IMMENS: Integrated Multi-Manager Environment for Naval Ships

Jonathan Sebastiaan Barnhoorn†* PhD MSc, Marleen Rakhorst-Oudendijk† MSc, Kim Veltman† MSc, Bas Holleman† BSc, Tjalling Haije† MSc, Jelle Wolbers† MSc, Johan Janssen† PhD MSc

† Research scientist at TNO (Dutch Organization for Applied Scientific Research), The Netherlands
* Corresponding Author. Email: jonathan.barnhoorn@tno.nl

Synopsis

The Royal Netherlands Navy is moving towards more autonomous ships to cope with the increasingly complex naval operations while sailing with a reduced manning. The strong expansion of automation and autonomous systems on board requires new solutions for optimal collaboration between the crew and the extensive set of ship systems.

We are developing a concept for an Integrated Multi-Manager Environment for Naval Ships: IMMENS. IMMENS aims to be an intelligent ship system that supports the crews of future ship classes, while enabling these crews to exercise and maintain meaningful human control. The ship's subsystems are integrated in a multi-agent system that formulates and proposes plans to the command team based on an internal representation of the ship's goals and the current situation. These plans consist of actions that are to be executed in order to achieve the mission goals. The system's computational design, combined with specific human-machine teaming functionality facilitates the crew and IMMENS to work as a joint cognitive system.

This paper shortly introduces IMMENS but reports mainly on the development and application of human-machine teaming concepts that support cooperation between the crew and IMMENS. The first concept offers interaction based on a human-machine language using Goals, Tasks, Constraints and Resources as shared knowledge elements. Since both human and machine understand these elements, human comprehension of the machine's inner workings is supported. The second concept provides the explication of relationships between goals, tasks, constraints, resources and/or external world as a means of intuitively explaining the machine's reasoning by providing transparency regarding, for instance, to which goals certain tasks contribute. Thirdly, we applied and extended on the concept of play-based delegation. A play is a plan-template that allows a user to quickly delegate a complex task to the systems, without the need to drill down, while leaving room for system intelligence to fill in the details.

These concepts have been applied in an integrated interactive demonstrator of IMMENS. An evaluation with subject matter experts from the Royal Netherlands Navy and senior human factors researchers was conducted based on a scenario of a frigate escorting a high-value asset. During the design, implementation and evaluation of the demonstrator, we learned that the first concept provided a workable shared knowledge model and indeed fitted with the IMMENS architecture on the one hand, and was intuitive to the users on the other hand. Explicating relationships provided the users with relevant insights, allowing them to understand, for instance, by means of which tasks and resources IMMENS expects to achieve the mission goals. Play-based delegation was found this a clear and valuable concept. Many opportunities for future research were identified, the most important of which concerns the feasibility of constructing a valid, complete and generic goal representation for IMMENS that takes into account the complexity of future missions and operational contexts, and includes the necessary world model.

Keywords: Human-machine Teaming, Explainable AI, Autonomous Systems, Utility-based Reasoning, Multi-Agent System, Transparency

1. Introduction

In the near future, the Royal Netherlands Navy (RNLN) will face specific and growing challenges at sea. Potential threats will increase in size, speed, and variety. The growing number of ship systems will become increasingly complex to manage. The number of tasks will multiply, while reaction times will shorten. At the same time, the number of on-board crew members will likely continue to decline. If the ships of the RNLN are to be truly future-proof, then increased levels of automation and autonomy, combined with new forms of cooperation between the crew and the ship's systems are required. We aim to confront this challenge by developing concepts that support collaboration between pillars of automation within a naval ship (e.g., combat- and platform management systems) and humans with the aim to optimally support the crew in performing their work. On the one hand, this is a technical challenge in which the ship's subsystems need to be connected to each other and act as a multi-agent system (MAS) with a high degree of autonomy according to a common goal. On the other hand, this is a human-machine teaming challenge in which the crew and the resulting system will need to function as a human-machine team, optimally utilizing machine- and human capabilities.

Over recent decades, human-machine teaming research has focused on topics like trust calibration (e.g., Schaefer et al., 2016), mental models of human and machine team members (e.g., Scheutz, DeLoach, Adams, 2017), explainable artificial intelligence (e.g., van der Waa et al., 2021), and much more. A topic that we focused on in this study is that of delegation (e.g., Castelfranchi & Falcone, 1998). With ever more machine autonomy and intelligence we will increasingly need methods for interacting with intelligent autonomous systems in ways that we trust and are familiar with. Delegation is the primary method for instructing systems that are intelligent, capable and effective yet remain subordinate. Research on delegation revolves around the question of how to delegate control to intelligent systems in a way that machine intelligence and autonomy are utilized in an optimal manner, without sacrificing human control about meaningful decisions. One may delegate by providing, for instance, goals, full or partial plans or constraints (Miller, 2005). A concept that has been developed to exert such delegation is the Playbook approach (Miller, 2005), revolving around the concept of plays, or plan-templates that can be used to specify in a broad sense what should happen, leaving room for the machine intelligence to fill in the details. Inspiring interfaces have been developed using the playbook concept (e.g., Calhoun, Ruff, Behymer, Frost, 2018). However, to our knowledge no study to date has focused on delegation to a system that primarily acts as a goal-based autonomous system. Instead, the reasoning about the optimal plan given the goals was always performed by the human. In our research, developing such a goal-based agent (operating in a MAS and utilizing a planning algorithm and a utility function to evaluate candidate plans) and its interaction with a naval crew are the main objectives. We call the system we are developing the Integrated Multi-Manager Environment for Naval Ships (IMMENS). While we approach these challenges in an integral manner and will shortly introduce the MAS in the next section, the main focus of the current paper will be on the human-machine teaming concepts that we developed. We will also illustrate how these concepts can be used in practice and, finally, present results of an intermediate evaluation.

