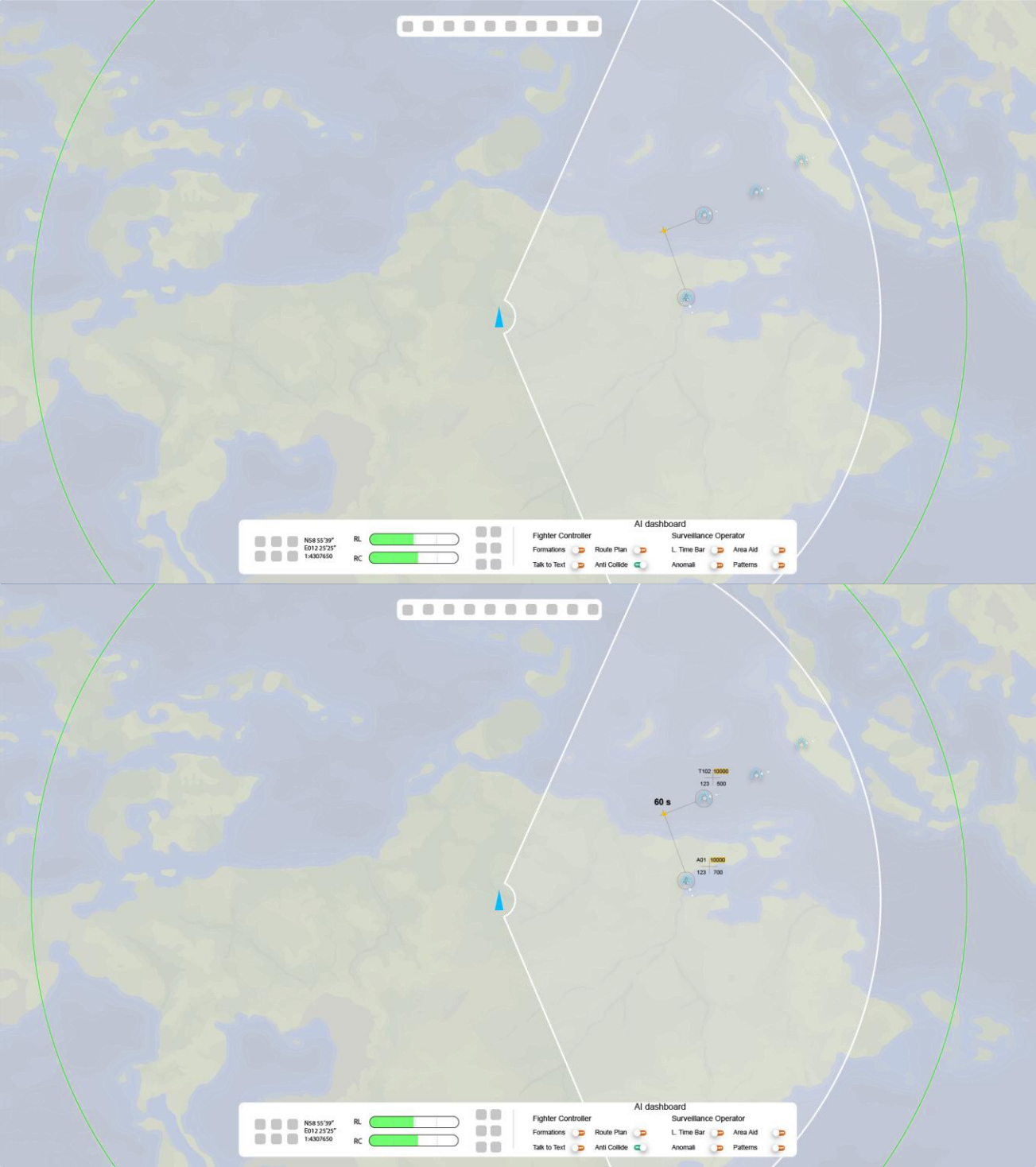
| | | | | |
|--|---|--|--|--|
| <p>NS8 55°39' E012 25'25" 14307650</p> <p>RL <input type="range"/></p> <p>RC <input type="range"/></p> | <p>Fighter Controller</p> <p>Formations <input type="checkbox"/></p> <p>Talk to Text <input type="checkbox"/></p> | <p>Surveillance Operator</p> <p>Route Plan <input type="checkbox"/></p> <p>Anti Collide <input type="checkbox"/></p> | <p>L Time Bar <input type="checkbox"/></p> <p>Anomali <input type="checkbox"/></p> | <p>Area Aid <input type="checkbox"/></p> <p>Patterns <input checked="" type="checkbox"/></p> |
|--|---|--|--|--|