2. IMMENS MULTI-AGENT SYSTEM

A naval vessel can be seen as a system of systems in which each subsystem is responsible for a specialized group of tasks (e.g. anti-air warfare). A subsystem may not be allowed to share all information with the other subsystems, for instance, because of security reasons. A subsystem may have its own properties, knowledge and reasoning capabilities. The individuality of these subsystems is one of the main motivations for our research to adopt the multi-agent paradigm (Wooldridge, 2009), in which each subsystem is treated as an autonomous entity (i.e. agent) with their own beliefs and goals (de Gier, Nijsten, Veltman, de Groot, 2020).

In IMMENS, a layer of agents, which we refer to as *managers*, is placed on top of the already existing subsystems (or software components) to create a Multi-Agent System (MAS). The aim of this layer is to provide a shared interface and to allow communication between the managers to reach a shared goal by planning for certain *capabilities*. In order to realize capabilities, managers need to communicate with other managers to request resources or to obtain information. *Resources* are owned by exactly one manager, but can be requested by other managers in the system (for example: an engagement manager may be able to execute an engagement, but to do that it needs tracking data from a radar manager). To prevent endless negotiations between the managers we introduced a special agent, the *director*, which guides the negotiation process between the managers and makes the final decision regarding which plans to propose to the command team.

In short, the aim of the MAS is to formulate and propose *plans* to the user(s). These plans consist of *actions* that are to be executed by the underlying systems of the managers and/or crew members in order to achieve the ship's current goals, which are defined by the *command aim*. During plan construction, the current command aim and the world state, consisting of the external and internal situation, are considered. Furthermore, certain *constraints* may exist, usually set by the user, which are also taken into account.

The main plan generation loop starts with the human command team defining the command aim. The digital director uses a *utility function* to evaluate the digitally available current world state in terms of the command aim. This function can assign a numerical value to each world state. To do this, low-level information about the world (e.g., originating from sensor values) is aggregated into high-level concepts via so-called utility *indicators*. These indicator values are evaluated using a (sub-)utility function, the value of that function guides the system towards which steps need to be taken because sub-utilities are linked to capabilities. For instance, the indicator crew injuries is linked to the sub-utility Safety of personnel, which is linked to the capability medical evacuation (medevac), allowing the system to predict that a medevac may improve the world state (or: contribute to achieving the command aim) in case the sub-utility Safety of personnel is low. The indicators and sub-utility functions also make the overall utility function more explainable because the overall utility score can be explained in terms of these.

The Director can request one or more capabilities from managers to achieve mission goals. These managers then use recipes to generate *sub-plans* that can achieve the desired capability. Because resources are shared over managers, implementing one sub-plan may affect other (currently running) sub-plans. Thus, the set of sub-plans, which we refer to as the plan, resulting from a director request needs to be evaluated as a whole. The director uses the utility function to do this and is then able to present the result of the planning process to the command team.

3. Human-Machine Teaming concepts

Machines outperform humans in many ways but at the same time, humans have unique knowledge and capabilities compared to machines. Humans bring general intelligence, can carry accountability, are aware of the social atmosphere on board and may have information that is not known to machines. Furthermore, meaningful human control over the vessel is a requirement to make deployment of artificially intelligent autonomous weapon systems acceptable from an ethical, legal, operational and diplomatic perspective (e.g. Roff, Moyes, 2016, van den Broek et al., 2020). This may be high-level control, but control nevertheless. Hence, essential requirements for IMMENS are that the command team can optimally contribute to the planning process through effective plan negotiation, and that humans can make the final decision regarding plans that are initiated. Such decision making needs to be conscious and informed (Scharre, Horowitz, 2015) and this poses requirements to the human-machine interface.

The socio Cognitive Engineering (sCE) method (e.g., Neerincx, 2019) has been used to structure and guide the design process. This method provides a design approach for human-centred designs and offers techniques to underpin the design with theory and empirical research. Here, we describe three concepts that we applied or developed during the design of our interface. We used design patterns, successful solutions for frequently occurring problems, to structure our research and support reusability of obtained insights when possible (Alexander, 1977, van Diggelen & Johnson, 2019, Van Diggelen, Neerincx, Peeters & Schraagen, 2019). Two of the three concepts are suitable for this format due to the specificity of the design problem and solution and are thus described in the form of a design pattern.

3.1. Interaction concept: Goal-based agent delegation with DASH

As mentioned, an essential property of IMMENS is that it includes a goal-based agent (the director). Based on a representation of the command aim and knowledge of the world, this agent proposes plans to achieve those goals. But how should IMMENS interact with the crew? Our aim is to combine human and machine intelligence and let them function as an integrated, hybrid team. However, machines are not capable of conversing about complex problems by means of human language. And humans are not able to process data with the same speed as machines. What are essential elements of IMMENS that the user should and can be made aware of, and how can we present this information to the user? To answer these questions and structure our interface we used DASH (Delegation to Autonomous Systems in Human-machine teaming, van Diggelen et al., 2019), an interaction concept that allows distributed delegation in a work system comprising one or more intelligent machines while maintaining meaningful human control. The main functions of DASH are:

- To provide situation awareness to humans about the current work system
- To allow humans to direct the work system

DASH offers human control and understanding of the machine's reasoning through common perspectives: Goals, Tasks, Constraints & Resources. These perspectives align well with the types of knowledge that IMMENS reasons with: goals are represented by the utility function and its indicators; tasks are represented by (aggregated) actions; and constraints and resources are known by the same name in IMMENS (note that for now, we focus on technical resources). In this way, the interactions between humans and machines are for the most part defined by a shared knowledge model or ontology. This ontology is aligned with both the human mental model of the work domain, as well as with the machine internal workings.