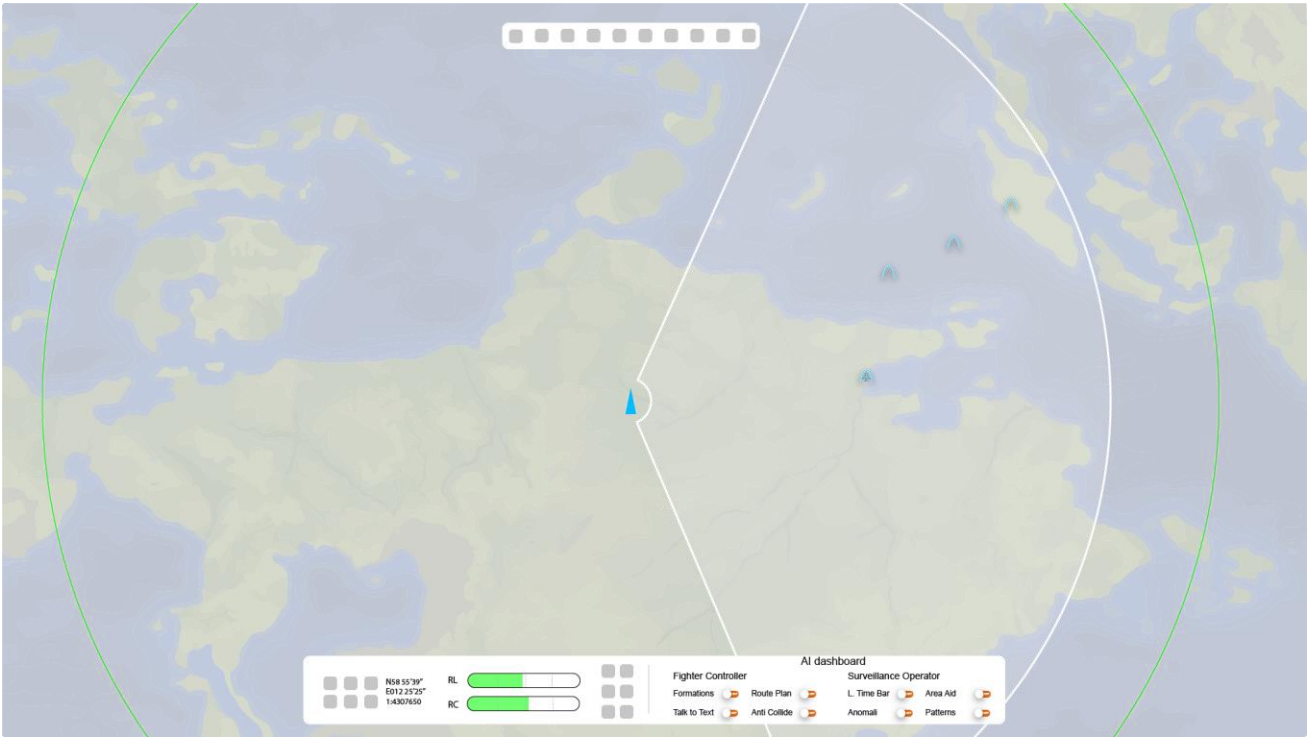




I. Appendix - Final concept: Collision warning

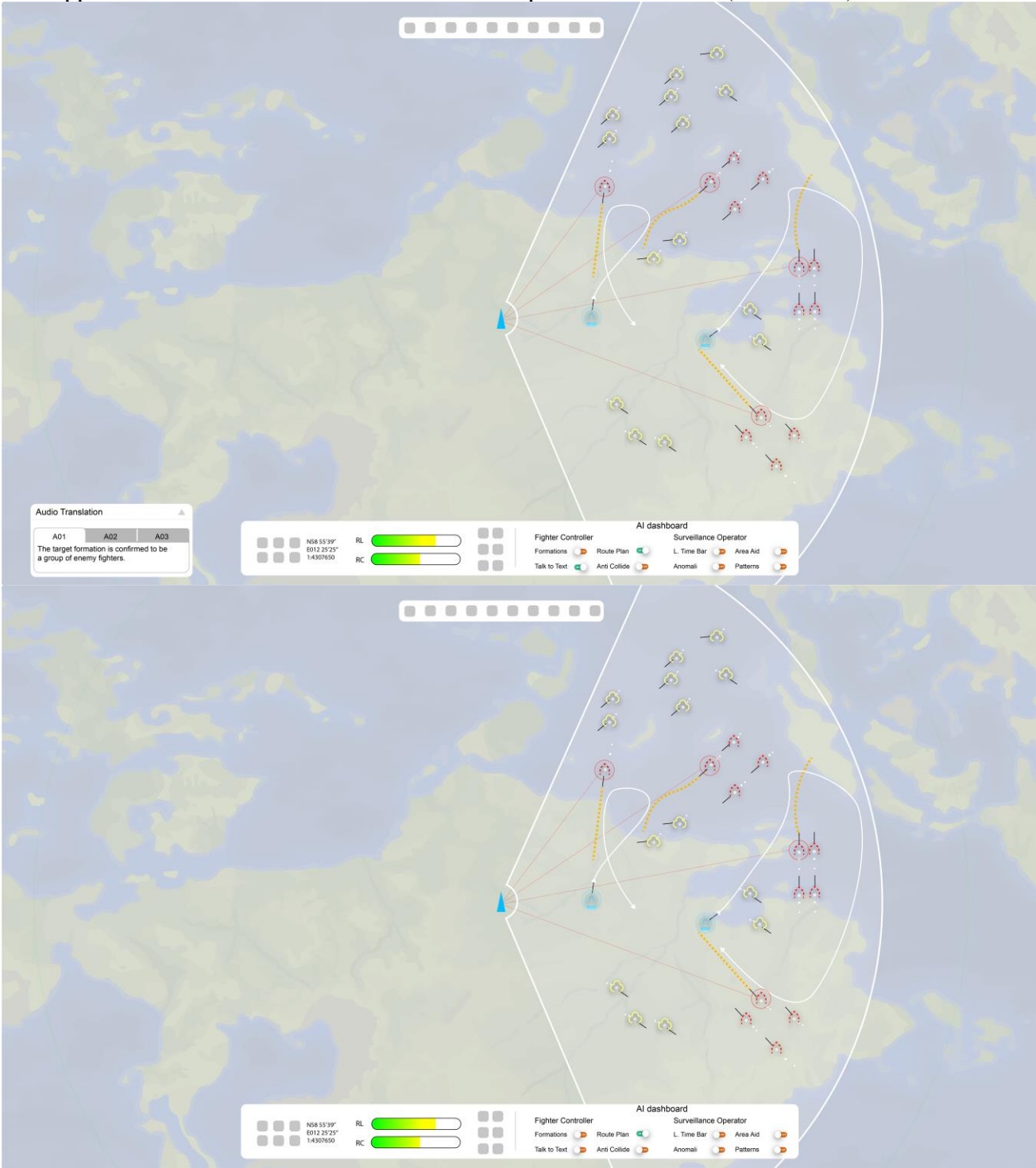
This appendix includes all frames of the final concept Collision warning.