Our interface represents these common perspectives, as is shown in Figure 1. At the left side, the Command Aim, related Mission goal(s), goals relating to the warfare domains and a selection of other goals together represent the utility function IMMENS is trying to optimise and how the ship is currently doing in achieving its goals. We are still working on the structure of the utility function and thus, the goal structure in the interface concept should be regarded as an early concept. At the right side, the currently active constraints and the status of the most important resources of the ship are presented. In the centre, new activities are suggested by IMMENS or can be delegated by the user to IMMENS, in order to optimise the achievement of the goals, and taking into account the applicable constraints and available resources. The tasks that are currently being performed are displayed at the bottom-left of the interface. Note that while the original DASH concept can be used to direct machines as well as humans in the work system, here, we focus on delegating to machines only for now.

 https://doi.org/10.24868/10732

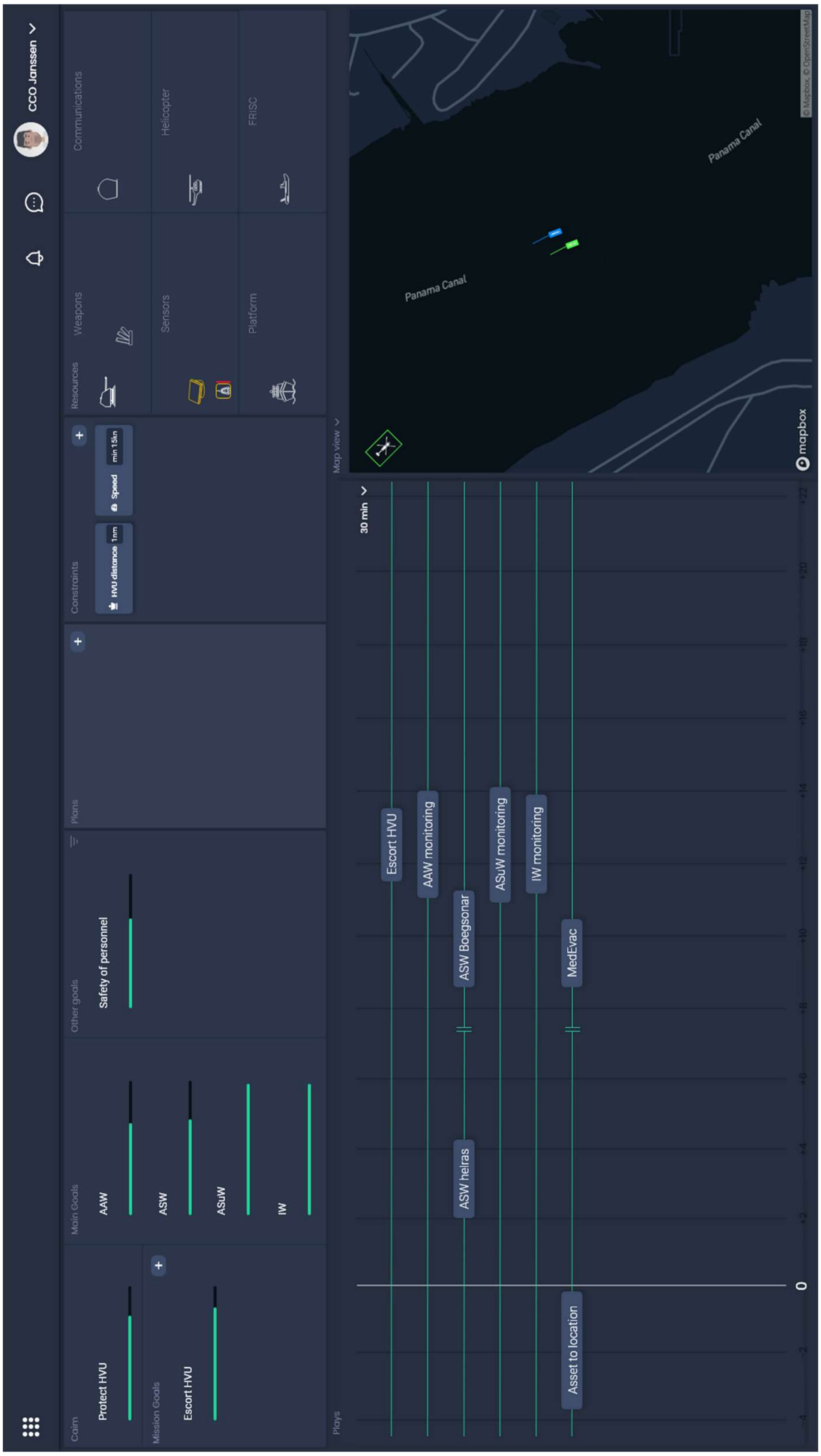

Figure 1: An overview of the interface. Note that all DASH elements are clearly discernible: Goals, Tasks (the plan overview), Constraints and Resources.