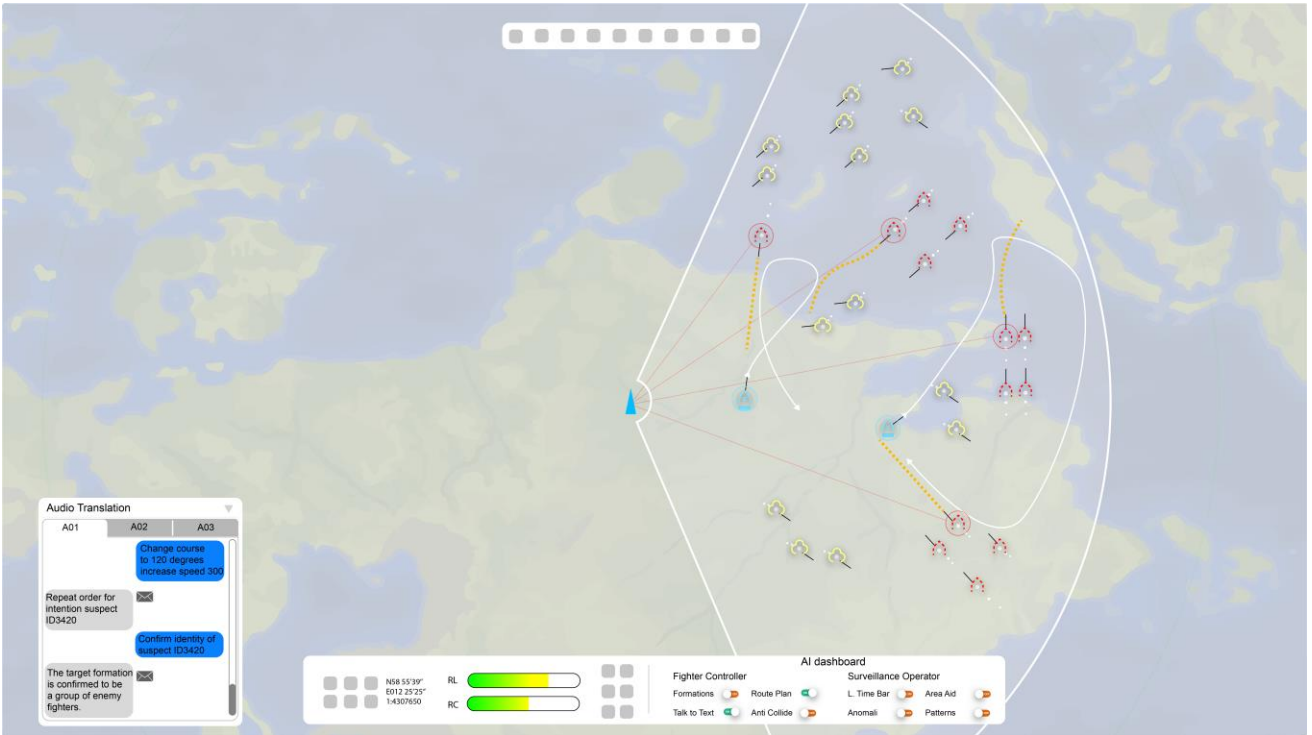




J. Appendix - Final concept: Communication (talk-to-text)

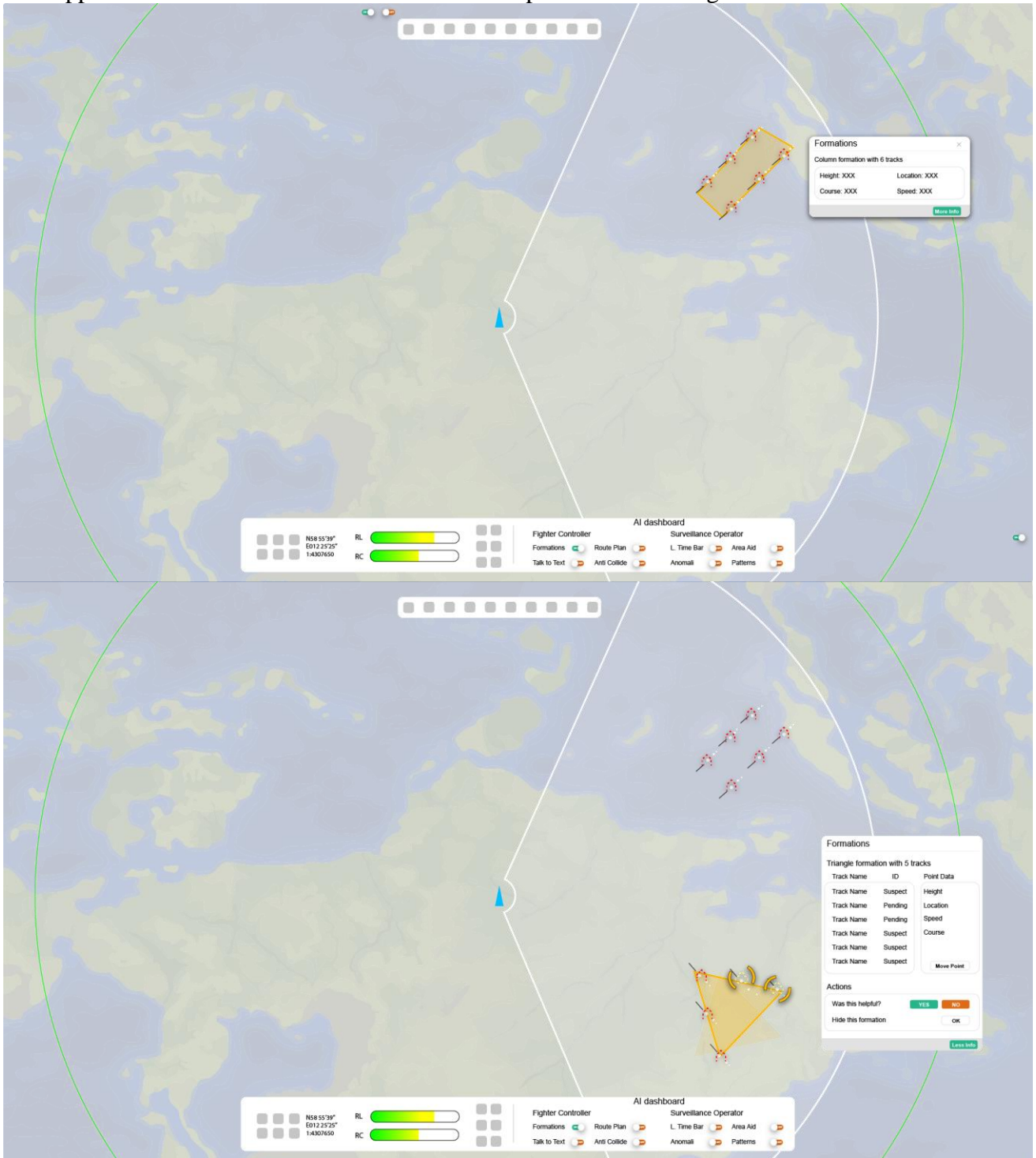
This appendix includes all frames of the final concept Communication (talk-to-text).