3.2. Interaction concept: Explication of relationships

As mentioned, decision making by the command team regarding the proposed plans needs to be conscious and informed. Thus, the command team needs to have insight in the reasoning underlying proposed plans. Our second interaction concept provides the explication of relationships between goals, tasks, constraints, resources and/or the external world, as a means of intuitive explainable AI of the machines' reasoning. Figure 2 shows an impression of this interaction concept. Explication of the relationships between the external world and the internal system status (in terms of goals, tasks, constraints, resources) increases a user's understanding of the reasoning by making visible, for instance, that a task is related to a certain goal according to IMMENS. It also allows understanding of how the internal and external battle relate to each other. For instance, a damaged resource may impact the proposed plan by delaying an important task. With the ability of IMMENS to explicate that this task depends on the damaged resource, the user can understand why the task is delayed. By explicating the system's (lack of) understanding of the relationships, the user's understanding of the system's functioning and the calibration of trust can be increased. Importantly, it would also allow the user to identify and act on missing relationships. Table 1 presents the design pattern of this human-machine teaming concept.

Table 1 Design Pattern ‘Explicating relationships between goals, tasks, resources, constraints and the external world’

Title	**Explicating relationships between goals, tasks, resources, constraints and the external world**
Design problem	Reasoning AI systems often consist of elements that represent a goal, tasks that can be performed to achieve this goal, resources that can be put to use to perform the tasks, and constraints that delineate the range of behaviours that is allowed. Such systems are in line with the DASH concept. However, how can an interface explain to a user what the reasoning underlying the system’s current or planned behaviour is? Simply presenting all afore mentioned elements to a user provides little insight, but generating human-like explanations (e.g. verbal) of behaviour is often too complex.
Design rationale	The pattern follows the DASH concept of using goals, tasks, constraints and resources as shared knowledge model aligning with both the human mental model and the machine’s inner workings. Given that humans have an intuitive understanding of how goals, tasks, resources and constraints relate to each other, simply highlighting relationships provides a basic understanding of the machine’s reasoning.
Design solution	When the user focuses on an element (a goal, task, resource, constraint or an element of the machine’s world model), the machine highlights which other elements are, according to the system, related. For instance, focusing on a goal may highlight the tasks that contribute to that goal, the resources used to perform the tasks, and so on. Focusing on a resource may highlight its location in the external world, constraints it is adhering to, and so on. Given that humans have an intuitive understanding of how goals, tasks, resources and so on relate to each other, highlighting relations as understood by the machine provides basic insight in its reasoning. It also allows the human to spot missing relationships. For instance, the machine may not be aware that a certain resource is needed for a task.
Example	See Figure 2 and Figure 3.
Positive effect	• Provides transparency of the machine’s internal functioning • Provides a basic understanding of the machine’s reasoning • Allows the human to identify whether relationships are missing

 https://doi.org/10.24868/10732

	• Allows the human to calibrate trust based on understanding of the machine's reasoning process
Negative effect	• The nature of the relationships are not explicated, therefore, the user may draw inaccurate conclusions. E.g., is a constraint related to a goal because it's essential for achieving it, or because it limits performance on that goal? • In cases where very few goals, tasks and resources are present, highlighting relationships may offer little value