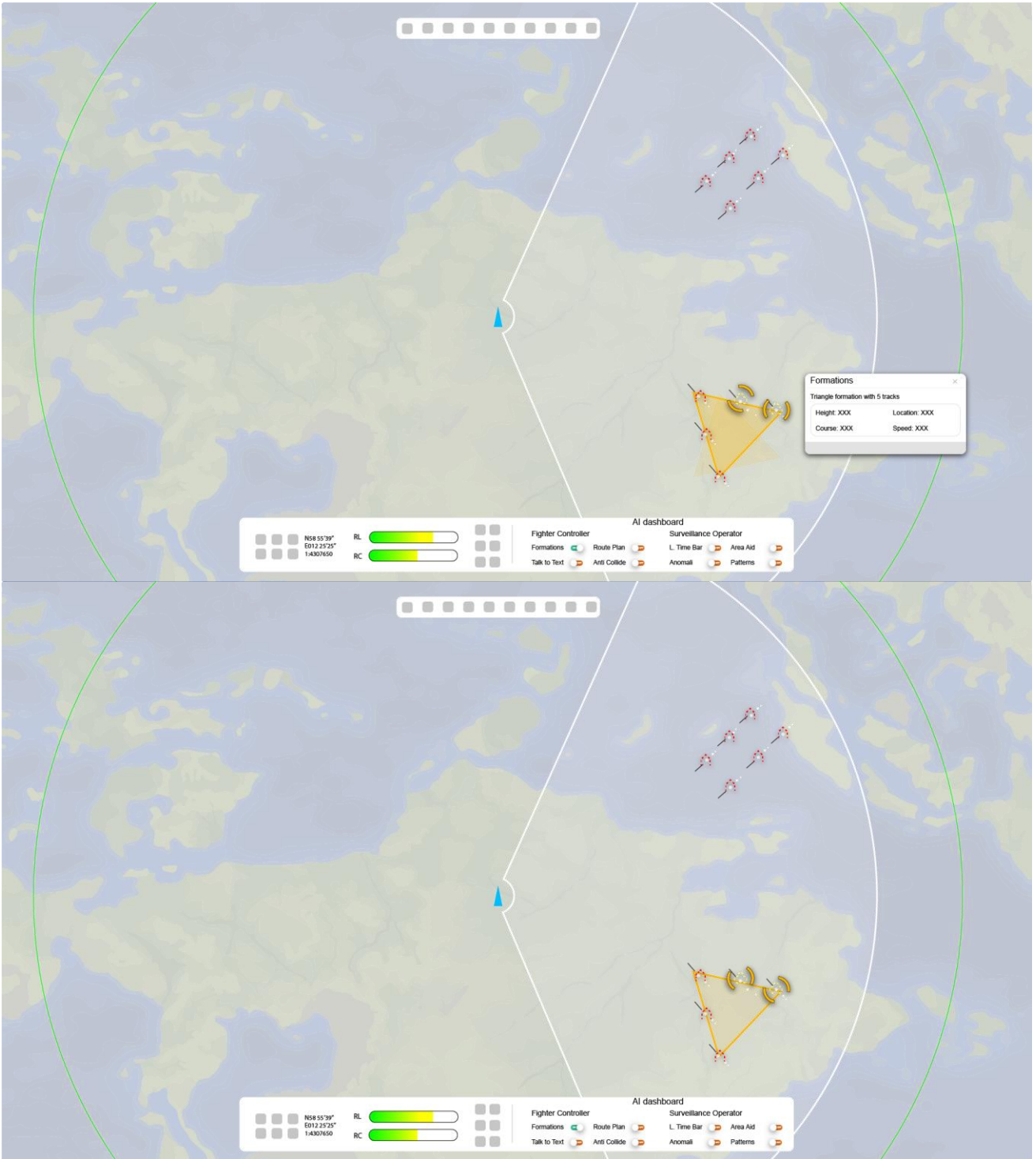


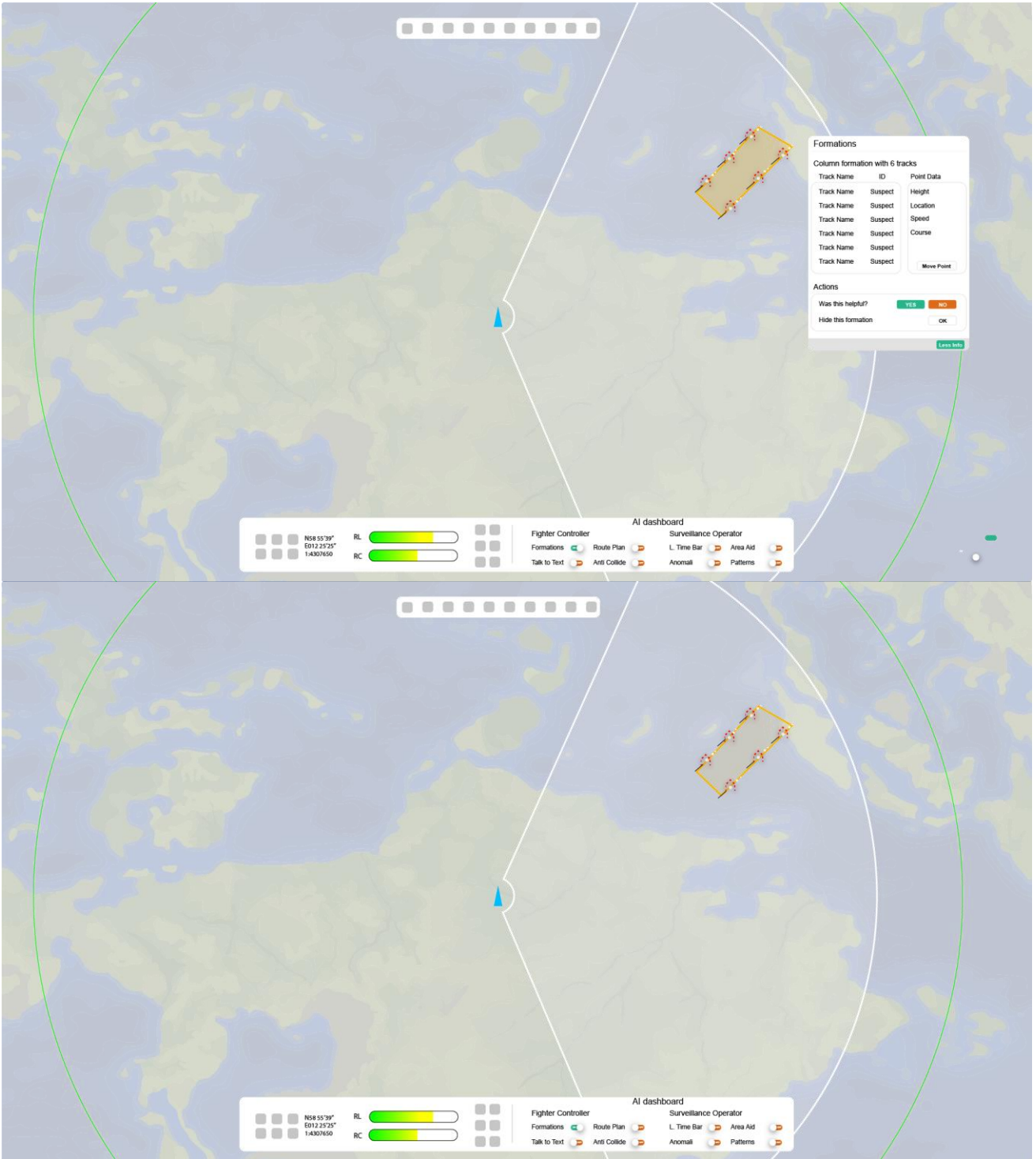


K. Appendix - Final concept: Formation recognition

This appendix includes all frames of the final concept Formation recognition.