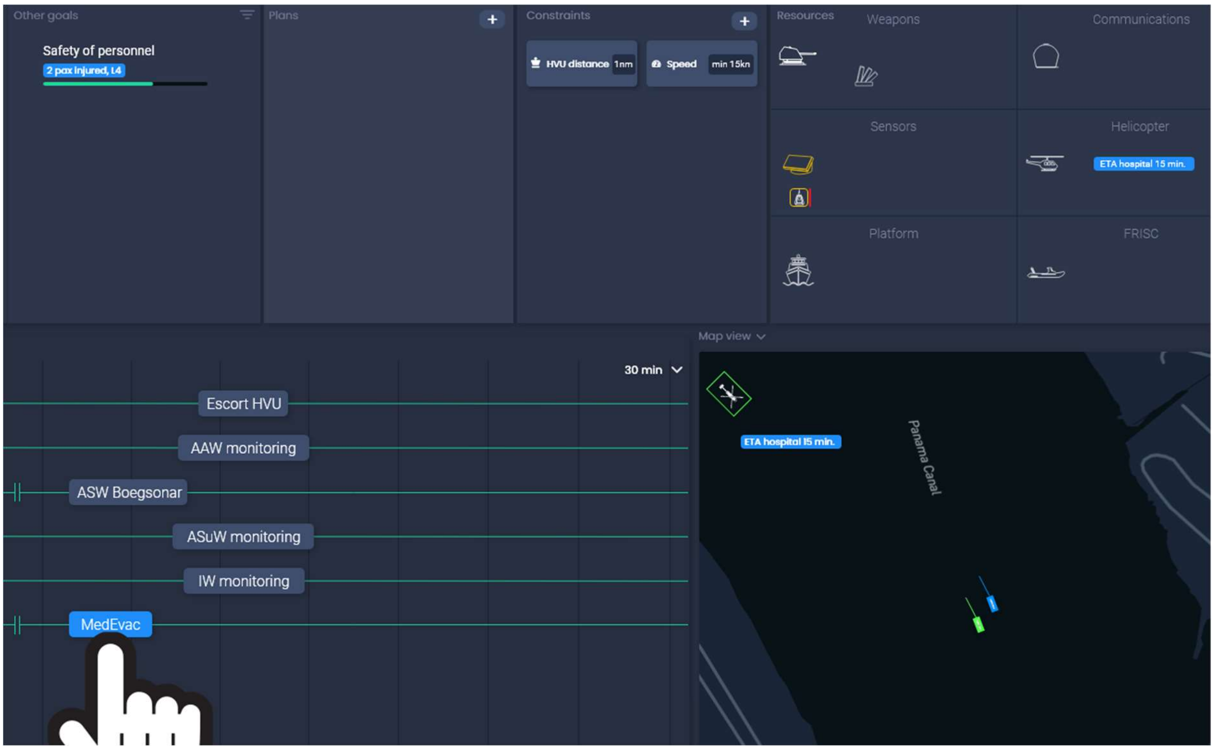

Figure 2: A member of the command team hovers over the 'Medevac' play. IMMENS explicates which relations exist between goals, tasks, constraints, resources and the external world. Namely, a medevac (task) is performed by the helicopter (resource) which has a position in the map (external world) in order to improve performance on the (sub)goal Safety of personnel.

3.3. Interaction concept: Play-based delegation

The third interaction concept is play-based delegation, with which we aim to enable effective, robust and adaptive plan negotiation and to support an appropriate human projection of the machines future behaviour. A play is a plan-template that allows the user to quickly delegate a complex task to the systems, without the need to drill down. Figure 3 shows a proposed play based on Tasks, Constraints and Resources. This play has already been selected (clicked on) by the user and the parameters that can be set and the settings IMMENS proposes for these parameters are displayed. In this way, the user has the ability to exert detailed control when desired, but can also choose to rely on the suggestions proposed by the system. Because plays have been used during training and operation preparation, the whole crew knows what to expect from the system when a certain play is initiated.

Table 2 presents the design pattern of the play concept. With the current work, we aim to further develop this concept by describing it in the form of a design pattern and presenting an instantiation and evaluation in the context of IMMENS.

Table 2 Design Pattern of the Plays concept.

Title	**Play-based delegation**
Design problem	When giving instructions to intelligent systems, the total set of possible commands and their implications is often too large for a human to fully understand and provide rapid instructions under time pressure. At the same time, access to the right commands supports human control and provides the flexibility needed to be effective in complex, dynamic environments. How can a user rapidly, but with sufficient accuracy, predictability and specificity, communicate elaborate sets of instructions to an intelligent system? How can the system be given more or less autonomy in different situations?
Design rationale	It is likely that, for quite some time, humans will need to be able to exert substantial control over intelligent systems. Plays provide users with control over highly autonomous systems by activating familiar behaviour(s) when needed (Miller, 2005). Although plays can leave it up to a system's intelligence to fill in the details, users are still able to drill down on specific play instructions when needed. During system design, it is difficult to envision all future situations and manners in which an intelligent system will be used. Allowing specification of desired behaviour before operations makes it possible to use recent, mission- and operation-specific knowledge to prepare behaviour sets that are optimal for the situations that will likely be encountered. These sets may, for instance, be adjusted between operations regarding the amount of autonomy the user is willing to give to the system.
Design solution	Plays are plan-templates that allow for fast delegation to intelligent systems without the need to drill down. They can comprise goals, tasks, resources, constraints and additional information like communication work agreements. Some plays only set an abstract goal (e.g., 'surveil area X'), others may specify detailed task sequences performed by specific resources within certain constraints. Plays offer a way to easily respond to situations that were foreseen during system development or mission preparation and are based on analysis of system capabilities and the work to be performed, using cognitive task analysis for example. During training the user and system cooperate on the basis of plays, such that a user can intuit a system's behaviour based on a selected play. If needed, plays can be adjusted before initialization.
Example	See Figure 3.
Positive effect	• Delegating work to a system while leaving room for it to fill in the details follows natural human behaviour (i.e. humans are familiar with delegation in daily tasks) • Provides the ability to rapidly respond to situations by instructing potentially complex behaviour sets quickly • Reduces need for micromanagement • Plays can be used as a common language during operation to allow all humans involved to understand the directions a system has received • Ability to optimize the behaviour a play describes, or the plays described, before a mission or operation • Offers control while simultaneously using a system's intelligence

Negative effect	• When a fitting play is not available in a situation, a user will need to fall back to more detailed directions • Determining when and when not to drill-down can be difficult • Requires extensive alignment and training to become familiar with plays • As more tools are made available to the user to improve control, the cognitive effort required to manage and switch attention among them increases • Instructions in the form of plays may be less flexible for adjustments at a later time

4. IMMENS demonstrator and example scenario

To validate the ideas underlying IMMENS, an integrated software demonstrator was developed incorporating both the MAS (as a back-end) and the user interface (front-end). As mentioned, the emphasis of our research is on finding the right structure for the MAS, utility function and human-machine interface. Therefore, we focused on supporting one specific scenario and a relatively low level of detail in terms of, for instance, the world model and implementation of functional managers. The back-end is a Python application running a custom MAS framework that communicates via a Representational State Transfer (REST) interface with the web-based front-end, which uses the Flask web framework, HTML, CSS and JavaScript.

The example scenario used to build this demonstrator involves a frigate escorting a High-Value Unit (HVU), e.g. an oil tanker, through a narrow strait. An air threat arises in this area in the form of fighter jets, with the frigate having the task of protecting the HVU and itself. A specific snapshot of the scenario was used to study human control, focussing on the crew's command team, in combination with the automatic planning process. In this snapshot, the vessel is currently moving through the narrow strait. The aftermath of a recent drone attack on the vessel is still being managed. At start of the scenario, medevac has already been requested for three wounded crew members; the helicopter is inbound and a few minutes away from the vessel. As the above is ongoing, two fighter jets are detected.

The snapshot was used to describe, in line with the sCE approach, the human-machine collaboration in this particular situation in three use cases: (A) 'Command team monitors overall situation' (see Figure 1 and Figure 2), (B) 'Assess threats and potentially adapt plan' (see Figure 3), and (C) 'Decide on play' (see Figure 3). Each use case consists of several components: first, an action sequence describes *how* the human-machine collaboration takes place in a specific situation or at a specific moment. Each step in the sequence comprises a *design solution* implementing the interaction. The functional requirements describe *what* the machine shall do to serve the objectives in the corresponding use cases. And, additionally, the claims specify the expected effects of the (situated) functional requirement to provide the justification (*why*). This approach structured the design process but also the evaluation.

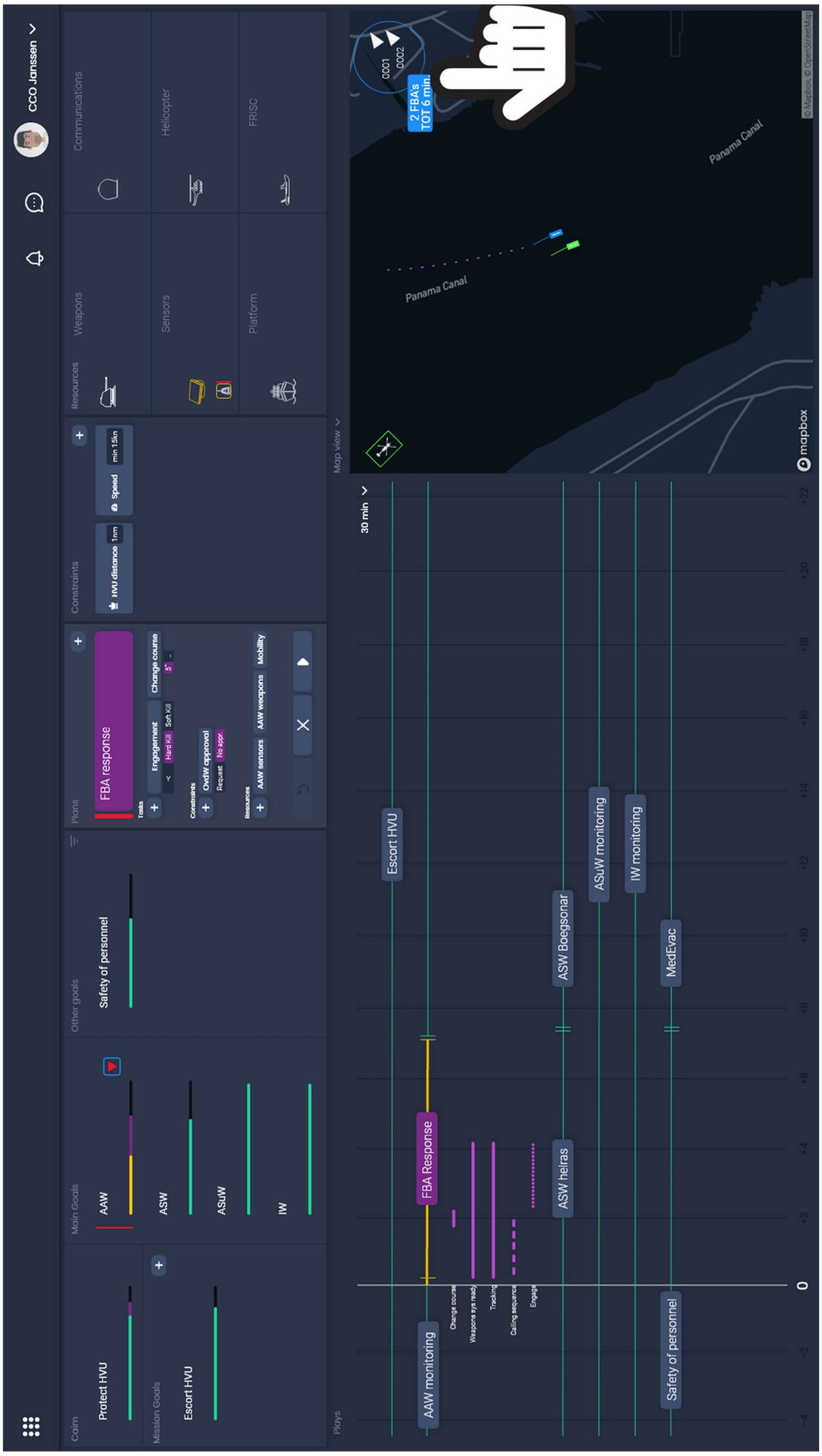

Figure 3: IMMENS has proposed performing the play ‘FBA response’ in response to the 2 fighters approaching with a Time On Top (TOP) of 6 minutes (the triangles denote that relevant contacts, the fighters, exists outside the map and in which direction). The effect of the event on the goal AAW is explicated. In this screenshot, the command team clicked ‘FBA response’ thereby entering a ‘simulation mode’ (purple means ‘simulated’ in which the new play is integrated in the current plan and the effect of initiating this play on the goal AAW is expressed.