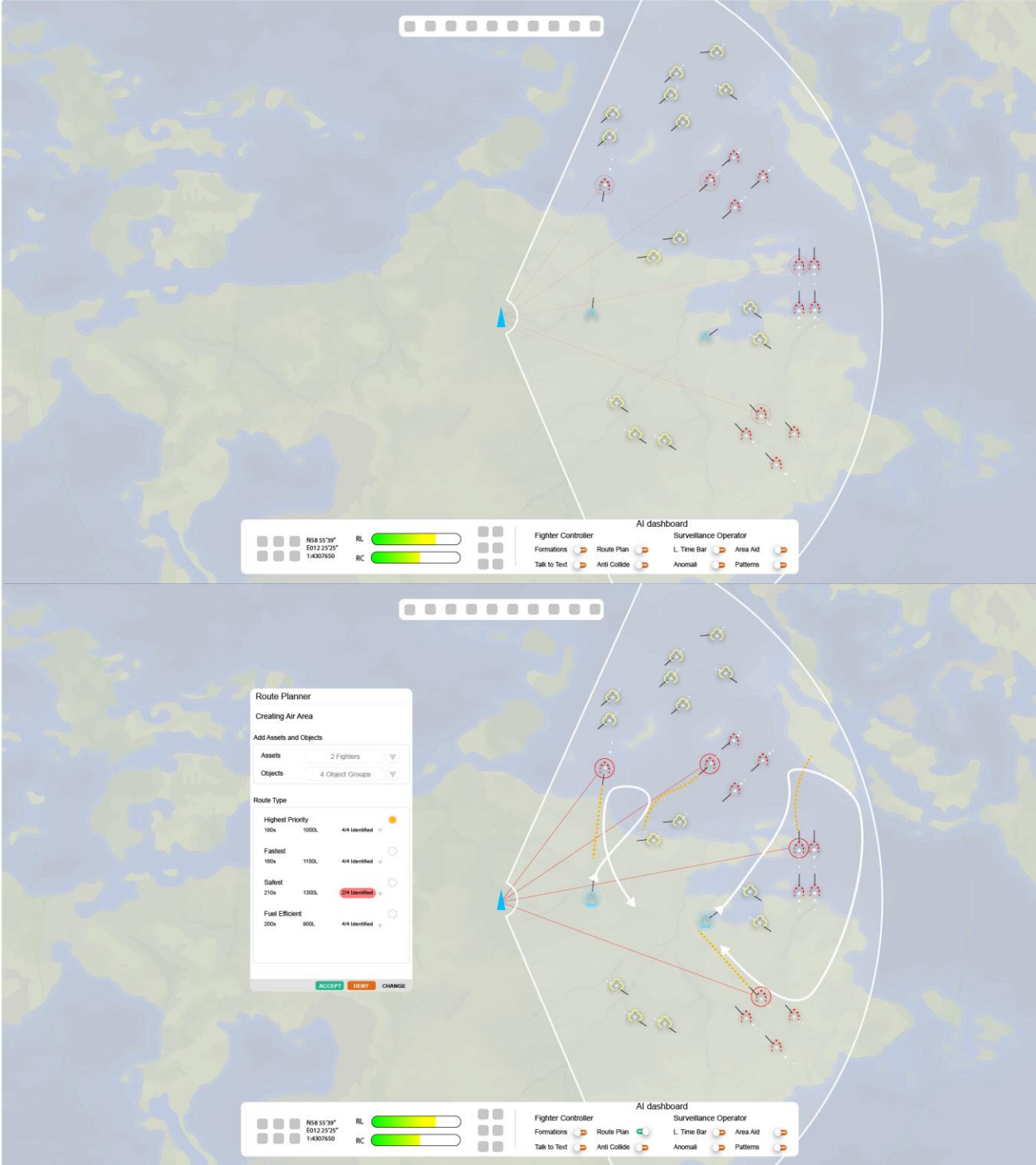


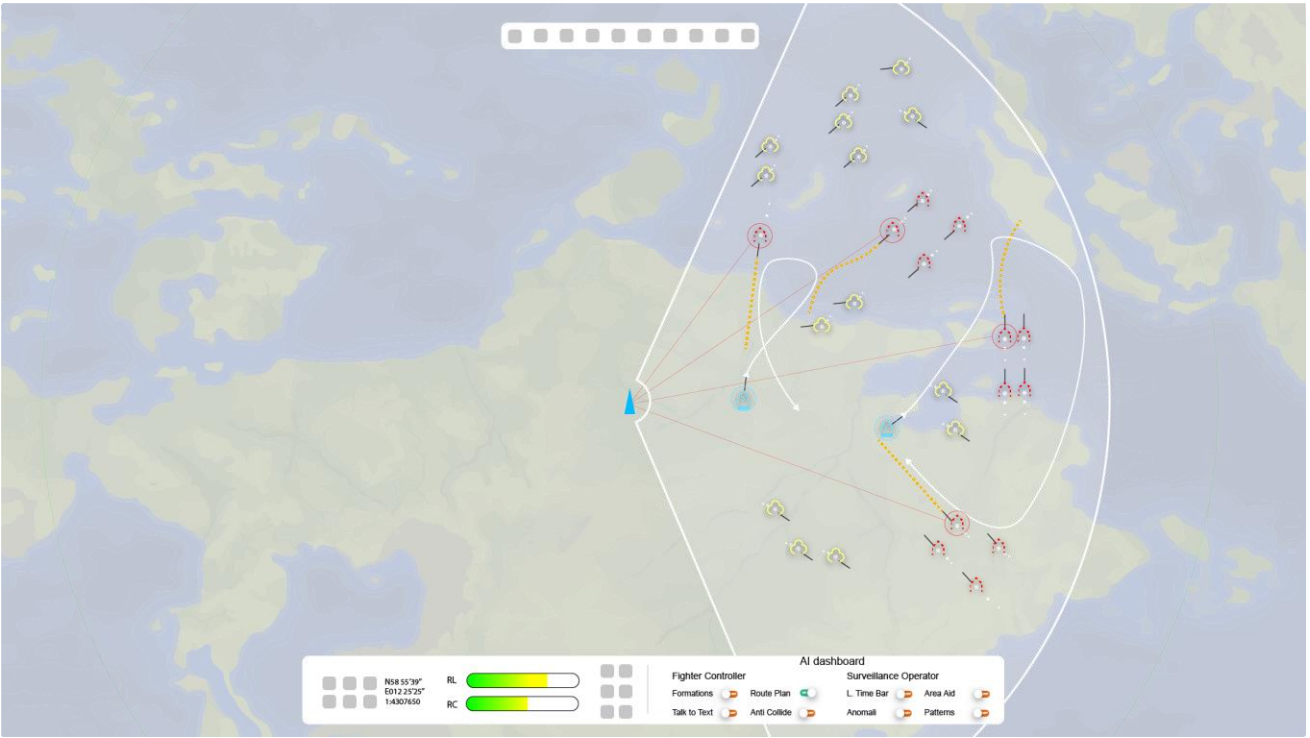




L. Appendix - Final concept: Route planner

This appendix includes all frames of the final concept Route planner.





DEPARTMENT OF INDUSTRIAL AND
MATERIALS SCIENCE
CHALMERS UNIVERSITY OF TECHNOLOGY

Gothenburg, Sweden 2021
www.chalmers.se



CHALMERS
UNIVERSITY OF TECHNOLOGY