5. Evaluation

Two activities were conducted to evaluate IMMENS's current status. A workshop with RNLN subject matter experts (SMEs) was conducted to evaluate operational value. A workshop with human factors experts was conducted to evaluate our results from a (scientific) human factors standpoint.

In the workshop from operational perspective, three experienced SMEs, from the Operations Department and from the Mechanical/Weapon Engineering Department joined. First, the SMEs discussed and explained how the scenario goals would be achieved when the current way of working has to be applied. Then, a walkthrough of the demonstrator was presented, after which questions were asked regarding the way the demonstrator contributes to the previously discussed current way of working. By explicitly discussing the current way of working first, impact of the demonstrator could be clearly discerned. In line with the sCE methodology, we presented the questions in the form of claims. Each claim consisted of a function that IMMENS should offer, a design solution that implemented the function, and an effect that the implemented function was expected to have. The group discussed advantages and disadvantages of the design solution and gave a score, based on consensus, to express whether the stated effect was achieved.

In the workshop with human factors experts, a human factors evaluation method that was developed in-house was used. This evaluation followed an 'outside-in' approach: the demonstrator was compared with a broad set of elements that compose a hybrid multi-team system (Schraagen, Barnhoorn, van Schendel, van Vught, 2021) with a focus on scientific value, taking the design patterns in account too.

In all, a broad set of results were collected by the evaluation activities. Note that extensively describing the evaluation methods and results is beyond the scope of the current paper. Here, we summarize the feedback concerning the three HMT concepts we proposed.

5.1. Goal-based agent delegation with DASH

The interface was designed according to the DASH concept, with goals, resources, constraints and tasks as main elements. The presentation of the goals was found visually calm, however, participants found the goals difficult to interpret and longed for more information regarding, among others, goal priority and weight. They mentioned that it is important to distinguish between performance (scores) on goals that reflect the current situation and scores that reflect the future or predicted situation. Furthermore, changes in goal scores were found to be difficult to interpret, especially given that different quantities are compared. As one participant said "How do I interpret negative or positive changes in the 'bars' compared to each other. For instance, with a certain plan performance on ASW may decrease strongly but Safety of Personnel may increase a little."

The resources were found to be presented in a good way, but scalability was mentioned as a point of attention because of the large number of systems present on naval ships.

The constraint functionality was deemed insufficient. Participants mentioned that such functionality should indeed be present but that it is unclear in the current implementation whether a) constraints are being met, b) what each constraints' value means, and c) whether the constraints are static or dynamic.

The presentation of tasks (presented as plays constituting the current plan) was found good and participants appreciated the possibility to drill-down on the plan to retrieve more detailed information. Currently, there is no (digital) representation of the 'plan' on a tactical level in the command centre, instead, the commander knows the whole plan and each crew member knows only a bit. Having a representation would be valuable.

5.2. Explication of relationships

As an addition to the DASH concept, we included functionality to show relationships between goals, tasks, resources, constraints and elements of the external situation. The design solution, highlighting these relationships, was found clear and informative. However, participants did not agree with the claim that this actually gives the user insight into the reasoning of IMMENS. Here, the meaning of the term reasoning was the subject of the discussion. Future work could benefit from a more detailed disentangling of which aspects of IMMENS should be provided insight into, and in which way.

5.3. Play-based delegation

The interface uses plays as a central interaction and planning concept. Participants found this a clear and valuable concept, however, it should be clearer how the plan and plays related to each other. The ability to simulate the effects of running a play was also appreciated. When applying plays, a proper balance needs to be found between allowing human control and applying machine intelligence. Furthermore, a limitation of the use of plays may be that this approach is less suitable for pro-active behaviour. For instance, tactical navigation in Anti-Submarine Warfare, in which the opponent is continuously challenged and responded to, may be difficult to capture in a play. Another challenge may be the transition of control when a suitable play is not available and a person needs to improvise. Finally, it was advised to address the temporal dimension of the play more clearly, i.e., that control is partly exerted during design and configuration of IMMENS.

6. Conclusions & Recommendations

We believe that increased levels of automation and autonomy, combined with new forms of cooperation between the crew and the ship's systems, are necessary to make the RNLN's ships future-proof. Therefore, in the research program 'Smart Ship', we are exploring how a command team can be supported by a system, IMMENS, that can assist in the planning process by proposing and negotiating about plans that achieve the command aim. In this paper, we presented three interaction concepts with which we aim to support human-machine teaming between the crew's command team and IMMENS. Here, we reflect on our approach and current results and discuss potential avenues for future work.

The general set-up of the interface, based on DASH and the use of play-based delegation, was found to be workable and intuitive. Numerous smaller issues were identified, concerning for instance the scalability of the resources overview, potential overlap between plays and plans and immaturity of the current implementation of working with constraints.

During the design, implementation and evaluation of our demonstrator we learned that the DASH concept provided a workable shared knowledge model and indeed fitted with the MAS architecture on the one hand, and was intuitive to the users on the other hand. We also learned that next to goals, tasks, resources and constraints, the external world is an indispensable part of providing transparency regarding IMMENS. Note that the 'internal world' is captured by the resources. Future work on DASH could consider giving the external world an explicit place in the concept.

The most pressing issue identified in the evaluation regards the feasibility of working with a utility function that represents goals at the highest level of a naval ship. As stated during one of the evaluations, "A fundamental question that remains is whether a valid, complete, generic utility model can be constructed for the complexity, diversity and dynamics of (near-)future missions and operational contexts […] including the necessary world model". To answer this question, technical challenges lie ahead but importantly, detailed analyses of the work context, tasks, and decision processes should also be performed since it is the combination of the work domain and technical capabilities that together determines the feasibility of working with utility functions in the way we envision. Besides constructing such a model, it is vital to find better ways to provide the command team with insight into and control over this function. The command team likely needs to be able to exert this control before but also during the operation as operation goals may change with changing circumstances.

Another general challenge that was identified during evaluation is that there is a risk that the plan as understood (and potentially, executed) by humans and as represented in IMMENS, deviates. For instance, when an officer gives an assignment via voice to a fellow crew member to perform a certain task, this is not known to IMMENS. Such a deviation between the human's mental model and the system's plan representation can quickly become problematic, especially in time-pressing situations. This stresses the need to find ways for the crew and IMMENS to interact as seamless and intuitively as possible.

A final challenge regards the integration of IMMENS in the navy's ways of working. Even with greatly improved automation, a crew will still exist of multiple distributed teams, each with their own responsibilities, knowledge and control over their own processes and resources. Execution of any high-level plan therefore relies on many crew members working across the ship, meaning that IMMENS and the entire crew have many mutual dependencies. How can intuitive and effective cooperation between crew and systems be achieved in the way of working across the ship?

In the current work we attempted to look ahead, exploring what human-machine collaboration could look like in the future in order to inform us about design choices we get to make on a short notice. Hence, many questions remain. Among other things, we currently plan to further investigate working with the utility function when the command aim changes and how to improve human-machine collaboration in dynamic circumstances where quick adjustments to the plan may be necessary.

Acknowledgements

The work described in this paper was funded by the Dutch Ministry of Defense. The authors wish to acknowledge in particular Agnieta Habben Jansen and Maarten Hartemink. The authors would like to thank many colleagues from TNO, among whom Zita, Martijn, Martijn, Chris, Jan, Jurriaan, Jasper, Mark, and many others.

References

Alexander C.: “A pattern language: towns, buildings, construction”, Oxford university press, 1977.

Calhoun, G. L., Ruff, H. A., Behymer, K. J., & Frost, E. M.: “Human-autonomy teaming interface design considerations for multi-unmanned vehicle control" Theoretical issues in ergonomics science, 2018.

Castelfranchi, C. & Falcone, R.: "Towards a theory of delegation for agent-based systems", Robotics and Autonomous systems, 1998.

van Diggelen, J., Barnhoorn, J., Post, R., Sijs, J., van der Stap, N., & van der Waa, J. "Delegation in Human-machine Teaming: Progress, Challenges and Prospects" Intelligent Human Systems Integration, 2021.

van Diggelen, J., & Johnson, M.: “Team design patterns”. Proceedings of the 7th International Conference on Human-Agent Interaction (pp. 118-126), September, 2019.

Van Diggelen, J., Neerincx, M. A., Peeters, M., & Schraagen, J. M. C.: “Developing Effective and Resilient Human-Agent Teamwork Using Team Design Patterns”, IEEE Intelligent Systems, 34(2), 15–24, 2019.

de Gier, J. M., Nijsten, L. M. J., Veltman, H. J., & de Groot, T. H.: “Autonomous Planning for Surface Engagement using a Multi-Agent System”, International Naval Engineering Conference & Exhibition (INEC), 2020.

Neerincx, M. A., Van Vught, W., Henkemans, O. B., Oleari, E., Broekens, J., Peters, R. … & Bierman, B.: “Socio-cognitive engineering of a robotic partner for child’s diabetes self-management”, Frontiers in Robotics and AI 6, November, 2019.

Miller, C. A., Funk, H., Wu, P., Goldman, R., Meisner, J., & Chapman, M.: “The playbook approach to adaptive automation”, Proceedings of the human factors and ergonomics society annual meeting (pp. 15–19), SAGE Publications, 2005.

Roff, H. M., & Moyes, R.: “Meaningful Human Control, Artificial Intelligence and Autonomous Weapons”, UN Convention on Certain Conventional Weapons, 2016.

Schaefer, K. E., Chen, J. Y. C., Szalma, J. L. & Hancock, P. A.: “A meta-analysis of factors influencing the development of trust in automation: Implications for understanding autonomy in future systems”, Human Factors, 58, 377–400, 2016.

Scharre, P., & Horowitz, M. C.: “Meaningful human control in weapon systems: A primer”, Working Paper, 2–16, 2015.

Scheutz, M., DeLoach, S. A., & Adams, J. A.: “A framework for developing and using shared mental models in human-agent teams”, Journal of Cognitive Engineering and Decision Making, 11(3), 203–224, 2017.

Schraagen, J. M., Barnhoorn, J. S., van Schendel, J., & van Vught, W. (2021). Supporting teamwork in hybrid multi-team systems. Theoretical Issues in Ergonomics Science, 1-22.

van den Broek, J., Griffioen, J. R. & van der Drift, M.: “Meaningful Human Control in Autonomous Shipping: An Overview”, IOP Conference Series: Materials Science and Engineering, Vol. 929, No.1, IOP Publishing, 2020.

van der Waa, J., Verdult, S., van den Bosch, K., van Diggelen, J., Haije, T., van der Stigchel, B., & Cocu, I.: “Moral Decision Making in Human-Agent Teams: Human Control and the Role of Explanations”, Frontiers in Robotics and AI, 8, 2021.

Wooldridge, M.: “An introduction to multiagent systems”, John Wiley & Sons, 2009.

 https://doi.org/10.24868/10732